Price of Productivity:

How Children Learn and Break Rules of Language

Charles Yang

University of Pennsylvania

Janurary 2016 for the MIT Press

Contents

Chapter 1

Border Wars

It seems preposterous to write a whole book about an equation (which can be found on page 16). So let me begin by laying out the problems that the equation is designed to solve, how these problems are important, and why the solution must be in the form of an equation.

There is no doubt that human language evolved as a biological capacity. It is also almost certainly the case that language emerged relatively recently and suddenly: perhaps no more than 150,000 years ago (Tattersall 2012), a blink of the eye in evolutionary terms. This puts a premium on *Darwin's Problem* (Hornstein 2009, Bolhuis et al. 2014): How to situate language, from the perspective of behavior, brain, and evolution, within the human cognitive and perceptual system which must be shared in part with other species and lineages?

Speculations abound within and without generative linguistics; see Hauser et al. 2014 for a critical assessment of the current literature. But no amount of evolutionary musing should distract us from the more traditional, and much more tangible, goals of language sciences. A theory of language needs to be sufficiently elastic to account for the complex patterns in the world's languages but at the same time sufficiently restrictive so as to guide children toward successful language acquisition in a few short years (Chomsky 1965). Only then does Darwin's Problem arise: as a statement of human biology, a theory of language can only include evolutionary innovations that would have been plausible in the extremely brief history of *Homo Sapiens*.

The current project deals with the boundary issues between language and cognition with an

eye on evolution. In one sense, it is the continuation of earlier research. The variational approach to language (Yang 2002, 2004) suggests that children use general learning mechanisms to navigate within the hypothesis space provided by Universal Grammar (UG); in doing so, it dispenses with domain-specific learning models long known to be problematic. Along a similar line, the present work develops a theory of linguistic representation, learning, and use that, once again, shifts the explanatory burden from Universal Grammar to factors external to language-–with specific considerations of computational complexity. In another sense, however, the current study provides an amendment to the variational framework. It investigates the extent to which key properties of language, rather than being built within, can be attributed to children's ability to derive generalizations from the linguistic data. A reduced load for the genetic endowment of language promises a more viable solution to Darwin's Problem but will inevitably exacerbates Plato's Problem (Chomsky 1986): How does the child acquire his/her knowledge of language, which is grossly under-determined by experience? The answer is partially provided by Universal Grammar but we cannot be asking for too much — at least no more than what evolution could conceivably offer in the very recent past.

Let's go straight to the heart of matter: the recursive composition of hierarchical structures ("Merge"; Chomsky 1995), the giant leap forward in the evolution of language and cognition (Hauser et al. 2002). What are the behavioral, and ultimately evolutionary, benefits of a combinatorial system over a finite inventory of fixed expressions, especially as we now know about brain's enormous capacity for storage? How does a simple and elegant computational procedure square with the manifest arbitrariness and idiosyncrasies across languages? The Société de Linguistique de Paris once issued a moratorium on the origin of language; indeed, a credible account of language and its place in cognition and evolution must be grounded firmly in empirical materials. As it happens, our evolutionary reflections turn up some old and unsettled scores in the study of language. And that's well and good. A minimalist UG is only convincing if it engages with, and provides convincing solutions to, the everyday problems that concern working linguists: Le biologiste passe, la grenouille reste.

1.1 How grammars leak

As Edward Sapir once famously noted, all grammars leak. In less colorful terms, all languages have exceptions that exist side by side with overarching rules and regularities. But lest this banal observation overshadow Sapir's main message: without a grammar, there will be no leaks to plug.

> It is obvious that a language cannot go beyond a certain point in this randomness. Many languages go incredibly far in this respect, it is true, but linguistic history shows conclusively that sooner or later the less frequently occurring associations are ironed out at the expense of the more vital ones. In other words, all languages have an inherent tendency to economy of expression. Were this tendency entirely inoperative, there would be no grammar. The fact of grammar, a universal trait of language, is simply a generalized expression of the feeling that analogous concepts and relations are most conveniently symbolized in analogous forms. Were a language ever completely grammatical, it would be a perfect engine of conceptual expression. Unfortunately, or luckily, no language is tyrannically consistent. All grammars leak. (Sapir 1928: p38-39)

The fact of exceptions, then, should not deter the linguist from formulating theories about the systematic properties of language. When we evoke labels such as *diacritics*, *irregularity*, and *lexicalization* — which can be found in every theorist's toolkit — we are simultaneously committing to a grammar, one which is associated with *basic* word order, *unmarked* forms, and *default* rules.

> It is quite obvious that many of the phonological rules of the language will have certain exceptions which, from the point of view of the synchronic description, will be quite arbitrary. This is no more surprising than the fact that there exist strong verbs or irregular plurals. Phonology, being essentially a finite system, can tolerate some lack of regularity (exceptions can be memorized); being a highly intricate system, resulting (very strikingly, in a language like English) from diverse and interwoven historical processes, it is to be expected that a margin of irregularity will persist in

> almost every aspect of the phonological description. Clearly, we must design our linguistic theory in such a way that the existence of exceptions does not prevent the systematic formulation of those regularities that remain. (Chomsky and Halle 1968: p172)

Not everyone agrees. Language scientists and engineers have been wrestling with leaky grammars ever since Sapir. If exceptions are idiosyncratic and must be somehow committed to memory, why not relegating all of language to storage, as the brain is capable of retaining vast quantities of information? To wit, the so-called past tense debate, to which we return repeatedly in the following pages, has been a struggle over whether *some* verbs (the irregulars; Pinker and Ullman 2002) or *all* verbs (the regulars as well; McClelland and Patterson 2002) are organized associative memory: the latter position would dispense with Sapir's "symbolized" rules altogether.

The controversy over exceptions intensified further when generative grammar moved from language specific rules and constructions to universal principles and constraints (Chomsky 1981). If exceptions already pose a serious challenge to the study of particular grammars, how do they figure into a theory that commits to an innate, universal, and invariant predisposition for language?

> ... it is reasonable to suppose that UG determines a set of core grammars and that what is actually represented in the mind of an individual even under the idealization of a homogeneous speech community would be a core grammar with a periphery of marked elements and constructions.
>
> Viewed against the reality of what a particular person may have inside his head, core grammar is an idealization. From another point of view, what a particular person has inside his head is an artifact resulting from the interplay of many idiosyncratic factors, as contrasted with the most significant reality of UG (an element of shared biological endowment) and the core grammar (one of the systems derived by fixing the parameters of UG in one of the permitted ways). (Chomsky 1981: p8)

Exceptions may have been put in their proper place — the periphery — but they have not exactly gone away. Chomsky's formulation brings into focus the problem of language acquisition. When constructing a theory of grammar, linguists have at their disposal a plethora of tools to disentangle the core from the periphery: grammaticality judgments, corpus statistics, historical documents, and an ever expanding arsenal of experimental methods. And they still disagree over the proper partitioning. How does a young child steer clear of the peripheral idiosyncrasies to acquire a core grammar, all in a few short years? Exceptions are defined in opposition of the grammar, but the acquisition data does not arrive wearing "core" or "periphery" on its sleeves: the learner seems to have a perfect chicken-and-egg problem.

The core vs. periphery problem was very much the focus of learnability research in the 1980s. We will review this work in the following pages but it is fair to say that no widely accepted solution has been produced. In a recent paper, Sag (2010) summarizes the state of the affairs from the perspectives of a skeptic:

> But how are we to know which phenomena belong to the core and which to the periphery? The literature offers no principled criteria for distinguishing the two, despite the obvious danger that without such criteria, the distinction seems both arbitrary and subjective. The bifurcation hence places the field at serious risk of developing a theory of language that is either vacuous or else rife with analyses that are either insufficiently general or otherwise empirically flawed. There is the further danger that grammatical theories developed on the basis of "core" phenomena may be falsified only by examining data from the periphery–data that falls outside the domain of active inquiry. (Sag 2010: p487)

A possible course to follow is to abandon the core vs. periphery distinction. There is a detectable continuity from Gross's taxonomy of French verbs (1975, 1979) to Sag's radically lexicalized treatment of movement dependencies (2010), from Lakoff's irregular syntax (1970) to present-day Construction Grammars, a network of "stored pairings of form and function" that constitutes the totality of linguistic knowledge (Goldberg 2003: p219). Similarly, according to

the usage-based approach to language acquisition, children do not make use of a systematic grammar; rather, "they sometimes have a set expression readily available and so they simply retrieve that expression from their stored linguistic experience" (Tomasello 2000b: p77). In a wide-ranging study, Culicover (1999) investigates numerous syntactic constructions that cannot be attributed to the core parameter system. (Or shouldn't be, for that would require an enormous number of parameters, defeating the very purpose of the parameter as compact descriptions of disparate phenomena.) He proposes that language acquisition follows inductive methods, where the learner draws generalizations over the entire range of language data. No dichotomy between the core and the periphery is supposed, and the child learner has no chicken-and-egg dilemma.

I for one am not quite ready to give up the core. Formal results have consistently shown that a constrained hypothesis space remains the most promising solution to the general problem of learning, of which language acquisition is a special case (Valiant 1984, Vapnik 2000, Nowak et al. 2002, Sakas and Fodor 2012). Additionally, when children's language deviates from the input, it nevertheless remains in a restrictive range of possibilities (Crain 1991, Yang 2002), which further supports the conception of the core as a highly structured system. As will be reviewed in Chapter 2, computational and quantitative studies of language suggest that the role of linguistic storage has been greatly exaggerated, and that there is clearcut evidence from child language for a categorical distinction between rules and exceptions, and between the core and the periphery.

But on the resolution of the boundary dispute, I am in agreement with the critics. It is no longer advisable to dodge the question. While some of the purported peripheral idiosyncrasies might only be apparent, it is no longer sufficient to point out how a core-less approach misses important empirical generalizations, or fails to provide a plausible solution to the problem of acquisition. Since not all aspects of language are plausibly innate — the "add -*d*" rule for English past tense, for instance — some kind of data-driven inductive learning is absolutely necessary. A positive answer must be given so that the boundary between the core and the periphery can be drawn, at least for theorists who would like to maintain such a distinction.

1.2 Where core meets periphery

Like most researchers, I started at the core only to be driven to the periphery. In Yang 2002, I developed the variational learning framework for language acquisition and change. The variational model holds language learning to be a probabilistic process: the child has a statistical distribution over the space of possible grammars (or parameter values), and it is this distribution that changes in response to linguistic data. As learning proceeds, the child will access the target grammar with increasing probability, while the non-target but linguistically possible grammars may still be used, albeit with decreasing probabilities. The competition scheme results in children's occasional but systematic deviation from the target grammar, which will be left standing in the end.

In many ways, the variational learning model is an improvement over traditional *transformational learning* models of which the triggering learning algorithm (Gibson and Wexler 1994) is the paradigm example. Under the transformational scheme, the learner is identified with a single grammar in the hypothesis space (Chomsky 1965, Wexler and Culicover 1980, Berwick 1985). The current grammar is abandoned if it fails to analyze an input utterance, and a new grammar is adopted instead. As pointed out by many researchers of child language (e.g., Bloom 1990, Randall 1990, Valian 1991, Niyogi and Berwick 1996), the triggering model is vulnerable to noise: after years of patient navigation, the learner's grammar may be undone by a single ungrammatical utterance. The variational model, which regards learning as probabilistic, can robustly countenance a certain level of noise. Instead of having a probability of 1 for a parameter value, the learner may settle on 0.99, reserving a noisy margin of 1%.

Variational learning is well equipped to handle noise — and only noise. It does not have the appropriate mechanism for distinguishing noise from exceptions. To take a concrete example, English ceased to be a verb raising language in the Middle English period (Ellegård 1953, Kroch 1989) and now employs periphrastic auxiliaries in question formation. Yet the primary linguistic data does contain instances of main verb raising. (1) is a well known nursery rhyme, which dates back to the 1700s when the loss of verb raising was already near completion.

(1) Baa baa black sheep have$_t$ you ___$_t$ any wool?

Suppose (1) and similar sentences appear in 1% of the utterances that a child receives. The variational model is straightforwardly applicable, except it will get the facts wrong. The child will converge on a stable combination that raises the main verb 1% of the time as in (1) while uses an auxiliary for the rest. But this is the correct numerical distribution but a wrong structural one. Unless the learner identifies that (1) is an exceptional pattern restricted to specific contexts (e.g., negative inversion, this particular nursery rhyme, etc.), it will raise the main verb across all contexts, albeit with a low probability of 0.01. No English-learning children ever go through a stage of main verb inversion: indeed, main verb inversion is completely unattested in child English.

