

What Momma Still *Had to* Teach You • Towards a Variability Model of Early Morphosyntactic Development*

Dominik Rus
(dr59@georgetown.edu)

1. INTRODUCTION

1.1 LANGUAGE AND ITS ACQUISITION IN A HISTORICAL CONTEXT

Language acquisition theories investigating human language development have always gone hand in hand with the developments in linguistic analyses and—more broadly— with the general philosophical and scientific thoughts of the era.

Systematic descriptive linguistic analyses by classic Indian and Greek grammarians, for example, were challenged by a Cartesian view (e.g., René Descartes and Galileo Galilei) that one of the most remarkable abilities in human communication is the ability to use linguistic signs to express freely-formed thoughts, marking the “true distinction” between man and animal, or man and machine, whether a machine is to be understood as the automata of the 17th centuries or modern computers that capture our imagination today (Chomsky 2000).

Cartesian thoughts on linguistic nativism and discrete infinity were pronounced spurious by empiricists, who in the 17th and 18th centuries revolutionized the scientific thought by claiming that all human knowledge derives from the senses, hence proposing that what makes us human is basically our experience (e.g., Francis Bacon, David Hume, and John Locke). Empiricists dominated the entire humanistic thought until the mid 20th century, initially surfacing mainly as physiologists or classical behaviorist psychologists (e.g., Ivan Pavlov and John B. Watson), promoting scientific research without recourse to inner mental states and human learning via trial-and-error, and denying any independent

* Many thanks to Donna Lardiere for detailed comments on the entire first draft of this paper. Special thanks also go to Pritha Chandra, Héctor Campos, Nina Hyams, Raffaella Zanuttini, Bob Frank, Charles Yang, and Susie Wurmbrandt for discussing with me some of the issues raised in this paper. All usual disclaimers apply.

significance for the mind. Later behaviorism became methodological (e.g., George Romanes and Edward R. Thorndike), acknowledging behavior to be either the only or the easiest method of observation and considering its significance for drawing conclusions about mental states. A reaction to this was radical behaviorism (e.g., B.F. Skinner), which—in contrast to the methodological behaviorist school—required no inter-observer agreement, postulated no mechanistic or reductionist learning, and argued against the idea that hypothetical mental states were simply causes of behavior, further claiming that the phenomena must be observable at least to the individual experiencing them (Baum 2005). Having expanded behavioral principals to processes within the organism and applying their lines of research to other cognitive domains—including language—radical behaviorism was directly responsible for viewing a language acquisition process as input-matching, comprising learning by trial-and-error, repetition of linguistic structures, and learning by analogy as the main inductive ways used in construing one’s language. In the context of language development, Paul, for example, spoke of novel forms in child grammars being “formation by analogy” (Paul 1891, cited in O’Grady 1997), while Bloomfield claimed that a regular analogy permits a speaker to utter speech-forms which he has not heard before and that the speaker utters new forms based on the (similar) properties heard before (Bloomfield 1933, cited in O’Grady 1997). Linguists at the time were interested primarily in the functional realms of language, mainly focusing on language description and typology (Chomsky 1986, 2000) and a “revolution” was about to happen.

The intellectual movement in the 1950s, generally referred to as “cognitive revolution” in the sciences and humanities, was not really a revolution, but a “counterrevolution” to revolutionary ideas of various behaviorist schools. It began in the modern context of greater interdisciplinary connections, particularly between (cognitive) psychology and linguistics, collectively now referred to as cognitive sciences, together with computer science, neuroscience, and philosophy. Cognitivists argued that the mind should be scientifically fully observable and that for its understanding, mental functions should be understood as the “internal”, rule-bound manipulation of symbols. Perception was no longer seen as discrimination, memory was no longer considered learning, language was believed to be creative, mental, and biological rather than verbal behavior,

and intelligence no longer meant intelligence tests scores (Miller 2003). Together with cognitivism came *linguistic nativist approaches* to language development, holding that children bring to the world hard-wired machinery containing pre-specified universal sets of principles that interact with the input, yielding adult grammatical competence. It comes as no surprise then that Chomsky has always noted that the study of language and its acquisition is in a way quite a young enterprise despite all the classic descriptive and typological work done prior to the birth of cognitivism (Chomsky 2000). In the so-called biolinguistic tradition (Chomsky 2000; Jenkins 2000), most of language is believed to be *selectional* rather than *instructional* (by analogy with the modern immunology theory that has been applied to theories of neurobiology and, ultimately, cognitive faculties; Gazzaniga 1992), *hard-wired* rather than *emergent*, and—most likely—subjected to *natural selection* (Gazzaniga 1992; Pinker 1997, 2005), all properties of *biological organisms* (Gazzaniga 1992; Piatelli-Palmarini 1989).

Methodological, biological, and particularly purposive behaviorism, absorbed into experimental and cognitive psychology today, is taking its revenge in the form of various information processing, connectionist, and other computational theories of human language and its acquisition (e.g., Elman 2005; Elman et al. 1996; Saffran et al. 1996; Sampson 2005). Still cognitive in nature, these computational or processing approaches to language and learnability draw from mathematical dynamic systems, stochastic psychology models of learning in animals and humans in linguistic and non-linguistic domains, as well as neuroscience and neuropsychological evidence from linguistic and non-linguistic domains, most notably motor control, vision, and memory (see Gazzaniga 1992; Pinker 1997). Most of these approaches explicitly acknowledge the role of biology (i.e., genetics) in language development (Elman 2005; Elman et al. 1996; Tomasello 2003), but denying Chomsky's epistemological (representational) type of innateness (Karmiloff-Smith et al. 1998; see also Fodor 2000; Jackendoff 2002;), and holding that innateness should be viewed *architecturally* rather than *representationally* and that language-specific brain mechanisms more like a “toolbox” with innate computation for language processing rather than machinery with (fully) genetically pre-specified parameters (see Carroll 2001; Jackendoff 2002).

In terms of language learnability theories then, the scientific community has been divided for decades into those who think that children are *imperfect* (smelly, dirty, noisy) creatures who make many “mistakes” in their slow and effortful course of language development (Crystal 1994, cited in Radford 2000) and those who think that children are *extraordinarily fast* and *efficient* language learners, conforming to the adult grammatical system extremely early without any explicit instruction and the need for negative evidence (Anderson and Lightfoot 2000; Crain, Gualmini and Pietroski 2005; Crain and Pietroski 2006; Gualmini and Crain 2005). The former—empiricist—approach has always tried to downplay the domain-specific *knowledge of language*, hence minimizing the role of biology in language acquisition and maximizing the role of the input as well as the learner’s sensitivity to it. On the other side of the fence have sat the rationalists, arguing for a richly specified *innate knowledge* of language, maximizing the genetics (and maturation) and minimizing the role of experience.

1.2 SETTING THE STAGE: A LESSON FROM A PERSONAL ANECDOTE

Though the “language acquisition debate” at the beginning of the 21st century is still very hot, it is by no means as black and white anymore as it was a few decades ago. A conversation between two graduate students at a BUCLD conference¹ who call themselves *cognitive interactionists* or, more precisely, *constructivists* (denoting a most recent subspecies of cognitive interactionists) makes this point more than obvious. In their 5-minute poster talk during which they will outline their general theory, as well as give you the specifics on their study, they will remind you that the *externally-driven* (or *input-driven*) approach to language learnability posits that the child comes equipped with *no* specific linguistic predisposition for language, assuming that the process of language acquisition arises as the interplay between general cognition and the linguistic input. This input-matching approach predicts language acquisition to be a slow(er), (more) effortful and piecemeal process, making use of memorization (such as schema- and rote-learning), as well as being heavily dependent on the input and its characteristics (e.g., frequency, consistency, saliency; see Tomasello 2003). Frequency is believed to predict the order of specific forms and constructions and (at least very early) language learning is posited to

¹ Boston University Conference on Language Development (BUCLD), held annually at Boston University in Cambridge, MA. The reference here refers to BUCLD30, held in November 2005.

be item-or construction-based, giving rise to a high number of errors and generalizations². The most interesting part of the discourse, however, comes when the two of them disagree on a number of issues regarding language acquisition after their mini talk is over and when the time seems to be ripe enough for more conceptual issues. While the first one argues that there is nothing like *syntax* in the mind of a speaker, but rather only a “meshed” level of *syntax-semantics* and that the learner initially constructs the grammar solely on the basis of linguistic experience, structuring language around these language-specific syntactic-semantic “pieces”, guided by some domain-general mechanisms, the second one ardently objects to this by saying that she believes there are both species-specific and domain-specific linguistic properties that humans are born with, including some syntax-specific properties. However, she explicitly disagrees with positing such properties as “this and that constraint” or “this and that parameter”. In her own words, linguistic-specific properties are much *more general* (e.g., building a binary tree-like phrase structure), possibly falling out of processing considerations, simply because the mammalian brain is wired the way it is (and possibly because of the way it has evolved).

The discourse continues when an avid (or so we thought) Chomskyan rationalist comes along and starts a conversation with the two constructivists. He disagrees with the idea that all acquisition can be explained by the input or the interaction between the input and human cognition, particularly since “human cognition” in interactionist theories is so vaguely-defined. He recites his graduate school manifesto that if the idea that language is an *empirical inquiry* is to be taken seriously, linguistics should aim to identify the abstract properties of the biological object we refer to as *human language* and hence language acquisition should be formalized as the scientific investigation that seeks for internalized knowledge in a speaker’s mind based on such biological property. He continues by saying that UG-based language acquisition theories rely on the assumption that at least some (if not most) of the language machinery is innately specified (*the Innateness Hypothesis*), thus facilitating the process of language acquisition, claiming that he cannot understand how the interactionists can even possibly conceive of language

² It follows from this “copy-paste” approach (Crain 2003) that a child will not be able to form (m)any “deep generalizations” about the linguistic material she is exposed to, since most (if not all) beginning language learning is “surface” learning (i.e., this is sort of a “what-you-see-is-what-you-get” approach (Crain 2003), or, in the constructivists’ terms, the “emergentist approach”, where particular structures are believed to *emerge* through the interplay between the input and rich human cognition.

learning being almost entirely dependent on the input. How very rationalist, right? Well, not quite so. He *does* add that the view that the child's grammar is viewed as fundamentally *the same* as the adult's grammar might not be on the right track and that the idea that when a child is making an error, she is simply entertaining one of the possibilities given by biologically-constrained UG and found in *any* natural language out there in the world, presumably, or at least in theory, either “dead” or “alive” (Crain, Goro and Thornton 2006; Gualmini and Crain 2005; Pietroski and Crain 2006; but see Newmeyer 2004). He adds that he believes that this sort of theorizing has handed down to new generations of generative acquisitionists like a *mantra*.

The two constructivists agree with such reasoning and are taken aback even more when the nativist announces that he does *not* believe in the classical construct of *parameters* either and that what generativists call a “parameter” may simply be a nice descriptive tool to account for language variation, and even that has been somehow unsuccessful considering the “algebraic” nature of generative grammar with +/- settings for a certain parameter (see Newmeyer 2004). Besides, as the nativist explains, the theory of parameters comes in wildly *different* flavors. While the “vanilla” parametric account à la Pinker and Bloom (1990), where the parameters are initially given with no actual settings specified, the “chocolate rocky road” flavor à la Chomsky (1986) account argues for a “small” number of parameters with binary options, one of which is initially specified as a *default* value (see also Hyams 1986); the extreme case, or the “double chocolate cinnamon swirl cookie dough” approach, as voiced in Baker (2003), for example, contrasts sharply in not only positing a large hierarchy of parameters and their interdependencies, but also argues for *overspecification*, i.e., the idea that parameters come with *all* values already given, in which case the child merely selects the appropriate value, based on the input (see also Tesan and Thornton 2004). But the theory of parameters saw no end to this. Every single issue of *Linguistic Inquiry*, for example, would propose some 10 new parameters and there simply seemed to be no upper bound of both, the *number* as well as the *nature* of these parameters. The generativist then concludes his speech by saying that a UG model is desirable at least on conceptual grounds (i.e., to explain some “universal”, “wired-in” language-specific properties, such as structure building via *Merge*, hierarchical structure-dependencies, and—possibly—

displacement), but in order to account for more explanatory power of such a model (i.e., the credo on which all empirically-driven interactionist theories rest), the “nurture” component has to be somehow incorporated into the “nature” component. Parameters might, in his view, only be certain innate biological biases (à la Kirby 2005), possibly language- and species-specific, though the researcher says that he wants to remain agnostic about this issue (see Section 4 below).

2. ACQUISITION VS. LEARNING IN CHILD MORPHOSYNTAX, OR WHAT IS ALL THIS FUSS ABOUT?

2.1 THE CASE OF ROOT INFINITIVES IN THE GENERATIVIST TRADITION

Let us unpack some of the conceptual claims from above in more detail and show how they have led to the development of some explicit mechanisms in empirically-driven acquisition contexts.

