

# Morphology and Language Acquisition

Constantine Lignos      Charles Yang  
University of Pennsylvania

**Keywords:** language acquisition, corpus linguistics, computational linguistics, frequency effects, productivity, analogy, English past tense

## 1 Introduction

Morphology is language specific and must be acquired on the basis of particular linguistic data. This chapter addresses three central issues for the understanding of morphological acquisition: the distributional properties of morphology, the mechanisms of learning from data, and the developmental patterns and trajectories in child language. First, corpus studies of child directed language reveal a systematic pattern of data sparsity: in general, only a (small) fraction of possible morphological forms will be attested in the acquisition data. The child learner, then, must be able to form wide ranging generalizations from early on, and on very partial corners of the morphological paradigm. Second, we review evidence from morphological development, especially the study of English past tense, and draw two broad conclusions: (a) morphologically complex forms appear not memorized by association but are learned in relation to their morpho-phonological classes, lexicalized or otherwise, and (b) as far as child language is concerned, morphological productivity appears categorical, where children often over-use productive processes but almost never commit analogical errors. Third, we situate morphological learning models in the context of language acquisition, where psychological considerations pose additional challenges and constraints. The child must proceed in an incremental and unsupervised fashion, which deprives them of assumptions (such as stem and derived pairs) that are taken for granted in theoretical analysis. Along the way, we discuss how language acquisition may provide useful insight for the theoretical study of morphology.

## 2 Distributions

One of the greatest challenges that a child learning morphology faces is the sparsity of morphological combination in the target language. We might imagine in the ideal case that a child would be presented with full paradigms when attempting to learn the morphology of a language, much like in a foreign-language course. For example, imagine that a learner

trying to learn a verbal system with four different inflectional categories (1-4) is exposed to four verbs (*A-D*). In optimal conditions, they would be exposed to every inflectional variant of every verb, seeing sixteen forms in total: *A1-4, B1-4, C1-4, D1-4*.

Morphological learning is relatively simple in this case. If the child were to receive every form for every lemma—a *dense* input where no evidence is absent—and is also told what features each form carries—*supervised* learning—rote memorization would result in perfect performance and any errors might be attributed to failures of memory or inference of incorrect generalizations. However, as we discuss in this section, neither of these assumptions are met and the process of morphological learning proceeds in an *unsupervised* fashion and on *sparse* input.

Because the input is sparse and distributed in a *Zipfian* fashion, a more realistic version of the learner’s input would contain a sparse set of unique forms—for example, only *A1, A3, A4, B1, B2, C1*, and *D2*, with all other combinations unseen. Additionally, children are not provided with pairs (e.g., *sing/sang, walk/walked*) that specify the morphological relations; while they must be able to eventually infer these pairs from the input, we cannot assume that they can do so perfectly and before they have identified the phonological relationship between items of the pair. Indeed, as we review in Section 4, the vast majority of the proposed models in the study of morphological acquisition provide paired input to the learner, which is tasked with identifying the mapping between these forms—by no means a trivial problem in its own right. The more formidable and less-studied problem is that in many cases of morphological learning, the learner will not observe anything close to a complete paradigm as discussed above: many entries are missing and related forms are not grouped together in the input.

## 2.1 Morphology in a Zipfian world

It turns out that the statistical properties of morphology are quite similar to the statistical properties of words (Zipf 1949); see Chapter 25 for background on corpus research. According to Zipf’s Law, relatively few words are used frequently—*very* frequently—while most words occur rarely, with many occurring only once in even large samples of texts. For instance, in the Brown Corpus (Kuçera and Francis 1967), about 43% of words occur only once, about 58% of words occur 1-2 times, 68% of words occur 1-3 times, etc. The frequencies of combinatorially formed units tail off even more precipitously. For instance, approximately 78% of Brown Corpus bigrams appear only once and 90% of them appear only once or twice. For trigrams, the singletons and the doubles make up 91% and 96% of the unique trigrams.

Morphological word formation, which can be viewed as the composition of morphological primitives such as morphemes,<sup>1</sup> show very similar statistical properties. In Figure 1, we present the distribution of verbal inflectional morphology in a 985,262 token Spanish corpus of child-directed speech assembled from CHILDES Spanish (MacWhinney 2000).

---

<sup>1</sup>We use terms such as morpheme, paradigm, lemma, and rule informally and descriptively to express the combinatorial nature of morphology, without making any specific commitment to their theoretical status.

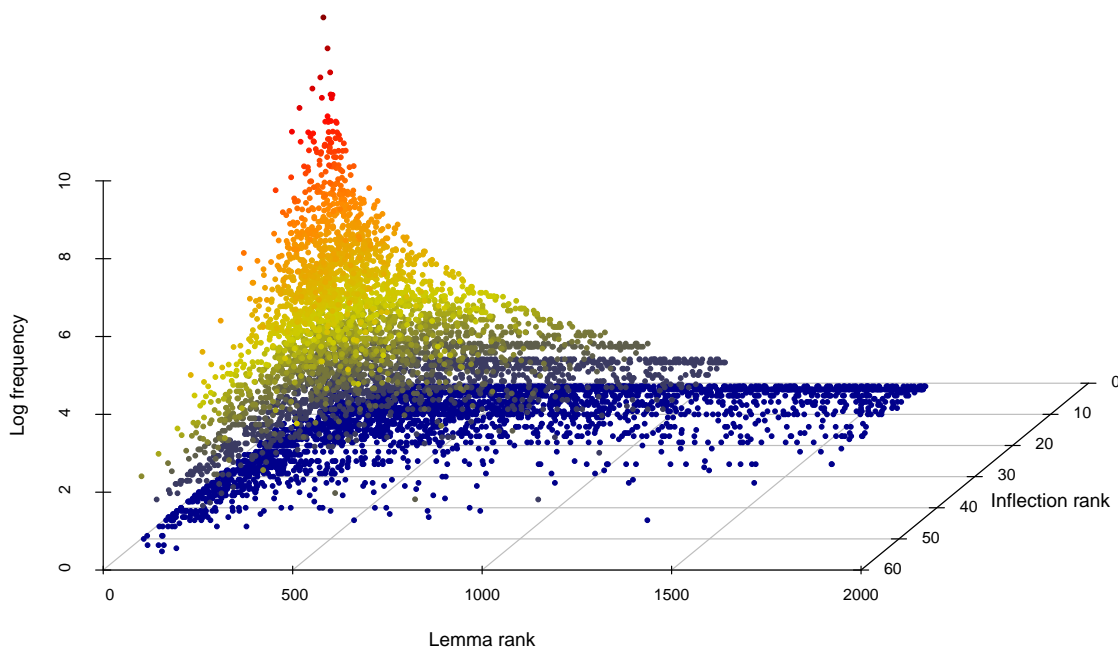


Figure 1: Frequencies of CHILDES Spanish lemmas across inflection categories.

The Spanish data has been analyzed using a part of speech tagger (Freeling; Atserias et al. 2006) which is especially suitable for the processing of Romance languages. There are 1,584 verbal lemmas, which altogether appear in 54 inflectional categories, the most frequent being the third person singular present indicative.<sup>2</sup> Lemmas are shown in order of increasing rank across the X axis, inflections show in order of decreasing rank across the Y axis, and  $\log_{10}$  frequency of an individual form is represented by color and the Z axis.

First, consider the pattern visible across lemmas. As described by Zipf, the most frequent lemmas are greatly more frequent than all others; the most frequent lemma *ver* is almost twice as frequent as the second-most frequent lemma *mirar*, and 521 lemmas—32.9% of all observed lemmas—only appear once. The ten most frequent lemmas account for 42.1% of all occurrences of verbs in the corpus.

A similar pattern is visible across inflectional categories; some appear with almost every lemma, but more appear with only a few lemmas. The most common inflectional category (third person singular present indicative) appears 37,573 times, while two inflectional categories appear only once each: the first and second person imperfect subjunctive. Thus we come to identify another Zipfian pattern; in addition to Zipf’s law applying across lemmas, it also applies across inflectional categories. Looking across the inflections within each lemma, we observe that while the most frequent lemmas appear in many inflectional cate-

<sup>2</sup>The copular and light verbs *estar*, *haber*, *hacer*, *ir*, *ser*, and *tener* are excluded from these counts and from Figure 1. An inflectional category is considered a unique combination of person, number, tense, mood, and aspect. Only active voice verbs are included.

gories, the least frequent lemmas appear in a few scattered inflectional categories. Visually, if the combination of lemmas and inflectional categories was dense and not sparse, there would be no trend in the Y axis and few gaps in the plot.