Without the ability to recognize exceptions, the learner will have difficulty setting the syntactic parameters. And this is not only a problem for the language acquisition specialist. In recent years, parameters have fallen on hard times because they do not appear as clean and elegant as originally conceived. In a well known critique of the parameter-based approach to language variation, Newmeyer (2004) considers exceptions to be an insurmountable challenge. For instance, while French generally places the adjective after the noun (*un livre noir* 'a black book'), there is a special class of adjectives that must appear before it, as shown especially clearly in the contrast between (2b) and (2c):

(2) a. une *nouvelle* maison 'a new house'

b. une *vieille* amie 'a friend for a long time'

c. une amie *vieille* 'a friend who is aged'

Additional parameters may be introduced to accommodate the mixed system in French, but the conceptual and the learning problem will not go away. The French facts can be described as one parameter plus lexical exceptions, or two parameters, one for the majority of adjectives and the other for a lexicalized subset, where there may be additional structural patterns within (e.g., Cinque 1994). Either way, the French-learning child needs to keep them somehow separate: the general pattern of nominal-adjective order should be established despite the counterexamples in (2b). Similarly, the English-learning child should not allow the occasional and contextually

restricted omission of the subject ("mix flour with spices", "had a rough day"; Haegeman 1990) to interfere with the setting of the obligatory-subject parameter. Even more challenging would be a system like Modern Hebrew, which is essentially a null-subject language for first and second person in past and future tense but an obligatory subject language for the rest of the person and tense combinations (Jaeggli and Safir 1989, Shlonsky 2009). Such mixed parametric systems are prima facie evidence against the parameter as overarching "global" properties of languages.

But it is not clear that abandoning parameters is going to help. Suppose one pursues, following Newmeyer (2004), a rule-based approach to syntax. The problem of exceptions remains exactly the same: How does the French learner acquire the *rule*, rather than the *parameter*, that adjectives in general follow the nominal except for those on a finite list (2)? Likewise, it only begs the question to reformulate parameters into a hierarchy of specificity (Baker 2001, Holmberg 2010, Biberauer et al. 2010, Roberts 2012) where some are general, some are construction specific and still others pertain to individual lexical items. Again, how does a French-learning child know which *vielle* — (2b) vs. (2c) — goes with the restricted parameter and which goes with the general parameter? If pursued to its logical limit, this approach becomes a completely lexicalized theory of language that lists everything (e.g., Sag 2010), which is neither theoretically satisfying nor, as we shall see in Chapter 2, empirically sufficient.

The alternative route, then, is to salvage the grammar from exceptions. That such a boundary is difficult to draw does not, of course, mean that it does not exist; see Cohn 2006 on the similar conundrum at the juncture between phonetics and phonology. One approach is to reinspect the exceptions: perhaps their idiosyncrasies ought to be assimilated to the core after all (see, e.g., Fodor 2001 for a direct response to some of "nut" cases studied by Culicover 1999). Another approach, one taken here, is to develop a principled demarcation between the core and the periphery, with the recognition that some exceptions are truly accidental and irreducible to general principles (Chomsky and Halle 1968). Such an approach is feasible because children are remarkably unfussed by the core vs. periphery problem that has troubled linguists. As will be extensively reviewed in Chapter 2 and throughout the book, children are very good at recognizing exceptions at every linguistic level, and manage to keep them separate from the core gram-

mar. To wit: Hebrew-learning children can partition the language into two parametric systems from the outset of language acquisition (Levy and Vainikka 2000). For the null subject component, they behave like children acquiring prototypical *pro*-drop languages such as Italian (Valian 1990). By contrast, for the obligatory-subject portion, they behave like children acquiring prototypical obligatory-subject languages like English (Valian 1990, Wang et al. 1992, Yang 2002), who go through the characteristic stage of subject omission for to 3 years.

It is children's remarkable mastery of language that gives hope to the theorist: there *must* be a principled division between the grammar and its leaky corners. The current study is a proposal of where the boundary should be drawn.

(3) **Tolerance Principle**:

If R is a productive rule applicable to N candidates, then the following relation holds between N and e, the number of exceptions that could but do not follow R:

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

The motivation for the Tolerance Principle will be laid out in detail in Chapter 3, where we develop a calculus for the price of productivity. The analogue to economics, and hence the title of the current work, is deliberate and I believe appropriate. Just as the price of goods is determined by the balance between supply and demand, we suggest that the price of linguistic productivity arises from the quantitative considerations of rules and exceptions.[1] Drawing extensively from the psycholinguistic literature, I show that exceptions to a rule impose costs to the real-time processing of language. Specifically, the learner postulates a productive rule only if it results in a more efficient organization of language, as measured in processing time, than listing everything in lexical storage. The Tolerance Principle asserts that for a rule to be productive, the number of exceptions must fall below a critical threshold.

We envision language learning as a search for productive generalizations. The child consid-

[1] Here I use the term "rule" to refer any kind of linguistic generalization that has the potential of open-ended application. In later chapters, I provide explicit formulation of rules, often the output of computational learning models, and quantify their productivity with the Tolerance Principle.

ers a rule R in her language and evaluates its productivity according to the associated numerical values: N and e, the number of items to which the rule is applicable, and the number of items which defy the rule. The rule is accepted as productive if e is sufficiently small; otherwise the learner formulates a revised rule (R') to obtain a new set of values (N' and e') and the Tolerance Principle is applied recursively, as illustrated in Figure 1.1. The core grammar must be able to tolerate a suitably small quantity of exceptions, which will be exiled to the periphery.

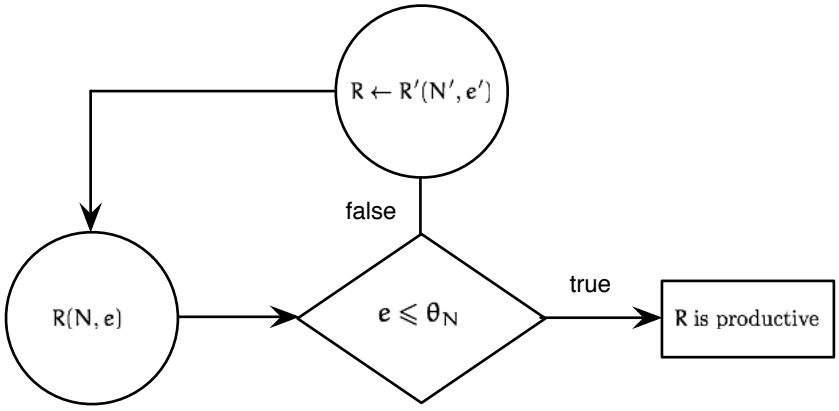

Figure 1.1: The Tolerance Principle as an evaluation measure in language acquisition.

1.3 Some outstanding problems

In lieu of a roadmap for the materials to come, let me highlight some representative case studies that fall under the purview of the Tolerance Principle.

The mystery of word formation

Just as there are infinitely many sentences, the number of words is also unbounded. Morphology makes compositional use of elemental units but some processes are clearly open-ended while others are severely restrictive. For instance, the English nominalization suffix *-ness* can apply to a broad range of adjectives (*red-redness*) while the suffix *-th* is restricted to only a handful of stems (e.g., *warm-warmth, wide-width*). Productivity has long been recognized as one of "the central mysteries" in morphology (Aronoff 1976: p35). From the perspective of language ac-

quisition, even innocuous cases become learning puzzles. Take the agentive suffix *-er*, which is unquestionably productive: *teach-teacher*, *drink-drinker*, and more recently, *blog-blogger*. But its productivity can mislead to a garden path once we view language learning as a mechanical process. Somehow the child learner must recognize that *rubber* is not for rubbing, *letter* is not someone who lets, and *counter* is (usually) not something that keeps track of numbers. In other words, to learn that *-er* is productive, the learner must tune out the spurious misapplications embedded in his/her linguistic experience.

When the majority doesn't rule

Thanks to its unique linguistic history, English has a metrical stress system that has become a fertile ground for phonological research (Chomsky and Halle 1968, Liberman and Prince 1977, Hayes 1982, Halle and Vergnaud 1987, Idsardi 1992, Burzio 1994, Halle 1998). Statistically, however, the English stress system is remarkably simple: about 85% of spoken English words place primary stress on the initial syllable (Cutler and Carter 1987, Legate and Yang 2013). This may tempt the learner to postulate a simple default rule that stresses the initial syllable, much like the "quantity insensitive" metrical systems found in languages such as Afrikaans and Chitimacha (Gordon 2002): the residual 15% or so can be lexically stored, resulting in a highly respectable batting average.

But evidently 85% isn't good enough, at least not for long. There is a brief, and very early, developmental stage during which English-learning infants take the stressed syllable as the beginning of a word (Echols et al. 1997, Jusczyk et al. 1999, Johnson and Jusczyk 2001), as if they treat the language as invariably stress-initial. However, production studies give unambiguous evidence that by no later than 2;5, children are already taking syllable weight into account when assigning stress to words (Kehoe 1997; see Fikkert 1994 for similar findings in Dutch.) Furthermore, no theoretical analysis of English stress seems to take the quantity-insensitive option, which would at least be a great statistical success. Behaviorally, both adults and children take syllable weights as well as lexical categories of words into account to stress novel words (Baker and Smith 1976, Kelly 1992, Guion et al. 2003, Oh et al. 2011). It is evident, then, an overwhelming statistical advantage does not necessarily translate into productivity — so what does?

Phase transitions in child language

In a celebrated experiment (Berko 1958), young children are shown to use inflectional morphology productively with nonce words (*wug-wugs, rick-ricked*). But these results need to be considered alongside findings that morphological rules do not become productive overnight. In an early study, MacWhinney (1975) shows that the development of morphology starts with rote-memorization before the child discovers the productive processes of word formation in their native language. In the acquisition of English, it has been observed that children typically follow a U-shape learning curve (Marcus et al. 1992, Pinker 1999): irregular verbs are inflected correctly (*hold-held*) early on before succumbing to over-regularization (*hold-*holded*) at a later point, which signals the onset of the productive "add -*d*" rule. Apparently children need time and data to accumulate enough regular verbs to counterbalance the irregular exceptions. The trajectory of English past tense learning is quite typical when considered in the cross-linguistic study of language development (Chapter 2). In many (but not all) cases of language acquisition, children show an initial stage of conservatism not beyond the input before the emergence of productive generalizations. How do they calibrate the balance between rules and exceptions? Where is the critical juncture at which children recognize the productivity of rules?

Defective words

If the *Wug* test puts the unbounded creativity of grammar in the spotlight, then the ineffables in language must be an awkward blemish. In a classic paper, Halle (1973) draws attention to morphological "gaps", the absence of inflected words for no apparent reason. For instance, there are about seventy verbs in Russian which lack an acceptable first person singular non-past form (data from Halle 1973: p7 and Sims 2006).

(4) *lažu 'I climb'

*pobežu (or *pobeždu) 'I conquer'

*deržu 'I talk rudely'

*muču 'I stir up'

*erunžu 'I behave foolishly'

There is nothing in the phonology or semantics of these words that could plausibly account for their illicit status, yet native speakers regard them as ill-formed. Indeed, defective paradigms such as (4) are hardly rare (Baerman et al. 2010), even in a morphologically impoverished language such as English: for example, speakers are unsure about the past participle of *stride* (Pullum and Wilson 1977, Pinker 1999). Missing inflections pose considerable challenges for the theories of morphology but a fundamental piece of the puzzle belongs to language acquisition: How does the learner know that *some* expected forms are impossible while the combinatorial use of language is in general unimpeded? Where are gaps expected to appear? In other words, upon the presentation of linguistic data, the child must deploy a decision procedure that detects productive regularities if present, and comes up empty handed when absent.

Learning what not to say

C. L. Baker, in a well-known study (1979), raises the problem of indeterminate inference in language learning. Of a range of examples he discusses, the English dative alternations have become most prominent.

(5) a. John gave a dish to Sam.

John gave Sam a dish.

b. John passed the salami to Fred

John passed Fred the salami.

c. John told a joke to Mary.

John told Mary a joke.

d. John donated a painting to the museum.

*John donated the museum a painting.

e. *John confessed the police the crime.

John confessed the crime to the police.

The double object and prepositional *to*-dative constructions seem interchangeable in the first three examples but the failure for *donate* and *confess* to do so is unexpected given the semantic similarities of the verbs. The absence of negative evidence in language acquisition (Brown

and Hanlon 1970, Marcus 1993) has led to a considerable body of literature on the acquisition of negative linguistic constraints: How does the child know that only some, but not all, unattested forms are ungrammatical? Furthermore, as will be reviewed in Chapter 6, the distribution of datives across languages (Chung 1998, Levin 2008) and the developmental trajectory of dative acquisition (Gropen et al. 1989, Conwell and Demuth 2007) suggest that these constructions cannot be accounted for solely by innate constraints of syntactic and semantic structures. First, children patiently accumulate evidence about the dative verbs and do not go beyond the adult input (Snyder and Stromswold 1997). Then, very much like the emergence of the "add *-d*" rule, they pounce on a productive rule: in fact, an overly general one which results in errors such as "I said her no" (Gropen et al. 1989, Bowerman and Croft 2008), where *say* was appropriated in the double object construction. These errors gradually disappear as the learner grasps the finer details of the datives. Again, we see a learning process in which the properties of specific items are extended to an entire class: "I texted him an apology" became available as soon as *text* became a verb. At the same time, the grammaticality contrast in (5) suggests that these generalizations must be appropriately constrained.