Take for example one of the most frequently-studied phenomena in the child morphosyntactic acquisition literature in the last decade or so – the phenomenon of *Optional (or Root) Infinitives (RIs)*, in which children between the ages of 1;5 and 3;0 cross-linguistically opt for an infinitive verb (or bare verb stem form) in matrix clause contexts, where the target (adult) system requires a tensed form, denoting tense and/or agreement. (1) below shows some examples of RIs from a variety of child languages:

- (1) a. Papa schoenen *wassen.* (Dutch; from Salustri and Hyams in press)
daddy shoes wash-inf
b. Mumma *ride* horsie. (English; from Guasti 2002)
c. Paula *play* ball. (English; from Radford 1990)
d. Auch Teddy fenster *gucken.* (German; from Salustri and Hyams in press)
also Teddy window look-inf
e. Kisa *finna* dúkkinni. (Icelandic; from Salustri and Hyams in press)
cat find-inf the-doll
f. Jag också *hoppa* där å där. (Swedish; from Salustri and Hyams in press)
I also hop-inf there and there

In the generative language acquisition research, the RI phenomenon has been captured under the umbrella of a broader research program, generally referred to as *the acquisition of finiteness*, and commonly concerned with the acquisition of the Verb Movement Parameter (e.g., Hornstein and Lightfoot 1994) and the functional categories, such as T(ense), Agr(eement), and C(omplementizers). Researchers have generally

investigated the distinction between “finite” and “nonfinite” verbs (e.g., Guasti 1993/1994; Poeppel and Wexler 1993) and the emergence of functional categories in early grammars (Poeppel and Wexler 1993; Rizzi 1993/94, 2005; Wexler 1994, 1998; for a detailed review see Guasti 2002). In this line of research, the term “finiteness” has been linked to the morphosyntactic elements that either directly express person or person/number and/or tense (e.g., main lexical verbs, copula verbs, auxiliary verbs in periphrastic (compositional) tenses, and modal verbs), or linguistic elements related to those (e.g., subject clitics or reflexive clitics).

A word on “finiteness” is in place here. Traditional terms *finite* and *nonfinite* are extremely difficult to define precisely and have led to some confusion in the field. Traditionally, a *finite* verb form has been described as the one expressing the grammatical category of *person* (hence, *agreement* or *phi-feature agreement* in today’s terms; see Chomsky 1995) whereas the nonfinite form the one lacking such agreement, though modern generative literature generally assumes that finiteness is a phenomenon associated with a lack or presence of a T head (or its featural specification, or even its lexical contents) in a phrase marker, rather than with the lexical form itself (see e.g., Chomsky 1995). For example, infinitives are traditionally “nonfinite”, but may show phi-feature agreement (e.g., inflected infinitives in Portuguese, infinitives inflected with clitic pronouns in Italian, etc.). Conversely, one can deal with a nonfinite clause in which T is present (e.g., in *to*-infinitival clauses in English). Hence finiteness does not really go hand in hand with the presence of T (or its lexical realization), much less with the presence of Agr (unless this is specified to be [+person] vs. [-person], in which case only the former would count as finite in traditional sense. But then the confusion arises as to how to treat merely Agr-related material (in the strict sense of the Split-Infl Hypothesis, arguing for the existence of both projections, T and Agr) in terms of the T head specification. Moreover, the term finiteness may well refer to language-specific lexical material and there may *not* be a universal treatment of the finiteness phenomenon at all (see e.g., Wurmbrandt 2005). Here I will use traditional terms “finite” and “nonfinite” in order to be consistent with the previous research, and particularly, when reporting the analysis and findings of previous studies that used these terms exclusively.

The acquisition of finiteness in this approach has also been directly related to the notion of Agree – a relation between the subject and the predicate of the clause (Chomsky 1995), and as such, the UG-based child language acquisition field has focused on the acquisition of *agreement paradigms* as well (for detailed reviews see Deen 2002; Guasti 2002; Hyams 2002). Under this approach, the Tense and/or Agreement node in a phrase structure is initially *underspecified* due to a different (immature) feature-checking option (Wexler 1994, 1998), *pruned* out of a phrase structure due to the operation of phrase structure truncation (Rizzi 1993/1994, 2005), or is *missing completely* due to “immature” grammar (Platzack 1992; Radford 1990, 1995). The immature or missing T/Agr properties are believed to be either learned in a piecemeal fashion, but (possibly) guided by UG (Varlokosta et al. 1998) or to arise via *biological maturation* (Radford 1990, 1995; Rizzi 1993/4, 2005; Wexler 1994; 1998). Yet another line of research has linked the RI phenomenon to (non-syntactic) performance limitations. Avrutin (1999), for example, argues that RIs result from children’s limited resources to carry out utterance planning and other pragmatic computation, similar to the ideas about early null subjects in child English (e.g., Bloom 1990).

Such *internally-driven* approach to the acquisition of (morpho)syntax predicts fast and effortless acquisition of finiteness with *very low* agreement error rates and the child's robust sensitivity to the morphosyntactic properties of the target system coming online extremely early in the course of language development, a phenomenon which Wexler (1998) refers to as *Very Early Knowledge of Inflection (VEKI)* and Hoekstra and Hyams (1998) as *Early Morphosyntactic Convergence (EMC)*.

The following two quotes below outline a common stance of researchers working in a UG-based model:

Indeed, there is substantial evidence that children acquire the specific morphosyntax of the target very quickly. Parameters such as V to I, V2, and so on are set very early. Early English is essentially English, early German is essentially German, and so on. These findings of “early morphosyntactic convergence” (Hoekstra and Hyams (H&H) 1998) seem to fly in the face of stage theories, which purport to show that there are universal syntactic stages in acquisition. Yet, there is also substantial evidence that children acquiring different languages show similar developmental effects.

(Salustri and Hyams in press: 1)

A large amount of recent research has shown that two-year-olds know much of the syntax of their language, particularly the system of inflection and verb movement.

To the extent that they produce non-adult forms, it has been argued that these are generated by the child's grammar, which differs in some minimal, circumscribed way from the adult grammar /.../.
(Schütze and Wexler 1996: 1)

It appears that the idea that kids are “perfect” language learners (Crain and Pietroski 2002, 2006; Montrul 2004; Pietroski and Crain 2005; Radford 2000) or even “more perfect than adults” in (morpho)syntax (Radford 2000) is so deeply-rooted in the generative linguistic mind that—I believe—has seriously hindered a lot of potentially new data-driven research. First, most generativists work with the *existing* data, generally found in the CHILDES database (MacWhinney 2000), so most researchers would not look at other languages, not documented on CHILDES. Second, most generative RI accounts persistently rely on the idea that RI is a *structural (syntactic)* deficit, usually entirely glossing over morphology or semantics (for some exceptions, see Phillips 1996 for the former and Gavrusseva 2005 for the latter). Third, none of the deterministic (undifferentialist) accounts has ever taken into account the nature of the input to a level beyond simply stating that a particular form is not expected to appear in a child's grammar since it is not attested in the target language or that early syntactic properties found in the input are important merely for *parameter setting* (and once this is done—pretty much in a snap (Lightfoot 1991; Pietroski and Crain 2002)—the kid's job is pretty much done). In other words, there has simply been no room for *learning* (in the old sense, coming from psychology and stochastic statistical models) in any of the recently entertained models of language acquisition. And we know that kids are both “growing” and “learning”, the biology interacting with the experience. The same way grammar (or, UG) is both *growing* and *learning*.

Interestingly, though I have worked in a UG-based paradigm ever since my first publication, I would always *disagree* in my *Future Research* sections or *Conclusions* with the “static” approach of UG-based accounts, adding my two cents to the more general discussion on externally- vs. internally-driven acquisition theories:

Last but not least, I would like to add a conceptual observation regarding the entire research program concerning the acquisition of early verb morphosyntax. We know that children are extremely sensitive to morphosyntactic and semantic properties of the target language from the earliest stage on, but we know that we do find empirical differences among child systems in terms of early finite and nonfinite forms. Hence, (a) we either have not been able to pinpoint the right (biological) mechanism responsible for these differences or (b) there might be some learning/statistical mechanism(s) at stake which divide child grammars into RI languages vs. non-RI languages (or even more narrowly, into BV [Bare Verb stem] languages vs. RI languages vs. BP [Bare Participle; see below] languages). Since UG-based accounts put forth an internally driven developing system rather than an externally driven one, factors such as frequency, consistency, and saliency in the input have been largely ignored.

(Rus in press: 15)

and:

And last but not least, since UG-based accounts put forth an internally-driven developing system rather than an externally-driven one, factors such as frequency, consistency, and saliency in the input have been largely ignored. I believe that the field needs to move beyond this slightly static approach and start looking at the findings of corpus linguistics which may provide some answers to the frequency and use of early root nonfinites and specific finite forms in child corpora. I am definitely not saying that the child is an input-matcher and cannot project beyond her experience, which is what a constructionist approach, which we briefly reviewed in the introductory remarks, would claim, but corpus linguistics (analyzing mother-child, child-child, as well as adult corpora) might give us some new insights into why children initially seem to prefer more/most polyfunctional verb forms (e.g., progressive *-ing* in English, perfective *-i* in Greek, possibly participial *-l* in Slovenian) and to what extent these early forms are conditioned by the input (if at all). /.../ Children are definitely sensitive to morphosyntactic and semantic properties of the target language from the earliest stage on, but since we do find empirical differences among child systems in terms of early finiteness, I believe that there must be some biological and/or input-dependent mechanisms at stake which divide child grammars into RI languages vs. non-RI languages (or even more narrowly into BV languages vs. RI languages vs. BP languages; see Rus in press). More research into early verb morphology is needed, particularly on morphologically complex languages and by combining the tools of generative linguistic theory and corpus linguistics.

(Rus submitted: 13)

2.2 EMPIRICISTS STRIKE BACK: *LEARNING* IS THE KEY

Interestingly, the above-mentioned study by Phillips (1996) has been one of the very few existing generative accounts that have attributed the RI phenomenon to the lack of *morphological knowledge*. Though Phillips still attributes the “deficit” to the core syntax (specifically, the failure of merging the Tense node with the Verb node), his study

explicitly relies on the assumption that this failure results from syntax interacting with immature morphological knowledge, or as he puts it:

Therefore I suggest that the cause of the children's problem is not a problem with syntactic or morphological knowledge, but difficulty they encounter with connecting the two systems. In particular, for some reason there's some difficulty associated with one-to-many relations between morphological forms and bundles of syntactic features. If children don't connect their verbs with inflection syntactically, then they don't have to face this morphological problem. The one-to-many access problem is clearly going to be more acute in less richly inflected languages, like English and Swedish, and therefore we expect RIs to be used more often and decline at a later age in children learning less highly inflected languages.
(Phillips 1996: 8)

Note that this implicitly assumes that the kid's morphology must be "correct" in order to drive syntactic derivations. What does not seem to be correct, or adult-like, however, is the child's connecting syntax with morphology. Phillips himself did not pursue this idea further empirically, nor did he spell out how *language-specific morphology* might matter in a broader learnability context, but he was quite explicit in saying that the children's problem is not an immature representation in the syntax, but rather the child's morphological system, which is—due to immature head movement—incapable of inserting the right (inflected) form. The result is simply a default/unmarked infinitive, which, according to Phillips, need *not* be an infinitive. Phillips concludes his account by stating:

I am assuming that there is *some* difference between the adults and the children—what I hope to have demonstrated is that the *syntax* of root infinitives has just the characteristics of a fairly widely distributed phenomenon in adult languages, and so root infinitives really don't show that there's any difference at all between children's and adults' syntax. This strengthens the conclusion from Section 2 that root infinitives are not due to a deficit in children's syntactic representations, and are realizations of entirely finite clauses.
(Phillips 1996: 10; emphasis in the original)

UG-based studies of early morphosyntax of morphologically complex languages such as Spanish, Italian, and Catalan have shown over and over again that the agreement error rates fall below 5% with no attested RIs and occasional bare present and past participles (see Montrul 2004 for Spanish and Catalan; see Deen 2002, Guasti 2002 and Hyams 2005 for detailed reviews on all early languages investigated so far). Wexler's (1998) account putting forth the Uniqueness Checking Constraint Theory (UCC) and VEKI, for example, concluded that children are "little inflection machines", knowing the

verb movement properties and hence the inflectional characteristics of their language before they reach a two-word stage. It has been hence taken for granted that children are *super fast* morphology learners, though *nothing* has ever been said about the learning itself. There is usually a period of at least a year before such learning becomes “perfect”, and in the case of morphologically more impoverished languages (e.g., English or Swedish), such a period can be prolonged to two years or beyond. That morphology—and hence *learning* in the stochastic sense—is definitely at stake here shows a study by Aguado-Orea and Pine (2005), which seems to directly contradict Wexler’s (1998) study. Aguado-Orea and Pine (2005) (henceforth, A&P) showed that although the agreement error rates in child Spanish (in the present tense) are extremely low (4.4% and 4.5% for the two children studied), the error rates vary *sharply* by the person/number *context*, with the error rates for 3rd plural (3p) present tense forms reaching more than 50%, as seen in the tables below, taken from A&P (p. 2):

(2a)

Error rates by context controlling for knowledge of relevant inflection (Juan)

Inflection	Context	Correct	Errors	%
1s	693	659	34	4.9
2s	146	131	15	10.2
3s	2001	1986	15	0.7
3p	249	173	76	30.5

Overall error rate hides very different error rates in different parts of the system

(2b)

Error rates by context controlling for knowledge of relevant inflection (Lucía)

Inflection	Context	Correct	Errors	%
1s	510	496	15	3.0
2s	118	96	22	22.9
3s	1023	1018	5	0.5
3p	30	13	17	56.7

Same pattern in Juan and Lucía’s data

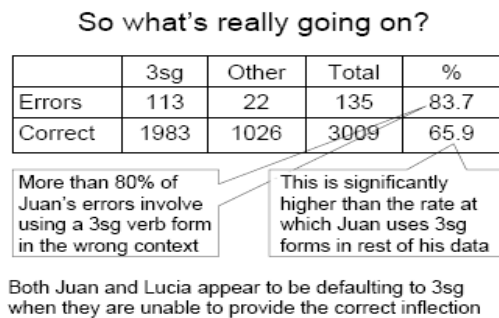
As seen from the tables above, the overall data (as generally reported by UG-based studies) get extremely skewed due to practically non-existing errors in the third person singular (3s) system of the paradigm³. While 2nd singular (2s) forms in Juan’s sample show a 10.2% error rate, Lucía’s error rate for these tokens raises to almost 23%. This is even higher in the case of 3p, where Lucía performs worse than at chance. Obviously there is a huge discrepancy between reporting a 96% and a 45% agreement correct rate. When statistically controlling for the outliers (by discarding 3s from the count), the

³ Note that this may indeed suggest that Spanish 3s is some kind of “default” form or “basic” form that gets spelled out when children do not know the correct inflected form, as claimed by Aguado-Orea & Pine (2005), but the status of 3s is still *not* known, particularly as there have been claims that 3s is, in fact, a root nonfinite form, analogous to RIs (Davidiak and Grinstead 2004; see below).

overall error rate rises to 15.2% for Juan and 27.5% for Lucia. Now this doesn't look that impressive anymore at all.