## 2.2 Quantifying sparsity

As a means of formalizing this sparsity, we can report the metric of inflectional *saturation* (Chan 2008). Saturation is computed per-lemma by dividing the total number of inflectional categories observed in the corpus by the number of inflectional categories observed within that lemma. A lemma with 100% saturation is seen in every inflectional category that appears in the corpus; it provides a complete paradigm for that verb. However, any verb with a lower saturation is missing forms, providing an incomplete paradigm full of accidental gaps.

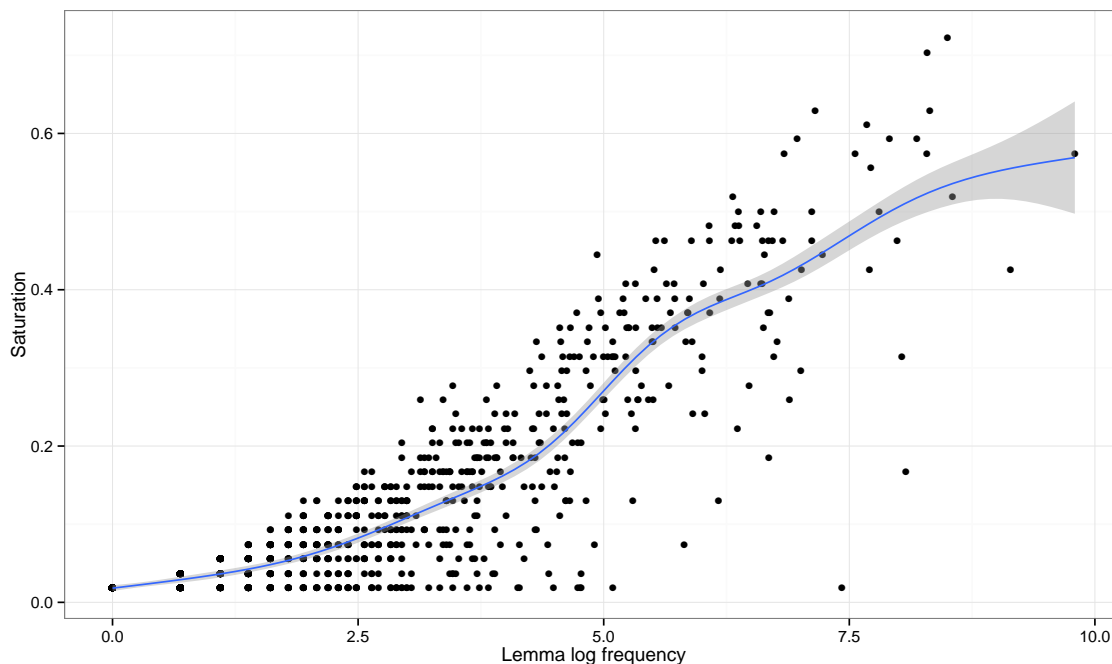


Figure 2: Saturation of CHILDES Spanish lemmas across lemma frequencies, with a GAM-derived fit line and standard error estimate.

The naïve expectation is that many verbs will show a high level of saturation, providing useful examples for a learner to use to infer the best candidates for the many unobserved lemma-inflectional category combinations. However, as shown in Figure 2, in the CHILDES Spanish data under examination, even the most frequent lemmas show relatively low saturation. The most saturated lemma is *decir*, with a saturation of 72.2%. The mean saturation across all lemmas is just 7.9%; thus the average verb appears in about one of

Corpus	Tokens (millions)	Infl. cat- egories	Max. infl. categories per lemma	Max. sat- uration
Brown Corpus	1.2	6	6	100.0
Wall Street Journal Corpus	1.3	6	6	100.0
Basque	0.6	22	16	72.7
Czech	2.0	72	41	56.9
Finnish	2.1	365	147	40.3
Greek	2.8	83	45	54.2
Hungarian	1.2	76	48	63.2
Hebrew	2.5	33	23	69.7
Slovene	2.4	32	24	75.0
Spanish	2.6	51	34	66.7
Swedish	1.0	21	14	66.7
Catalan	1.7	45	33	73.3
Italian	1.4	55	47	85.5
CHILDES Spanish	1.4	55	46	83.6
CHILDES Catalan	0.3	39	27	69.2
CHILDES Italian	0.3	49	31	63.3

Table 1: Saturation in a variety of languages, adapted from Chan 2008.

every thirteen inflectional categories observed in the corpus. A particular combination of lemma and inflectional category is far more likely to be missing than observed.

The phenomenon of low saturation is not specific to Spanish; Table 1 presents the analysis of the inflectional morphologies of several languages and genres as computed by Chan (2008).<sup>3</sup> However, excluding English corpora, regardless of corpus size the max saturation does not approach 100%; the common case is that the learner does not see a complete paradigm for any verb even with relatively large amounts of input.

### 2.3 Implications for learning

These simple statistical investigations of morphology have direct implications for the acquisition of morphology. First, it is unrealistic to expect the full paradigm of any particular stem to be available to the child learner (e.g., Albright 2005). Language acquisition takes a finite, and in fact quite modest, amount of input (Hart and Risley 2003) which, true to Zipf’s Law, contains numerous instances of few linguistic types: the learner must be able to generalize fairly aggressively to “fill in” the paradigm table for unattested forms. Second,

<sup>3</sup>Chan finds higher numbers for saturation than in the CHILDES Spanish example that we give in this section because his study does not exclude light verbs. While light verbs appear in many inflectional variants, they are generally irregular and do not provide much value as exemplar paradigms.

Subjects	1 form	2 forms	3 forms	4 forms	5 forms	6 forms	S/N
Italian children	81.8	7.7	4.0	2.5	1.7	0.3	1.53
Italian adults	63.9	11.0	7.3	5.5	3.6	2.3	2.54
Spanish children	80.1	5.8	3.9	3.2	3.0	1.9	2.23
Spanish adults	76.6	5.8	4.6	3.6	3.3	3.2	2.61
Catalan children	69.2	8.1	7.6	4.6	3.8	2.0	2.10
Catalan adults	72.5	7.0	3.9	4.6	4.9	3.3	2.34

Table 2: Verb agreement distributions in child and adult Italian, Spanish, and Catalan.

the sparsity of morphological distribution must be taken into account when we assess the empirical properties of children’s morphology.

As we shall see in Section 3, children’s inflectional morphology is in general quite good. The errors tends to be those of omission (bare or infinitive when a tensed form is required) or overuse of a general form (e.g., the over-regularization of irregular verbs), and the rates of errors are generally low. But it has been pointed out that low rates of morphological errors do not necessarily imply children’s mastery of morphology, especially its combinational aspects. In an influential line of research in child language, *usage*-based theorists such as Tomasello note that children’s morphology may well result from the storage of lexically specific morphological forms in the input: if the retrieval mechanism is generally reliable, children will also make few morphological errors. Indeed, the combinatorial *diversity* in child morphology appears quite low, which seems to support lexically specific morphological learning. For instance, Pizzuto and Caselli (1994) found that in a corpus of child Italian speech, 47% of all verbs used by 3 young children (1;6 to 3;0) were used in 1 person-number agreement form, and an additional 40% were used with 2 or 3 forms, where six forms are possible (3 person  $\times$  2 number). Only 13% of all verbs appeared in 4 or more forms. The low level of combinatorial diversity has been a major source of evidence for an item-based theory of language learning (Tomasello 2000).

Here the sparsity of child morphology must be viewed in the context of morphological sparsity in general. Table 2 summarizes the results from the corpus analysis of all of child and child-directed data in Italian, Spanish, and Catalan that are currently available in the CHILDES database (MacWhinney 2000). The morphological data is again analyzed with the Freeling tagger, and only tensed forms are counted. Each cell represents the percentage of verb stems that are used in 1, 2, 3, 4, 5, and 6 person and number forms.