Variation, stability, and change

A common (non)response to the problem of exceptions is to appeal to individual variation: as Labov (1972a: p292) laments, "'My dialect' turns out to be characterized by all the sentence types that have been objected to by others." Some exceptional patterns in language may indeed be a matter of individual and/or dialect variation but they still require a principled explanation. Everyone learns the "*-d*" rule of English past tense as a child, but some learners do so a full year ahead of others (Maratsos 2000, Yang 2002). We cannot understand variability in language acquisition, which has become a major focus in recent years (Hart and Risley 1995), unless we understand how children learn languages in the first place.

Much like the debate over core vs. periphery, there has been a long-standing controversy on the role of rules and exceptions in language change. Is the rise of a linguistic form due to the reorganization of a rule system that systematically applies across the board (Kiparsky 1965, Kroch 1989, Labov 1994), or does it proceed by an item-by-item and construction-by-construction ba-

sis (Wang 1969, 1979, Hudson 1997, Bresnan and Ford 2010)? If there is a division between rules and exceptions, then there must be mechanisms that allow the boundary to blur, as languages are always in a flux of variation and change.

******

These and many other problems will form the empirical ground of the present study. We chose these topics not only because they are amenable to numerical analysis — a necessity for the Tolerance Principle, and the equation in (3) will be used almost 100 times — but also because they are traditionally treated with UG-internal solutions. For instance, morphological productivity has been connected to general syntactic principles (e.g., Marantz 2001), the dative constructions are proposed to follow universal syntactic and semantic constraints mediated by innate linking relations (e.g., Pinker 1989), paradigmatic gaps are produced by shielding specific words from the general rules of language (e.g., Halle 1973). We will not have space to provide detailed assessment of these proposals: rather, we show that they are dispensable and indeed should be dispensed with. The empirical problems are well handled, and indeed unified, by an independently motivated principle of learning: we can do more with less UG, thereby taking a step closer to both Darwin's and Plato's Problem.

But first, let us understand why the core is worth saving, and why it is ill-advised to focus on exceptions at the expense of rules.

Bibliography

Akhtar, N. and Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental psychology*, 33(6):952–965.

Albright, A. (2002). Islands of reliability for regular morphology: Evidence from Italian. *Language*, 78(4):684–709.

Albright, A. (2003). A quantitative study of Spanish paradigm gaps. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, pages 1–14. Cascadilla, Somerville, MA.

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, Oxford.

Albright, A. (2006). Cautious generalization of inflectional morphology, and its role in defectivity. Paper presented at Defective Paradigms.

Albright, A. (2009). Lexical and morphological conditioning of paradigm gaps. In Rice, C. and Blaho, S., editors, *Modeling ungrammaticality in Optimality Theory*, pages 117–164. Equinox, London.

Albright, A., Andrade, A., and Hayes, B. (2001). Segmental environments of Spanish diphthongization. In Albright, A. and Cho, T., editors, *UCLA Working Papers in Linguistics 7: Papers in Phonology 5*, pages 117–151. Cascadilla, Somerville, MA.

Albright, A. and Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2):119–161.

Alegre, M. and Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40(1):41–61.

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.

Ambridge, B., Kidd, E., Rowland, C. F., and Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of child language*, 42(02):239–273.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98(3):409.

Anderson, S. R. (1969). *West Scandinavian vowel systems and the ordering of phonological rules*. PhD thesis, MIT.

Anderson, S. R. (1974). *The organization of phonology*. Academic Press, New York.

Anderson, S. R. (1988a). Morphological change. In Newmeyer, F. J., editor, *Linguistics: The Cambridge survey*, pages 146–191. Cambridge University Press, Cambridge.

Anderson, S. R. (1988b). Morphology as a parsing problem. *Linguistics*, 26(4):521–544.

Anderson, S. R. (1992). *A-morphous morphology*. Cambridge University Press, Cambridge.

Anderson, S. R. (2010). Failing one's obligations: Defectiveness in Rumantsch reflexes of DĒBĒRE. In Baerman, M., Corbett, G. G., and Brown, D., editors, *Defective paradigms: Missing forms and what they tell us*, pages 19–34. Oxford University Press, Oxford.

Anderson, S. R. (2015). Morphological change. In Bowern, C. and Evans, B., editors, *The Routledge Handbook of Historical Linguistics*, pages 264–285. Routledge, New York.

Anderwald, L. (2009). *The morphology of English dialects: Verb-formation in non-standard English*. Cambridge University Press.

Anderwald, L. (2013). Natural language change or prescriptive influence?: Throve, dove, pled, drug and snuck in 19th-century american english. *English World-Wide*, 34(2):146–176.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis, with commentary by George A. Miller and Pamela C. Wakefield. *Monographs of the society for research in child development*, pages i–186.

Angluin, D. (1980). Inductive inference of formal languages from positive data. *Information and control*, 45(2):117–135.

Anshen, F. and Aronoff, M. (1988). Producing morphologically complex words. *Linguistics*, 26(4):641–656.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory*, 20(1):1–42.

Armstrong, S. L., Gleitman, L. R., and Gleitman, H. (1983). What some concepts might not be. *Cognition*, 13(3):263–308.

Aronoff, M. (1976). *Word formation in generative grammar*. MIT Press, Cambridge.

Aronoff, M. (1994). *Morphology by itself: stems and inflectional classes*. MIT Press, Cambridge.

Aronoff, M. (2015). Competition and the body: Two non-computational factors in the organization of linguistic morphological systems. In *Mayfest workshop on morpholgy*, College Park, MD. University of Maryland.

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.

Baayen, R. H. (1989). *A corpus-based approach to morphological productivity: Statistical analysis and psycholinguistic interpretation*. PhD thesis, Vrije Universiteit Amsterdam.

Baayen, R. H. (2009). Corpus linguistics in morphology: morphological productivity. In Ludeling, A. and Kyto, M., editors, *Corpus linguistics. An international handbook*, pages 900–919. Mouton de Gruyter.

Baayen, R. H., Dijkstra, T., and Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37(1):94–117.

Baayen, R. H. and Lieber, R. (1991). Productivity and English derivation: a corpus-based study. *Linguistics*, 29(4):801–843.

Baayen, R. H., McQueen, J. M., Dijkstra, T., and Schreuder, R. (2003). Frequency effects in regular inflectional morphology: Revisiting Dutch plurals. In *Morphological structure in language processing*, pages 355–390. Mouton, Berlin.

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1996). CELEX2. Linguistic Data Consortium: LDC96L14.

Baayen, R. H. and Renouf, A. (1996). Chronicling the times: Productive lexical innovations in an english newspaper. *Language*, pages 69–96.

Baerman, M. (2008). Historical observations on defectiveness: the first singular non-past. *Russian Linguistics*, 32(1):81–97.

Baerman, M. and Corbett, G. G. (2010). Defectiveness: Typology and diachrony. In Baerman, M., Corbett, G. G., and Brown, D., editors, *Defective paradigms: Missing forms and what they tell us*, pages 19–34. Oxford University Press, Oxford.

Baerman, M., Corbett, G. G., and Brown, D., editors (2010). *Defective paradigms: Missing forms and what they tell us*. Oxford University Press, Oxford.

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry*, 10(4):533–581.

Baker, M. (1988). *Incorporation: A theory of grammatical function changing*. University of Chicago Press, Chicago.

Baker, M. (2001). *The atoms of language: The mind's hiddne rules of grammar*. Basic Books, New York.

Baker, R. and Smith, P. (1976). A psycholinguistic study of english stress assignment rules. *Language and Speech*, 19(1):9–27.

Bakovic, E. (2013). *Blocking and complementarity in phonological theory*. Equinox Publishing.

Balota, D. A. and Duchek, J. M. (1988). Age-related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging*, 3(1):84.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., and Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3):445–459.

Baptista, B. O. (1984). English stress rules and native speakers. *Language and speech*, 27(3):217–233.

Baroni, M. (2009). Distributions in text. In Lüdeling, A. and Kyöto, M., editors, *Corpus linguistics: An international handbook*, pages 803–821. Mouton de Gruyter.

Baroni, M. and Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, 36(4):673–721.

Baronian, L. (2005). *North of phonology*. PhD thesis, Stanford University.

Baronian, L. and Kulinich, E. (2012). Paradigm gaps in whole word morphology. *Irregularity in morphology (and beyond). Studia typologica*, 11:81–100.

Bartke, S., Rösner, F., Streb, J., and Wiese, R. (2005). An ERP-study of German 'irregular' morphology. *Journal of Neurolinguistics*, 18(1):29–55.

Bauer, L. (1983). *English word-formation*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.

Bauer, L. (2001). *Morphological productivity*. Cambridge University Press, Cambridge.

Beard, R. (1995). *Lexeme-morpheme base morphology*. SUNY Press, Albany.

Beck, S. and Johnson, K. (2004). Double objects again. *Linguistic Inquiry*, 35(1):97–124.

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.

Bellugi, U. and Brown, R. (1964). *The acquisition of language: Report of the Fourth Conference sponsored by the Committee on Intellective Processes Research of the Social Science Research Council*. Monographs of the Society for Research in Child Development Volume 29. Wileyy.

Berko, J. (1958). The child's learning of English morphology. *Word*, 14(2–3):150–177.

Berwick, R. (1985). *The acquisition of syntactic knowledge*. The MIT Press, Cambridge, MA.

Berwick, R. C., Pietroski, P., Yankama, B., and Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science*, 35(7):1207–1242.

Biberauer, T., Holmberg, A., Roberts, I., and Sheehan, M., editors (2010). *Parametric variation: Null subjects in minimalist theory*. Cambridge University Press.

Bickerton, D. (1995). *Language and Human Behavior*. University of Washington Press, Seattle, WA.

Bierwisch, M. (1967). Syntactic features in morphology: General problems of so-called pronominal inflection in German. In *To honor Roman Jakobson: essays on the occasion of his 70th birthday*, volume 1, pages 239–270. Mouton, The Hague.

Bikel, D. M. (2004). Intricacies of Collins' parsing model. *Computational Linguistics*, 30(4):479–511.

Bittner, D. (2000a). Gender classification and the inflectional system of german nouns. *Trends in linguistics studies and monographs*, 124:1–24.

Bittner, D. (2000b). Sprachwandel durch Spracherwerb? Pluralwerb. In Bittner, A., Bittner, D., and Köpcke, K.-M., editors, *Systemorganisation in Phonologie, Morphologie und Syntax*, pages 123–140. Omls, Hildesheim.

Bloch, B. (1947). English verb inflection. *Language*, 23(4):399–418.

Bloom, L. (1970). *Language development: Form and function in emerging grammar*. MIT Press, Cambridge.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic inquiry*, pages 491–504.

Blum, L. and Blum, M. (1975). Toward a mathematical theory of inductive inference. *Information and control*, 28(2):125–155.

BNC Consortium (1995). British National Corpus. University of Oxford.

Bobaljik, J. (2000). The ins and outs of contextual allomorphy. *University of Maryland Working Papers in Linguistics*, 10:35–71.

Bobrow, S. A. and Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, 1(3):343–346.

Bolhuis, J. J., Tattersall, I., Chomsky, N., and Berwick, R. C. (2014). How could language have evolved? *PLoS Biol*, 12(8):e1001934.

Bolinger, D. L. M. (1971). *The phrasal verb in English*. Harvard University Press, Cambridge, MA.

Bolozky, S. (1999). *Measuring productivity in word formation: The case of Israeli Hebrew*, volume 27. Brill.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., and Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science*, 16(4):298–304.

Botha, R. P. (1969). *The function of the lexicon in transformational generative grammar*. Mouton.

Bouldin, J. M. (1990). *The syntax and semantics of postnominal adjectives in English*. PhD thesis, University of Minnesota.

Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish*, volume 11. CUP Archive.

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 ?–? Cambridge University Press, New York.

Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In Hawkins, J. A., editor, *Explaining language universals*, pages 73–101. Basil Blackwell, Oxford.

Bowerman, M. and Croft, W. (2008). The acquisition of the english causative alternation. *Crosslinguistic perspectives on argument structure: Implications for learnability*, pages 279–307.

Boyd, J. K. and Goldberg, A. E. (2011). Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language*, 87(1):55–83.

Boyé, G. and Cabredo Hofherr, P. (2010). Defectivity as stem suppletion in French and Spanish verbs. In Baerman, M., Corbett, G. G., and Brown, D., editors, *Defective paradigms: Missing forms and what they tell us*, pages 35–52. Oxford University Press, Oxford.

Braine, M. D. (1963). The ontogeny of english phrase structure: The first phase. *Language*, pages 1–13.

Braine, M. D. (1971). On two types of models of the internalization of grammars. In Slobin, D. I., editor, *The ontogenesis of grammar: A theoretical symposium*, pages 153–186. Academic Press, New York.

Brame, M. K. and Bordelois, I. (1973). Vocalic alternations in Spanish. *Linguistic Inquiry*, 4(2):111–168.

Brants, S., Dipper, S., Eisenberg, P., Hansen-Schirra, S., König, E., Lezius, W., Rohrer, C., Smith, G., and Uszkoreit, H. (2004). Tiger: Linguistic interpretation of a german corpus. *Research on Language and Computation*, 2(4):597–620.