Morphological learning also seems to be implicated in the choice of the children's first verbs. In Juan's sample, according to A&P, *quiero* (want.1s pres) accounts for 53.7% of all verb tokens in the data, with *puedo* (can.1s pres) at 10.8%, *tengo* (have.1s pres) at 8.2% and *pongo* (put. 1s pres) at 4.4%. Roughly 75% of all verbs are *highly frequent* verbs, which, again, is a piece of data that will not do any good to any UG-based account (and even less so if these are *rote-learned* and appear only in a very limited number of contexts or "schemas"; see Goldberg 1995; Tomasello 2003). Lucia shows exactly the same developmental pattern, with as much as 71% of *quiero*, 9% of *puedo*, and 3.5% of *tengo*. What is more, A&P show that as most frequent verb tokens are systematically excluded from the analysis (similar to an outlier situation from the data above), the error rate drastically grows. Both Juan and Lucia show a steep rising error curve going from *quiero* to the most infrequently-used verbs such as *vuelo* (fly. 1s pres) and *parezco* (appear. 1s pres) in the case of Juan and *saco* (gather, collect, take.1s pres) and *salto* (jump, hop. 1s pres) in the case of Lucia. On the basis of all this evidence, A&P provide the following chart (p. 3):

(3) The rate of Juan's errors broken down to different "systems"



A&P conclude that the 3s pres form is by far the *most common* in child Spanish and thus appears to be the simplest/least-marked finite form (p.3). They hypothesize that the reason why 3s is so much more accessible than the infinitives (or the 3s pres form in other languages, for that matter) is the position of finite verbs in the input. They do not explicate this issue further, but claim that their count shows that 74% of all finite verbs in

Spanish appear in an utterance-final position (vs. 13% in Dutch, for example). They further also propose *gradual acquisition* of compound (compositional/periphrastic) forms in child Spanish and the tendency to default to the simplest/most frequent form when children are unable to produce the correct form.

I wanted to see whether the data that Rus and Chandra have been working with have been skewed this way too, particularly due to the fact that they were collapsed across 15 children (Rus and Chandra 2005, in press; Rus submitted, in press). I first looked at the agreement correct (Agr corr.) rates for finite verbs in the present tense (which is the only synthetic tense in the language, the future and past tenses being periphrastic/compositional), but this time broken down for each person/number cell of the paradigm. (4a) is what was reported by Rus and Chandra (2005) and (4b) was reported in Rus submitted:

(4a) Agr correct (across all 15 children) (4b) Pres tense finite verbs (across all 15 children)

Total finite Vs	#	187
Agr corr. (“real” Agr errors)⁴		183 (=97.9%)

	SG	DU	PL
1	10	N/A	N/A
2	4	N/A	N/A
3	168	N/A	1

We reported 183 verb tokens for finite verbs in the present tense paradigm across all children, but we quickly see from (4b) that 168 (or 91.8%) of all verb tokens are 3sg tokens. This, of course, raises the issue of whether the children truly *know* the agreement. Rus (submitted), following Hyams (2002, 2005), writes that knowing agreement does *not* entail that the child will need to have all of the paradigm forms in her grammar, or:

Slovenian children seem to acquire the verb paradigm very early. However, DU forms as well as 1PL and 2PL forms are not found in the data, with 3SG being the most frequent. This conforms to the previous findings in the field, namely that the singular inflection is the most common and that plural inflections appear later than singular in the course of language development. We found only one case of 3PL with finite verbs. /.../ However, 3PL inflection appears quite frequently with COPBE [Copula BE], so it is probably not true that crosslinguistically early verbs will generally appear in the singular only. Also, DU forms might not be found simply because there might not be any DU contexts. From the transcript alone, it is hard to determine whether the child is addressing one or two interlocutors. The children reported here most often refer to themselves (1SG), address their speaker (2SG), or talk about the third person/object (either present or absent) (3SG).

⁴ The term “real” agreement errors refers to the adjusted agreement rate after controlling for the common feature where the child is using a 3 sg verb instead of the 1 sg one (which, in the acquisition literature has always never been taken as an error since caretakers often talk to the children in 3rd person; Hyams 2005; p.c.).

We have a double-sword situation here. An empiricist would definitely argue that the figures above do *not* show that children truly *know* agreement, but a generativist may blame it on the context or maturation and get away with it. The complication arises, however, when one starts thinking of a 3sg (pres) form as being a *default/easiest finite form*, simply spelled out when the child does not know the “correct” inflection, as argued by A&P. Interestingly, the high suppliance rate of 3sg pres forms implies that it may be that it is this form that is analogous to an RI form found in other child languages, which is what has been claimed for child Spanish (Davidiak and Grinstead 2004), where the 3sg verb in the present tense paradigm for *-ar* and *-er* verbs bears *no* phonological tense or agreement marker, but merely *a thematic vowel*, which in both descriptive and generative linguistic accounts have always been taken as part of the verb stem⁵. However, what is going on in child Spanish is probably (quite) different from the situation found in child Slovenian. Children acquiring Spanish do make an error by producing a 3sg pres form in 1sg and 2sg contexts with overt personal pronouns (e.g., *Yo habla*, I speak.3sg.pres), which suggests that it might be the case that 3sg pres is truly some RI analog (or more broadly, an early root nonfinite). In child Slovenian, such cases are not attested, though it is often not obvious at all from the context whether the child is referring to herself (1sg) or someone else (requiring 3sg), especially since overt subjects are hardly attested (Rus and Chandra 2005; Rus in press). Generally, 3sg pres correctly appears in 3sg context, though I agree that we need more data to see if Davidiak and Grinstead (2004) may be right about bare verb forms (containing only a thematic vowel with no overt tense/agreement) being RIs, which would be the same scenario in child Slovenian (see below, particularly (13) and (14) for Slovenian tense/agreement morphophonology). Ironically, however, constructionist accounts (e.g., the one outlined by A&P above)—to my knowledge—have never claimed that these may be *nonfinites*, but the

⁵ Campos (p.c.) reminds me that it has been argued that a Spanish *-a/-e* verb form (e.g., “nada” from “nad-a-r” or “come” from “com-e-r”) has two functions, or two different syntactic representations associated with it, namely the “real” 3sgpres form, denoting a [present][3][sg] form with *-a/-e* as the thematic vowel to which \emptyset is added for T/Agr, and a non-finite(-like) *-a/-e* form, with a \emptyset thematic vowel (and no T/Agr marking; below marked as Bare Verb, bareV), found in certain expressions that can be productive (in the sense that it is found in certain schematic or formulaic expressions that can take any verb from the Spanish lexicon), e.g.:

El competidor *nada* que *nada* por alcanzar la meta.
 the competitor swim.bareV what swim.bareV for achieve the goal
 ‘The competitor swims and swims (=swims a lot/all the time) in order to achieve his goal.’

Cf.

Juan *nada* en esta piscina todos los días.
 Juan swim.3sgpres in this pool all the days
 ‘Juan swims in this (swimming) pool every day.’

“easiest”/default finites whose high suppliance rate reflects the raw frequency with which these forms occur in the input as well as final-utterance “position saliency”. This may result from their learnability approach/theory itself where the term “finiteness” does not really have any theory-internal status, since grammar is taken to be more like a system of schemas, patterns, and “constructions” denoting various syntax-semantic mappings (e.g., Goldberg 1995; Tomasello 2003).

I calculated the error rates in the present tense verb tokens for five children from whom we have the most data, which provides us with a more reliable measurement than collapsing the data across all 15 children⁶. The following table shows the result:

(5a) Agr error rates for present tense verb forms for the five children studied

CHILD	finite Vs
Lenart	1/67 (1.5%)
Katja	2/45 (4.4%)
Tomaz	N/A
Vesna	N/A
Gordana	1/5 (20%)

The error rates are so low that it is hard to make any predictions. When looking at the actual verbs that appear with incorrect agreement, we get the following:

(5b) Finite verbs with incorrect Agr for the five children studied

CHILD	finite Vs
Lenart	<i>kak</i> (make a pooh; bare stem instead of 3sg)
Katja	<i>brisa</i> (mop; 3sg as conjugated in <i>-am</i> verbs instead of <i>-em</i> verbs plus no consonant stem change), <i>telefonira</i> (phone; bare stem instead of 3sg) ⁷
Tomaz	N/A
Vesna	N/A
Gordana	<i>poje</i> (sing; 3sg as conjugated in <i>-am</i> verbs instead of <i>-em</i> verbs)

The present tense verbs that appear in the incorrect form are indeed extremely low-frequency verbs in the data (with the exception of *peti*, “to sing”, no other verb occurs more than once). Is it possible that Slovenian children show the same trend in exhibiting errors with different percentage rates in different systems? It is possible. We definitely

⁶ Some children do not speak much, so it is hard to make any generalizations. The five children reported here speak a lot and most of the data (roughly 65%) come from these five children alone.

⁷ See below for details on the verb paradigms and stem changes in (adult) Slovenian, which will be relevant for our later discussion on morphological learning.

need more data to make a final conclusion here, but the trend that we observe reminds of the scenario found in child Spanish.

Rus and Chandra (2005), Rus (submitted, in press) and Rus and Chandra (in press) all reported *perfect* or *near-perfect* agreement on (bare) past participles, present tense finite verbs, and imperatives in child Slovenian respectively. They do mention, though, that the rate of *BE* omission—both as copula *BE* (COPBE), or auxiliary *BE*; (AUXBE)—is quite high, but do not provide any numbers as they concentrate on other verb forms in their studies. Rus (submitted: 16) concludes:

In sum, Slovenian children are sensitive to morphosyntactic and semantic properties of the target language from the earliest utterances on. They are faithful to the subject setting, word order (e.g., second position clitics), and Nominative subject case marking. They also distinguish between finite and nonfinite verbs, as seen in the knowledge of finite verb inflections and nonfinite complementation. They also show productive use of T/Agr verb morphology.

and:

We can claim with certainty that Slovenian children show perfect knowledge of subject-verb agreement. Rus and Chandra (2005) also found that these same children show perfect subject-participle (gender/number) agreement in past and future tense constructions involving active past participles. Hence we cannot but conclude that Slovenian children must have a mental representation of functional material as early as we can test them.

True, children may have (and probably *do* have) a mental representation of functional material, since it appears over and over again in the data. But how *perfect* children really are when it comes to *learning*? In other words, how good of *morphological learners* children really are? Is learning verb morphology and morphosyntax really a snap, like Hoekstra and Hyams (1998), Lightfoot (1991), Pietroski and Crain (2002) and Wexler (1998) suggest?