As can be observed in Table 2, Spanish and Catalan children and adults show very similar, and very sparse, agreement usage diversity. Italian children use somewhat more stems in only one form than Italian adults (81.8% vs. 63.9%), but this can be attributed to the measure in the last column, which is the ratio between the total number of inflected forms ( $S$ ) over the total number of stems ( $N$ ). This ratio provides the average number of opportunities for a stem to be inflected: when the ratios are comparable between children and adults as in the case of Spanish and Catalan, we observe similar diversities of agreement

morphology. In the case of Italian, the adults had roughly two thirds more opportunities to use a stem than the children, which would account for the modest discrepancy in the frequency of one-form verbs.

Of course, these comparisons do not conclusively show that children’s grasp of agreement morphology is identical to adults. To do so, one must calculate the statistically predicted morphological diversity under the assumption of an adult-like morphological system, while taking the Zipf-law distribution of stems and inflectional categories into account (Figure 1 and Table 1), and compare it against the empirically attested usage diversity. (For an assessment of syntactic combination in early child language, see Yang 2013.) Likewise, an item-based approach also needs to be precisely formalized with quantitative predictions to be verified against the empirical patterns in child and adult language. Nevertheless, it is premature to conclude that low usage diversity necessarily implicates an under-developed morphological system: *all* morphological systems are sparsely represented. In what follows, we review how children’s morphological systems develop (Section 3) and how children may overcome these statistical challenges in morphological acquisition (Section 4).

### 3 Development

It is impossible to review the acquisition of morphology without an extended discussion of the English past tense, one of the most extensively studied topics in all of linguistics and cognitive science (Pinker 1999). The generally impoverished morphology of English may not seem like a representative case for the complexity in the world’s languages, but the theoretical issues it raises have important and quite general ramifications for the methodology and theories of morphology. So we begin.

#### 3.1 English past tense: Words vs. rules

In the traditional linguistics literature, the past tense of English is treated as a computational system that takes the stem (e.g., *walk* and *think*) as the input and produces the past tense (e.g., *walked* and *thought*). Bloch 1947, for instance, divides the verbs into inflectional categories: the verb *think* falls into a category that adds the *-t* (his category B3, p413) and follows the morphophonemic alternation (his Type II, p415) that changes the vowel to /ɔ/. According to this view, both regulars and irregulars are generated with input-output systems: it is just that the latter requires special instructions since their categories and alternations are not predictable. Generative grammar has largely preserved this tradition from SPE on (Chomsky and Halle 1968): the main difference can be found in the more abstract formulations than those found in Bloch’s treatment, whose classification generally pertains to the surface pattern. In the present discussion, we will use the term “rule” to describe the computational process that maps an input to its morphological derived form (cf. *schema* used by Bybee and Slobin 1982). Crucially, when we speak of rules for irregular verbs, we are, à la Bloch 1947, referring to the lexicalized rules that apply to a fixed list of verbs but do not extend beyond them.

The major point of contention in the past tense debate (see Clahsen 1999; Pinker 1999; Pinker and Ullman 2002; McClelland and Patterson 2002 for general overviews) concerns the treatment of the regular verbs: the two sides are in general agreement that the irregular verbs are acquired as paired associations between the stem and the past form. The connectionist position, starting with Rumelhart and McClelland 1986, holds that the regulars are stored and learned as associative pairs similarly to the irregulars, whereas the Words and Rules model, starting with Pinker and Prince 1988, holds that the regular verbs are handled by a symbolic rule (“add *-d*”) in a decidedly non-associative fashion. This debate has spilled into the morphological processing literature; see Chapter 29 for discussion. Hence, our review highlights a third approach (Yang 2002; Albright and Hayes 2003), briefly considered but not pursued by Pinker and Prince 1988, one which may be of more interest to the linguistic theorists: irregular verbs are also handled by rules albeit lexicalized ones, and they are not learned by direct word association.

What kind of evidence would bear on the nature of irregular past tense learning—by word association or rules? The strongest evidence for associative storage comes from frequency effects in child language acquisition: in general, irregular past tense forms that are more frequently used in the input tend to be acquired more accurately by children.<sup>4</sup>

Logically, children may make three kinds of errors regarding the use of past tense.<sup>5</sup>

- (1) a. Over-regularization: an irregular verb is inflected with the *-ed* rule (e.g., *hold-helded*).
- b. Over-irregularization:
  - i. Regulars: a regular verb is inflected with an irregular form (e.g., *heal-helt* along the lines of *feel-felt*).
  - ii. Irregulars: an irregular verb is inflected with an incorrect, non-target, irregular form (e.g., *sting-stang* along the lines of *sing-sang*).

Over-regularization is extremely well known and has become the focus of virtually all empirical studies of past tense acquisition. By contrast, over-irregularization errors, which are sometimes referred to as analogical errors, are frequently alluded to (e.g. Pinker 1999) but rarely studied systematically. These turn out to be virtually absent; we return to this matter in Section 3.2 when we discuss the role of productivity and analogy in child morphology. Earlier studies of past tense find an overregularization rate of 4.2% (Marcus et al. 1992; Pinker 1995), but later studies have found somewhat higher rates. For instance, Yang 2002 reports 10% out of approximately 7,000 tokens on the basis of four large longitudinal

---

<sup>4</sup>The measurement of learning accuracy is by no means a simple matter. On the one hand, young children under the age of 4 often omit tense marking altogether, characteristic of their general syntactic development (Phillips 1995; Legate and Yang 2007). On the one hand, there are methodological issues concerning the inclusion and exclusion of verbal forms, which has resulted in quite different reports in the literature; see Maratsos 2000 for discussion.

<sup>5</sup>By errors, we mean forms that deviate from the “standard” adult usage form: unfortunately, the acquisition data presently available does not offer any opportunity to systematically study the acquisition of dialectal variation in past tense formation (Anderwald 2009).



corpora in the public domain, Maslen et al. 2004 reports 7.8% out of approximately 1,300 tokens from a single child. In general, however, a frequency effect has been found in the rate of over-regularization: irregular verbs that appear in past tense more frequently in adult speech tend to have lower rates of over-regularization errors. Marcus et al. 1992 report that across all irregular verbs, children's rates of over-regularization for irregulars and the (log) frequencies of their past tense forms in child directed input show a coefficient of correlation of -0.33. Most of the evidence provided for the associative account of irregular verbs draw on various forms of frequency effects (e.g., Pinker 1995).

However, -0.33 is a fairly low correlation as far as statistical analyses go. Partly motivated by this, Yang 2002 explores the plausibility of the traditional rule-based approach to irregular past tense. According to this view, the learning of irregular verbs consists of several components:

- (2) a. The learner must construct rules for the irregular verbs such as those found in Bloch 1947.
- b. The learner must associate, or lexicalize, each irregular stem with the corresponding irregular rule.
- c. To inflect an irregular past tense, the learner needs to locate the associated irregular rule and then apply it.
- d. The failure to either retrieve the irregular rule or to apply the irregular rule results in the use of the default rule of “add -d” and thus over-regularization.

It is important to note that the rule-based approach does not dispense with the role of associative memory altogether. The irregular verbs are unpredictable and have always been treated with some type of lexicalization. However, the rule-based approach makes significantly different predictions from the word-based approach. For instance, the learner receives evidence for the “ought” rule, which applies to *bring*, *buy*, *catch*, *seek*, *teach*, and *think*, whenever any of the six verbs appears in past tense. This leads to the notion of a *rule frequency*, which is the sum of the frequencies of the words that follow the rule. The word-rule association, of course, can only be established upon exposure to the specific words: *brought* will contribute to the lexicalized association between *bring*, and only *bring*, and the “ought” rule.

The bipartite nature of the rule-based approach contrasts with the word-based associative approach. Consider two irregulars that are comparable in their word frequency but one belongs to a very frequent irregular rule and the other belongs to a relatively infrequent irregular rule. For the word-based theory, the two verbs are expected to be learned at a comparable level of accuracy. For the rule-based theory, the verb belonging to the more frequent rule will be expected to be learned better because the rule, being more abundantly attested in the input, will be used more reliably. The irregular verbs *taught* and *flew* are two such examples. They are comparable in frequency (81 and 117 in 3.6 million words of English spoken by the mothers in the CHILDES database), but children produced the over-regularized *flied* 26% of the time (23/89) but *teached* only 12% of the time (3/25). (Many additional examples can be found in Yang 2002.) This performance disparity is hard

to reconcile with the word-based theory but is straightforward under the rule-based theory: *teach* belongs to the “ought” rule which counts far more frequent verbs such as *thought*, *brought*, *bought* and *caught* (with frequencies of 1516, 416, 367, and 288 in the same corpus). By contrast, the most frequent member of the rule that applies to *fly* is *knew*, which only appears 390 times. This sort of free rider effects goes a long way to ameliorate the weak statistical correlation between input frequencies and overregularization rates.