Bresnan, J. and Ford, M. (2010). Predicting syntax: Processing dative constructions in american and australian varieties of english. *Language*, 86(1):168–213.

Bresnan, J. and Nikitina, T. (2009). On the gradience of the dative alternation. In Wee, L. and Uyechi, L., editors, *Reality explorations and discovery: pattern interaction in language and life*, pages 161–184. CSLI Publications, Stanford, CA.

Brill, E. (1995). Transformation-based error-driven learning and natural language processing: A case study in part of speech tagging. *Computational Linguistics*, 21(4):543–565.

Bromberger, S. and Halle, M. (1989). Why phonology is different. *Linguistic Inquiry*, 20(1):51–70.

Brooks, P. J. and Tomasello, M. (1999). How children constrain their argument structure constructions. *Language*, pages 720–738.

Brown, D. and Hippisley, A. (2012). *Network morphology: A defaults-based theory of word structure*, volume 133. Cambridge University Press.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press, Cambridge.

Brown, R. and Bellugi, U. (1964). Three processes in the child's acquisition of syntax. *Harvard educational review*, 34(2):133–151.

Brown, R. and Fraser, C. (1963). The acqisition of syntax. In Cofer, C. and Musgrave, B., editors, *Verbal behavior and learning: Problems and processes*, pages 158–201. McGraw-Hill.

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, New York.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology*, 55(1):1–5.

Bruening, B. (2010). Double object constructions disguised as prepositional datives. *Linguistic Inquiry*, 41(2):287–305.

Bruening, B. (2011a). A-adjectives again: Response to Goldberg. Available at http://lingcomm.blogspot.com/.

Bruening, B. (2011b). A-adjectives are PPs, not adjectives. Available at http://lingcomm.blogspot.com/.

Brysbaert, M. and New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4):977–990.

Burzio, L. (1994). *Principles of English stress*. Cambridge University Press, Cambridge.

Bush, R. R. and Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, 68(3):313–323.

Butt, J. and Benjamin, C. (1988). *A new reference grammar of modern Spanish*. Edward Arnold, London.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins, Philadelphia.

Bybee, J. and Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 59(2):251–270.

Bybee, J. and Pardo, E. (1981). Morphological and lexical conditioning of rules: Experimental evidence from Spanish. *Linguistics*, 19:937–968.

Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5):425–455.

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.

Cacciari, C. and Tabossi, P. (1988). The comprehension of idioms. *Journal of memory and language*, 27(6):668–683.

Campbell, A. L. and Tomasello, M. (2001). The acquisition of English dative constructions. *Applied Psycholinguistics*, 22(2):253–267.

Campbell, L. (2004). *Historical linguistics: An introduction*. MIT Press, Cambridge, 2nd edition.

Caprin, C. and Guasti, M. T. (2009). The acquisition of morphosyntax in italian: A cross-sectional study. *Applied psycholinguistics*, 30(01):23–52.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive neuropsychology*, 14(1):177–208.

Caramazza, A., Laundanna, A., and Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28(3):297–332.

Carey, S. and Bartlett, E. (1978). Acquire a single new word. *Child Language Development*, 15(17-29).

Carnegia Mellon University (2008). The CMU pronunciation dictionary 0.7. http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

Cerella, J. and Fozard, J. L. (1984). Lexical access and age. *Developmental Psychology*, 20(2):235.

Chan, E. (2008). *Structures and distributions in morphology learning*. PhD thesis, University of Pennsylvania.

Charniak, E. and Johnson, M. (2005). Coarse-to-fine n-best parsing and maxent discriminative reranking. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 173–180. Association for Computational Linguistics.

Chater, N. and Vitányi, P. (2007). Ideal learning of natural language: Positive results about learning from positive evidence. *Journal of Mathematical psychology*, 51(3):135–163.

Chomsky, N. (1951). Morphophonemics of Modern Hebrew. Master's thesis, University of Pennsylvania. Published by Garland, New York, 1979.

Chomsky, N. (1955). The logical structure of linguistic theory. Ms., Harvard University and MIT. Revised version published by Plenum, New York, 1975.

Chomsky, N. (1958). [Review of Belevitch 1956]. *Language*, 34(1):99–105.

Chomsky, N. (1959). A review of B.F. Skinner's *Verbal Behavior*. *Language*, 35(1):26–58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press, Cambridge.

Chomsky, N. (1970). Remarks on nominalization. In Jacobs, R. A. and Rosenbaum, P., editors, *Readings in English transformational grammar*, pages 184–221. Ginn, Waltham, MA.

Chomsky, N. (1975). *Reflections on language*. Pantheon, New York.

Chomsky, N. (1981). *Lectures in government and binding*. Foris, Dordrecht.

Chomsky, N. (1986). *Knowledge of language: Its nature, origins, and use*. Greenwood Publishing Group.

Chomsky, N. (1995). *The minimalist program*. MIT Press, Cambridge, MA.

Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M., editor, *Ken Hale: A life in language*, pages 1–52. Cambridge, MA: MIT Press.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1):1–22.

Chomsky, N. (2013). Problems of projection. *Lingua*, 130:33–49.

Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. MIT Press, Cambridge.

Chung, S. (1998). *The design of agreement: Evidence from Chamorro*. University of Chicago Press.

Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. In Koster, J., Pollock, J.-Y., Rizzi, L., and Zanuttini, R., editors, *Paths toward Universal Grammar*, pages 85–110. Georgetown University Press, Georgetown.

Cinque, G. (2010). *The syntax of adjectives: a comparative study*. MIT Press, Cambridge, MA.

Clahsen, H. (1997). The representation of participles in the german mental lexicon: evidence for the dual-mechanism model. *Yearbook of Morphology*, pages 73–95.

Clahsen, H. (1999). Lexical entries and rules of languge: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22:991–1069.

Clahsen, H., Aveledo, 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., Eisenbeiss, S., and Sonnenstuhl, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguistics*, 23(?):201–249.

Clahsen, H., Hadler, M., and Weyerts, H. (2004). Speeded production of inflected words in children and adults. *Journal of Child Language*, 31(3):683–712.

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 acquisiiton of verb placement*, pages 181–234. Kluwer, Dordrecht.

Clahsen, H. and Rothweiler, M. (1993). Inflectional rules in children's grammars: Evidence from german participles. *Yearbook of Morphology*, pages 1–34.

Clahsen, H., Rothweiler, M., Woest, A., and Marcus, G. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, 45:225–255.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In MacWhinney, B., editor, *Mechanisms of language acquisition*, pages 1–33. Erlbaum, Hillsdale, NJ.

Clark, E. V. and Cohen, S. R. (1984). Productivity and memory for newly formed words. *Journal of Child Language*, 11(03):611–625.

Clark, E. V. and Hecht, B. F. (1982). Learning to coin agent and instrument nouns. *Cognition*, 12(1):1–24.

Coady, J. A. and Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *Journal of Experimental Child Psychology*, 89(3):183–213.

Cohen, W. W. (1995). Fast effective rule induction. In *Proceedings of the Twelfth International Conference on Machine Learning, Lake Tahoe, California*.

Cohn, A. (2006). Is there gradient phonology? In Fanselow, G., Féry, C., Vogel, R., and Schlesewsky, M., editors, *Gradience in grammar: Generative perspectives*, pages 25–44. Oxford University Press, Oxford.

Colé, P., Beauvillain, C., and Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and language*, 28(1):1–13

Coleman, J. and Pierrehumbert, J. (1997). Stochastic phonological grammars and acceptability. In Coleman, J., editor, *Proceedings of the 3rd meeting of the ACL Special Interest Group in Computational Phonology*, pages 49–56, Somerset, NJ. Association for Computational Linguistics.

Collins, M. (1999). *Head-driven statistical models for natural language processing*. PhD thesis, University of Pennsylvania.

Conwell, E. and Demuth, K. (2007). Early syntactic productivity: Evidence from dative shift. *Cognition*, 103(2):163–179.

Cooreman, A. M. (1987). *Transitivity and discourse continuity in Chamorro narratives*. Walter de Gruyter.

Coppock, E. (2008). *The logical and empirical foundations of Baker's Paradox*. PhD thesis, Stanford University.

Corbett, G. G. and Fraser, N. M. (1993). Network morphology: a datr account of russian nominal inflection. *Journal of Linguistics*, 29(01):113–142.

Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, 14(4):597–612.

Culicover, P. W. (1999). *Syntactic nuts: hard cases, syntactic theory, and language acquisition*, volume 1. Oxford University Press.

Cutler, A. (1996). Prosody and the word boundary problem. *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, pages 87–99.

Cutler, A. and Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2(3–4):133–142.

Dąbrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. *Journal of Child Language*, 28(3):545–574.

Dąbrowska, E. (2005). Productivity and beyond: mastering the polish genitive inflection. *Journal of child language*, 32(01):191–205.

Dąbrowska, E. and Szczerbinski, M. (2006). Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity. *Journal of Child Language*, 33:559–597.

Daelemans, W., Gillis, S., and Durieux, G. (1994). The acquisition of stress: A data-oriented approach. *Computational Linguistics*, 20(3):421–451.

Daelemans, W., Zavrel, J., van der Sloot, K., and van den Bosch, A. (2009). TiMBL: Tilburg Memory-Based Learner. Technical Report ILK 09-01, Induction of Linguistic Knowledge Research Group, Tilburg University.

Dagum, P. and Luby, M. (1993). Approximating probabilistic inference in bayesian belief networks is np-hard. *Artificial intelligence*, 60(1):141–153.

Dale, P. S. and Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods*, 28(1):125–127.

Davies, M. (2008). Corpus of Contemporary American English (COCA): 410+ million words, 1990-present. available at http://www.americancorpus.org/. COCA.

Davies, M. (2010). The corpus of historical american english: 400 million words, 1810-2009. *http://corpus. byu. edu/coha/>, retrieved*, 24:2011.

Deen, K. U. (2005). *The acquisition of Swahili*, volume 40. John Benjamins Publishing.

Demuth, K. (1996). The prosodic structure of early words. In Morgan, J. L. and Demuth, K., editors, *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, pages 171–186. Psychology Press.

Demuth, K. (2003). The acquisition of Bantu languages. In *The Bantu languages*, pages 209–222. Curzon Press Surrey,, United Kingdom.

Di Sciullo, A.-M. and Williams, E. (1987). *On the definition of word*. MIT PRess, Cambridge, MA.

Downing, P. A. (1996). *Numeral classifier systems: The case of Japanese*. John Benjamins Publishing, Amsterdam.

Dresher, B. E. and Kaye, J. (1990). A computational learning model for metrical phonology. *Cognition*, 34:137–195.

Dressler, W. U. (1999). Why collapse morphological concepts? *Behavioral and Brain Sciences*, 22(6):1021.

Dromi, E. (1987). *Early lexical development*. Cambridge University PressUniversity Press, Cambridge.

Echols, C. H., Crowhurst, M. J., and Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of memory and language*, 36(2):202–225.

Eddington, D. (1996). Diphthongization in Spanish derivational morphology: An empirical investigation. *Hispanic Linguistics*, 8(1):1–13.

Ellegård, A. (1953). *The auxiliary* do*: The establishment and regulations of its use in English*. Gothenberg Studies in English. Almqvist and Wiksell, Stockholm.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1):71–99.

Elsen, H. (2002). The acquisition of german plurals. In *Morphology 2000: Selected Papers from the 9th Morphology Meeting, Vienna, 25-27 February 2000*, volume 218, page 117. John Benjamins Publishing.

Embick, D. and Halle, M. (2005). On the status of stems in morphological theory. In Geerts, T., van Ginneken, I., and Jacobs, H., editors, *Romance Languages and Linguistic Theory 2003: Selected papers from Going Romance 2003*, pages 37–62. John Benjamins, Amsterdam.

Embick, D. and Marantz, A. (2008). Architecture and blocking. *Linguistic Inquiry*, 39(1):1–53.

Eythórsson, T. (2002). Changes in subject case marking in icelandic. In Lightfoot, D., editor, *Syntactic effects of morphological change*, pages 196–212. Oxford University Press.

Fanselow, G. and Féry, C. (2002). Ineffability in grammar. In Fanselow, G. and Féry, C., editors, *Resolving conflicts in grammars: Optimality Theory in syntax, morphology, and phonology*, pages 265–307. Helmut Buske, Hamburg.

Feldman, J. (2000). Minimization of boolean complexity in human concept learning. *Nature*, 407(6804):630–633.

Feldman, N. H., Griffiths, T. L., Goldwater, S., and Morgan, J. L. (2013). A role for the developing lexicon in phonetic category acquisition. *Psychological review*, 120(4):751.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., and Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, pages i–185.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. PhD thesis, Leiden University.

Firth, J. (1957). A synopsis of linguistic theory 1930–1955. In Palmer, F. R., editor, *Selected papers of J.R. Firth*, pages 1–32. Longman, London.

Fleischhauer, E. and Clahsen, H. (2012). Generating inflected word forms in real time: Evaluating the role of age, frequency, and working memory. In Biller, A., Chung, E., and Kimball, A., editors, *Proceedings of the 36th annual Boston University Conference on Language Development*, volume 1, pages 164–176, Somerville, MA. Cascadilla Press.