To investigate this, I looked at the children's performance on *all* "agreement markers" (i.e., all markers that express finiteness, as outlined above in Section 2). I computed the number and frequencies of all verb tokens for the same five children from whom we have the most data, comparing these to the agreement error rates in all the verb forms that these five children produced. (6a) shows the raw numbers of verb tokens and the frequencies associated with them:

(6a) Verb tokens and their frequencies for all verbs for the 5 children studied

C H I L D	#All V utterances	# (%) of PastParts	# (%) of finite Vs	# (%) of IMPs	# (%) of BE (COP or AUX)	# (%) All Vs w/finite markers excl IMPs
L	394	81 (20.5%)	67 (17%)	187 (47.5%)	59 (15%)	126/394 (32%)
K	356	31 (8.7%)	45 (12.7)	161 (45.2%)	119 (33.4%)	164/356 (46%)
T	81	38 (47%)	10 (12.2%)	21 (26%)	12 (14.8%)	22/81 (27%)
V	69	10 (14.5%)	21 (30.4%)	20 (29%)	18 (26%)	39/69 (56.5%)
G	58	10 (17.2%)	5 (8.7)	31 (53.4%)	12 (20.7%)	17/58 (29.3)

(6b) below shows the agreement error rates (by collapsing the figures from all of the existing Rus and Rus and Chandra studies on bare past participles (BPastParts), finite verbs in the present tense (finite Vs) and imperatives (IMPs)) across all 15 children, which are—again—impressively *low*, but compare these with the ones in (6c), which now show the agreement error rates on *all* verbs that denote finiteness (as assumed in the system put forth by UG-based accounts and outlined in Section 2 above):

(6b) Agr error rates for bare past participles, finite verbs, and imperatives across all 15 children

	BPastParts	Finite Vs	IMPs
Error rates	4.5%	2.1%	0.9%

(6c) Agr error rates for all verb forms in the data for the five children studied

CHILD	BPastParts	Finite Vs	IMPs	BE (COP/AUX)	All finiteness markers excl. IMPs
Lenart	4/81 (4.9%)	1/67 (1.5%)	3/187 (1.6%)	55/59 (93.2%)	56/126 (44.4%)
Katja	2/31 (6.4%)	2/45 (4.4%)	2/161 (1.2%)	55/119 (46.21%)	57/164 (34.7%)
Tomaz	N/A	N/A	N/A	5/12 (41.67%)	5/12 (41.67%)
Vesna	N/A	N/A	N/A	12/18 (66.6%)	12/18 (66.6%)
Gordana	1/58	1/5	N/A	2/12	3/17

	(1.7%)	(20%)		(16.6%)	(17.6%)
--	--------	-------	--	---------	---------

Though the data in some systems (e.g., in the case of Gordana’s finite verbs in the present tense, for example) are very scarce, (6c) clearly shows that the agreement error rates *vary* considerably across the children. While the error rate spans from 1.7% to 6.4% in bare past participles, it spans from 1.4% to 20% in finite verbs and from 16.6% to as high as 93.2% in COPBE and AUXBE constructions⁸. The overall error rate on all verbs denoting finiteness is as low as 17.6%, but skyrockets to 66.6% in the case of Vesna, for example. Now clearly not all looks that perfect!

In the computation above, I excluded the instances of IMPs as “finiteness” markers in order to be the most conservative possible. Though we will see later that the Legate and Yang’s (2005) system treats them as “tense-based” and hence *important* for the learning of tense morphology, I would like to be more cautious here, since it has been claimed not only in the syntactic literature (see Rupp 2003 and Rus 2005 for reviews), but also in the acquisition literature (see Bohnacker 1999 for a detailed review) that these are *nonfinite* constructions, lacking TP/AgrP (and AspP), on a par with infinitives⁹.

The plethora of the accounts of early root nonfinites and—particularly—the confusion about some of the most basic assumptions of what counts as a finite or nonfinite form makes it extremely difficult to make any definite predictions about the children’s knowledge of “finiteness”. If 3sg present tense forms in child Slovenian, which exhibits very similar verb paradigm characteristics as child Spanish, are indeed nonfinite, then what would one make of imperatives? If even imperatives are nonfinite, as suggested by Salustri and Hyams (2003) (but see a reply by Rus and Chandra (submitted), then what *does* count as finite? I believe that the term “finite” is extremely confusing and may well depend on language-specific properties than some more general linguistic markers (Wurmbrandt 2005; p.c.).

Suppose children really were so sensitive to the finite vs. non-finite distinction (or any other distinction, for that matter). What any theory of language learnability must account then is to show *why* children sometimes *fail* and *how* they fail. In other words,

⁸ Almost all of these errors are *omissions* rather than *commissions* (i.e., actual agreement errors), which is common across all early languages investigated so far (see Deen 2002; Guasti 2002).

⁹ Salustri and Hyams (2003, in press), for example, crucially build their entire argument on this assumption and without this premise their entire account, trying to capture the cross-linguistic differences in the nature and the rate of RIs, loses its merit completely.

developmental psycholinguistics needs to explain *where* certain errors come from and *why* they arise in early systems. A first general guess would be that children are simply faithful to the input, at least in the very early stages (e.g., before a two-word stage) before any kind of deep(er) generalizations can be made. Empirically then, one can resort to *statistical distributions* found in the input (i.e., child-directed speech) and compare them to the actual child data to test how much of a child's language really is simply based on *input-matching*?

2.3 IS LANGUAGE DEVELOPMENT *MERELY* INPUT-MATCHING?

I showed above why careful quantification of the data *is* important in acquisition studies and why *less deterministic* accounts may be desirable on at least a couple of grounds. First, as we saw above, careful quantificational analyses may reveal the information that otherwise gets skewed in more deterministic (“undifferentalist”) accounts that purport to show the “all or nothing” outcome, rather than some more probabilistic outcome with (a lot of) variability. Second, quantificational analyses may inform us about where the weaknesses and the strengths of a particular system lie and how learning (and hence, language *development*) actually *proceeds* in a *gradual* fashion. This, however, does *not* need to discredit the generativist Continuity Hypothesis, since continuity may refer to some basic, core (syntactic or computational, if you want) mechanism (following, e.g., Chomsky 2000), which may be continuous and operate *qualitatively* the *same* way in children as it does in adults. What I am proposing here is that what *does* change (and is hence *non-continuous*) is the learning that is associated with this core linguistic capacity¹⁰.

Several recent studies in learnability have shown mathematically, computationally, or via rigorous demonstration of the data that statistics is not all that is at stake in the process of language development (see Bertolo 2001; Marcus 1998; Yang 2002, 2004), but to test the emergentist claim about the child being extremely *faithful* to simple statistical distributions in the input data (at least in the earliest stages), I computed

¹⁰ I am fully aware of the fact that probabilistic accounts may be merely exercised at a *different level of analysis* when compared to deterministic (e.g., generative) accounts. However, I believe it is fair to point out that even for a deterministic account to be as explanatory as possible, it must be able show *how* the system gets perfect. In other words, positing innate knowledge of a certain mechanism on a merely conceptual level is *not* satisfactory. Any kind of learnability model must not only be *formally sufficient*, but also *developmentally-compatible*. We will all agree that no matter how much innate knowledge one wants to equip the children with, language is still acquired from the experience (see also Yang 2002 for similar ideas).

certain verb form proportions in my own child-directed data (from the corpus reported in the Rus and Chandra 2005 study, though never before analyzed) and compared them to the actual child data (from the same corpus). The table below sums up the frequency rates for each of the verb tokens for child-directed data, for the child data across all 15 children, as well as for the five children that I studied above:

(7) Frequency rates for each verb token in Slovenian child-directed and child speech

	PastParts	Finite Vs	IMPs	BE (COP/ AUX)	All finiteness markers excl IMPs
Child-directed (178 verb tokens)	15.7%	29.2%	6.7%	46%	75.2%
the five children studied above	21.44%	16.2%	40.22%	21.98%	38.16%
all 15 children (1205 verb tokens from Rus and Chandra 2005)	16.3%	15.5%	56.4%	11.8%	27.3%

The total number of verb utterances is as low as 178; hence, we cannot generalize much about this sample, though we can certainly observe some tendencies. Moreover, there may be something else going on with the *high discrepancy* in the case of imperatives. It seems that this is one of the most frequent verb forms found in early Slovenian, yet hardly attested in child-directed speech. This is exactly what Salustri and Hyams (2003, in press) found in early Italian and Spanish (and this is what their account would predict for child Slovenian, too, since it is geared towards null subject languages with “rich” tense/agreement morphology that show neither an RI stage nor a bare verb stem stage).

Based on some preliminary results from above and the results shown in Salustri and Hyams (2003), I agree that a child is *not* simply an input-matcher and that she diverges a great deal from child-directed speech. However, much more data are of course needed to make such a claim a generalization. Furthermore, I did not pursue any additional statistical analyses here, such as chi square or correlation tests, since any kind of analysis using such a small number of tokens would be pretty much useless here. Additionally, using a correlation test, for example, is not even possible. Though we have percentages of certain verb forms/utterances for each child of the 15 children, we *cannot*

reliably compute such proportions as for the input to each child, since obviously more than one child receives such input in any case when the kindergarten teacher and the examiner interact with the children. In other words, based on audio data alone, it is by no means clear which of the 15 children has heard a certain utterance and which hasn't. Ideally, we would need the figures showing input as well as output verb tokens for each child.

But still, approaches that rest merely on frequency and distribution would definitely have a hard time explaining the kind of *discrepancy* we see in the data above, and the question arises as to *who* and *what* exactly the kids are modeling in these cases. Moreover, probabilistic-only accounts tell us *very little* (if anything) about the nature of early linguistic forms (e.g., precise syntactic and semantic properties of early verbs and their constructions), and *much less* about the earliest syntactic and semantic generalizations that children make across these, since all they can capture is the raw tokens and frequencies of the data. Emergentist (constructivist) approaches may fare better, but it is far from obvious what predictions an acquisition model with a “rich human cognition” component would yield in such an analysis and how “cognition” would *empirically* account for the observed frequencies, correlations, and—particularly—discrepancies. In other words, such accounts do not offer much explanation of *which form* is correctly-used (e.g., or correctly-inflected, for example) in some contexts but not the others (rather than just making connection to the input), and—particularly—*why* this is the case. It seems that we need to push for some “golden middle” account that would seriously take formal sufficiency and continuity, along with probabilistic learning, dependent on the child's linguistic experience. This is what Legate and Yang (2005), modeled after Yang (2002), have tried to do in their *Variational Model*, which we turn to in the next section.

3. UG AND LEARNING IN ONE: THE VARIABILITY MODEL OF LEARNABILITY

3.1 LEARNABILITY AND “VARIATION”

Legate and Yang's (2005) approach to the RI phenomenon stems from the work of Yang (2002) and is the most recent model in a series of language learnability models, and the only existing generative approach so far that has—to my knowledge—*seriously* (and

formally) taken into account the child's exposure to the linguistic data (i.e., *the input*) as a variable in the acquisition process. The technical details about the model and the background from evolutionary biology that have given rise to the conceptual arguments on which the model rests are beyond the scope of this paper, but let me briefly summarize how the model came about and what its main claims are.

Yang (2002) argues that language—when taken to be a biological entity—fits into a “variational” model, observed at many levels of biological, developmental and ecological characteristics. I will refer to such a possible model (for the time being still a conceptual and tentative one) as a *variability* model, not to confuse “variability” in terms of the grammars’ outcomes with “language variation” as in predominantly sociolinguistic research, but suggesting some implementation of some “variational” or “variationist” approach. This “variational” thinking is in line with Darwinian’s evolutionary remarks, where in any scientific analysis non-uniformity is taken to be interpreted as a collection of *distinct* individuals. Yang (2002) argues that whereas transformational learning models identify the learner with a *single* hypothesis, which directly changes as the input is processed, his model is *variational*, suggesting a change in the *distribution* of *I-language* grammars (i.e., the options given and permissible by UG). The Variability Model of language acquisition (henceforth, VALA) hence crucially depends on grammar competition, where “grammar” refers to a possible human grammar and the number of such grammars is assumed to be finite, given by UG¹¹.

Yang’s model may be taken as a theory-free model, spiced up with the conceptual “generative spirit” in terms of continuity, parametric differences, and the role of UG. However, even the role of UG (when compared to the models that basically equate UG

¹¹ The “competition” metaphor is not new in the learnability literature at all (see Piatelli-Palmarini 1989; Lightfoot 1991), but Yang argues that it has *never* been formalized and pursued *systematically* as directly related to *quantitative* data in language development (Yang 2002: 53). What Yang means here is that the previous models were purely theoretical (building on the statistical and mathematical notions), rather than drawing conclusions from the actual child-directed and child speech corpora. The first ideas of grammar competition were voiced by phonologists (e.g., Jakobson 1941, cited in Yang 2002, interpreted “errors” in children’s phonological acquisition as possible phonological forms found in other non-target adult languages; current research on early phonetic discrimination (e.g., Kuhl et al. 1992) also hints at the variation model, in which the hypothesis space goes from “more” to “less” via competition). In syntactic learnability models, it has been proposed that children may entertain several grammatical options for a certain syntactic phenomenon, but generally a lot of these models come with theory-internal assumptions and mechanisms which have not yet been (satisfactorily) tested for empirically (e.g., Roeper’s 2000 claim that the children will prefer grammars with less complex structural representations akin to the Minimalist Program). Yang’s model comes closest to Clark’s (1992) Genetic Algorithm (GA), in which candidate grammars are evaluated against the input data with some independent measure of fitness is explicitly defined, though Clark’s GA posits that the learner first computes the degree of parsibility for all grammars over a large sample of sentences and then determine the differential reproduction that leads to the next generation of grammars. Yang believes that GA results in a model with *too much computational cost* and this is why he introduces the use of probabilities/weights to capture the cumulative effects of discriminating linguistic evidence (cf. Yang 2002: 55).