Yang 2002 has identified another pattern that further suggests that past tense acquisition is not a matter of building word associations. A subset of English irregular verbs has long been recognized as following more general morphophonological patterns. For instance, verbs such as *hide-hid*, *feel-felt*, and *say-said* have been analyzed as instances of the vowel shortening process triggered by the suffix (null, *-t*, and *-d* respectively), similar to alternations in derivational word formation such as *deep-depth* and *divine-divinity* (Halle and Mohanan 1985; Myers 1987). Intriguingly, all vowel shortening verbs, including those with very low input frequencies, are used very well by child learners, with an overall overregularization rate of only 2% (Yang 2002). The learning mechanisms for and the representation of these vowel shortening verbs are by no means clearly understood but it is plausible that vowel shortening, being a fairly general process in English morphophonology, will be supported by data beyond the realm of past tense, which in turn benefits the acquisition of the irregular verbs that make use of the same process.

Needless to say, these investigations of English past tense acquisition are only beginning to unravel the underlying complexity of this very small corner of English morphology. But we believe that morphological theories have much to offer to the psychological studies of lexical learning and processing. Indeed, the rule-governed nature of irregular verbs reviewed here goes hand in hand with recent findings in behavioral and neurological studies that the irregular past tense is compositionally formed (see e.g., Allen and Badecker 2002; Stockall and Marantz 2006; Morris and Stockall 2012); abstract morphological postulates seem to have observable psychological correlates; see Chapter 29 for a review of morphological processing literature.

The present discussion also raises some important questions. How are rules such as “add *-d*” as well as the lexicalized irregular rules acquired? Some of the existing work is reviewed in Section 4, but much remains unclear. Additionally, how does the learner know that lexicalized rules are lexicalized and do not extend to novel items while the “add *-d*” rule is general and open ended. Notice that this issue also needs to be resolved by the Words and Rules model: without knowing what the default rule is, the learner does not know whether to commit a verb to associative storage or to apply a rule without needing to commit any additional memory. In other words, how does the child distinguish the regular rules in her language from the lexicalized ones?

### **3.2 Productivity and analogy in child morphology**

The acquisition of morphological productivity can be summarized very succinctly: children draw a sharp, nearly categorical, distinction between productive and unproductive pro-

cesses. These findings should be integrated into the study of productivity, a central topic in morphological research (Chapter 4).

### 3.2.1 The Wug Test

The simplest and most direct assessment of productivity is the celebrated Wug test. In a landmark study, (Berko 1958) introduced young children to a wide range of novel words including nouns, verbs and other categories:

- (3) This is a wug.  
Here is another one.  
These are two \_\_\_\_\_

The Wug test is now extensively used in a wide range of studies (see e.g., Bybee and Slobin 1982; Bybee and Moder 1983; Marcus et al. 1992; Clahsen 1999; Hahn and Nakisa 2000; Albright and Hayes 2003; Hahn and Nakisa 2000; Hayes et al. 2009; Zhang and Lai 2010; Becker et al. 2011) which subsequently have had considerable influence in linguistic theorizing (Pinker 1989; Tomasello 2003; Taylor 2003; Hay and Baayen 2005). Yet the Wug test by no means provides a direct window into morphology. It turns out that children are far from perfect on their Wug tests (Berko 1958: 160). For instance, in the inflection of regular past tense, first graders' success rate for novel (regular) past tense ranges from 25% to 85%, producing responses such as *rick-ricked* and *spow-spowed*. Yet most if not all English learning children have acquired the "add -ed" rule by the age of three (Marcus et al. 1992), as indicated by their spontaneous use of over-regularized forms. Thus a gap of a minimum of three to four years exists between the knowledge of regular past tense and the appropriate use of it in a specific experimental design that involves learning a novel word; failure at the Wug test does not entail the language user's morphological deficiency.

When children do produce a response in the Wug test, their behavior points to a categorical distinction between regular and irregular processes. For instance, when children fail to produce *spowed* for *spow*, they produced no response at all rather than, say, *spew*, which would follow the analogy of *know-knew*. The role of analogy was systematically investigated by Berko, who also created novel verbs such as *bing* and *gling* that bear great similarity to existing irregular verbs, corresponding to the *i-a/u* class that seems to have the greatest potential for analogical extension:

- (4) This is a man who knows how to GLING.  
He's GLINGING.  
(Picture of a man exercising.)  
He did the same thing yesterday.  
What do he do yesterday?  
Yesterday he \_\_\_\_\_.

Children, however, are overwhelmingly conservative in their response: only one of the 86 children in Berko's study supplied the analogical form *bang* and *glang*; the rest all produced

the *-ed* forms. It is possible that adults treat the Wug test differently, and Berko did notice that adult subjects are far more likely to use the irregular form with at least half of the subjects supplying *ba/ung* and *gla/ung*. But given the task complexity with the Wug test (see Schütze 2005 for extensive discussion), results obtained from adult performance, while interesting and potentially informative about the organization of morphology, cannot be directly used as evidence for the acquisition of morphology.<sup>6</sup>

### 3.2.2 Regularization vs. irregularization in English

The use of naturalistic data, which is now abundantly available in the public domain, sidesteps the methodological issues inherent in behavioral tasks such as the Wug test and provides very clear evidence for how children deal with morphological productivity. This is only made possible because past tense forms are highly frequent and an enormous amount of time has been accumulated in the public domain (MacWhinney 2000). As reviewed above, children frequently attach *ed* to irregular verbs. Likewise, when new words such as *google* were introduced into the English language, its (regularly inflected) past tense (*googled*) was instantly available.

The lexicalized irregular rules, by contrast, are almost never overgeneralized by children beyond the lexicalized lists, despite frequent anecdotal references to these forms in child language, such as *bite-bote*, *wipe-wope*, *think-thunk* etc. (Bowerman 1982; Bybee 1985; Pinker and Prince 1988; Pinker 1999). But the empirical evidence for irregularization is very slim: there is not a single attested example of these in the entire CHILDES database of about 5 million words of child English. The most comprehensive empirical study of analogical errors (Xu and Pinker 1995) in fact refers to analogical errors as “weird past tense errors” on the basis of their rarity. Xu and Pinker examined all the past tense forms produced by nine children in the CHILDES database, about 20,500 in all. Only 40 over-irregularization errors (0.02%) were identified, which is at least an order of magnitude lower than the rate of over-regularization. A closer examination of the irregularization errors, exhaustively listed in the Xu and Pinker study, suggests an even lower rate of irregularization. Of the 40 attested examples, at least ten are very likely speech errors (e.g., “fit-feet”, “say-set”, “fight-fooed”, “bite-bet”) or potentially transcription errors. For instance, “slep” is included as an analogical error of “sleep” but is more likely due to the word final t/d deletion process in spoken English (Labov 1989). The only systematic pattern is *ing*→*a/ung* pattern, and the verb *bring* is the only item that is frequently over-irregularized. It is possible that children form a productive rule that changes /ɪ/ to /æ/ before ŋ: for children’s small vocabulary, this pattern is consistent with two out of the three verbs that fit this structural description (*ring*, *sing* and *bring*). But it is also possible that *brang* is present in the input as a matter

<sup>6</sup>We do not discuss the use of rating tasks in assessing morphological productivity (e.g., Albright and Hayes 2003). Rating, by definition a gradient task on a scale, guarantees a gradient outcome. It is useful to gain a refined understanding of phenomena that are independently shown to be gradient but is not suitable for determining whether a process or class is categorical or gradient in the first place (Armstrong et al. 1983). Nevertheless, when adult subjects were given both a regular form and an irregularized competitor (e.g., *glinged* vs. *gla/ung*) in a rating task, the regular form is still preferred (Prasada and Pinker 1993).

of dialect variation (Herman and Herman 2014): in fact, *brang* can be found in the child directed speech collected in CHILDES (MacWhinney 2000). Nevertheless, the drastically different rates of over-regularization and over-irregularization suggest that there is a (near) categorical distinction with respect to productivity between the regular rule and the irregular rules.