Fodor, J. A., Bever, T. G., and Garrett, M. F. (1974). *The psychology of language*. McGraw-Hill, New York.

Fodor, J. D. (2001). Parameters and the periphery: Reflections on syntactic nuts.

Fodor, J. D. and Crain, S. (1987). Simplicity and generality of rules in language acquisition. In MacWhinney, B., editor, *Mechanisms of language acquisition*, pages 35–63. Lawrence Erlbaum Associates.

Fodor, J. D. and Sakas, W. G. (2005). The subset principle in syntax: Costs of compliance. *Journal of Linguistics*, 41(3):513–569.

Forster, K. I. (1976). Accessing the mental lexicon. In Wales, R. and Walker, E., editors, *New approaches to language mechanisms*, pages 257–287. North-Holland, Amsterdam.

Forster, K. I. (1992). Memory-addressing mechanisms and lexical access. In Forst, R. and Katz, L., editors, *Orthography, phonology, morphology, and meaning*, pages 413–434. Elsevier, Amsterdam.

Fraser, B. (1974). *The verb-particle combination in English*. Taishukan, Tokyo.

Fruchter, J. and Marantz, A. (2015). Decomposition, lookup, and recombination: Meg evidence for the full decomposition model of complex visual word recognition. *Brain and language*, 143:81–96.

Fruchter, J., Stockall, L., and Marantz, A. (2013). Meg masked priming evidence for form-based decomposition of irregular verbs. *Frontiers in Human Neuroscience*, 7(798).

Gagliardi, A. and Lidz, J. (2014). Statistical insensitivity in the acquisition of tsez noun classes. *Language*, 90(1):58–89.

Gambell, T. and Yang, C. (2005). Mechanisms and constraints in word segmentation. Ms., Yale University.

Gawlitzek-Maiwald, I. (1994). How do children cope with variation in the input? the case of German plurals and compounding. In Tracy, R. and Lattey, E., editors, *How tolerant is Universal Grammar? Essays on language learnability and language variation*, pages 225–266. Niemeyer, Tübingen.

Gerken, L. (1994a). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of child language*, 21(03):565–584.

Gerken, L. (1994b). Young childrenâĂŚs representation of prosodic phonology: Evidence from english-speakersâĂŚ weak syllable productions. *Journal of memory and language*, 33(1):19–38.

Gerken, L. and McIntosh, B. J. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29(3):448–457.

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.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8(2):149–156.

Gibbs, R. W. and Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition*, 20(3):243–259.

Gibson, E. and Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25(3):407–454.

Gildea, D. (2001). Corpus variation and parser performance. In *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing*, pages 167–202.

Gildea, D. and Jurafsky, D. (2002). Automatic labeling of semantic roles. *Computational linguistics*, 28(3):245–288.

Gillette, J., Gleitman, H., Gleitman, L., and Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73(2):135–176.

Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition*, 1(1):3–55.

Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 10:447–474.

Goldberg, A. E. (1995). *Constructions*. University of Chicago Press.

Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in cognitive sciences*, 7(5):219–224.

Goldberg, A. E. (2011). Are a-adjectives like afraid prepositional phrases underlying and does it matter from a learnability perspective? Unpublished manuscript. Princeton University.

Goldsmith, J. (2001). Unsupervised learning of morphology of a natural language. *Computational Linguistics*, 27(2):153–198.

Goldsmith, J. (2015). Towards a new empiricism for linguistics. In Chater, N., Clark, A., Goldsmith, J., and Perfors, A., editors, *Empiricism and language learnbility*, pages 58–105. Oxford University Press.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., and Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of child language*, 14(01):23–45.

Goodman, N. (1983). *Fact, fiction, and forecast*. Harvard University Press, Cambridge, 4th edition.

Gordon, M. (2002). A factorial typology of quantity-insensitive stress. *Natural Language & Linguistic Theory*, 20(3):491–552.

Gorman, K. (2013). *Generative phonotactics*. PhD thesis, University of Pennsylvania.

Graham, R. L., Knuth, D. E., and Patashnik, O. (1989). *Concrete Mathematics*. Addison-Wesley.

Green, G. M. (1974). *Semantics and syntactic regularity*. Indiana University Press.

Grimshaw, J. (1990). *Argument structure*. the MIT Press.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., and Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language*, 65(?):203–257.

Gross, M. (1975). *Méthodes en syntaxe: régime des constructions complétives*, volume 1365. Hermann.

Gross, M. (1979). On the failure of generative grammar. *Language*, pages 859–885.

Guasti, M. T. (1993). Verb syntax in italian child grammar: Finite and nonfinite verbs. *Language acquisition*, 3(1):1–40.

Guasti, M. T. (2004). *Language acquisition: The growth of grammar*. The MIT Press.

Guion, S., Clark, J., Harada, T., and Wayland, R. (2003). Factors affecting stress placement for english nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech*, 46(4):403–427.

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.

Haegeman, L. (1990). Understood subjects in english diaries: On the relevance of theoretical syntax for the study of register variation. *Multilingua*, 9(2):157–199.

Hahn, U. and Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, 41(?):313–360.

Hale, K. and Keyser, S. J. (2002). *Prolegomenon to a theory of argument structure*. MIT Press.

Hall, W. S., Nagy, W. E., and Linn, R. (1984). *Spoken words: Effects of situation and social group on oral word usage and frequency*. Lawrence Erlbaum, Hillsdale, NJ.

Halldórsson, H. (1982). On dative substitution (in icelandic). *Íslenskt mál og almenn málfræði*, 4:159–189.

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, 4(1):3–16.

Halle, M. (1978). Knowledge unlearned and untaught: What speakers know about the sounds of their language. In Halle, M., Bresnan, J., and Miller, G. A., editors, *Linguistic theory and psychological reality*, pages 294–303. MIT Press, Cambridge.

Halle, M. (1997). Distributed morphology: Impoverishment and fission. *MIT Working Papers in Linguistics*, 30:425–449.

Halle, M. (1998). The stress of English words 1968-1998. *Linguistic Inquiry*, 29(4):539–568.

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 Marantz, A. (1994). Some key features of distributed morphology. *MIT working papers in linguistics*, 21:275–288.

Halle, M. and Mohanan, K. P. (1985). Segmental phonology of Modern English. *Linguistic Inquiry*, 16(1):57–116.

Halle, M. and Vergnaud, J.-R. (1987). *An essay on stress*. MIT Press, Cambridge.

Hankamer, J. (1992). Morphological parsing and the lexicon. In Marslen-Wilson, W. D., editor, *Lexical representation and process*, pages 392–408. MIT Press, Cambridge.

Harley, H. (2002). Possession and the double object construction. *Language Variation Yearbook*, 2:29–68.

Harley, H. (2014). On the identity of roots. *Theoretical linguistics*, 40(3-4):225–276.

Harris, J. W. (1969). *Spanish phonology*. MIT Press, Cambridge.

Harris, J. W. (1977). Remarks of diphthongization in Spanish. *Lingua*, 41(3-4):261–305.

Harris, Z. S. (1951). *Methods in structural linguistics*. University of Chicago Press, Chicago.

Harris, Z. S. (1954). Distributional structure. *Word*, 10(23):146–162.

Harris, Z. S. (1955). From phoneme to morpheme. *Language*, 31(2):190–222.

Hart, B. and Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hart, B. and Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, 27(1):4–9.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *science*, 298(5598):1569–1579.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., and Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in psychology*, 5:doi: 10.3389/fpsyg.2014.00401.

Hay, J. and Baayen, R. H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences*, 9(7):342–348.

Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry*, 13(2):227–276.

Hayes, B. (1995). *Metrical stress theory*. University of Chicago Press, Chicago.

Hayes, B. and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3):379–440.

Hayes, B., Zuraw, K., Siptár, P., and Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*, 85(4):822–863.

Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms*. cambridge university Press.

Herman, L. and Herman, M. S. (2014). *American dialects: A manual for actors, directors, and writers*. Routledge.

Hetzron, R. (1975). Where the grammar fails. *Language*, 51(4):859–872.

Hirsch-Pasek, K., Treiman, R., and Schneiderman, M. H. (1984). Brown and hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, 11(?):81–88.

Hockett, C. F. (1942). English verb inflection. *Studies in Linguistics*, 1(2):1–8.

Holmberg, A. (2010). Parameters in minimalist theory: The case of scandinavian. *Theoretical Linguistics*, 36(1):1–48.

Horn, L. (1989). *A natural history of negation*. University of Chicago Press, Chicago.

Horning, J. J. (1969). A study of grammatical inference. Technical report, Stanford University, Stanford, CA.

Hornstein, N. (2009). *A theory of syntax: Minimal operations and universal grammar*. Cambridge University Press.

Hovav, M. R. and Levin, B. (2008). The english dative alternation: The case for verb sensitivity. *JOURNAL OF LINGUISTICS*, 44(1):129.

Howes, D. H. and Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, 41(6):401–410.

Hu, Q. (1993). Overextension of animacy in chinese classifier acquisition. In *The proceedings of the twenty-fifth annual child language research forum*, volume 25, page 127. Center for the Study of Language (CSLI).

Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic inquiry*, pages 531–574.

Huddleston, R. and Pullum, G. K. (2001). *The Cambridge grammar of the English language*. Cambridge University Press, Cambridge.

Hudson, R. (1997). The rise of auxiliary do-verb raising or category-strengthening? *Transactions of the Philological Society*, 95(1):41–72.

Hudson, R. (2000). *I amn't. *Language*, 76(2):297–323.

Hudson-Kam, C. L. and Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive psychology*, 59(1):30–66.

Hurford, J. R. (2011). *The origins of grammar: Language in the light of evolution II*, volume 2. Oxford University Press.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., and Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental psychology*, 27(2):236.

Huttenlocher, J. and Smiley, P. (1987). Early word meanings: The case of object names. *Cognitive Psychology*, 19(1):63–89.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Reidel, Dordrecht.

Idsardi, W. J. (1992). *The computation of prosody*. PhD thesis, MIT.

Inkelas, S. and Zoll, C. (2005). *Reduplication: Doubling in morphology*. Cambridge University Press, Cambridge.

Itô, J. and Mester, A. (2003). *Japanese morphophonemics: Markedness and word structure*. MIT Press, Cambridge.

Jackendoff, R. (1990a). On larson's treatment of the double object construction. *linguistic Inquiry*, pages 427–456.

Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. MIT Press.

Jackendoff, R. S. (1975). Morphological and semantic regularities in the lexicon. *Language*, 51(3):639–671.

Jackendoff, R. S. (1990b). *Semantic structures*. MIT Press, Cambridge.

Jackendoff, R. S. (1996). *The architecture of the language faculty*. MIT Press, Cambridge.

Jaeggli, O. and Safir, K. J. (1989). The null subject parameter and parametric theory. In *The null subject parameter*, pages 1–44. Springer.

Janda, R. D. (1990). Frequency, markedness and morphological change: On predicting the spread of noun-plural-s in modern high german and west germanic. In *ESCOL*, volume 90, pages 136–153.

Jarmulowicz, L. (2002). English derivational suffix frequency and children's stress judgements. *Brain and Language*, 81(1–3):192–204.

Jelinek, E. and Carnie, A. (2003). Argument hierarchies and the mapping principle. In Carnie, A., Harley, H., and Willie, M., editors, *Formal approaches to function in grammar*, pages 265–296. John Benjamins, Amsterdam.

Jelinek, F. (1998). *Statistical methods for speech recognition*. MIT Press, Cambridge.

Jespersen, O. (1942). *A Modern English grammar on historical principles. Part VI: Morphology*. Ejnar Munksgaard, Copenhagen.

Johnson, E. K. and Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44(4):493–666.

Johnson, M., Griffiths, T. L., and Goldwater, S. (2006). Adaptor grammars: A framework for specifying compositional nonparametric bayesian models. In *Advances in neural information processing systems*, pages 641–648.

Jónsson, J. G. (1997). Verbs taking oblique subjects (in icelandic). *Íslenskt mál og almenn málfræði*, 19-20:11–43.

Jónsson, J. G. (2003). Not so quirky: On subject case in icelandic. In Brandner, E. and Zinsmeister, H., editors, *New Perspectives in Case Theory*, pages 127–163. Center for the Study of Language and Information.

Jónsson, J. G. and Eythórsson, T. (2005). Variation in subject case marking in insular scandinavian. *Nordic Journal of Linguistics*, 28(02):223–245.

Jónsson, J. G. and Eythórsson, T. (2011). Structured exceptions and case selection in insular scandinavian. In Simon, H. J. and Wiese, H., editors, *Expecting the unexpected: Exceptions in grammar*, volume 216, pages 213–241. Walter de Gruyter.

Jung, Y.-J. and Miyagawa, S. (2004). Decomposing ditransitive verbs. In *Proceedings of SICGG*, pages 101–120.

Jusczyk, P. W. and Aslin, R. N. (1995). Infant's detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 46(1):65–97.

Jusczyk, P. W., Cutler, A., and Redanz, N. J. (1993a). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64(3):675–687.