with a language acquisition device)—as I understand it—is somehow downplayed, since it not only interacts a great deal with the input, but its shape (“growth”, if you want) actually depends directly on the linguistic experience. UG in this model is crucially viewed as an *initial* acquisition state, i.e., what is initially given by biology and what will grow on the basis of the input. Hence, “Yang’s UG” does not come with the sort of “baggage” we were used to in the 1980s and 1990s with hundreds and hundreds of constraints on structure, but a system that simply *delimits* the space of biologically possible grammars, where linguistic variation is brought about along the line of a handful of some broader (“universal”) parametric dimensions. Though Yang *explicitly rejects* the older parameter-switching metaphor (e.g., Chomsky 1986; Hyams 1986), his model still crucially relies on the notion of parameters in the “classic sense”, where parameters come (ideally) in binary opposition¹². These are needed for the sake of *constraining* the learnability space. Though one may conceptually disagree with such a notion of “parameters”, one simply *must* posit certain constraints on learning, whatever their source is and however language-specific or human-specific they are. Not doing so would result in no possible theorizing about learnability at all. We know that there are constraints on any human learning and behavior – a problem that has sometimes been referred to as the “frame problem” in the artificial intelligence research, which has always tried to *constrain* a possible learning space when “feeding” pre-specified parameters into an artificial learning device (Gazzaniga 1992; see also Pinker 1997). Call your constraint “Parameter P”, “Cue C”, “Condition A”, “Bias Y”, “Preference 682”, “I-will-not-be-spelled-out-at-LF-feature”, or “what-should-we-call-it-constraint”, it will still be some information about what is *possible* and what is *not possible* in human grammars (and hence, what children do and what they do not do) – simply this and nothing more. In other words, when getting rid of conceptual baggage about parameters and their exact

¹² What exactly all these parameters are (i.e., which grammatical subsystem (or interface system) they refer to, either to phonology, morphophonology, morphosyntax, narrow syntax, semantics, etc.) and how many one needs to posit is still far from obvious, but according to some estimates (e.g., Baker 2001; Clark 1992), there must be between 30 and 40 (binary) parameters in the UG space. Yang’s (2002) study investigates four of them, namely the WH Parameter (with [+WH] vs. [-WH] values), the Verb Raising Parameter (with [+raising] vs. [-raising] values), the V2 Parameter (with [+V2] vs. [-V2] values), and the Null Subject Parameter (again, with a binary option). However, Yang is very explicit about his model not treating grammars as *wholes* (in which case the child would need to consider between 2^{30} and 2^{40} grammar weights, which seems to be (psychologically and computationally) extremely implausible), but rather as *partial systems*. Ideally, all parameters would be *independent* of each other, which would considerably cut down the possible parametric space. However, Yang admits that not all parameters are (probably) binary, hence “additional resources” must be postulated to solve the problem of parametric interference (Yang 2002: 41). His study involves independent parameters only, but discusses (and offers) an additional mathematical/computational approach to the interference problems (see Yang 2002: 39).

sources, Yang's learning model becomes pretty much *theory-neutral* as it is dependent on purely mathematical notions, taking into account statistical notions and distributions that explain the non-uniformity and the gradualness associated with language acquisition. His learnability process is schematized below in (8):

- (8) For an input sentence s , the child
 - a. with probability P_i selects a grammar G_i ,
 - b. analyzes s with G_i
 - c.
 - if successful, reward G_i by increasing P_i
 - otherwise punish G_i by decreasing P_i

The following are the key features and results of VALA, (78) in Yang (2002: 123):

- (9)
 - a. Language acquisition can be modeled as a selectionist process in which variant grammars compete to match linguistic evidence.
 - b. Under the condition of explanatory continuity, the irregularity in child language and the gradualness of language development can be attributed to a probabilistic combination of multiple grammars, rather than to imperfect exercise of a single grammar.
 - c. Formal sufficiency and development compatibility can be simultaneously met in the variational model, for which the course of acquisition is determined by the relative compatibilities of the grammars with input data; such compatibilities, expressed in penalty probabilities, are quantifiable and empirically testable.

3.2 RIs AND MORPHOLOGICAL LEARNING IN THE VARIABILITY MODEL OF GRAMMATICAL ACQUISITION

Legate and Yang (2005) (henceforth, L&Y) argue that the explanations of RIs which maturational (generative) acquisition accounts put forth are not (entirely) satisfactory since the mechanisms of biological maturation of linguistic ability are still not well-understood. They further argue that biological maturation abandons the Continuity Hypothesis—the hypothesis that children's competence system is not qualitatively different from the adults'—which, according to them, has served well in the investigation of child language and cognitive development in general. They also remind the reader that it is useful to bear in mind the *gradient* distribution of RIs not only *across* languages but also *within* a language. In other words, the extent of RI use across languages varies quantitatively along a broad spectrum, as seen below in the tables in (10a) and (10b):

(10a) Proportions of RIs across some Germanic child languages and French¹³

Child	Lang	Age	%RIs	Source
Eve	English	1;6-1;10	78	Deen 2002 after Sano and Hyams 1994
Sarah	English	1;8-2;3	90	Poeppel and Wexler 1993
Maarten	Flemish	1;11	51	Deen 2002 after Krämer 1993
Nathalie	French	1;7-2;1	76	Deen 2002 after Rasetti 2000
Simone	German	2;0-2;6	52	Deen 2002 after Beherens 1993
Embla	Swedish	1;8-1;10	61	Deen 2002 after Platzack 1990

(10b) Proportions of RIs across child Romance languages

Child	Language	Age	%RIs	Source
Claudia	Italian	1;4-2;4	3	Deen 2002 after Pizzuti and Caselli 1992
Diana	Italian	1;10-2;6	1.5	Guasti 1993/1994
Marti	Catalan/Spanish	1;9-2;5	0.56	Torrens 1995
Josep	Catalan/Spanish	1;9-2;6	3	Torrens 1995
Gisela	Catalan	1;10-2;6	1.2	Torrens 1995

As mentioned above, in child languages with relatively long RI stages, the percentage of RI use declines gradually. The following example is from Clahsen and Penke's (1992) case study using the Simone corpus where we can clearly see a decrease in RIs and an increase in finite verbs:

(10c) An example of a gradual decline of an RI stage

Simone (German)	Age	% RIs	% finite Vs
from	2;0	72%	28%
Clahsen	2;6	32%	68%
and Penke	3;0	16%	84%
(1992)	4;0	4%	96%

Moreover, even during the RI stage, the child does not use RIs *exclusively*, but also the correctly-inflected verb forms clothed with tense and/or agreement morphemes (even on the very *same* verb in the very same context (see e.g., Kazman 1991) – hence the “optionality” All these facts seem to indicate that the RI phenomenon *cannot* be taken as a categorical or universal deficiency in children's grammar. Nor do these facts support the classic view of grammatical development as being an on-or-off process of switching

¹³ Following L&Y, I am including early English here too, though some researchers (e.g., Hoekstra and Hyams 1998; see also Deen 2002) have argued that English should be analyzed separately from other Germanic languages, since it is the bare verb stem that surfaces in the RI-type of clauses rather than the infinitive (with the usual infinitival marker “to”).

parameter values (e.g., “triggering” in the sense of Gibson and Wexler 1994; see also Yang 2002 for an overview). Finally, it also seems very unlikely then that the phenomenon is entirely due to the processing limitations or the lack of pragmatic/discourse resources (e.g., the above-mentioned Avrutin’s 1999 study), since, for instance, there seems to be no independent reason to assume that Spanish children’s processing or pragmatic capacity is significantly less limited than French children’s capacity. Given the strong inverse correlation between the morphological *richness* of the language (however one wants to categorize “richness”; possibly in the sense of the complexity of verb conjugation paradigms; see Hyams 2005) and the productivity of RIs, it appears that morphological learning must play a crucial (and quantitative) role in the explanation of the RI phenomenon. L&Y thus propose an alternative approach to the RI phenomenon, one that is equipped with a concrete model of acquisition in which morphological learning across languages is connected to the underlying grammatical system of RIs (i.e., VALA, à la Yang 2002).

Concretely then, L&Y argue that the optionality in RIs in languages like English reflects the presence of a grammar such as Chinese, which does not manifest (overt) tense marking, and that the gradual elimination of such grammar is facilitated by the learning of the morphosyntactic system of the target language. Nothing is said about the actual morphosyntactic system from the point of view of a clausal structure, though. What L&Y have in mind is simply the presence of absence of overt tense/agreement morphology¹⁴. Under this approach, quantitative differences among morphosyntactic systems result in the cross-linguistic variability in the RI phenomenon. In such variability learnability framework, there are a number of features that distinguish it from the traditional conception of parameter setting. First, the rise of the target grammar is gradual, as its probability gradually approaches 1 (this is characteristic of language development in general). Second, the demise of non-target grammars is *also* gradual: before these are eliminated, they will be accessed (*probabilistically*), which leads to a principled

¹⁴ Hence, one needs to fill in the (conceptual) observations from the UG-based literature and Yang’s (2002) assumptions about the Continuity Hypothesis regarding the phrase structure. L&Y build their model on the basis of the “activeness” of the Tense feature, which—in their study—roughly corresponds to the presence of overt tense/agreement morphology and overt morphology dependent on Tense/Agreement (see below), generally captured under a Tense node in a clausal representation (e.g., that of Chomsky 1995), though conceptually, they would probably not want to claim that a clausal representation in a child’s systems lacks the Tense node. It is not clear how they would capture “covert” (or no) Tense features in a Chinese-style grammar, where—as they put it—Tense is not an active morphosyntactic feature. However, this issue hardly matters empirically in a (more) probabilistic system such as the one they propose.

interpretation of “errors” in child grammar as potential adult grammars made available by UG (in line with the guiding principle of the Continuity Hypothesis). Lastly, the speed with which a grammar (or a parametric value) rises to dominance is *correlated* with how incompatible its competitor is with the input, which can be determined by syntactic investigation of the cross-linguistic variations in the parameters framework. The frequency of such input evidence can be directly estimated from child-directed speech corpora, which in turn allows one to make quantitative assessments of longitudinal trends in the child’s syntactic development¹⁵.

To invoke the parametric approach in line with Yang (2002), L&Y distinguish between [+Tense] and [-Tense] languages, the former group referring to languages in which T(ense) is an *active morphosyntactic feature*, the latter group to languages where T is *not active* (i.e., where temporal relations are expressed in a different way, generally by the use of temporal adverbs or certain aspectual morphemes). Hence, in their model, a child acquiring a [+T] language must rule out a [-T] option; when the child fails to do this, a matrix clause will surface with an infinitive (or bare verb) instead of a tensed form. In probabilistic learning, the [-T] value can be accessed non-zero times before it is eliminated, where elimination is based entirely on the input. Put it simply, the child rules out the [-T] grammar on the basis of overt morphology¹⁶.

A word on [+T] morphology is in place here. [+T] in the L&Y system does *not* refer to the morphological markers expressing tense itself only, but also the markers that *depend* on tense, even though they may *not overtly express* it. For example, take 3sg present tense -s ending in English. It has been shown that children know the features that are morphologically associated with the realization of it before they can actually use it

¹⁵ One can already see one problem with this approach, though, namely how to reconcile this premise with the previously-mentioned finding about early imperatives in child Italian, Spanish, and Slovenian. I am not sure how VALA could handle early imperatives, where it has been shown that children are *not* faithful to the frequencies and distributions in the input at all (Rus and Chandra submitted, in press; Salustri and Hyams 2003, in press). The two existing lines of accounts of early imperatives differ in terms of both the structure and status of early imperatives. If imperatives truly turn out to be early root nonfinites, lacking TP/AgRP/AspP, as suggested by Salustri & Hyams (2003, in press), then VALA might be compatible with such an account (in a similar fashion to the account outlined here for RIs with the optionality between an active and a non-active Tense/Agreement feature), though note that in L&Y’s VALA account of RIs, it is assumed that imperatives are “full” clauses and hence imperative morphology as indication for the presence of Tense (“activeness” of Tense); see below for further details on this issue).

¹⁶ Though L&Y argue that the child needs to “rule out” the [-T] option on the basis of overt [T] morphology, I do *not* think that this implies that [-T] is the initial, *default* setting. In fact, as mentioned above, Yang (2002) rejects the idea of parameter switching and he seems to reject the idea of “default parameter setting”, too (contra Hyams 1986). I believe that the [-T] option is simply accessed probabilistically when the child encounters it in the input, which seems to be plentiful in an English type of grammar (see below for quantitative details on this). Hence, English-speaking children, for example, encounter a lot of [-T] tokens and may start with this grammatical setting, switch to [+T] in some cases and stick to [-T] in others (“competition” with parallel grammars), and finally, gradually reach the [+T] only option (though L&Y seem to totally gloss over the threshold values in this process; see below for this issue).

consistently (see Deen 2002 and Guasti 2002 for detailed reviews). It has been further shown that though English-speaking children omit *-s* in very high proportions (e.g., Schütze and Wexler 1996), they never misinterpret it when it is overt. In other words, children's errors are almost always entirely those of omission rather than those of commission (i.e., incorrect use of *-s* in non-3sg contexts), as mentioned above already. Hence, a token such as a 3sg pres verb in English will count as a [+T] token, hence rewarding a [+T] grammar and punishing a [-T] grammar. Furthermore, consider verb tokens with no overt realization of either tense or tense-dependent morphology (e.g., a verb form with a zero morpheme, if one postulates zero morphology in the grammar at all). These would be compatible with both, a [+T] as well as with a [-T] grammar. As L&Y argue, zero morphology need not be posited in the beginning, and if the child's acquisition device attempts to analyze it with a [-T] grammar, the analysis will be successful and the form will incorrectly punish the [+T] grammar. Later the [+T] grammar must win over the [-T] one, of course, on the basis of other, independent verb tokens. The point is that even if we posit null tense morphology, the verb forms that express neither tense directly nor are they tense-dependent will be liable to punish the [+T] grammar. All these *language-specific* facts have to be considered when basing the model on input token frequency.