### 3.2.3 Productivity across languages

The productivity/analogy asymmetry has been observed in many case studies of children's language, mostly in the arena of morphological acquisition but can be found in syntax as well. When children make mistakes, they almost always employ a default or productive form (e.g., *thinked*) or omit the appropriate form altogether: they almost never substitute with an inappropriate form. Of course, the very notion of a default raises complex theoretical questions (see Chapter 12) and we will also briefly the acquisition of what appears to defective, or default-less morphological systems (Section 3.3). But first, an overview of previous findings on productivity from the cross-linguistic research of child morphology.

In a study that targets the agreement affixes *-st* (2nd singular) and *-t* (3rd singular), Clahsen and Penke 1992 find that while a child ("Simone") supplied an agreement affix in obligatory context only 83% of time, the rest are errors of agreement omission. When an agreement affix is on the offer, it is almost always the appropriate one (over 98% of the time); inappropriate use (e.g., substituting a *-t* for *-st*) is thus virtually absent. Similar patterns can be observed in the acquisition of Italian. In a cross sectional study (Caprin and Guasti 2009: 31), children in all age groups use a diverse and consistent range of tensed forms. Furthermore, the use of person and number agreement is essentially error free throughout, reaching an overall correct percentage of 97.5%, consistent with previous reports (Guasti 1993; Pizzuto and Caselli 1994). Children's impressive command of agreement is best demonstrated in the acquisition of languages with considerable morphological complexity. In a study of morphosyntactic acquisition in Xhosa (Gxilishe et al. 2007), children gradually expand the use of subject agreement across both verbs and noun classes between 24 and 36 months. The rate of marking in obligatory contexts as well as the diversity of the morphological contexts themselves steadily increased. In a process best described as probabilistic, the children often alternate between marking a verb root in one instance and leaving it bare in another, very much like the use/omission alternation pattern reviewed earlier. Crucially, however, virtually all agreement errors are those of omission: 139 out of 143 or 97.2% to be precise. Substitution errors are again very rare, confirming previous research on similar languages (Demuth 2003; Deen 2005).

We now turn to several case studies that focus more specifically on the contrast between regular and irregular morphologies in children's naturalistic speech. This type of evidence has been accumulating from research on the Words and Rules model of morphology (Pinker 1999; Clahsen 1999), for which a categorical distinction between regular and irregular processes is of central importance. The evidence is unequivocal.

The German participle system consists of a productive default *-t* suffix (*fragen-gefragt*

“ask-asked”), as well as an unpredictable set of irregulars taking *-n* (*stehlen-gestohlen* “steal-stolen”) (Wiese 1996). In a series of studies, Clahsen and colleagues (Clahsen and Rothweiler 1993; Weyerts and Clahsen 1994; Clahsen 1999) find that children across all age groups overapply the *-t* suffix to the irregulars, where the reverse usage is virtually absent. Their longitudinal data contains 116 incorrect participle endings, out of which 93% are *-t* errors (\**gekommt* instead of *gekommen* “come”). The rest are irregularization errors such as \**geschneien* for *geschneit* (“snowed”). According to the authors, the overall rate of *-t* regularization is 10% of all usage, which results in the *-n* irregularization rate of 0.75%. The acquisition of German past participles, therefore, is quite analogous to that of English past tense reviewed earlier, as both point to the productive asymmetry between regulars and irregulars.

The inflection of Spanish verbs provides a complex but suitable case for exploring productivity in child language. In Spanish, stems generally consist of theme vowels and roots, which are then combined with affixes for inflection. For instance, a finite form of the verb *hablar* (“to talk”) is *habl-á-ba-ais*, which represent the root (*habl* “speak”), the theme vowel (*á*), the past tense (*ba*) and the 2nd personal plural (*ais*). The theme vowels form three conjugations, with the 1st (*a*) being the numerically most dominant class which has 9,706 members, followed by the 2nd (712) and 3rd (730), according to Real Academia Española 1992 (cited in Clahsen et al. 2002). The irregularity in Spanish inflection comes in two broad classes concerning the stem and the suffix respectively. There are some 30 verbs that are highly irregular with the insertion of a velar stop in certain inflections. These examples include *tener* (“to have”), *poner* (“put”) or *salir* (“go out”), whose first person singular forms are *tengo*, *pongo* and *salgo*, respectively. The majority of irregulars undergo a well known morphophonemic alternation known as diphthongization, a process which is not limited to verbal morphology per se (Harris 1969; Eddington 1996). For these verbs, the mid vowel is diphthongized in stressed syllables. (5) shows the pattern for the present indicative of the verbs *comenzar* (“begin”) and *contar* (“count”), where the graphic accent indicates phonological stress:

- (5)    *comiénto*    *comiéntas*    *comiénta*    *comenzámos*    *comenzáis*    *comiéntan*  
          *cuénto*        *cuéntas*        *cuénta*        *contámos*        *contáis*        *cuéntan*

While it has been suggested that the *form* of diphthongization is predictable (Harris *ibid*), the identity of the verbs that undergo diphthongization is arbitrary and must be lexicalized. It is possible to find minimal pairs such as *contar-montar* where the former contains the diphthong (*cuento*) but the latter does not (*monto*). And there are a few common verbs that show both diphthongization and velar insertion in some forms. For instance, *tener* (“to have”) and *venir* (“to come”) show velar insertion in the present subjunctive and 1st person singular of the present indicative and [ie] diphthongization in 2nd singular and 3rd person of the present indicative. While inflectional irregularities in Spanish mostly affect the stem, the suffixes are affected as well. For the stem *querer* “to want,” for instance, the 1ps past tense is *quise*, which involves the stem change noted earlier while taking the irregular suffix

rather than the regular suffix, which would have resulted in *\*quisí*. The suffix in the 3ps past tense *puso* “s/he/it put” is *-o* and the regular suffix would have formed *\*pusió*.

Clahsen et al. (2002) analyzed the verbal inflections of the 15 Spanish-speaking children and found strong evidence for the categorical distinction between regular and irregular suffixes as well as stem alternations.

- (6) a. The irregulars: children produced a total of 3,614 irregular verb tokens, out of which 168 (4.6%) are incorrect either in stem formation or suffixation.
  - i. Of the 120 stem formation errors (see below), 116 are over-regularizations and only one is analogical irregularization.
  - ii. Of the 133 suffixation errors, 132 are over-regularizations with no occurrence of irregularization.
- b. The regulars: children produced 2073 regular verb tokens, only 2 of which are the inappropriate use of irregular suffixes.

Clahsen et al.’s study did not include errors regarding diphthongs; all the stem formation errors are failures to use a diphthong when required, and no instance of diphthongization when unnecessary is found. Although this broadly supports the notion of diphthongization as a lexicalized process, it does not consider the possibility of *inappropriate* use of diphthongization, e.g., the child produces [ie] alternation when the correct diphthong is [ue]. To address this issue, Mayol 2007 provides a finer-grained investigation of inflectional errors focusing more specifically on the types of stem formation errors and their underlying causes. The naturalistic speech transcripts of six Spanish speaking children revealed a total of 345 stem formation errors out of a total 1782 tokens, but importantly, not a single case of inappropriate form was found.

Much more data, both in volume and also in linguistic variety, is needed to refine our understanding of morphological acquisition. Unfortunately, there is little research on the acquisition of derivational morphology which should provide a fertile ground for investigations of productivity. The primary reason for this is that derivational morphology is acquired quite late and much of the vocabulary is introduced in a school setting (see e.g., Jarmulowicz 2002). Nevertheless, the best available evidence from child morphology suggests that children are capable of drawing a categorical distinction between regular and irregular morphologies, as part of their excellent grasp of the structural properties of morphological alternations across the languages.

### 3.3 When productivity fails

We would be remiss if we didn’t address some perennial challenges in the study of morphology. As children quickly and accurately grasp the productive aspects of morphology, it becomes especially puzzling that there are corners of the grammar where productivity unexpectedly fails. In a classic paper, Halle (1973) draws attention to the existence of paradigmatic gaps; for instance, there are around 75 Russian verbs, all belonging to the second conjugation, that lack a first person singular (1sg.) non-past form, a fact which cannot

be predicted on frequency or phonotactic grounds alone (see Sims 2006 for additional discussion of the Russian data and Baerman et al. 2010 for a recent general survey). Although we are not aware of any acquisition study of these classic cases of paradigmatic gaps, the issue of non-productivity has turned up in the child morphology literature and we briefly review a prominent case.