Jusczyk, P. W., Friederici, A. D., Wessels, J. E., Svenkerud, V., and Jusczyk, A. (1993b). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32(3):402–420.

Jusczyk, P. W. and Hohne, E. A. (1997). Infants' memory for spoken words. *Science*, 277:1984–1986.

Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3-4):159–207.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. PhD thesis, MIT. Published by Garland, New York, 1980.

Kam, C. H. and Newport, E. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1(2):151–195.

Kehoe, M. (1997). Stress error patterns in english-speaking children's word productions. *Clinical linguistics & phonetics*, 11(5):389–409.

Kehoe, M. and Stoel-Gammon, C. (1997). The acquisition of prosodic structure: An investigation of current accounts of children's prosodic development. *Language*, 73:113–144.

Kelly, M. H. (1992). The role of phonology in grammatical category assinments. *Psychological Review*, 99(?):349–364.

Keshava, S. and Pitler, E. (2006). A simpler, intuitive approach to morpheme induction. In *Proceedings of 2nd Pascal Challenges Workshop*, pages 31–35.

Kiparsky, P. (1965). *Phonological change*. PhD thesis, MIT.

Kiparsky, P. (1966). Über den deutschen Akzent. *Studia Grammatica*, 7:69–97.

Kiparsky, P. (1973). Elsewhere in phonology. In Anderson, S. R. and Kiparsky, P., editors, *A festschrift for Morris Halle*, pages 93–106. Holt, Rinehart and Winston, New York.

Kiparsky, P. (1974). Remarks on analogical change. In Anderson, J. M. and Jones, C., editors, *Historical linguistics II: Theory and description in phonology*, pages 257–276. North-Holland. Reprinted in Kiparsky 1982a, 199-215.

Kiparsky, P. (1982). Lexical morphology and phonology. In Yang, I.-S., editor, *Linguistics in the Morning Calm: Selected Papers from SICOL 1981*, pages 3–91. Hanshin, Seoul.

Köpcke, K.-M. (1998). The acquisition of plural marking in English and German revisited: Schemata versus rules. *Journal of Child Language*, 25(2):293–319.

Kottum, S. E. (1981). The genitive singular form of masculine nouns in polish. *Scando-Slavica*, 27(1):179–186.

Krifka, M. (1999). Manner in dative alternation. In *West Coast Conference on Formal Linguistics*, volume 18, pages 260–271.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1(3):199–244.

Kučera, H. and Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press, Providence.

Kuczaj, S. A. (1976). -ing, -s, *and* -ed*: A study of the acquisition of certain verb inflections*. PhD thesis, University of Minnesota.

Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*, 16(5):589–600.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044):606–608.

Kwisthout, J., Wareham, T., and van Rooij, I. (2011). Bayesian intractability is not an ailment that approximation can cure. *Cognitive Science*, 35(5):779–784.

Laaha, S., Ravid, D., Korecky-Kröll, K., Laaha, G., and Dressler, W. U. (2006). Early noun plurals in German: Regularity, productivity or default? *Journal of Child Language*, 33(2):271–302.

Labov, W. (1972a). *Language in the inner city: Studies in Black English Vernacular*. University of Pennsylvania Press.

Labov, W. (1972b). *Sociolinguistic patterns*. University of Pennsylvania Press, Philadelphia.

Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, 1(1):85–97.

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Blackwell, Oxford.

Labov, W. (2007). Transmission and diffusion. *Language*, 83(2):344–387.

Labov, W., Ash, S., and Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology, and sound change*. Mouton de Gruyter, Berlin.

Ladefoged, P. and Fromkin, V. (1968). Experiments on competence and performance. *IEEE Transactions on Audio and Electroacoustics*, 16(1):130–136.

Lakoff, G. (1970). *Irregularity in syntax*. Holt, Rinehart and Winston, New York.

Landau, B. and Gleitman, L. R. (1984). *Language and experience: Evidence from the blind child*, volume 8. Harvard University Press.

Landauer, T. K., Foltz, P. W., and Laham, D. (1998). An introduction to latent semantic analysis. *Discourse processes*, 25(2-3):259–284.

Larson, R. K. (1988). On the double object construction. *Linguistic Inquiry*, 19(3):335–391.

Larson, R. K. (1990). Double objects revisited: Reply to jackendoff. *Linguistic Inquiry*, pages 589–632.

Larson, R. K. and Marušič, F. (2004). On indefinite pronoun structures with aps: Reply to Kishimoto. *Linguistic Inquiry*, 35(2):268–287.

Legate, J. A. (2008). Morphological and abstract case. *Linguistic Inquiry*, 39(1):55–101.

Legate, J. A. and Yang, C. (2002). Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review*, 19:151–162.

Legate, J. A. and Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, 14(3):315–344.

Legate, J. A. and Yang, C. (2013). Assessing child and adult grammar. In Berwick, R. and Piattelli-Palmarini, M., editors, *Rich languages from poor inputs: In honor of Carol Chomsky*, pages 168–182. (Longer version available from the authors.). Oxford University Press.

Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1):1–75.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press.

Levin, B. (2008). Dative verbs: A crosslinguistic perspective. *Lingvisticæ Investigationes*, 31(2):285–312.

Levine, M. (1975). *A cognitive theory of learning: Research on hypothesis testing*. Lawrence Erlbaum.

Levy, Y. and Vainikka, A. (2000). The development of a mixed null subject system: a cross-linguistic perspective with data on the acquisition of hebrew. *Language Acquisition*, 8(4):363–384.

Li, W. (1992). Random texts exhibit Zipf-law-like word frequency distributions. *IEEE Transactions in Information Theory*, 38(6):1842–1845.

Liberman, M. and Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8:249–336.

Lieber, R. (1980). *On the organization of the lexicon*. PhD thesis, MIT.

Lieber, R. (1981). Morphological conversion within a restrictive theory of the lexicon. In Moortgat, M., van der Hulst, H., and Hoekstra, T., editors, *The scope of lexical rules*, pages 161–200. Foris.

Lightfoot, D. W. (1979). *Principles of diachronic syntax*. Cambridge University Press, Cambridge.

Lignos, C. (2011). Modeling infant word segmentation. In *Proceedings of the 15th Conference on Computational Language Learning*, pages 28–38.

Lignos, C. (2013). *Modeling words in the mind*. PhD thesis, University of Pennsylvania, Philadelphia.

Lignos, C., Chan, E., Marcus, M., and Yang, C. (2010). Evidence for a morphological acquisition model from development data. In Franich, K., Iserman, K. M., and Keil, L. L., editors, *Pro-*

ceedings of the 34th annual Boston University conference on language development, volume 2, pages 269–280.

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*.

Lignos, C. and Gorman, K. (2012). Revisiting frequency and storage in morphological processing. *Proceedings of CLS*, 48:447–461.

Long, R. B. (1969). *The sentence and its parts*. University of Chicago Press.

MacWhinney, B. (1975). Rules, rote, and analogy in morphological formations by hungarian children. *Journal of Child Language*, 19(2):65–77.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum, Mahwah, NJ, 3rd edition.

Maiden, M. and O'Neill, P. (2010). On morphomic defectiveness: Evidence from the Romance languages of the Iberian Peninsula. In Baerman, M., Corbett, G. G., and Brown, D., editors, *Defective paradigms: Missing forms and what they tell us*, pages 103–124. Oxford University Press, Oxford.

Malkiel, Y. (1966). Diphthongization, monothongization, metaphony: studies in their interaction in the paradigm of the Old Spanish -IR verbs. *Language*, 42(2):430–472.

Mandelbrot, B. (1954). Structure formelle des textes et communication. *Word*, 10(1):1–27.

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Dimitriadis, A., Siegel, L., Surek-Clark, C., and Williams, A., editors, *Penn Working Papers in Linguistics 4.2: Proceedings of the 21st annual Penn Linguistics Colloquium*, pages 201–225. Penn Linguistics Club, Philadelphia.

Marantz, A. (2000). Case and licensing. In Reuland, E. J., editor, *Arguments and case: Explaining Burzio's generalization*, pages 11–30. John Benjamins, Amsterdam.

Marantz, A. (2001). Words. In *20th West Coast Conference on Formal Linguistics, University of Southern California*.

Marantz, A. (2013). Locality domains for contextual allomorphy across the interfaces. In Matushansky, O. and Marantz, A., editors, *Distributed Morphology today: Morphemes for Morris Halle*, pages 95–115. MIT Press.

Maratsos, M. (2000). More overregularizations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen and Xu. *Journal of Child Language*, 27:183–212.

Maratsos, M. and Chalkley, M. A. (1980). The internal language of children's syntax: The nature and ontogenesis of syntactic categories. In Nelson, K., editor, *Children's language*, volume 2, pages 127–214. Gardner.

Marchand, H. (1969). *The categories and types of present-day English word-formation: A sychronic-diachronic approach*. C.H Beck'sche, München, 2nd edition.

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, Chicago.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, 46(1):53–85.

Marcus, G. F. (1995). The acquisition of the english past tense in children and multilayered connectionist networks. *Cognition*, 56(3):271–279.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., and Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29:189–256.

Marcus, M. P., Santorini, B., Marcinkiewicz, M. A., and Taylor, A. (1999). Treebank-3. Linguistic Data Consortium: LDC99T42.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1–2):71–102.

Marslen-Wilson, W. D. and Tyler, L. K. (1997). Dissociating types of mental computation. *Nature*, 387:592–594.

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.

Mateo, F. and Rojo Sastre, A. (1995). *El arte de conjugar en español*. Hatier, Paris.

Matthews, P. H. (1974). *Morphology: An introduction to the theory of word-structure*. Cambridge University Press, Cambridge.

Mausch, H. (2003). Current alternations in inflection of polish masculine inanimate nouns in the singular: A pilot study. *Investigationes Linguisticae*, 9:1–21.

Mayol, L. (2007). Acquisition of irregular patterns in Spanish verbal morphology. In Nurmi, V. and Sustretov, D., editors, *Proceedings of the twelfth ESSLLI Student Session*, pages 1–11, Dublin.

Mazurkewich, I. (1984). The acquisition of the dative alternation by second language learners and linguistic theory. *Language learning*, 34(1):91–108.

Mazurkewich, I. and White, L. (1984). The acquisition of the dative alternation: Unlearning overgeneralizations. *Cognition*, 16(3):261–283.

McCarthy, J. J. and Prince, A. (1995). Faithfulnesss and reduplicative identity. In Beckman, J., Dickey, L. W., and Urbancyzk, S., editors, *Papers in Optimality Theory*, pages 249–384. GLSA, Amherst, MA.

McClelland, J. L. and Bybee, J. (2007). Gradience of gradience: A reply to jackendoff. *The Linguistic Review*, 24(4):437–455.

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.

McClelland, J. L. and Rumelhart, D. (1981). An interactive activation model of context effects on letter perception: Part 1. an account of the basic findings. *Psychological Review*, 88(?):375–407.

McClosky, D., Charniak, E., and Johnson, M. (2010). Automatic domain adaptation for parsing. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 28–36. Association for Computational Linguistics.

McNeill, D. (1966). The creation of language by children. In *Psycholinguistic Papers*, pages 99–132. Lyons, John and Wales, Roger.

Medin, D. L. and Schaffer, M. M. (1978). Context theory of classification learning. *Psychological review*, 85(3):207.

Medina, T. N., Snedeker, J., Trueswell, J. C., and Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, 108(22):9014–9019.

Melançon, A. and Shi, R. (2015). Representations of abstract grammatical feature agreement in young children. *Journal of child language*, pages 1–15.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Miller, G. A. (1957). Some effects of intermittent silence. *American Journal of Psychology*, 70(2):311–314.

Miller, G. A. (1991). *The science of words*. Scientific American Library.

Miller, G. A. and Chomsky, N. (1963). Finitary models of language users. In Luce, R. D., Bush, R. R., and Galanter, E., editors, *Handbook of mathematical psychology. Volume II*, pages 419–491. Wiley, New York.

Mills, A. (1986). *The acquisition of gender: A study of English and German*. Springer, Berlin.

Mitchell, T. M. (1982). Generalization as search. *Artificial intelligence*, 18(2):203–226.

Molnar, R. A. (2001). "Generalize and sift" as a model of inflection acquisition. Master's thesis, MIT.

Morris, J. and Stockall, L. (2012). Early, equivalent erp masked priming effects for regular and irregular morphology. *Brain and language*, 123(2):81–93.

Mugdan, J. (1977). *Flexionsmorphologie und Psycholinguistik*. Narr, Tübingen.

Mulford, R. (1985). Comprehension of icelandic pronoun gender: semantic versus formal factors. *Journal of Child Language*, 12(02):443–453.

Murray, W. S. and Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, 111(3):721–756.

Myers, S. (1987). Vowel shortening in English. *Natural Language and Linguistic Theory*, 5(4):485–518.

Nagy, W. E. and Anderson, R. C. (1984). How many words are there in printed school english? *Reading Research Quarterly*, 19(3):304–330.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of child language*, 17(02):357–374.

Naigles, L. R. and Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? effects of input frequency and structure on children's early verb use. *Journal of Child Language*, 25(01):95–120.

Neue, F. (1866). *Formenlehre der lateinischen Sprache*. H. Lindemann, Stuttgart.