Morphological evidence for tense morphology in Spanish is trickier and way more interesting than English. Tense itself as realization of an inflectional feature is only realized overtly in the past imperfective tense (in all other tenses, tense and agreement are fused into one single (portmanteau) morpheme). Compare the present and the past imperfective conjugation paradigms in Spanish below in (11) (not as represented and assumed by Y&L; see below):

(11)a. Present tense of *cantar*, "to sing"

1sg	cant o	-Ø
2sg	cant a	-s
3sg	cant a	-Ø
1pl	cant a	-mos
2pl	cant á	-is
3pl	cant a	-n

[stem: **canta**; with *-a* as the thematic vowel, followed by agreement morphemes; note the thematic vowel change in 1sg; the thematic vowel is generally assumed to be part of the stem in traditional as well as generative accounts (e.g., Montrul 2004); though Campos (p.c.) reminds me that in some accounts 1sg has

the underlying form of “canta-o”, preserving the thematic vowel *-a*, which gets deleted to avoid the adjacency of two vowels (*VV; $a \rightarrow \emptyset / _ V$), in which case *-o* counts as a tense/agreement marker; L&Y analyze it as “part of the stem” to be the most “conservative”, though they do note that it may be an agreement marker; see also Note 17 below]

(11b) Past Imperfect of *cantar*, “to sing”

1sg	cant	a	ba	-Ø
2sg	cant	a	ba	-s
3sg	cant	a	ba	-Ø
1pl	cant	a	ba	-mos
2pl	cant	á	ba	-is
3pl	cant	a	ba	-n

[stem: **canta**; with *-a* as the thematic vowel, followed by *-ba*, the tense morpheme, which, in turn, is followed by the same agreement morphemes as found in the present tense conjugation; see (10a) above]

L&Y argue that all agreement-based forms in the present tense paradigm with the exception of 3sg count as [+T] tokens, since the inflections that they exhibit do not appear on any other untensed (basically, “nonfinite”) form, i.e., infinitives, gerunds, and present and past participles (infinitive: *canta-r*; present participle/gerund: *canta-ndo*; past participle: *canta-do*)¹⁷. *Canta*, for example, is an ambiguous token since—from the adult (descriptive) perspective at least—it can be either an instance of a verb stem or a 3sg present tense form with a zero morpheme. In addition, Spanish *-ar* verbs form the imperatives with *-a* (or *-Ø*, again, depending on what you take as a stem and what as agreement). 3sg pres tokens as well as *-a* imperatives may thus potentially incorrectly punish a [+T] grammar, hence they are all included in the count of [-T] tokens in order to

¹⁷ Note that 1sg *-o* stem change is now taken to be marked [+T], since in the L&Y model, any stem vowel change will be tense-dependent, though this is *not* spelled out explicitly in their study for this particular process (1sg verb form) at all. Furthermore, I notice another big inconsistency here. Though it is true that the majority of Spanish stem-changing verbs do not exhibit a phonological stem change in nonfinite forms (e.g., *colgar*→*colgando*, *colgado* (1sgpres: *cuelgo*); *despertar*→*despertando*, *despertado* (1sgpres: *despierto*), a lot of verbs do, in fact (e.g., *volver*→*vuelto*, *pedir*→*pidiendo*, *pidido*). In their characterizing the Spanish [+T] tokens, L&Y seem to miss this fact entirely. These stem-changing forms would all be, of course, [-T] since they are present or past participles or gerunds, yet, they do exhibit stem-changing, just like in the present tense conjugation. Furthermore, as Campos (p.c.) reminds me, tensed (“finite”) verb forms in Spanish show a different syntactic behavior when compared to untensed (“nonfinite”) verb forms, particularly in terms of cl(itic) position, with cl+V for tensed verbs and V+cl for untensed verbs, for instance:

<u>L</u> a	canta.	vs.	Cántala
it.sgfem-cl sing.3sgpres			sing.imp-it.sgfem-cl
‘S/he sings it.’			‘Sing it!’ (imperative; considered untensed, similarly to Salustri and Hyams 2003 for Italian)

Similarly for participles:

Se	fue	cantandola.	(and same V+cl word order for the phenomenon with the bare verb stem <i>-a/-e</i>
Se-refl	left.3sg	singing-it.sgfem-cl	form, described above in Note 5; e.g., /.../ cantala que cantala)
‘S/he left singing it.’			

This piece of evidence suggests that not all that is at stake in the categorization of tense and agreement in tensed and untensed forms may lie in the verb’s morphological marking or the marking “dependent” on it.

yield the most conservative learnability scenario possible. The third option of [+T] morphology is when the verb undergoes a phonological stem change. Spanish has several verbs that behave this way:

(12a) Spanish verbs with a phonological stem change

- i. o → ue (e.g., colgar, dormir, llover, mover, volver)
- ii. e → ie (e.g., despertar, cerrar, empezar, pensar, tener, venir, querer)
- iii. e → i (e.g., decir, pedir, reñir, seguir)

(12b) Present tense of Spanish *poder*, “can/may/be able to” (o → ue stem change)

- 1sg pued o -Ø
- 2sg pued e -s
- 3sg pued e -Ø
- 1pl pod e -mos
- 2pl pod é -is
- 3pl pued e -n

Hence, in the L&Y system, agreement-based tokens not overlapping morphologically with nonfinite forms, as well as phonological verb stem changes, although not expressing tense (directly), require tense for their realization and thus count as [+T] tokens, rewarding a [+T] grammar¹⁸.

Crucially then, according to L&Y, the length of the RI stage *correlates* with the number of *ambiguous morphological* [+T] tokens in the input data. The more input that rewards the [+T] grammar there is for a child, the less inclined the child will be to use RIs. L&Y quantify this with an analysis using the inflectional paradigms of three languages, sampled from the RI spectrum mentioned above, namely Spanish, French, and English. In the Spanish case, RIs are hardly attested; they generally fall between 1% and 5%. In French, the percentage varies between 30% and 70%, while in English, this percentage is even higher, usually between 75% and 90%. The following table sums up their findings:

¹⁸ I personally prefer the terms *T-dependent* (or *T-based*) tokens and *T-dependent* (*T-based*) morphology, since I believe that L&Y’s terms such as [+T] tokens and [+T] morphology might be misleading, as we are definitely not talking about overt tense realization in all cases. Also, keep in mind from the previous note that in some cases it will be the *syntax* that will dictate the “finiteness” or “nonfiniteness” of a specific form in an utterance, e.g., in the case of clitic placement, so the issue falls out of the realms of “mere morphology” (or the morphological component of the grammar, if you want).

(13) Advantages of a [+T] grammar over a [-T] grammar for child Spanish, French, and English

Spanish 1194 child-directed utterances from 9 files from the Fernandez and Aguado corpus from CHILDES	[+T] Vs	877
	[-T] Vs	200
	77.2% (677/877 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 22.8% consistent w/ [-T]; [+T] grammar has a significant advantage of 54.4%	
French 1046 child-directed utterances from the Phillippe corpus from CHILDES	[+T] Vs	1049
	[-T] Vs	316
	69.9% (733/1049 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 30.1% consistent w/ [-T]; [+T] grammar has an advantage of 39.8%	
English 58447 child-directed utterances from the Brown corpus from CHILDES	[+T] Vs	27198
	[-T] Vs	12405
	54.4% (14793/27198 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 45.6% consistent w/ [-T]; [+T] grammar has an insignificant advantage of 8.8%	

The results above clearly show that while the Spanish input gives plenty of opportunity for children to “set” the grammar to the [+T] value, English-speaking kids are getting extremely mixed signals about “finiteness” from their environment. French kids fall somewhere in between, and this is exactly what the RI studies investigating the frequency in the child data have shown. Based on the morphological complexity and very robust one-to-one form-function mapping in the case of verb morphology, the existing formal acquisition models would predict a very low rate of RIs in child Slovenian, which has been confirmed in our own studies reported above. Y&L’s model would further predict that Slovenian-speaking children may have an even greater advantage of [+T] over [-T] when compared to Spanish-speaking children, since Slovenian is even more complex morphologically and shows even less RIs than Spanish. In other words, Slovenian input probably provides more tokens that would favor a [+T] grammar than Spanish, French, or English do. I investigate this claim empirically in the next section.

3.3 VALA PUT TO TEST: EVIDENCE FROM CHILD SLOVENIAN

As a short digression, let us briefly sketch the morphosyntactic properties of adult Slovenian which are relevant for the present discussion. Slovenian has one synthetic tense only, namely the present tense. Other tenses, i.e., the past tense, the future tense, and the “old-fashioned” pluperfect (not used in colloquial speech anymore) are all compositional/periphrastic, composed of the auxiliary verb *biti*, “to be”, and the active past participle¹⁹. The productive present tense suffixes carry both T and Agr features that cannot be teased apart morphologically (portmanteau morphemes), just like in the Spanish present tense outlined above. Verbs (in the present tense) are inflected according to the schema in (14) below, adapted from Toporišič (2000):

(14) [Root + Thematic Vowel]_{stem} + suffix (Tense/Person/Number)

Slovenian infinitives carry the *-ti* morphological ending, which can be reduced to *-t* in colloquial speech and always in supine forms (i.e., when the infinitive follows a verb of movement, expressing purpose). The morphosyntactic system comprises three persons (1, 2, 3), three genders (MASC, FEM, NEUT) and three numbers (SG, DU, PL), but since gender has no (separate) overt morphological instantiation in the present tense conjugation, each verb in the paradigm has nine cells. The table in (15) below shows a paradigm for the verb *igrati*, “to play”, in the present tense:

(15) Present tense verb paradigm for Slovenian *igrati*, “to play”

	SG	DU	PL
1	igra- m	igra- va	igra- mo
2	igra- š	igra- ta	igra- te
3	igra- ∅	igra- ta	igra- jo

[stem: **igra**; with *-a* as the thematic vowel, followed by the agreement morphemes; see the Spanish paradigm in (11) above]

According to the L&Y system, all present tense verb forms but 3sg would count as [+T] tokens (expressing overt agreement). 3sg would be *excluded* from the count since it neither realizes tense nor agreement morphology, but is rather simply a stem with a final thematic vowel (same as in the Spanish scenario we saw above). According to the Slovenian traditional descriptive grammar (e.g., Toporišič 2000), there are five classes of verbs regarding the stem's morphophonological changes, summarized in the table in (16)

¹⁹ The periphrastic forms fit into the [+T] vs. [-T] analysis with the auxiliary and copula *BE*, of course, which has always been assumed to be associated with the T node in a clausal structure, as outlined in Section 2 above.

below, with a representative example verb listed in each class and conjugated in all three persons in the sg (adapted from Toporišič 2000):

(16) 1sg, 2sg, and 3sg forms of Slovenian verbs in the present tense, classified by phonological stem changes

	-am <i>igrati</i> ("to play")	-im <i>narediti</i> ("to make/do")	-jem <i>piti</i> ("to drink")	-em <i>pasti</i> ("to fall")	-m <i>hoteti</i> ("to want")
1	igram	naredim	pijem	padem	hočeš
2	igraš	narediš	piješ	padeš	hočeš
3	igra	naredi	pije	pade	hoče

As one can see, the last three conjugation classes generally all exhibit not only thematic-vowel changes, but also consonant stem changes, not found in the infinitive and the participle. Furthermore, as noted above, all instances of *BE*, either as a COPBE or AUXBE, count as [+T] tokens. *BE* is a highly irregular verb in Slovenian, similarly to the scenario common in most European languages.

The imperatives are *much* trickier. Though L&Y assume that they are tense-dependent (they even explicitly say that they “arguably involve present tense”), Salustri and Hyams (2003, in press)—as mentioned above already—take them as non-tensed forms on a par with infinitives. The following table sums up the verbal forms found in our child-directed data, slightly more detailed than the representation in (7) above:

(17) Frequency rates for each verb token in Slovenian child-directed speech

	#(%) of all V tokens	#(%) of finite Vs (pres tense)	#(%) of IMPs	#(%) of COPBE	#(%) of Past Parts	#(%) of AUX BE	#(%) of INF
CHILD- DIR. SPEECH	178	52 (29.2%)	12 (6.7%)	54 (30.3%)	28 (15.7%)	28 (15.7)	4 (2.2%)

I did two counts, first with the imperatives (IMPs) and the second one without them. Even in the most conservative scenario with the past participles that show correct agreement counted as [-T] (and AUXBE as [+T] in participial constructions in periphrastic tenses) and without the IMPs (i.e., IMPs counted as [-T]), the advantage of a [+T] setting over a [-T] one is *much higher* when compared to the ratio reported for child

French and child English, and very similar to the one reported for child Spanish (18a). After including the IMPs in the count, however, the advantage of a [+T]-based grammar over a [-T] one is even higher than in the case of early Spanish (18b):

(18a) [+T] vs. [-T] grammar in child-directed Slovenian (excluding IMPs)

Slovenian 178 child-directed V tokens from the Kranjc (1999) corpus	[+T] Vs	134
	[-T] Vs	44
	75.3% (134/178 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 24.7% consistent w/ [-T]; [+T] grammar has a significant advantage of 50.6%	

(18b) [+T] vs. [-T] grammar in child-directed Slovenian (including IMPs)

Slovenian 178 child-directed V tokens from the Kranjc (1999) corpus	[+T] Vs	146
	[-T] Vs	32
	82% (146/178 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 18% consistent w/ [-T]; [+T] grammar has a significant advantage of 64%	

Hence, to sum up, after counting all the [+T] and [-T] tokens in child-directed speech, based strictly on the L&Y system (i.e., as in their count for early Spanish), we find that in Slovenian, 146 (or 82 %) of the 178 tokens punish the [-T] grammar and reward the [+T] grammar, while only 32 (or 18%) of the tokens show the opposite trend. As I hypothesized above, in Slovenian child-directed speech, a [+T] grammar has a *significant advantage* over a [-T] grammar, with almost 8 times higher advantage when compared to child-directed English. True, we only have 178 utterances of child-directed speech in our corpus and it would be a mistake to make any big generalizations, but we can definitely see a trend here. For the ease of exposition, compare these results to the findings reported above for child-directed Spanish, French and English:

(18c) Advantages of a [+T] grammar over a [-T] grammar for child Slovenian, Spanish, French, and English

Slovenian	82% (146/178 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 18% consistent w/ [-T]; [+T] grammar has a significant advantage of 64%
Spanish	77.2% (677/877 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 22.8% consistent w/ [-T]; [+T] grammar has a significant advantage of 54.4%
French	69.9% (733/1049 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 30.1% consistent w/ [-T]; [+T] grammar has an advantage of 39.8%
English	54.4% (14793/27198 Vs) of input [+T] Vs w/ unambiguous implication for the [+T] grammar; 45.6% consistent w/ [-T]; [+T] grammar has an insignificant advantage of 8.8%

These results suggest that Slovenian children should have *less* trouble in producing overt T/Agr morphology than Spanish-speaking children, and *much less* than French-speaking and English-speaking children. The data confirm that this is indeed true; Slovenian early grammar seems to exhibit neither RIs nor bare verb stem forms (Rus and Chandra 2005 report <1% of RIs; Rus in press reports 0.5% of bare verb stems and 4.8% of infinitives, none of which seem to be “real” RIs, but correct, target-like nonfinite complements to finite verbs, with the finite verb missing). However, VALA does not (straightforwardly) explain a high percentage of bare past participles in early Slovenian (Rus and Chandra 2005), which according to the variational model outlined here, would lack the “finiteness marker” (or [+T] setting), namely auxiliary *BE* (Rus and Chandra 2005)²⁰.

²⁰ I can think of two possibilities here. One possibility would be to treat past participles as [+T] tokens due to the morphophonological stem changes which are similar to the ones found in the present tense conjugation and, more convincingly, perhaps, due to *overt phi-feature* (gender, number) *agreement* (this way Slovenian participles would be analyzed on a par with Italian participles, for example, but different from English-type of participles with no overt gender and number morphology). As outlined above, what counts as a [+T] token is *language-specific* and it may turn out that Slovenian and Italian past participles count as [+T] whereas English and Spanish participles count as [-T]. The idea that past participles with productive (subject-participle) agreement morphology are more than just VPs (i.e., with no functional elements/projections attached to them) is *not far-fetched* and far from new in the syntactic literature. Rus & Chandra (2005) concluded that at least in child Slovenian they must be *T-based*, full-fledged CPs (entailing that they have TPs), based on the presence of overt *-wh* markers, reflexive and possessive clitics, as well as perfect phi-feature (gender, number, though [-person]) agreement, but with a missing *BE* under the T node (which, in fact, remains unexplained across several child languages in general). A second possibility would be to treat bare past participles as early root nonfinites, on a par with RIs (à la Varlokosta et al. 1998). If this turns out to be true, then it would be necessary to treat adult past participles as [-T] tokens (as VALA does for English, German and Spanish and the same way I did for Slovenian above), and see if there is a gradual development of auxiliary *BE* in participial constructions (in this case the presence of past participles would punish the [+T] grammar and maybe this is what confuses the Slovenian children in the use of [-T] in COPBE and AUXBE constructions. Note also that child-directed Slovenian is full of past participles found in periphrastic tenses (though, of course, these are (almost) never bare). Further research awaits the investigation of adult-directed bare and non-bare participial constructions in Slovenian.

4. CONCLUSION: IS VALA *THE RIGHT ANSWER* OR WHERE DO WE GO FROM HERE?

When language acquisition is put in a larger context, we see that Slovenian-speaking children are not merely input-matchers, but that the scenario is much more complex than this, with some forms reflecting the input, some others not. More research is definitely needed in order to compute the advantage or disadvantage of the (certain type of) input for an individual child, compared to the individual output on a much larger dataset. Furthermore, there might be some other, more deterministic mechanisms at stake that would account for the observed discrepancies that have not been proposed yet between the child-directed data and the actual child data. All in all, the language acquisition research needs to team up with quantitative research. Statistics cannot explain everything, but can yield some neat results when applied, pointing out a characteristic of the data that might be otherwise overlooked or hidden in more deterministic models. In regards to the variability model of learnability, I would say that it may indeed be on the right track, though as represented in L&Y, it is built on quite a few assumptions that have been scrutinized by some researchers. Let me spell them out one by one, starting with empirical inconsistencies first, moving to some more general theoretical foundations the model is built on, and finally, pointing at some conceptual observations about the model and concluding with some general steps a model like Yang's should take to account for all the observations that I have laid out in this paper.

First, L&Y's characterization of [+T] vs. [-T] seems too simplistic, ignoring many adult morphological and syntactic facts coming from the works of theoretical linguistics. For example, the structure and status of imperatives in both linguistic as well as acquisition theories is still a great matter of dispute. While L&Y unequivocally take them as tense-based forms, rewarding a [+T] grammar, Salustri and Hyams (2003, in press) argue that they are RIs on a par with infinitives in adult languages. Imperatives may not make a big difference in the quantitative analysis of child English, but they represent almost half (!) of the "verbal" data in our child Slovenian data, so we need to make sure our count is valid in the first place. 3sg present tense finite verbs pose a similar problem. In child Spanish, for example, they have been claimed to be "regular" finites (e.g., Montrul 2004; Torrens 1995), default finites, spelled-out when children do not know the right form (Aguado-Orea and Pine 2005), and even RI-analog early nonfinites (Davidiak

and Grinstead 2004), thus making it almost impossible to know whether to categorize them as [+T] or [-T] either. Furthermore, the very recent accounts on the status and use of (bare) past participles are not in agreement either. They have been analyzed as RI-analogous early root nonfinites (Varlokosta et al. 1998), RI-analogue bare perfectives (Hyams 2005), as well as full-fledged finite clauses in constructions with null auxiliaries (Rus and Chandra 2005). In addition, the role of aspect has been ignored in most studies on early RIs (for the exceptions to this, see Hoekstra and Hyams 1998 and Gavrusseva 2005), and we simply do not know how early aspect interacts with early tense development and how this would be quantified in Yang's current variability model.

I further also doubt that a verb's vowel stem changing might always imply tense-based morphosyntax and that a simple lack of morphological overlap between finite and nonfinite forms implies the presence or absence of tense. As I argued and showed above, L&Y's analysis entirely glosses over the instances of Spanish verbs which show stem changing in both the present tense verb conjugation as well as in the participles and gerunds. To complicate things even further, tense-dependency may fall out of the realms of morphology and may be deduced from the syntax (e.g., clitic placement with tensed vs. untensed verbs in Italian or Spanish)—something that Y&L's system doesn't capture at all either. In sum, we need some *more precise* and *explicit* measurements of what counts as tense-based and what not (and, independently properly, *why* not), since the status of imperatives, past participles, and 3sg present tense finite verbs have been so wildly different that it is nowadays extremely hard (if not even impossible) to conclude what counts as "finite" and what as "nonfinite" in the traditional sense, or tense-based ([+T]) and not tense-based ([-T]), if these are the right theoretical notions after all.

Third, the reason why children may sometimes seem to be so insensitive to tense and/or agreement may truly lie in the fact that their input is too confusing, not definite and unambiguous enough for early/earlier convergence to adult grammar, and such convergence may well rely on the individual child's performance on morphological learning, which might explain the well-documented differences observed among children. However, neither Yang (2002) nor L&Y (2005) shows how a learner might *learn* a distinction between [+T] and [-T] tokens from the (raw) data. They owe us an account of *how* a learner would learn and categorize them appropriately. In other words, L&Y's

model simply assumes that kids will “magically” learn the appropriate form-[±T] mappings involved in tense and agreement morphosyntax and start computing the probability weights right from the beginning. The process is viewed from an adult perspective taking into account the adult “insight” into what grammar is and what it does, yet we know that kids need to learn from scratch a lot of things before they can start categorizing verb tokens and building their overall grammar: they first need to learn that a certain word is a verb, they then need to parse the root and the affixes of this verb and/or notice/learn a vowel stem change (if necessary), then somehow organize the “tense-based” properties and generalize them across a larger sample, and only then can they start computing whether a verb is a subset of either a [+T] or a [-T] grammar.

Though Yang’s model captures both, “nature” and “nurture”, what is—as I argued in the introductory section—a desirable scenario in any language acquisition model, I would now also like to add my two cents to some (more) conceptual issues regarding the model.

Although I strongly agree with Yang in positing *certain* kinds of constraints on form and function, which may well turn out to be both domain-specific as well as species-specific properties, falling from biology, I am very *skeptical* about the classic view of parameters as being binary and as “clean” as Yang and L&Y present them. For example, in his (2002) study, Yang distinguishes between [+V raising] and [-V raising] languages. Though English and French may be “clean” representative samples of these two values (and even this has been disputed in English; see a review in Rus 2003), Slovenian verb raising, for example, is extremely *messy*. Not only do certain verbs raise and others do not, but certain verbs raise as high as to the Aspectual head(s), whereas some raise to Tense (and obligatorily to the Complementizer in interrogatives, possibly also in imperatives). The issue gets even more complicated when one realizes that some verbs raise only in some contexts, but never in others, where “contexts” are not only syntax- but also pragmatic-dependent (see Rus 2003 for a review of studies). Legate and Yang’s “new” notion of [+T] vs. [-T] might not fare any better, for that matter. In other words, we need some (stronger) motivation for positing or ascribing importance to [+T] and [-T] to begin with—again, to make sure that this is a *theoretically valid* distinction, and if so, where in the grammar such a distinction would be localized.

Hence, if this program is to be taken seriously—and I believe it should be—further research should at least (a) determine important statistical distributions of certain verb tokens via rigorous statistical tests or computationally performed on child-directed and child speech corpora and to show that children do not learn from simple distributions only; (b) determine via a rigorous linguistic analysis what exactly could count as [+T] and what as [-T], possibly on a sample of more languages, taking into account the insights of descriptive (typological) as well as generative linguistics; (c) show a computer simulation of [+T]/[-T] parameter learning to see how an artificial learning device would deal with this issue when presented with a huge number of verbs (to see what theoretical validity Yang's parameter has and show how a learner/system would learn the parametric distinction (if at all); (d) show what threshold each form or all [+T] forms collectively need to reach for learning to be successful when we stop seeing the switch between a Chinese-type and an English-type of grammar.

If one wants to go even more radically, one would be even suspicious about the agenda of positing (all) parameters as innate universal constraints on form as being UG-dependent in the first place. It may turn out that these are simply certain *learning biases* (Kirby 2005), the source of which I remain agnostic about here (i.e., they may fall from some more general, learning mechanisms or be more language-specific). I am *not* saying that the generative approach is not on the right track, but simply that a lot of assumptions and conceptual claims about UG and its characteristics (e.g., parameters) need more empirical support since only this way will the field be able to fight off those who posit no innate biases for language learning to begin with. Furthermore, we will all agree that language acquisition is not all about the setting of some 30 or 40 parameters of syntax. Hence, I wonder *how* we could incorporate UG-based parametric dimensions into the learning component when it comes to “less universal” (and more language-specific) properties of grammar. It is obvious that in the learning of verb morphology, for example, one can definitely see much more interaction between syntactic “parameters” and the actual morphological learning than in something like learning noun case systems, reflected in noun declensional classes. In other words, one of the most challenging issues for a variability model at this point is to find out what exactly universally-given “parameters” could be so that they could be both, general enough to apply to all

languages, and at the same time, explanatory enough to account for cross-linguistic differences. It may turn out that constraints are simply *semantically-based* instructions on form, possibly grounded in the human cognition (e.g., Tomasello 2006). For example, time anchoring in linguistics has been described in terms of the mapping between the functional heads computing tense, aspect and agreement, and the semantics of tense. Our Variability Model would then compute the association between form and time anchoring (e.g., a past tense marker for “pastness” and a present tense marker for “non-pastness”, to put it in a very simplified way).

And last but not least, I would like to see a variability model that would not only capture “broader linguistic variables” (e.g., null subject vs. non-null-subject), but could actually pin down *more specific* morphosyntactic properties of early forms on the basis of *specific* forms that the child has encountered. For example, I would like to know how 10^{12} instances of active past participles in the input to a child yields a morphosyntactic difference (if any) in the acquisition of participle raising (call it a participle raising or a (lower) verb raising operation/bias, or simply, participle placement) when compared to 10^4 instance of the same form in the input to another child. In other words, we should strive to make our learnability model even more faithful to input-sensitive learning by *still* leaving some room open for innate biases/parameters (in order not to be equated with those positing entirely unsupervised language learning, based solely on the statistical distribution in the raw data). I think this is the most exciting area to explore in the current variability models as well as to pin down precise biases or parameters that have possibly not been entertained yet in the syntactic and acquisition literature.

Whatever the answers to the questions and arguments posed throughout the paper—and—particularly in the last section—may turn out to be, quantitative analysis in child language acquisition *are* needed and should be incorporated into (more) deterministic models of language acquisition, since I believe that only by teaming up linguistic theories with corpus linguistics and computational modeling will we be able to explain the complex nature of language acquisition.

5. REFERENCES

- Aguado-Orea, C. and J. Pine. 2005. Testing Wexler's Unique Checking Constraint with data from early child Spanish. To appear in *BUCLD30: Proceedings of the 30th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Anderson, S. and D. Lightfoot. 2000. The Human Language Faculty as an Organ. *Annual Review of Physiology*, 62: 697-722.
- . 2002. *The Language Organ: Linguistics as Cognitive Physiology*. Cambridge: Cambridge University Press.
- Avrutin, S. 1999. *Development of the Syntax-Discourse Interface. Studies in Theoretical Psycholinguistics*. Dordrecht: Kluwer.
- Baker, M. 2001. *The Atoms of Language*. London: Basic Books.
- . 2003. Linguistic differences and language design. *Trends in Cognitive Sciences* 7: 349-453.
- Baum, W. 2005. *Understanding Behaviorism: Behavior, Culture and Evolution*. New York: Blackwell.
- Bertolo, S. 2001. A Brief overview of learnability. In *Language Acquisition and Learnability*, edited by S. Bertolo, 1-14. Cambridge, UK: Cambridge University Press.
- Bohnacker, U. 1999. Icelandic plus English: Language differentiation and functional categories in a successively bilingual child. Ph.D. dissertation, University of Durham.
- Bloom, P. 1990. Subjectless sentences in child language. *Linguistic Inquiry* 21: 491-504.
- Carroll, S. 2001. *Input and Evidence: The Raw Material of Second Language Acquisition*. Amsterdam: John Benjamins.
- Chomsky, N. 1986. *Knowledge of Language. Its Nature, Origin, and Use*. New York: Praeger.
- . 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- . 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Clark, R. 1992. The selection of syntactic knowledge. *Language Acquisition* 2: 83-149.

- Clahsen, H. and M. Penke. 1992. The acquisition of agreement morphology and its syntactic consequences. New evidence on German child language from the Simone corpus. In *The Acquisition of Verb Placement, Functional Categories and V2 Phenomena in Language Acquisition*, edited by J. Meisel. Dordrecht: Kluwer.
- Crain, S. 2003. *De-constructing constructivism*. Talk delivered at University of Maryland at College Park.
- Crain, S., A. Gualmini and P. Pietroski. 2005. Brass tacks in linguistic theory: Innate grammatical principles. In *The Innate Mind: Structure and Content*, edited by P. Carruthers, S. Laurence and S. Stich, 175-197. Oxford: Oxford University Press.
- Crain, S. and P. Pietroski. 2006. Is Generative Grammar deceptively simple or simply deceptive? *Lingua* 116: 64-68.
- Crain, S., T. Goro and R. Thornton. 2006. Language acquisition is language change. To appear in *Journal of Psycholinguistic Research* 35.
- Davidiak, E. and J. Grinstead. 2004. Root non-finite forms in child Spanish. Talk delivered at GALANA (Generative Approaches to Language Acquisition) 1 at University of Hawaii. Handout.
- Deen, Ud K. 2002. *The acquisition of Nairobi Swahili: the morphosyntax of inflectional prefixes and subjects*. Ph.D. dissertation, UCLA.
- Elman, J. 2005. Connectionist models of cognitive development: Where next? *Trends in Cognitive Science* 9: 111-117.
- Elman, J., E. Bates, M. Johnson, A. Karmiloff-Smith, D. Parisi and K. Plunkett. 1996. *Rethinking Innateness. A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.
- Fodor, J. 2000. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.
- Gavruseva, E. 2005. Semantic changes in root infinitives: A cross-linguistic perspective. To appear in *BUCLD30: Proceedings of the 30th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Gazzaniga, M. 1992. *Nature's Mind. The Biological Roots of Thinking, Emotions, Sexuality, Language, and Intelligence*. New York: Basic Books.
- Gibson, E. and K. Wexler. 1994. Triggers. *Linguistic Inquiry* 25: 407-454.
- Goldberg, A. [1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.](#)

- Gualmini, A. and S. Crain. 2005. The structure of children's linguistic knowledge. *Linguistic Inquiry* 36: 463-474.
- Guasti, M. T. 1993/1994. Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition* 3:1-40.
- Guasti, M. T. 2002. *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press.
- Hoekstra, T. and N. Hyams. 1998. Aspects of root infinitives. *Lingua* 106: 81-112.
- Hornstein, N. and D. Lightfoot, eds. *Verb Movement*. Cambridge, UK: Cambridge University Press.
- Hyams, N. 1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: D.Reidel.
- , 2002. Clausal structure in child Greek. *The Linguistic Review* 19: 225-269.
- , 2005. Child non-finite clauses and the mood-aspect connection: Evidence from child Greek. In *Aspectual Inquiries*, edited by R. Slabakova and P. Kempchinsky. New York: Springer.
- Jackendoff, R. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jenkins, L. 2000. *Biolinguistics*. Cambridge, UK: Cambridge University Press.
- Karmiloff-Smith *et al.* 1998. Linguistic dissociations in Williams syndrome: evaluating receptive syntax in on-line and off-line tasks. *Neuropsychologia* 36: 343-351.
- Kazman, R. 1991. Why Do Children Say 'Me do it'? In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 455-460. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kirby, S. 2005. *The adaptive mechanisms in the evolution of language*. Talk delivered at the Alice V. and David H. Morris International Symposium on Language and Communication. Symposium on the Evolution of Language. October 2005.
- Kranjc, S. 1999. Razvoj govora predšolskih otrok [Language development of pre-school children]. Ljubljana: Znanstveni Institut Filozofske Fakultete Ljubljana.
- Kuhl, P., K. Williams, F. Lacerda, K. Stevens and B. Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255: 606-608.

- Legate, J. and C. Yang. 2005. Morphosyntactic learning and the development of tense. Ms, Yale University.
- Lightfoot, D. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press/Bradford Books.
- MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk. Volume 1: [Transcription format and programs](#). Volume 2: [The Database](#)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Marcus, G. 1998. Can connectionism save constructivism? *Cognition* 66: 153-182.
- Miller, G. 2003. The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences* 7: 141-144.
- Montrul, S. 2004. *The Acquisition of Spanish. Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and Adult L2 Acquisition*. [Language Acquisition and Language Disorders](#) 37. Amsterdam: John Benjamins.
- Newmeyer, F. 2004. Typological evidence and Universal Grammar. In *What Counts as Evidence in Linguistics. The Case of Innateness*, edited by M. Penke and A. Rosenbach. Special issue of *Studies in Language* 28: 527-548.
- Phillips, C. 1996. Root infinitives are finite. In *BUCLD 20: Proceedings of the 20th Annual Boston Conference on Language Development*, edited by A. Stringfellow, D. Cahana-Amitay, E. Hughes and A. Zukowski, 588-599. Somerville, MA: Cascadilla Press.
- O'Grady, W. 1997. *Syntactic Development*. Chicago: Chicago University Press.
- Piatelli-Palmarini, M. 1989. Evolution, selection, and cognition. *Cognition* 31: 1-44.
- Pietroski, P. and S. Crain. 2002. Why Language Acquisition is a Snap. *Linguistic Review* 19: 163-183.
- . 2005. Innate Ideas. In *The Cambridge Companion to Chomsky*, edited by J. McGilvray. Cambridge: Cambridge University Press.
- Pinker, S. 1997. *How the Mind Works*. New York: Norton.
- . 2005. So How Does the Mind Work? *Mind and Language* 20: 1-24.
- Pinker, S. and P. Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13: 707-784.

- Platzack, C. 1992. Functional categories and early Swedish. In *The Acquisition of Verb Placement, Functional Categories and V2 Phenomena in Language Acquisition*, edited by J. Meisel, 63-82. Dordrecht: Kluwer.
- Poepfel, D. and K. Wexler. 1993. The full competence hypothesis of clause structure in early German. *Language* 69: 1-33.
- Radford, A. 1990. *Syntactic Theory and the Acquisition of the English Syntax*. Oxford: Blackwell.
- . 1995. *Phrase structure and functional categories*. In *The Handbook of Child Language*, edited by P. Fletcher and B. MacWhinney, 483-507. Cambridge, MA: Blackwell.
- . 2000. Children in search of perfection: Towards a minimalist model of acquisition. In *Essex Research Reports in Linguistics* 34. Essex, UK: Department of Language and Linguistics, Essex University.
- Rizzi, L. 1993/1994. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3: 371-393.
- . 2005. On the grammatical basis of language development: A case study. In *The Oxford Handbook of Comparative Syntax*, edited by G. Cinque and R. Kayne. Oxford: Oxford University Press.
- Roeper, T. 2000. Universal bilingualism. *Language and Cognition* 2: 169-186.
- Rupp, B. 2003. *The Syntax of Imperatives in English and Germanic: Word Order Variation in the Minimalist Framework*. Basingstoke: Palgrave MacMillan.
- Rus, D. 2003. Quirky verb movement in Slovenian. Ms, Georgetown University.
- . 2005. Slovenian embedded imperatives. In *Georgetown Working Papers in Theoretical Linguistics*, Vol 4, edited by C. Brandstetter and D. Rus, 153-183. Washington, DC: Department of Linguistics, Georgetown University.
- . In press. Early root nonfinites and the acquisition of finiteness in child grammar: Evidence from early child Slovenian. To appear in *Proceedings of Formal Approaches to Slavic Linguistics (FASL) XIV: the Princeton Meeting*. Ann Arbor: Michigan Slavica Publications.
- . Submitted. The Acquisition of finiteness in early child grammars: Some novel evidence from early child Slovenian. To appear in *Lecturas Sobre Psicolingüística*. Mexico City: Universidad Autónoma de México.

- Rus, D. and P. Chandra. 2005. Bare participles are *not* root infinitives: Evidence from early child Slovenian. In *BUCLD 29: Proceedings of the 29th annual Boston University Conference on Language Development 29*, edited by A. Brugos, M. Clark-Cotton and S. Ha, 493-503. Somerville, MA: Cascadilla Press.
- . In press. Questioning the “Imperatives as Root Infinitives” Hypothesis. To appear in *BUCLD30: Proceedings of the 30th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- . Submitted. Child Language Imperatives: Root Infinitive analogues? To appear in *Proceedings of Generative Approaches to Language Acquisition (GALA): the Siena Meeting*. Cambridge, UK: Cambridge Scholars Press.
- Saffran, J., R. Aslin and E. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274: 1926-1928.
- Salustri, M. and N. Hyams. 2003. Is there an analogue to the RI stage in the null subject languages? In *BUCLD27: Proceedings of the Boston University Conference on Language Development 27*, 692-703. Somerville, MA: Cascadilla Press.
- Salustri, M. and N. Hyams. In press. Looking for the universal core of the RI Stage. To appear in *The Acquisition of Syntax in Romance Languages*, edited by V. Torrens and L. Escobar. Amsterdam: John Benjamins.
- Sampson, G. 2005. *The Language Instinct Debate*. London/New York: Continuum International.
- Schütze, C. and K. Wexler. 1996. Subject case licensing and English root infinitives. In *BUCLD 20: Proceedings of the 20th Annual Boston Conference on Language Development*, edited by A. Stringfellow, D. Cahana-Amitay, E. Hughes and A. Zukowski, 670-681. Somerville, MA: Cascadilla Press.
- Tesan, G. and R. Thornton. 2004. Overspecification in early English. In *Proceedings of Generative Approaches to Language Acquisition (GALA): the Utrecht Meeting*, 469-480.
- Toporišič, J. 2000. *Slovenska slovnica* [Slovenian Grammar]. Maribor: Založba Obzorja.
- Tomasello, M. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- . 2006. Acquiring linguistic constructions. In *Handbook of Child Psychology*, edited by D. Kuhn and R. Siegler. New York: Wiley.
- Torrens, V. 1995. The Acquisition of the functional category inflection in Spanish and Catalan. In *MIT Working Papers in Linguistics* 26, edited by C. Schütze et al.

- Varlokosta, S., B. Rohrbacher and A. Vainikka. 1998. Functional projections, markedness, and 'Root Infinitives' in early child Greek. *Linguistic Review* 15: 187-207.
- Wexler, K. 1994. Optional infinitives, head movement and economy of derivation. In *Verb Movement*, edited by N. Hornstein and D. Lightfoot. Cambridge, MA: Cambridge University Press.
- Wexler, K. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106: 23-79.
- Wurmbrandt, S. 2005. *Infinitives are tenseless*. Talk delivered at Georgetown University. November 2005.
- Yang, C. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.
- , 2004. Universal Grammar, statistics or both? *Trends in Cognitive Sciences* 8: 451-456.