According to the dual route model, there is a default rule for every morphological context; see Pinker 1999 for a discussion of how the learner may identify that rule on the basis of statistical and distributional properties of words. In Polish, however, masculine nouns in the genitive singular (gen.sg.) either take an *-a* or *-u* suffix, but neither is productive when evaluated against the standard suite of tests for productivity (Dąbrowska 2001). The distribution of these two suffixes is widely studied but has stubbornly resisted systematic classification (e.g., Westfal 1956). In contrast, the genitive plural (gen.pl.) for masculines is unproblematic: the default suffix is *-ow* with a small number of exceptional nouns taking *-i/-y*. Previous studies of Polish acquisition (Dąbrowska 2001: , Table 8 and 12) have found that children make very few errors in the gen.sg. (mean rate 1% for three children), and frequently overextend *-ow* in the gen.pl (mean rate 11.3%). The significant difference in error rates suggests that, similar to the many cases reviewed earlier, the marking of the singular is lexicalized and does not generally extended beyond the lists of fixed words whereas the plurals contain a productive rule that triggers over-regularization.

The case of Polish masculine genitives, especially the singulars, is of considerable theoretical interest. It suggests that the child learner should not presuppose the existence of a default rule. Likewise, the absence of a default poses challenges to competition-based theoretical frameworks such as Distributed Morphology (Halle and Marantz 1993; Chapter 16) and Optimality Theory (Prince and Smolensky 2004; Chapter 21) under which a winning form generally emerges; see the collections of articles in Baerman et al. 2010. Furthermore, the default rule cannot be established on the basis of statistical dominance (*pace* Pinker 1995). According to a word list based on child directed Polish and made available in CHILDES, *-a* and *-u* consume 62% and 38% of the stems (837 and 516) but this does not give rise to the default status for *-a* as the child acquisition data as well as traditional descriptions clearly indicate.<sup>7</sup>

At the same time, it is useful to recall that statistical minority does not entail the absence of productivity. Despite considerable disagreement over the characterization of the German noun plural system (Clahsen 1999; Wiese 1999; Wunderlich 1999), the productivity of the *-s* suffix, which applies to the smallest number of noun stems, is not controversial and indeed can be observed in children's over-regularization (Marcus et al. 1995). How children assess the statistical and structural factors in the world's language will keep the theorist occupied in the years to come.

---

<sup>7</sup>We thank Margaret Borowczyk for the corpus analysis.



## 4 Learning mechanisms

Ideally, the modeling of morphological learning would be faithful to the process of language acquisition, featuring the development of models that predict the type of learning behaviors shown by child learners and models that operate on the type of input children receive. As discussed in this section, there is much progress to be made to develop linguistically and psychologically faithful models. The majority of existing models make the assumption that the input is both dense and supervised, contrary to the evidence given in Section 2, and those that do attempt to model properties of child language acquisition struggle to balance the mechanisms required for learning of irregular forms with those required for categorical productivity.

### 4.1 English past tense learning models

As with studies of development, the creation of computational models of morphological learning has focused heavily on the English past tense. Rumelhart and McClelland (1986) provide the first well-known computational model that attempted to model the learning of morphology in a manner aligned with research in child development. At a high level, this connectionist model exhibits the same overregularization phase (e.g., Ervin and Miller 1963; Marcus et al. 1992) during training as children learning the English past tense.

But while it replicates a salient development pattern, the Rumelhart and McClelland model makes unpredictable errors that are unattested in child learning when trying to handle regular forms, such as *membled* for the past tense of *mail* (see Pinker and Prince 1988 for details). While the specifics of these problems have been addressed with further improvements to this class of models (Plunkett and Juola 1999), the core tension between capturing the minimal generalizations required for correct handling of both irregular and regular forms remains.

As discussed in Section 3, rule-only theoretical models such as the Rules over Words model (Yang 2002; see also the computational work of Mooney and Califf 1995; Yip and Sussman 1997) address this problem by expressing both regulars and irregulars as rules, with irregular rules only applying to memorized lexical entries without further generalization. The Words and Rules model (Pinker 1999) presents a mixture of rule-based and associative models, for handling default rules and classes of irregular forms, respectively, but does not provide a validated mechanism for assigning individual forms to the correct model. Albright and Hayes (2002) present a model that induces rules by identifying minimal generalizations and go beyond many previous studies by learning full paradigms and attempting to align the output of their model with (adult) speaker intuitions.

Finally, to provide an accurate account of English past tense acquisition, a model should occasionally over-regularize the irregular verbs but should almost never over-irregularize at all, mirroring the developmental asymmetry reviewed earlier. Unfortunately, the absence of analogical errors is almost never addressed in the computational modeling work, perhaps because this finding has been overshadowed by the attention devoted to over-regularization.

For instance, a recent study (O'Donnell 2011) tests a variety of models for past tense, by examining the models' output when presented with both existing and novel words. Most of the models are reasonably successful at passing the Wug test for regular verbs, but they frequently generate analogical forms from the existing irregulars. The best model overall (O'Donnell *ibid*), a Bayesian model that attempts to provide a balance between lexicalized storage and rule-based computation (Johnson et al. 2007b), produces 10% of over-irregularization patterns on novel items, which is on par with children's *over-regularization* rate and an order of magnitude higher than the over-irregularization rate. In other words, it is as if this model treats every novel instance as an irregular verb. In general, the problem of productivity has not been systematically studied in the modeling work; how to bridge this gap between the computational and empirical studies of morphological learning deserves fuller attention in future research (see (Yang 2005) for a preliminary investigation).

## 4.2 The distributional learning of morphology

While the past tense learning models vary in their specificity—for example, Rumelhart and McClelland 1986 and Albright and Hayes 2002 test actual implementations, while Pinker 1999 provides only a description of a plausible model—they uniformly make the assumption that the input to the learner is neatly organized into pairs of forms representing a single inflectional change. The input to these models consists of entries like *go/went*, *make/made*, *bake/baked*. The closest thing to these pairs that may be observed in natural input are situations like this one from the Adam data in CHILDES (Brown 1973):

- (7) Adam: Mary. Mary go? Mary go, mommy?  
 Mother: Who?  
 A: Where Mary go?  
 M: Mary?  
 A: Yeah.  
 M: She's gone home. She went home in July.  
 A: Mary gone.

To use a word pair-based approach in acquisition, a child learner would need to discover these words pairs from input, a process that is largely unexplored. As the above dialog shows, such a process would also be non-trivial, especially in cases of suppletion. While there is evidence that automatic approaches can identify sets of morphologically related words (e.g., Parkes et al. 1998; Wicentowski 2002), little work has been done to attempt to extract the high-accuracy word pairs required by supervised single-inflection learners. Thus while cognitively-motivated models have had success in replicating phenomena related to language acquisition, they have thus far not done so in an unsupervised setting.

Unsupervised learning of morphology is well-studied in engineering-oriented models; however, these models generally make no attempt to align themselves with developmental studies. One approach is using sequences of characters to identify morpheme boundaries—an idea first presented by Harris (1955)—either by an algorithmic approach (Hafer and

Weiss 1974; Keshava and Pitler 2006) or developing n-gram-based statistical models (e.g., Spiegler et al. 2009). Often segmentation-based approaches rely on minimum description length to guide the appropriate amount of segmentation (Brent et al. 1995; Creutz and Lagus 2005), or organize the segmentations learned into paradigms (Goldsmith 2001; Monson 2008). While the use of paradigms creates what appears to be a useful organization of the learned rules, recent work questions the learnability of paradigms or item-based representations from realistic input due to the quantitative sparseness of data available to learners (Chan 2008). Bayesian approaches (e.g., Goldwater et al. 2006; Naradowsky and Goldwater 2009; Johnson et al. 2007a) use a prior distribution to guide the learner toward a desirable target, often explicitly taking into account information regarding the frequency distributions of stems and affixes. Lignos et al. (2009) use heuristic approaches to incrementally learn affixes and the stems they can be applied to, directly taking advantage of the statistical properties of morphology discussed in Section 2.

A major benefit of studies that evaluate models using corpus input—a requirement in modern computational linguistics venues, but often optional elsewhere—is that their performance on naturalistic and sparse data can be directly evaluated. As discussed in Section 2, learning models that assume the input contains full paradigms will fail to find them in the input.