New, B., Brysbaert, M., Segui, J., Ferrand, L., and Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language*, 51(4):568–585.

Newmeyer, F. J. (2004). Against a parameter-setting approach to typological variation. *Linguistic Variation Yearbook*, 4(1):181–234.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science*, 14(1):11–28.

Niemi, J., Laine, M., and Tuominen, J. (1994). Cognitive morphology in finnish: Foundations of a new model. *Language and Cognitive Processes*, 9(3):423–446.

Niyogi, P. (2006). *The computational nature of language learning and evolution*. Mit Press Cambridge.

Niyogi, P. and Berwick, R. (1996). Formaling triggers: A learning model for finite space. *Linguistic Inquiry*, 27(4):605–622. Revised version of MIT AI Laboratory Memo #1449, 1993.

Nosofsky, R. M., Palmeri, T. J., and McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological review*, 101(1):53.

Nowak, M. A., Komarova, N. L., and Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature*, 417(?):611–617.

Nowenstein, I. E. (2015). Acquiring intra-speaker variation: The case of icelandic dative substitution. To appear in the Proceedings of the Generative Approach to Language Acquisition in North America (GALANA). University of Maryland, College Park.

Noyer, R. (1992). *Features, positions, and affixes in automous morphological structure*. PhD thesis, MIT. Published by Garland, New York, 1997.

O'Donnell, T. (2011). *Productivity and reuse in language*. PhD thesis, Harvard University.

Oehrle, R. T. (1976). *The grammatical status of the English dative alternation*. PhD thesis, Massachusetts Institute of Technology.

Oh, G., Guion-Anderson, S., and Redford, M. A. (2011). Developmental change in factors affecting stress placement in native english-speaking children. In *Proceedings of the 17th International Congress of Phonetic Sciences*, pages 1522–1525.

Orgun, C. O. and Sprouse, R. (1999). From MParse to CONTROL: Deriving ungrammaticality. *Phonology*, 16(2):191–224.

Osherson, D. N. and Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9(1):35–58.

Osherson, D. N., Stob, M., and Weinstein, S. (1986). *Systems that learn: An introduction to learning theory for cognitive and computer scientists*. The MIT Press, Cambridge, MA.

Parducci, A. and Perrett, L. F. (1971). Category rating scales: Effects of relative frequency of stimulus values. *Journal of Experimental Psychology*, 89(2):427–452.

Park, T. (1978). Plurals in child speech. *Journal of Child Language*, 5:237–250.

Pearl, L. and Lidz, J. (2009). When domain-general learning fails and when it succeeds: Identifying the contribution of domain specificity. *Language Learning and Development*, 5(4):235–265.

Pearl, L. and Sprouse, J. (2013). Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition*, 20(1):23–68.

Penke, M. and Krause, M. (2002). German noun plurals: A challenge to the dual-mechanism approach. *Brain and Language*, 81(1–3):301–311.

Penny, R. (2002). *A history of the Spanish language*. Cambridge University Press, Cambridge, 2nd edition.

Pérez-Pereira, M. (1991). The acquisition of gender: What spanish children tell us. *Journal of Child Language*, 18(03):571–590.

Perfors, A., Tenenbaum, J. B., and Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, 118(3):306–338.

Perfors, A., Tenenbaum, J. B., and Wonnacott, E. (2010). Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language*, 37(3):607–642.

Pertsova, K. (2005). How lexical conservatism can lead to paradigm gaps. In Heinz, J., Martin, A., and Pertsova, K., editors, *UCLA Working Papers in Linguistics 11: Papers in Phonology 6*, pages 13–30. UCLA Linguistics Department, Los Angeles.

Pesetsky, D. (1977). Russian morphology and lexical theory. Ms., MIT.

Pesetsky, D. (1995). *Zero syntax: Experiencer and Cascade*. MIT Press, Cambridge, MA.

Pesetsky, D. (2015). Complementizer-trace effects. In Eveaert, M. and van Riemsdijk, H., editors, *Companion to Syntax, 2nd edition*, page (To appear). Wiley.

Peters, A. M. (1983). *The units of language acquisition*. Cambridge University Press, Cambridge.

Phillips, C. (1995). Syntax at age two: Cross-linguistic differences. *MIT Working Papers in Linguistics*, 26:325–382.

Phillips, C. (2013). On the nature of island constraints. i: Language processing and reductionist accounts. In Sprouse, J. and Hornstein, N., editors, *Experimental syntax and island effects*, pages 132–157. Cambridge University Press, Cambridge.

Pierrehumbert, J. (1994). Syllable structure and word structure: A study of triconsonantal clusters in English. In Keating, P. A., editor, *Phonological structure and phonetic form: Papers in Laboratory Phonology III*, pages 168–188. Cambridge University Press, Cambridge.

Pierrehumbert, J. (2003). Probabilistic phonology: Discrimination and robustness. In Bod, R., Hay, J., and Jannedy, S., editors, *Probabilistic Linguistics*, pages 177–228. MIT Press, Cambridge.

Pine, J. M. and Lieven, E. V. (1997). Slot and frame patterns and the development of the determiner category. *Applied psycholinguistics*, 18(02):123–138.

Pinker, S. (1984). *Language learnability and language development*. Harvard University Press, Cambridge.

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

Pinker, S. (1991). Rules of language. *Science*, 253:530–535.

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, Cambridge.

Pinker, S. (1999). *Words and rules: The ingredients of language*. Basic Books, New York.

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 Levy, Y., editor, *Other children, other languages*, pages 137–187. Lawrence Erlbaum Associates.

Plag, I. (2003). *Word-formation in English*. Cambridge University Press, Cambridge.

Plag, I. and Baayen, H. (2009). Suffix ordering and morphological processing. *Language*, 85(1):109–152.

Plag, I., Dalton-Puffer, C., and Baayen, R. H. (1999). Morphological productivity across speech and writing. *English Language and Linguistics*, 3(2):209–228.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and cognitive processes*, 12(5-6):765–806.

Plunkett, K. and Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perception: Implications for child language acquisition. *Cognition*, 38(1):43–102.

Poser, W. J. (1992). Blocking of phrasal constructions by lexical items. In Sag, I. and Szabolsci, A., editors, *Lexical matters*, pages 111–130. CSLI, Stanford.

Prasada, S. and Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, 8(1):1–56.

Prasada, S., Pinker, S., and Snyder, W. (1990). Some evidence that irregular forms are retrieved from memory but regular forms are rule-generated. Paper presented at the 31st International Meeting of the Psychonomics Society.

Prince, A. and Smolensky, P. (1993). Optimality Theory: Constraint interaction in generative grammar. Technical Report TR-2, Rutgers University Center for Cognitive Science. Published by MIT Press, Cambridge, 2004.

Prince, A. and Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. MIT Press, Cambridge.

Pullum, G. K. and Wilson, D. (1977). Autonomous syntax and the analysis of auxiliaries. *Language*, 53(4):741–788.

Quine, W. V. O. (1960). *Word and object*. MIT press, Cambridge, MA.

Raffelsiefen, R. (1996). Gaps in word formation. In Kleinhenz, U., editor, *Interfaces in phonology*, pages 194–209. Akademie-Verlag, Berlin.

Raffelsiefen, R. (1999). Phonological constraints on English word fomation. In Booij, G. and van Marle, J., editors, *Yearbook of Morphology 1998*, pages 225–287. Kluwer, Dordrecht.

Raffelsiefen, R. (2004). Absolute ill-formedness and other morphophonological effects. *Phonology*, 21(1):91–142.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, 45(1):45–94.

Randall, J. (1990). Catapults and pendulums: the mechanics of language acquisition. *Linguistics*, 28(6):1381–1406.

Randall, J. H. (1980). -ity: A study in word formation restrictions. *Journal of Psycholinguistic Research*, 9(6):523–534.

Randall, J. H. (1987). *Indirect positive evidence: Overturning overgeneralizations in language acquisition*. Indiana University Linguistics Club.

Rastle, K., Davis, M. H., and New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, 11(6):1090–1098.

Rauh, G. (1993). On the grammar of lexical and non-lexical prepositions in english. In Aurnague, M. and Vieu, L., editors, *The semantics of prepositions*, pages 99–150. Walter de Gruyter.

Real Academia Española (1992). *Diccionario de la lengua española*. Real Academia Española, Madrid, 21st edition.

Rebrus, P. and Törkenczy, M. (2009). Covert and overt defectiveness in paradigms. In Rice, C. and Blaho, S., editors, *Modeling ungrammaticality in Optimality Theory*, pages 195–236. Equinox, London.

Redington, M., Chater, N., and Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22(4):425–469.

Regel, S., Opitz, A., Müller, G., and Friederici, A. D. (2015). The past tense debate revisited: Electrophysiological evidence for subregularities of irregular verb inflection. *Journal of cognitive neuroscience*, 27(9):1870–1885.

Rescorla, L. A. (1980). Overextension in early language development. *Journal of Child Language*, 7(02):321–335.

Rice, C. (2005). Optimal gaps in optimal paradigms. *Catalan Journal of Linguistics*, 4:155–170.

Rice, C. and Blaho, S. (2009). Modeling ungrammaticality. In Rice, C. and Blaho, S., editors, *Modeling ungrammaticality in Optimality Theory*, pages 1–16. Equinox, London.

Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5):465–471.

Rivest, R. (1976). On self-organizing sequential search heuristics. *Communications of the ACM*, 2(63-67).

Rizzi, L. (1982). *Issues in Italian syntax*, volume 11. Walter de Gruyter.

Rizzi, L. (1986). Null objects in italian and the theory of pro. *Linguistic inquiry*, pages 501–557.

Roberts, I. (2012). Macroparameters and minimalism: A programme for comparative research. In Galves, C., Cyrino, S., Lopes, R., Sandalo, F., and Avelar, J., editors, *Parameter theory and linguistic change*, pages 320–335. Oxford University Press.

Roeper, T. and Williams, E. (1987). *Parameter setting*. Springer.

Rubenstein, H. and Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM*, 8(10):627–633.

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.

Saffran, J. R., Aslin, R. N., and Newport, E. (1996). Statistical learning by by 8-month-old infants. *Science*, 274:1926–1928.

Sag, I. A. (2010). English filler-gap constructions. *Language*, 86(3):486–545.

Sakas, W. G. and Fodor, J. D. (2012). Disambiguating syntactic triggers. *Language Acquisition*, 19(2):83–143.

Salkoff, M. (1983). Bees are swarming in the garden: a systematic synchronic study of productivity. *Language*, 59(2):288–346.

Sankoff, G. and Blondeau, H. (2007). Language change across the lifespan: /r/ in Montreal French. *Language*, 83(3):560–588.

Sapir, E. (1928). *Language: An introduction to the study of speech*. Harcourt Brace, New York.

Scarborough, D. L., Cortese, C., and Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 7(1):3–12.

Schieffelin, B. B. and Ochs, E. (1986). *Language socialization across cultures*. Cambridge University Press.

Schlesinger, I. M. (1971). Production of utterances and language acquisition. In Slobin, D., editor, *The ontogenesis of grammar*, pages 63–101. Academic Press.

Schmid, H. (1995). Improvements in part-of-speech tagging with an application to German. In *Proceedings of the ACL SIGDAT-Workshop*, pages 47–50.

Scholes, R. J. (1966). *Phonotactic grammaticality*. Mouton, Berlin.

Schuler, K., Davis, J., Yang, C., and Newport, E. (2015). Testing the Tolerance Principle for rule productivity in an artificial grammar. In *The Cognitive Development Society conference*, Columbus, OH.

Schütze, C. T. (2001). On the nature of default case. *Syntax*, 4(3):205–238.

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, Berlin.

Schütze, C. T. (2011). Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(2):206–221.

Schütze, C. T. R. (1997). *INFL in child and adult language: Agreement, case and licensing*. PhD thesis, MIT.

Sebastián, N., Cuetos, F., Martí, A., and Carreiras, M. (2000). *LEXESP: Léxico informatizado del español*. Edicions de la Universitat de Barcelona, Barcelona.

Seidl, A. and Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development*, 1(3–4):289–316.

Sharoff, S. (2005). Methods and tools for development of the Russian Reference Corpus. In Archer, D., Wilson, A., and Rayson, P., editors, *Corpus linguistics around the world*, pages 167–180. Rodopi, Amsterdam.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820):1317–1323.

Shi, R. and Melançon, A. (2010). Syntactic categorization in french-learning infants. *Infancy*, 15(5):517–533.

Shi, R., Werker, J. F., and Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72(2):B11–B21.

Shipley, E. F., Smith, C. S., and Gleitman, L. R. (1969). A study in the acquisition of language: Free responses to commands. *Language*, pages 322–342.

Shlonsky, U. (2009). Hebrew as a partial null-subject language. *Studia Linguistica*, 63(1):133–157.

Siegel, D. C. (1974). *Topics in English morphology*. PhD thesis, MIT. Published by Garland, New York, 1979.

Sigurjónsdóttir, H. (2002). Case marking in child language: How do icelandic children learn to use case? Master's thesis, University of Iceland, Reykjavik.

Sims, A. D. (2006). *Minding the gap: Inflectional defectiveness in a paradigmatic theory*. PhD thesis, Ohio State University.

Sinclair, J. (1987). *Collins COBUILD, Collins Birgmingham University International Language Database: English language dictionary*. London: Collins.

Skinner, B. F. (1957). *Verbal behavior*. Appleton Century Crofts, New York.

Skousen, R., Lonsdale, D., and Parkinson, D. B., editors (2002). *Analogical modeling: An exemplar-based approach to language*. John Benjamins, Amsterdam.

Sleator, D. D. and Tarjan, R. E. (1985a). Amortized efficiency of list update and paging rules. *Communications of the ACM*, 28(2):202–208.

Sleator, D. D. and Tarjan, R. E. (1985b). Self-adjusting binary search trees. *Journal of the ACM (JACM)*, 32(3):652–686.

Slobin, D. I. (1971). On the learning of morphological rules: A reply to Palermo and Eberhart. In Slobin, D. I., editor, *The ontogenesis of grammar*. Academic Press, New York.

Sluijter, A. and van Heven, V. (1996). Spectral balance as an acoustic correlate of lingustic stress. *Journal of the Acoustical Society of America*, 100(4):2471–2485.

Smith, H. (1994). "dative sickness" in Germanic. *Natural Language and Linguistic Theory*, 12(4):675–736.

Smith, L. and Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3):1558–1568.

Smith, N. V. and Tsimpli, I.-M. (1995). *The mind of a savant: Language learning and modularity*. Blackwell Publishing.

Smoczyńska, M. (1985). The acquisition of Polish. In Slobin, D. I., editor, *The corsslinguistic study of language acquisition. Volume 1: The data*, pages 595–686. Erlbaum, Hillsdale, NJ.

Smolensky, P. (1996). The initial state and 'richness of the base' in Optimality Theory. Technical Report JHU-CogSci-96-4, Department of Cognitive Science, Johns Hopkins University.

Smolka, E., Zwitserlood, P., and Rösler, F. (2007). Stem access in regular and irregular inflection: Evidence from german participles. *Journal of Memory and Language*, 57(3):325–347.

Snyder, W. (2007). *Child language: The parametric approach*. Oxford University Press.

Snyder, W. and Stromswold, K. (1997). The structure and acquisition of english dative constructions. *Linguistic Inquiry*, pages 281–317.

Sober, E. (1975). *Simplicity*. Oxford University Press, New York.

Solomyak, O. and Marantz, A. (2010). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience*, 22(9):2042–2057.

Sonnenstuhl, I. and Huth, A. (2002). Processing and representation of German *-n* plurals: A dual mechanism approach. *Brain and Language*, 81(1–3):276–290.

Stefanowitsch, A. (2008). Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics*, 19(3):513–531.

Steriade, D. (1997). Lexical conservatism. In *Linguistics in the morning calm: Selected Papers from SICOL 1997*, pages 157–179. Hanshin, Seoul.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57(4):421–457.

Stevens, J., Trueswell, J., Yang, C., and Gleitman, L. (2015). The pursuit of word meanings. Manuscript in submission. University of Pennsylvania.

Stockall, L. and Marantz, A. (2006). A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, 1(1):85–123.

Studdert-Kennedy, M. (1998). The particulate origins of language generativity: from syllable to gesture. In Hurford, J., Studdert-Kennedy, M., and Knight, C., editors, *Approaches to the Evolution of Language*, pages 202–221. Cambridge University Press, Cambridge.

Stump, G. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge University Press.

Stvan, L. S. (1998). *The semantics and pragmatics of bare singular noun phrases*. PhD thesis, Northwestern University.

Suppes, P. (1974). The semantics of children's language. *American Psychologist*, 29(1):103–114.

Sweet, H. (1892). *A new English grammar: logical and historical. 1. Introduction, phonology and accidence*. Clarendon Press, Oxford.

Swinney, D. and Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18(5):523–534.

Szagun, G. (2001). Learning different regularities: The acquisition of noun plurals by German-speaking children. *First Language*, 21:109–141.

Szagun, G. (2004). Learning by ear: on the acquisition of case and gender marking by german-speaking children with normal hearing and with cochlear implants. *Journal of Child Language*, 31(01):1–30.

Tabossi, P., Fanari, R., and Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition*, 37(4):529–540.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7(?):263–272.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology*, 57A(4):745–765.

Taft, M. and Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6):638–647.

Taft, M. and Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15(6):607–620.

Tardif, T., Shatz, M., and Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of english, italian, and mandarin. *Journal of Child Language*, 24(3):535–565.

Tattersall, I. (2012). *Masters of the Planet: the Search for our Human Origins*. Macmillan.

Taylor, A. (1994). Variation in past tense formation in the history of English. In Izvorski, R., Meyerhoff, M., Reynolds, B., and Tredinnick, V., editors, *Penn Working Papers in Linguistics 1*, pages 143–158. Penn Linguistics Club, Philadelphia.

Taylor, J. R. (2003). *Linguistic categorization*. Oxford University Press.

Tenenbaum, J. B. and Griffiths, T. L. (2001). Generalization, similarity and bayesian inference. *Behavioral and Brain Sciences*, 24(4):629–640.

Terrace, H. S., Petitto, L.-A., Sanders, R. J., and Bever, T. G. (1979). Can an ape create a sentence? *Science*, 206(4421):891–902.

Tesar, B. and Smolensky, P. (2001). *Learnability in Optimality Theory*. MIT Press, Cambridge.

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., and Weniger, D. (2004). Neural correlates for the acquisition of natural language syntax. *NeuroImage*, 17:700–709.

Thráinsson, H. (2005). Íslensk tunga iii: Setningar [icelandic language iii: Sentences]. *Reykjavík: Almenna bókafélagið*.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Harvard University Press, Cambridge.

Tomasello, M. (2000a). Do young children have adult syntactic competence? *Cognition*, 74(3):209–253.

Tomasello, M. (2000b). First steps toward a usage-based theory of language acquisition. *Cognitive linguistics*, 11(1/2):61–82.

Tomasello, M. (2003). *Constructing a language*. Harvard University Press, Cambridge.

Topping, D. M. (1973). *Chamorro Reference Grammar*. University of Hawaii Press, Honolulu. reprinted in 1980.

Trabasso, T. and Bower, G. H. (1975). *Attention in learning: Theory and research*. Krieger Pub Co.

Trammell, R. L. (1978). The psychological reality of underlying forms and rules for stress. *Journal of Psycholinguistic Research*, 7(2):79–94.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, 35:566–585.

Tulving, E. (1972). Episodic and semantic memory. In Tulving, E. and Donaldson, W., editors, *Organization of memory*, pages 381–402. Academic Press, New York.

Twain, M. (1880). *A Tramp Abroad*. American Publishing Company, Hartford, CT.

Tyler, A. and Nagy, W. (1989). The acquisition of english derivational morphology. *Journal of memory and language*, 28(6):649–667.

Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, 22(4):562.

Valian, V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition*, 35(2):105–122.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, 40(1-2):21–81.

Valian, V., Solt, S., and Stewart, J. (2009). Abstract categories or limited-scope formulae? the case of children's determiners. *Journal of Child Language*, 36(4):743–778.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142.

Van Lancker, D., Canter, G. J., and Terbeek, D. (1981). Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech, Language, and Hearing Research*, 24(3):330–335.

Van Marle, J. (1992). The relationship between morphological productivity and frequency: A comment on baayen's performance- oriented conception of morphological productivity. In Booij, G. and van Marle, J., editors, *Yearbook of Morphology 1991*, Yearbook of Morphology, pages 151–163. Springer Netherlands.

Vapnik, V. (2000). *The nature of statistical learning theory*. Springer, Berlin.

Vihtnan, M. M., DePaolis, R. A., and Davis, B. L. (1998). Is there a "trochaic bias" in early word learning? evidence from infant production in english and french. *Child development*, 69(4):935–949.

Villavicencio, A., Idiart, M., Berwick, R. C., and Malioutov, I. (2013). Language acquisiiton and probabilistic models: Keeping it simple. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 1321–1330.

Vollmann, R., Sedlak, M., Müller, B., and Vassilakou, M. (1997). Early verb inflection and noun plural formation in four austrian children: the demarcation of phases and interindividual variation. *Papers and studies in contrastive linguistics*, 33:59–78.

Walker, J. (1936). *Walker's rhyming dictionary*. Dutton, New York.

Wallenberg, J. C., Ingason, A. K., Sigurðsson, E. F., and Rögnvaldsson, E. (2011). Icelandic parsed historical corpus (icepahc). *Version 0.9. Size*, 1.

Wang, Q., Lillo-Martin, D., Best, C. T., and Levitt, A. (1992). Null subject versus null object: Some evidence from the acquisition of chinese and english. *Language acquisition*, 2(3):221–254.

Wang, W. S.-Y. (1969). Competing changes as a cause of residue. *Language*, 45(1):9–25.

Wang, W. S.-Y. (1979). Language change: A lexical perspective. *Annual Review of Anthropology*, 8:353–371.

Weinreich, U., Labov, W., and Herzog, M. (1968). Empirical foundations for a theory of language change. In Lehmann, W., editor, *Directions for historical linguistics: A symposium*, pages 95–195. University of Texas Press, Austin.

Weir, R. H. (1962). *Language in the crib*, volume 14. Mouton.

Weist, R. M. and Witkowska-Stadnik, K. (1986). Basic relations in child language and the word order myth. *International Journal of Psychology*, 21(1–4):363–381.

Weist, R. M., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E., and Konieczna, E. (1984). The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language*, 11(2):347–374.

Werker, J. F. and Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1):49–63.

Westermann, G. (2000). A constructivist dual-representation model of verb inflection. In Gleitman, L. and Joshi, A., editors, *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*, pages 977–982. Lawrence Erlbaum.

Westfal, S. (1956). *A study in Polish morphology: The genitive singular masculine*. Mouton, 's-Gravenhage.

Wexler, K. and Culicover, P. (1980). *Formal principles of language acquisition*. MIT Press, Cambridge, MA.

Weyerts, H. and Clahsen, H. (1994). Netzwerke und symbolische regeln im spracherwerb: experimentelle ergebnisse zur entwicklung der flexionsmorphologie. *Linguistische Berichte*, 154(4):430–460.

Wiese, R. (1996). *The phonology of German*. Clarendon, Oxford.

Wiese, R. (1999). On default rules and other rules. *Behavioral and Brain Sciences*, 22(6):1043–1044.

Williams, E. (2007). Dumping lexicalism. In Ramchand, G. and Reiss, C., editors, *The Oxford handbook of linguistic interfaces*, pages 353–382. Oxford University Press, Oxford.

Wolf, M. and McCarthy, J. J. (2009). Less than zero: Correspondence and the null output. In Rice, C. and Blaho, S., editors, *Modeling ungrammaticality in Optimality Theory*, pages 17–66. Equinox, London.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language & communication*, 18(1):47–67.

Wu, J., Cheng, Z., and Pan, S. (2014). *New Age Chinese-English Dictionary*. Commerical Press, Beijing.

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.

Xu, F. and Tenenbaum, J. B. (2007). Word learning as bayesian inference. *Psychological review*, 114(2):245.

Yamamoto, K. and Keil, F. (2000). The acquisition of japanese numeral classifiers: Linkage between grammatical forms and conceptual categories. *Journal of East Asian Linguistics*, 9(4):379–409.

Yang, C. (2000). Dig-dug, think-thunk: Review of steven pinker's *words and rules*. *The London Review of Books*, 22(16):42–43.

Yang, C. (2002). *Knowledge and learning in natural language*. Oxford University Press, Oxford.

Yang, C. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences*, 8(10):451–456.

Yang, C. (2006a). *The infinite gift: How children learn and unlearn the languages of the world*. Scribner, New York.

Yang, C. (2006b). A stochastic model of morphological change. Ms., Yale University.

Yang, C. (2012). Computational models of syntactic acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(2):205–213.

Yang, C. (2013a). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, 110(16):6324–6327

Yang, C. (2013b). Who's afraid of george kingsley zipf? or: Do children and chimps have language? *Significance*, 10(6):29–34.

Yang, C. (2015). Negative knowledge from positive evidence. *Language*, 91(4):938–953.

Yang, C., Ellman, A., and Legate, J. A. (2015). Input and its structural description. In Ott, D. and Gallego, A., editors, *50th anniversary of Noam Chomsky's Aspects of the Theory of Syntax*. MITWPL.

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.

Yip, M., Maling, J., and Jackendoff, R. (1987). Case in tiers. *Language*, pages 217–250.

Zaenen, A., Maling, J., and Thráinsson, H. (1985). Case and grammatical functions: The icelandic passive. *Natural Language & Linguistic Theory*, 3(4):441–483.

Zaliznjak, A. A., editor (1977). *Grammatičeskij slovar' russkogo jazyka: Slovizmenenie*. Russkij jazyk, Moskva.

Zhang, J. and Lai, Y. (2010). Testing the role of phonetic knowledge in mandarin tone sandhi. *Phonology*, 27(01):153–201.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley, Cambridge.