Chan (2008: ch. 3) provides a clear demonstration of the effects of sparsity on models that require full paradigms to infer a reasonable representation of the morphology of the language. When the Wall Street Journal portion of the Penn Treebank (Marcus et al. 1999) is provided as input to the *Linguistica* system (Goldsmith 2006), a minimum description length (MDL) morphology learning system that attempts to derive the most compact characterization of the data, it learns a set of signatures, structures which contain a set of stems and the suffixes they take. For example, it identifies 604 stems (e.g., *alarm*) that take the set of suffixes  $\{-\emptyset, -ed, -ing, -s\}$ .<sup>8</sup> This is a rather good representation of English verbal morphology. But what else is learned? One signature of 285 stems takes the suffixes  $\{-\emptyset, -ed, -ing\}$ , one of 242 stems takes the suffixes  $\{-\emptyset, -ed\}$ , and one of 2624 stems, the largest signature of all, takes the suffixes  $\{-\emptyset, -s\}$ .

These partial paradigms are a symptom of the Zipfian distributions described in Section 2. Despite the fact that all regular verbs in English should be members of the  $\{-\emptyset, -ed, -ing, -s\}$  signature—modulo any orthographic oddities—most stems are only seen with some of the possible inflections. A minimum description length learner whose goal is to provide a compression of the input cannot infer the common pattern when the majority of relevant data points are missing.<sup>9</sup> A useful morphological learning model must be able to look beyond the sparsity of the input, much like the child learner must look beyond the poverty of

<sup>8</sup>Many computational models are evaluated using orthographic input. While this provides a relatively poor analog to the input a child learner receives, the properties of sparsity under discussion here are the same regardless of the representation of the input.

<sup>9</sup>The specific issue with Goldsmith’s model is partially addressed by Monson (2008) using a more complex search strategy, but the core problem of inducing a full paradigm when one is not present in the input still remains.

the stimulus. In this sense, productivity is at the heart of all morphological learning models (Yang 2005): the child must know where to generalize and when to lexicalize.

It is clear that children excel at drawing useful generalization from distributional information. In addition to their generally excellent command of morphology which we reviewed earlier, there is now some suggestive evidence on how they can carry out distributional learning. For instance, Gerken et al. (2005) finds that infants of age 1;5 succeed in learning syntactic category generalizations from distributional cues, one finding in a constantly-growing set of studies that demonstrate that infants and young children possess powerful associative and statistical learning capabilities; see also Gómez and Lakusta 2004. Yet it still remains an open question whether distributional learning mechanisms demonstrated in a laboratory setting can scale up to actual languages with realistic complexity (Yang 2004). A most fruitful direction in the computational modeling of morphological learning will likely see close engagement with the experimental study of child language acquisition, while taking the full range of cross-linguistic developmental findings into account.

## 5 Conclusion

To summarize, children's acquisition of morphology provides a wealth of empirical data to bear on the theoretical issues in morphology. The fact that their morphological knowledge is generally secure is especially impressive when considered in a psychological setting of language learning. Linguists have access to a wide range of evidence when studying the morphological system of a language. For instance, theorists can make use of a large lexical corpus that far exceeds the quantity of data available to the child learner. Native speakers may provide judgment involving ungrammatical examples, which are not available in language acquisition (Brown and Hanlon 1970; Marcus 1993). Behavioral experiments such as the Wug test and rating can also prove informative despite their limitations. The child learner, by contrast, has none of these tools available. To understand the mechanisms of morphological acquisition will require collaborative efforts among theoretical linguists, developmental psychologists and computer scientists in the years to come.

## References

- Albright, A. (2005). The morphological basis of paradigm leveling. In Downing, L. J., Hall, T. A., and Raffelsiefen, R., editors, *Paradigms in phonological theory*, pages 17–43. Oxford University Press.
- Albright, A. and Hayes, B. (2002). Modeling English past tense intuitions with minimal generalization. In *Proceedings of the ACL-02 Workshop on Morphological and Phonological Learning*, pages 58–69.

- Albright, A. and Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2):119–161.
- Allen, M. and Badecker, W. (2002). Inflectional regularity: Probing the nature of lexical representation in a cross-modal priming task. *Journal of Memory and Language*, 46(4):705–722.
- Anderwald, L. (2009). *The morphology of English dialects: Verb-formation in non-standard English*. Cambridge University Press.
- Armstrong, S. L., Gleitman, L. R., and Gleitman, H. (1983). What some concepts might not be. *Cognition*, 13(3):263–308.
- Atserias, J., Casas, B., Comelles, E., González, M., Padró, L., and Padró, M. (2006). Freeling 1.3: Syntactic and semantic services in an open-source nlp library. In *Proceedings of LREC*, volume 6, pages 48–55.
- Baerman, M., Corbett, G. G., and Brown, D., editors (2010). *Defective paradigms: Missing forms and what they tell us*. Oxford University Press, Oxford.
- Becker, M., Ketrez, N., and Nevins, A. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*, 87(1):84–125.
- Berko, J. (1958). The child’s learning of English morphology. *Word*, 14:150–177.
- Bloch, B. (1947). English verb inflection. *Language*, 23(4):399–418.
- Bowerman, M. (1982). Reorganizational process in lexical and syntactic development. In Wanner, E. and Gleitman, L. R., editors, *Language acquisition: The state of the art*, pages 319–346. Cambridge University Press.
- Brent, M., Murthy, S., and Lundberg, A. (1995). Discovering Morphemic Suffixes: A Case Study in Minimum Description Length Induction. In *Proceedings of the fifth international workshop on artificial intelligence and statistics*, pages 264–271.
- Brown, R. (1973). *A first language: The early stages*. Harvard University Press.
- Brown, R. and Hanlon, C. (1970). Derivational complexity and the order of acquisition in child speech. In Hayes, J. R., editor, *Cognition and the development of language*, pages 11–53. Wiley.
- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.
- Bybee, J. and Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59(2):251–270.

- Bybee, J. L. and Slobin, D. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58(2):265–289.
- Caprin, C. and Guasti, M. T. (2009). The acquisition of morphosyntax in Italian: A cross-sectional study. *Applied Psycholinguistics*, 30(01):23–52.
- Chan, E. (2008). *Structures and distributions in morphology learning*. PhD thesis, University of Pennsylvania.
- Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. MIT Press.
- Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22(6):991–1013.
- Clahsen, H., Avelledo, F., and Roca, I. (2002). The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language*, 29(3):591–622.
- Clahsen, H. and Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone corpus. In Meisel, J., editor, *The acquisition of verb placement*, pages 181–234. Kluwer.
- Clahsen, H. and Rothweiler, M. (1993). Inflectional rules in children’s grammars: Evidence from German participles. In Booij, G. and Marle, J., editors, *Yearbook of Morphology 1992*, pages 1–34. Springer.
- Creutz, M. and Lagus, K. (2005). Unsupervised morpheme segmentation and morphology induction from text corpora using Morfessor 1.0. Technical report, Helsinki University of Technology.
- Dąbrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. *Journal of Child Language*, 28(3):545–574.
- Deen, K. U. (2005). *The acquisition of Swahili*, volume 40 of *Language Acquisition & Language Disorders*. John Benjamins Publishing.
- Demuth, K. (2003). The acquisition of Bantu languages. In *The Bantu languages*, chapter The acquisition of Bantu languages, pages 209–222. Curzon Press.
- Eddington, D. (1996). Diphthongization in Spanish derivational morphology: An empirical investigation. *Hispanic Linguistics*, 8(1):1–13.
- Ervin, S. M. and Miller, W. R. (1963). Language development. *Child Psychology*, 62(1):108–143.
- Gerken, L., Wilson, R., and Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32(2):249–268.

- Goldsmith, J. (2001). Unsupervised Learning of the Morphology of a Natural Language. *Computational Linguistics*, 27(2):153–198.
- Goldsmith, J. (2006). An algorithm for the unsupervised learning of morphology. *Natural Language Engineering*, 12(4):353–371.
- Goldwater, S., Griffiths, T., and Johnson, M. (2006). Interpolating between types and tokens by estimating power-law generators. In *Advances in Neural Information Processing Systems 18*, pages 459–466.
- Gómez, R. L. and Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, 7(5):567–580.
- Guasti, M. T. (1993). Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition*, 3(1):1–40.
- Gxilishe, S., de Villiers, P., de Villiers, J., Belikova, A., Meroni, L., and Umeda, M. (2007). The acquisition of subject agreement in xhosa. In *Proceedings of the Conference on Generative Approaches to Language Acquisition (GALANA)*, volume 2, pages 114–23.
- Hafer, M. A. and Weiss, S. F. (1974). Word segmentation by letter successor varieties. *Information storage and retrieval*, 10(11):371–385.
- Hahn, U. and Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, 41(4):313–360.
- Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, 4(1):3–16.
- Halle, M. and Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Hale, K. and Keyser, S. J., editors, *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, pages 111–176. MIT Press, Cambridge.
- Halle, M. and Mohanan, K. P. (1985). Segmental phonology of Modern English. *Linguistic Inquiry*, 16(1):57–116.
- Harris, J. W. (1969). *Spanish phonology*. MIT Press, Cambridge.
- Harris, Z. S. (1955). From phoneme to morpheme. *Language*, 31(2):190–222.
- Hart, B. and Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, 27(1):4–9.
- Hay, J. and Baayen, R. H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences*, 9(7):342–348.
- Hayes, B., Zuraw, K., Siptár, P., and Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 85(4):822–863.

- Herman, L. and Herman, M. S. (2014). *American dialects: A manual for actors, directors, and writers*. Routledge.
- Jarmulowicz, L. (2002). English derivational suffix frequency and children’s stress judgments. *Brain and Language*, 81(1–3):192–204.
- Johnson, M., Griffiths, T., and Goldwater, S. (2007a). Bayesian inference for pcfgs via Markov chain Monte Carlo. In *Proceedings of NAACL HLT 2007*, pages 139–146.
- Johnson, M., Griffiths, T. L., and Goldwater, S. (2007b). Adaptor grammars: A framework for specifying compositional nonparametric bayesian models. In *Advances in Neural Information Processing Systems 19*, pages 641–648.
- Keshava, S. and Pitler, E. (2006). A simpler, intuitive approach to morpheme induction. In *Proceedings of The 2nd Pascal Challenges Workshop*, pages 31–35.
- Kuçera, H. and Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press.
- Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, 1(1):85–97.
- Legate, J. A. and Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, 14(3):315–344.
- Lignos, C., Chan, E., Marcus, M. P., and Yang, C. (2009). A rule-based unsupervised morphology learning framework. In *Working Notes for the CLEF 2009 Workshop*.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum.
- Maratsos, M. (2000). More overregularizations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen and Xu. *Journal of Child Language*, 27(1):183–212.
- Marcus, G., Pinker, S., Ullman, M. T., Hollander, M., Rosen, J., and Xu, F. (1992). *Overregularization in language acquisition*. Monographs of the Society for Research in Child Development. University of Chicago Press.
- Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, 46(1):53–85.
- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., and Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29(3):189–256.
- Marcus, M. P., Santorini, B., Marcinkiewicz, M. A., and Taylor, A. (1999). Treebank-3. Linguistic Data Consortium: LDC99T42.



- Maslen, R., Theakston, A. L., Lieven, E. V., and Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language and Hearing Research*, 47(6):1319–1333.
- Mayol, L. (2007). Acquisition of irregular patterns in Spanish verbal morphology. In *Proceedings of the twelfth ESSLLI Student Session*, pages 1–11.
- McClelland, J. L. and Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, 6(11):465–472.
- Monson, C. (2008). *ParaMor: From Paradigm Structure to Natural Language Morphology Induction*. PhD thesis, Carnegie Mellon University.
- Mooney, R. J. and Califf, M. E. (1995). Induction of first-order decision lists: Results on learning the past tense of English. *Journal of Artificial Intelligence Research*, 3(1):1–24.
- Morris, J. and Stockall, L. (2012). Early, equivalent ERP masked priming effects for regular and irregular morphology. *Brain and language*, 123(2):81–93.
- Myers, S. (1987). Vowel shortening in English. *Natural Language and Linguistic Theory*, 5(4):485–518.
- Naradowsky, J. and Goldwater, S. (2009). Improving morphology induction by learning spelling rules. In *Proceedings of the International Joint Conference on Artificial Intelligence*.
- O’Donnell, T. (2011). *Productivity and reuse in language*. PhD thesis, Harvard University.
- Parkes, C., Malek, A., and Marcus, M. (1998). Towards unsupervised extraction of verb paradigms from large corpora. In *Proceedings of the Sixth Workshop on Very Large Corpora*, pages 110–117.
- Phillips, C. (1995). Syntax at age two: Cross-linguistic differences. *MIT Working Papers in Linguistics* 26, pages 325–382.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. MIT Press.
- Pinker, S. (1995). Why the child holded the baby rabbit: A case study in language acquisition. In Gleitman, L. R. and Liberman, M., editors, *An invitation to cognitive science, Vol. 1: Language*, pages 107–133. MIT Press.
- Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books.
- Pinker, S. and Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28(1):73–193.

- Pinker, S. and Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Science*, 6(11):456–463.
- Pizzuto, E. and Caselli, M. C. (1994). The acquisition of Italian verb morphology in a cross-linguistic perspective. In *Other children, other languages: Issues in the theory of language acquisition*, pages 137–187. Lawrence Erlbaum Associates.
- Plunkett, K. and Juola, P. (1999). A connectionist model of english past tense and plural morphology. *Cognitive Science*, 23(4):463–490.
- Prasada, S. and Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8(1):1–56.
- Prince, A. and Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. MIT Press.
- Real Academia Española (1992). *Diccionario de la lengua española*. Real Academia Española, Madrid, 21st edition.
- Rumelhart, D. E. and McClelland, J. L. (1986). On learning the past tenses of English verbs. In McClelland, J. L., Rumelhart, D. E., and the PDP Research Group, editors, *Parallel distributed processing: Explorations into the microstructure of cognition. Volume 2: Psychological and biological models*, pages 216–271. MIT Press.
- Schütze, C. T. (2005). Thinking about what we are asking speakers to do. In Kepser, S. and Reis, M., editors, *Linguistic evidence: Empirical, theoretical, and computational perspectives*, pages 457–485. Mouton de Gruyter.
- Sims, A. D. (2006). *Minding the gap: Inflectional defectiveness in a paradigmatic theory*. PhD thesis, Ohio State University.
- Spiegler, S., Golénia, B., and Flach, P. (2009). Promodes: A probabilistic generative model for word decomposition. In *Working Notes of the 10th Workshop of the Cross-Language Evaluation Forum (CLEF2009)*.
- Stockall, L. and Marantz, A. (2006). A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, 1(1):85–123.
- Taylor, J. R. (2003). *Linguistic categorization*. Oxford University Press.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3):209–253.
- Tomasello, M. (2003). *Constructing a language*. Harvard University Press.
- Westfal, S. (1956). *A study in Polish morphology: The genitive singular masculine*, volume 8 of *Slavic Pringints and Reprintings*. Mouton and Co.

- Weyerts, H. and Clahsen, H. (1994). Netzwerke und symbolische Regeln im Spracherwerb: experimentelle Ergebnisse zur Entwicklung der Flexionsmorphologie. *Linguistische Berichte*, 154:430–460.
- Wicentowski, R. (2002). *Modeling and learning multilingual inflectional morphology in a minimally supervised framework*. PhD thesis, Johns Hopkins University.
- Wiese, R. (1996). *The phonology of German*. Clarendon Press.
- Wiese, R. (1999). On default rules and other rules. *Behavioral and Brain Sciences*, 22(6):1043–1044.
- Wunderlich, D. (1999). German noun plural reconsidered. *Behavioral and Brain Sciences*, 22:1044–1045.
- Xu, F. and Pinker, S. (1995). Weird past tense forms. *Journal of Child Language*, 22(3):531–556.
- Yang, C. (2002). *Knowledge and learning in natural language*. Oxford University Press.
- Yang, C. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences*, 8(10):451–456.
- Yang, C. (2005). On productivity. *Linguistic Variation Yearbook* 5, pages 333–370.
- Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, 110(16):6324–6327
- Yip, K. and Sussman, G. J. (1997). Sparse representations for fast, one-shot learning. In *Proceedings of the National Conference on Artificial Intelligence*, pages 521–527.
- Zhang, J. and Lai, Y. (2010). Testing the role of phonetic knowledge in mandarin tone sandhi. *Phonology*, 27(1):153–201.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley.