Andreas Trotzke and Josef Bayer (Eds.)
Syntactic Complexity across Interfaces

Interface Explorations

Editors
Artemis Alexiadou and T. Alan Hall

Volume 30

Syntactic Complexity across Interfaces

Edited by
Andreas Trotzke and Josef Bayer

DE GRUYTER
MOUTON

ISBN 978-1-61451-785-6
e-ISBN (PDF) 978-1-61451-790-0
e-ISBN (EPUB) 978-1-5015-0101-2
ISSN 1861-4167

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: PTP-Berlin, Protago TEX-Production GmbH
Printing and binding: CPI books GmbH, Leck
♾Printed on acid-free paper
Printed in Germany

www.degruyter.com

Contents

Andreas Trotzke and Josef Bayer

1 Syntactic complexity across interfaces

1.1 Introduction

Syntactic complexity has always been a matter of intense investigation in formal linguistics. Since complex syntax is clearly evidenced by sentential embedding and since embedding of one sentence in another is taken to signal recursivity of the grammar, the capacity of computing syntactic complexity is of central interest to the recent hypothesis that syntactic recursion is *the* defining property of natural language (Hauser et al. 2002). In the light of more recent claims according to which complex syntax is not a universal property of all living languages (Everett 2005), the issue of how to detect and define syntactic complexity has been revived with a combination of classical and new arguments (Nevins et al. 2009).

The existing collections on the nature of syntactic complexity either deal with syntactic complexity from a functional-typological perspective (Miestamo et al. 2008; Sampson et al. 2009) or place a premium on the property of syntactic recursion (van der Hulst 2010; Sauerland and Trotzke 2011; Roeper and Speas 2014). In contrast, the current volume makes a new contribution to the ongoing debate by taking into account the recent development in linguistic theory to approach UG 'from below' by referring to both grammar-internal and grammar-external interfaces when explaining design features of the human language faculty (Chomsky 2007). According to this shift in perspective, it is reasonable to assume that UG only contains properties such as recursive Merge, binary branching structure, and the valued-unvalued feature distinction. All other properties of grammar might follow from the interaction between UG and other components within the model of grammar (the phonological and the semantic component; i.e. grammar-internal components) and from the interplay between UG and grammar-external components such as the performance and acquisition systems. As for the interaction with grammar-internal components, the new division of labor among the components of the model of grammar raises new issues for defining and detecting syntactic complexity. In particular, the question of the complexity of grammar has to be answered separately for 'narrow syntax' and for the grammar as a whole, including the interface components (Trotzke and Zwart 2014). As for the interaction with grammar-external components, Trotzke et al. (2013) show that systematic properties of performance systems (the 'performance interface', according to their terminology) can play an important role within the research program outlined by Chomsky (2005, 2007). In particular, investigations of the performance interface can revise current conceptions of UG by relegating widely assumed grammatical constraints to properties of the performance systems, as recently argued, for instance, by Bever (2009) for the Extended Projection Principle or by Hawkins (2013) for the Final-Over-Final Constraint (Biberauer et al. 2014).

Given this conceptual background of approaching the issue of syntactic complexity from the perspective of recent linguistic theory, the volume starts with two contributions that deal with the formal complexity of natural languages in terms of the Chomsky hierarchy, the most prominent complexity measure in formal language theory. These two contributions set the scene for the volume by discussing general aspects of grammar architecture and by turning to the question of whether languages can vary as to their formal complexity. The two papers are followed by three contributions that address specific issues of clausal embedding (small clauses, parentheses, peripheral adverbial clauses, and right dislocation/afterthought constructions). The last part of the volume contains three papers that provide accounts of how to address topics revolving around syntactic complexity in terms of grammar-external interfaces in the domain of language acquisition.

1.2 Syntactic complexity and formal language theory

In contrast to the recent typological-functional literature, the comparative complexity of languages is not an issue in formal language theory. The question relevant in this context is where the grammar of natural language is to be placed in the 'Chomsky hierarchy', a complexity hierarchy of formal languages. In the 1950s, Noam Chomsky developed formal language theory as a mathematically precise model of language. Chomsky established that behaviorist accounts of language were insufficient to account for the computational properties of natural languages, whereas the phrase structure grammars Chomsky introduced stood a chance to be sufficient. In particular, Chomsky (1956, 1959) showed that the property of self-embedding involves the kind of complexity that requires (at least) context-free grammars, rather than less complex types of grammar (specifically, finite-state devices). Following the lead of Chomsky, theoretical linguists developed concrete phrase structure grammars for specific languages. Crucially, and as should be clear from the above, the discussion in formal language theory focuses on general computational properties of 'narrow syntax', a core component of the model of grammar that can be equated with the faculty of language in the narrow sense as defined in Hauser et al. (2002). In addition to this component that applies simple rules merging elements, the model of grammar includes interface components dealing with sound and meaning. Accordingly, the question of the complexity of the grammar has to be answered separately for the grammar as a whole and for the individual components (including narrow syntax); with different answers forthcoming in each case. In recent literature, it is an open question which phenomena are to be associated with which component of the grammar, with current proposals relocating seemingly narrow syntactic phenomena such as head movement and inflectional morphology to the interface with phonology (e.g. Chomsky 2001). By discussing notions of formal language theory, the following two contributions investigate which properties of the grammar should be relegated to the

interface components and which features of natural language should be considered as belonging to narrow syntax and, therefore, should be evaluated according to the Chomsky hierarchy.

In his contribution "Against complexity parameters," **Uli Sauerland** addresses the recent proposal that languages can vary concerning their formal complexity in terms of the Chomsky hierarchy. According to Sauerland, such accounts are essentially proposing that this variation is a parameter choice – the 'complexity parameter'. Sauerland argues that parameterizing languages in this regard is unwarranted and not supported by the evidence. Based on a discussion of languages such as Swiss German, Standard German, and English, Sauerland makes two claims. First, he argues that certain word order differences between these languages should not be addressed in terms of the Chomsky hierarchy. Instead, as Sauerland argues, these variations can be addressed by independently established word-order parameters, belonging to the domain of the phonological interface and not to narrow syntax. After relegating this issue to variation in the domain of linearization, Sauerland turns to a second argument against complexity parameters by referring to the semantics interface. He claims that the semantics of a non-context-free language would need to radically differ from that of a context-free language. Specifically, he argues that the semantics of language is inherently context-free, and, as a consequence, the standard semantics of scope requires at least a memory system that supports context-free grammars. Since Sauerland takes it for granted that the semantics of natural languages should not vary, he concludes that these properties of the semantics interface provide important evidence against complexity parameters.

Jan-Wouter Zwart also takes the Chomsky hierarchy as a starting point. In his paper, "Top-down derivation, recursion, and the model of grammar," he adopts the strong focus on the role of the interfaces from recent minimalist literature and argues that the issue of syntactic complexity of the grammar has to be answered separately for the grammar as a whole and for the individual components (including 'narrow syntax'). Given this theoretical background, he claims that linguistic recursion should be understood as the interface-related treatment of a complex string as a single item within another complex string. In particular, he demonstrates that this simplex/complex ambiguity is due to separate derivational sequences ('derivation layers'). He argues that the grammar creating those strings ('narrow syntax') may be of the minimal complexity of a finite-state grammar. Zwart claims that competing views suffer from the unmotivated assumption that the rules and principles of grammar are fed by a homogeneous set of symbols. In contrast, he proposes that the symbols in the alphabet/numeration may themselves be the output of separate derivations. Based on this clarification, he concludes that arguments against the finite-state character of generating phrase structure lose their force. As a consequence, the complexity of natural language should not be addressed, in the first place, in terms of the types of grammar rules, but in terms of interaction among derivation layers, crucially involving the interfaces.

1.3 Syntactic complexity and clausal embedding

The following three contributions address specific issues of clausal embedding: small clauses, parentheses, peripheral adverbial clauses, and right dislocation/afterthought. The three papers ask to what extent grammar-internal interface conditions and properties can help to detect and define syntactic complexity. Do interface properties concur with the syntactic complexity ascribed to the phenomena in question? Or do interface-related features of the data even exclude an analysis in terms of syntactic complexity?

Leah S. Bauke deals with the issue of small clauses, a prominent case for which syntactic complexity is notoriously difficult to define. Working with a minimalist perspective, she focuses on the question of how basic syntactic operations are determined by interface conditions. In her paper "What small clauses can tell us about complex sentence structure," she argues for a revised analysis of small clauses. In particular, she claims that agreement between the small clause constituents can be established directly upon Merger and need not be mediated by a functional head. However, within minimalist theory, cases of XP-XP Merger are considered problematic because they pose labeling ambiguities. As a consequence, the input to the operation Merge is suggested to be constrained with the effect that at least one element must be or must count as a lexical item. Bauke demonstrates that this constraint poses no problem for her analysis, in which small clauses are generated by direct Merger of the two constituents that make up the small clause. She adopts the approach that complex syntactic objects already merged in the course of the derivation can be shrunk to lexical items, and, based on this account, she proposes an analysis of so far unaccounted for extraction and subextraction patterns in Russian and English small clauses.

The contribution by **Werner Frey and Hubert Truckenbrodt** focuses on the syntax-phonology interface. In their paper "Syntactic and prosodic integration and disintegration in peripheral adverbial clauses and in right dislocation/afterthought," they analyze different clausal dependencies in German by bringing together their respective work on peripheral adverbial clauses and on right dislocation and afterthought constructions. Frey and Truckenbrodt analyze these phenomena within a single set of analytical assumptions that relate to the notions of 'integration' and 'root sentence'. In the first part of their paper, they demonstrate that peripheral adverbial clauses require high syntactic attachment. Put more technically, peripheral adverbial clauses are either in the specifier of their host clause or are adjoined to their host clause. The authors show that this converges with phonological evidence. Both prosody and information structure of peripheral adverbial clauses reflect their borderline status between integration and disintegration. In the second part, they show that right dislocated or afterthought constituents are 'added' to the clause in the sense that they do not occupy a thematic position in their clausal host. However, these constituents show c-command relations like the elements they resume ('connectedness effects'). Based on evidence from the prosody and information structure

of right dislocation and afterthought constructions, the authors show that a syntactic adjunction analysis, if it aims at generalizing across right dislocation and afterthought constructions, cannot represent the properties of disintegration in a principled way. As an alternative, they propose a deletion analysis, which captures both the property of disintegration and the connectedness effects.

Marlies Kluck starts her contribution with the observation that syntactic complexity that does not involve subordination, such as coordinate structures and parentheticals, is still poorly understood. In her paper "On representing anchored parentheses in syntax," she turns to the questions of how and where 'anchored parentheses' are represented in grammar. By 'anchored' parentheses, Kluck refers to parentheses that are not freely attached somewhere in their host, but are attached at the constituent-level to an 'anchor'. In this sense, nominal appositions, nominal appositive relative clauses, amalgams, and sluiced parentheticals belong to this category. Contra 'orphan' approaches, which put parentheticals outside the domain of syntax, Kluck argues on both conceptual and empirical grounds for anchored parentheticals as represented at the level of syntax. In particular, she provides two reasons for this claim: First, anchored parentheticals must be part of syntax under common assumptions about the model of grammar. Parentheticals are linearized in their hosts and interpreted relative to their hosts. Second, since anchored parentheses are related to a specific constituent in their host, namely the anchor, their integration should not take place at a post-syntactic level.

1.4 Syntactic complexity and the acquisition interface

While the contributions sketched above deal with the interaction between syntax and grammar-internal interfaces, the following three papers focus on grammar-external interfaces in the domain of language acquisition. To keep UG as slim and simple as possible, these interfaces have recently been analyzed as external third factor effects (Trotzke et al. 2013), and include, according to Chomsky (2005: 6), "(a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints [...] including principles of efficient computation." Following Bever (2009: 280), we use the term 'acquisition interface' to refer to these grammar-external conditions in the context of language acquisition (i.e. to specific constraints on learnability). Given this interface notion, the following three contributions ask to what extent grammar-external acquisition processes can contribute to the debate on how to detect and define syntactic complexity. The contributions deal in particular with (i) different acquisition processes operative in the development of syntactic subordination, (ii) the identification of different scales of syntactic complexity by means of acquisition devices such as semantic bootstrapping and (iii) the application of minimalist economic principles to the acquisition problem.

Tonjes Veenstra argues that the pidgin-creole cycle provides crucial insights into the development of subordination in natural language. In his paper "The development of subordination," he starts with the observation that interlanguage varieties and pidgins both lack subordinative structures, and that creoles, by contrast, do exhibit such structures. On the basis of a comparison between Saramaccan (an English/Portuguese-based creole spoken in Suriname) and Fòngbe (its major substrate language), Veenstra shows that the creole patterns cannot be accounted for by substrate influence alone. Concentrating on sentence-embedding predicates and their associated syntax, he argues that the mismatches between the creole and its substrate are due to processes of incipient (second) language learning. Given different acquisition processes operative in creole genesis, Veenstra claims that incipient learning, i.e. early second language acquisition, plays a substantial role. Specifically, he argues that incipient learning accounts both for variable selection patterns of clausal embedding and for the absence of morphologically marked verb forms in embedded contexts. On the other hand, non-incipient learning, i.e. more advanced second language acquisition, accounts for the appearance of an unspecified subordinator. Relexification can explain the specific properties that this subordinator exhibits.

Tom Roeper also focuses on the acquisition interface. In his paper "Avoid Phase: How interfaces provide critical triggers for *wh*-movement in acquisition," he discusses the claim that both small clauses and infinitives lack a CP. More specifically, he focuses on the fact that Germanic languages generally do not permit a *wh*-word in an indirect infinitival question, while English, as an exception, does. In the context of evidence from first language acquisition, Roeper argues that the child acquiring English can posit a zero scope-marker in the matrix CP and need not posit a new infinitival CP. In support of this view, Roeper presents experimental evidence according to which the child more often interprets the medial *wh*- as covertly moved to the matrix CP with an infinitive than with a tensed clause. In addition, he refers to the notion of 'acquisition efficiency' and postulates 'Avoid Phase' as a general economic principle that converges with assumptions in minimalist theory. Adopting the notion of periodic 'Transfer' of syntax to a semantic interface interpretation, he claims that periodic Transfer is psychologically costly, and, consequently, the child should limit the number of Transfers by limiting the set of phases (thus the term 'Avoid Phase'). In other words, the child will maintain a more economical representation by positing as few phases as possible. Given this background, Roeper argues that tensed clauses initiate an interface transfer, while the default representation of infinitives, lacking a CP, does not.

Like Roeper, **Misha Becker** deals with data from first language acquisition. In her paper "Learning structures with displaced arguments," she starts with the assumption that sentences in which an argument has been displaced with respect to the position associated with its thematic/semantic role are more complex than sentences without such displacement. By focusing on this issue, Becker looks at how children acquire two constructions that involve such a more complex alignment of thematic relations: 'raising-to-subject constructions' and '*tough*-constructions'. Becker claims that inan-

imate subjects provide a clue to the detection or postulation of a complex structure. Essentially, she extends the scope of semantic bootstrapping and argues that not only can an animate NP serve as a cue to canonical subjecthood, but an inanimate subject can serve as a cue that the subject is displaced, and therefore that the structure is more complex. Becker supports her claims with two types of data: (1) naturalistic input data (child-directed speech from the CHILDES corpus) and (2) controlled experimental input data in simulated learning tasks with both children and adults.

Acknowledgement: We thank the participants at the workshop on *Complex Sentences, Types of Embedding, and Recursivity* (05–06 March 2012, University of Konstanz). Andreas Trotzke acknowledges financial support from the DFG Excellence Initiative (University of Konstanz, project no. 65411). We are grateful to all the scholars who volunteered as reviewers of the contributions to this volume: Markus Bader, Nicole Dehé, Marcel den Dikken, Doreen Georgi, Jeff Good, Jeremy Hartman, Katharina Hartmann, Fabian Heck, Riny Huybregts, Göz Kaufmann, Vincenzo Moscati, Andreas Pankau, Martin Salzmann, and Eva Wittenberg. Last but not least, we thank Uwe Braun for editorial help.

References

Bever, Thomas G. 2009. Remarks on the individual basis for linguistic structures. In *Of minds and language: A dialogue with Noam Chomsky in the Basque country*, Massimo Piattelli-Palmarini, Juan Uriagereka and Pello Salaburu (eds.), 278–295. Oxford: Oxford University Press.

Biberauer, Theresa, Anders Holmberg and Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45: 169–225.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions of Information Theory IT-2* 3: 113–124.

Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2: 137–167.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, Michael J. Kenstowicz (ed.), 1–52. Cambridge, MA: MIT Press.

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

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, Uli Sauerland and Hans-Martin Gärtner (eds.), 1–29. Berlin/New York: Mouton de Gruyter.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology* 46: 621–634.

Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch. 2002. The faculty of Language: What is it, who has it, and how did it evolve? *Science* 298: 1569–1579.

Hawkins, John A. 2013. Disharmonic word orders from a processing-efficiency perspective. In *Theoretical approaches to disharmonic word orders*, Theresa Biberauer and Michelle Sheehan (eds.), 391–406. Oxford: Oxford University Press.

van der Hulst, Harry (ed.). 2010. *Recursion and human language*. Berlin/New York: Mouton de Gruyter.

Miestamo, Matti, Kaius Sinnemäki and Fred Karlsson (eds.). 2008. *Language complexity: Typology, contact, change*. Amsterdam/Philadelphia: Benjamins.

Nevins, Andrew, David Pesetsky and Cilene Rodrigues. 2009. Pirahã exceptionality: A reassessment. *Language* 85: 355–404.

Roeper, Tom and Margaret Speas (eds.). 2014. *Recursion: Complexity in cognition*. Dordrecht: Springer.

Sampson, Geoffrey, David Gil and Peter Trudgill (eds.). 2009. *Language complexity as an evolving variable*. Oxford: Oxford University Press.

Sauerland, Uli and Andreas Trotzke. 2011. Biolinguistic perspectives on recursion: Introduction to the special issue. *Biolinguistics* 5: 1–9.

Trotzke, Andreas, Markus Bader and Lyn Frazier. 2013. Third factors and the performance interface in language design. *Biolinguistics* 7: 1–34.

Trotzke, Andreas and Jan-Wouter Zwart. 2014. The complexity of narrow syntax: Minimalism, representational economy, and simplest Merge. In *Measuring grammatical complexity*, Frederick J. Newmeyer and Laurel B. Preston (eds.), 128–147. Oxford: Oxford University Press.

Uli Sauerland

2 Against complexity parameters

Recent work in linguistics proposes that languages can vary as to their formal complexity in terms of the Chomsky hierarchy. In this paper, I argue that this proposal for the parametrization of languages is unwarranted and not supported by the evidence. I address three types of languages: 1) languages where there is evidence for a complexity greater than context-free (Swiss German, Dutch), languages where there is no evidence that their complexity exceeds that of context-free grammars (Standard German, English) and languages where there is no evidence that their complexity exceeds that of regular grammars (possibly Pirahã (Everett 2005) and Teiwa (Klamer 2010), but also large fragments of English). However, all differences concern only weak generative capacity. I first argue that the differences between the three types of languages can be easily addressed by established word-order parameters rather than a complexity parameter. Secondly, I argue that, if one still wanted to maintain the claim that there are languages parametrized to be less complex than context-free, the semantics of such a non-context-free language would need to radically differ from that of a context-free language. Specifically, I argue that the semantics of embedding and other types of scope is inherently linked to non-finite notions of memory such as push-down automata. In this sense, the semantics of language is inherently context-free. This observation makes it easy to test for context-freeness.

2.1 Introduction

The study of the formal complexity of language has been an important topic within generative linguistics since its inception (see Trotzke 2008; Lobina Bona 2012; Fitch and Friederici 2012 for recent reviews). Initially, formal complexity considerations have been invoked to rule out theories of language because the theories in question predicted a specific level of formal complexity to be impossible for language (Chomsky 1957; Huybregts 1984; Shieber 1985). More recent work in this tradition compares directly humans and other species using artificial grammar learning. This work is also oriented at finding the level of complexity the human language faculty can attain (Fitch and Hauser 2004; Gentner et al. 2006; Abe and Watanabe 2011; Beckers et al. 2012). Research in neurolinguistics that compares neural activity in linguistic and artificial grammar learning tasks underscores the relation between these two types of work (Friederici et al. 2011).

A second line of argument using the term complexity has emerged mostly within historical lingustics (Deutscher 2005; Givón 2009; Heine and Kuteva 2007; Sampson

et al. 2009; Wray and Grace 2007), though also in some synchronic work (Everett 2005).[1] McWhorter (2001) proposes furthermore that creole languages lack complexity (but see DeGraff 2003; Bane 2008; Veenstra this volume). At least some of this work targets the notion of formal complexity characterized by the Chomsky hierarchy just as the work discussed in the previous paragraph (see also Trotzke and Zwart 2014 on the relation between the two lines of research). However, the work just mentioned claims that individual languages are parameterized for different levels of formal complexity. So, this discussion also assumes that all humans are capable of learning all human languages. But it proposes that specific languages vary as to which level of the complexity hierarchy they occupy and furthermore that this variation is not just coincidental, but a parameter choice – the complexity parameter.

The view that formal complexity is a parameter of language is quite different from the use of formal complexity only to rule out some models of insufficiently rich models of language like regular grammars. In the following, I use the term *Complexity Parameter View* for this proposal. I attribute this view to Deutscher (2005); Everett (2005); Givón (2009); Heine and Kuteva (2007); Wray and Grace (2007) as all seven authors claim that substantial, interesting variation between languages can be explained by appeal to a notion of syntactic complexity according to which self-embedding of sentences is more complex than concatenation of two independent sentences.[2] In this sense, the work differs from other work in syntax using only intuitive notions of linguistic complexity (Dahl 2004; Hawkins 2004), which are difficult to evaluate. Self-embedding, however, was shown by Chomsky (1959) to be the crucial property distinguishing context-free languages from regular languages.[3] Some of the works may be reinterpreted to view formal complexity not as a parameter, but instead claim that variation in formal complexity is an epiphenomenon of other parameter settings, for instance word-order settings – the view I advocate below. But, as far as I can see, this is not the view actually taken: At least one author (Everett 2005) explicitly claims a restriction that can be characterized directly by the notion of recursion. We argue in Sauerland and Trotzke (2011) that the discussion about recursion following Hauser et al. (2002) only makes sense if we understand the term *'recursion'* as self-embedding. Furthermore, the seven papers mentioned appeal to a progressive trajectory of increasing formal complexity in natural language syntax in the course of human develop-

1 In language acquisition, too, formal complexity has been invoked. Frank (1998) proposes that children's language acquisition can be insightfully modeled by different levels of the complexity hierarchy. Roeper and Snyder (2005); Roeper (2011) argue that recursion in compounds must be learned. These proposals are not addressed in this paper, since they primarily concern maturation of language, not the parametrization of mature language.

2 See also work by Jackendoff and Wittenberg (2014).

3 Recall that Chomsky (1959) defines self-embedding as a type of center-embedding on string basis, not in phrase structural terms.

ment.[4] As far as I can see, the idea of such a trajectory requires a commitment to a complexity parameter.

The relevant concept of formal complexity for this paper is that of the Chomsky hierarchy (Chomsky 1956, 1959).[5] In effect, only three classes of the hierarchy are relevant for the following: 1) regular grammars (also called finite state grammars, Chomsky and Miller 1958) represent the level of lowest complexity on the hierarchy. 2) Context-free grammars (also called phrase structure grammars) are at the level of intermediate complexity. Finally, 3) context-sensitive grammars represent the highest level of complexity relevant for the discussion. The level of context-sensitive grammars may be restricted to the mildly context-sensitive grammars that Joshi (1985) introduced (see Joshi 2003; Graf 2013). Concerning the lower limit of human ability on the complexity, Chomsky (1957) using English evidence showed that language could not be analyzed using only regular grammar formalisms, but required at least context-free grammars. Furthermore, Huybregts (1976, 1984) and Shieber (1985) used Dutch and Swiss German evidence to show that language goes beyond the context-free level of complexity. Finally, Joshi (1985) hypothesized an upper limit; namely, that mildly context-sensitive grammars represent the highest level of complexity that language can attain (see Michaelis and Kracht 1997; Bhatt and Joshi 2004 for discussion).

For identifying the three levels of formal complexity, it is useful to keep the three formal grammars illustrated in (1) in mind. The $(ab)^n$ grammar is the regular grammar used in much of the work in artificial grammar learning starting with Fitch and Hauser (2004). But for comparison with the grammars in (1b) and (1c), it is more conspicuos to use a grammar that produces sequences of pairs of two identical terminals. This process is also clearly finite state, as it requires only memory for one symbol. In contrast to this grammar, the mirror grammar in (1b) requires an unlimited push-down memory and is therefore at the context-free level of the Chomsky hierarchy. Finally, the copy grammar (1c) is beyond context-free, but mildly context-sensitive.[6]

(1) a. $(xx)^n$: *aa*, *aabb*, *aabbcc*, *aabbccdd*, …
b. mirror grammar: *aa*, *abba*, *abccba*, *abcddcba*, …
c. copy grammar: *aa*, *abab*, *abcabc*, *abcdabcd*, …

4 Evidence of a superficially similar trajectory exists in other domains: Hay and Bauer (2007) report a positive correlation of phonemic inventory of a language and the size of the population of speakers of the language. Since all languages in prehistoric times were spoken only by a small group, these results entail a positive correlation of historic development and phoneme inventory size, which can be regarded as a form of complexity though Hay and Bauer (2007) do not use this term. Furthermore, Maddieson (2005a,b, 2011) reports a positive correlation of consonant inventory and syllable complexity.

5 Zwart (this volume) assumes a slightly different conception of the distinction between phrasal and terminal symbols, but also derives at the finite state vs. context-free grammar distinction.

6 One reviewer points out that the notation I use here for the pair grammar, $(xx)^n$, is not standard. However, there is no standard comprehensive formula for this grammar as far as I know.

Consider now the variation in the expression of infinitival complement clauses in West Germanic: (1) and (2) show representative examples from three West Germanic languages: Swiss German, Standard German, and English (cf. Wurmbrand 2004; Schmid 2005). (We follow the frequent practice to present Swiss and Standard German examples in the word order that they occur in an embedded clause, abstracting away from the effect of verb-second in the matrix clause. This is indicated by the initial ellipsis dots in the following.)

(2) a. Swiss German (Shieber 1985: 334)

… *mer em Hans es huus hälfed aastriiche*
… we the.DAT Hans the house helped paint

b. Standard German

… *wir dem Hans das Haus anzustreichen halfen*
… we the.DAT Hans the house paint helped

c. English
We helped John to paint the house.

(3) a. Swiss German (Shieber 1985: 335)

… *mer d'chind em Hans es huus lönd hälfe*
… we the children.ACC the.DAT Hans the house let helped
aastiiche
paint

b. German

… *wir die Kinder dem Hans das Haus anstreichen*
… we the.ACC children the.DAT Hans the house paint
helfen ließen
help let

c. We let the children help John paint the house.

In both Swiss German and Standard German, there is a syntactic relationship of case assignment between the verb and its object as shown graphically in (4) for the examples in (2): specifically, *hälfe/helfen* ('help') assigns dative case to its nominal object, while *aastriche/anzustreichen* ('paint') does not and therefore the object receives accusative case. Case assignment therefore provides a purely syntactic argument that the grammars underlying the Swiss German and Standard German data must be essentially the copy grammar for Swiss German with its nested dependencies and something like the mirror grammar for the Standard German data with its nested dependencies. Though English has no Dative case marking, even if it did, a regular grammar of the $(xx)^n$ type would be sufficient to generate the grammatical English strings involving infinitival complementation as the representation in (4c) shows.

(4) a. we [the Hans]$_{DAT}$ [the house]$_{ACC}$ helped paint

b. we [the Hans]$_{DAT}$ [the house]$_{ACC}$ paint helped

c. we helped John paint the house

For adherents of the complexity parameter view, the natural explanation for the variation in (2) and (3) is to link it to the complexity parameter. In fact, it is often assumed that children start language acquisition with a default parameter setting which has to be the most restrictive setting. Then, children in the face of positive evidence for a different parameter setting in the target language change parameters to a less restrictive setting (Crain 1991). If this view is combined with the complexity parameter, the default setting of the complexity parameter would have to be the one restricted to regular grammars. Children acquiring Standard German would then at some point encounter examples such as those in (2b) and (3b) and reset their complexity parameter to allow context-free grammars. Swiss German speakers would also start out with the regular grammar restriction, and then possibly first move on to the context-free setting, but ultimately would realize that their target language shows evidence for construction beyond context-free. Of course, in English grammar some phenomena exist that also require at least context-free complexity as Chomsky (1957) showed (involving for example relative clauses, *either … or* , or *if … then*). However, the relevant types of examples seem rather rare. For the following argument, I assume (possibly counterfactually) that some English children never encounter such sentences. So, in a sense, we are actually considering a language that shares the syntax of infinitival complementation with English, but does not share the English syntax of relative clauses and other constructions that provide evidence for context-freeness in English. In the following, I will call this hypothetical language the Regular Fragment of English – i.e. a subset of English sentences such that it can be weakly generated by regular rules and includes infinitival complements.

In the two following sections, I address two potential applications of the complexity parameter view. In the first section, I consider the difference between languages where there is no evidence against a restriction to context-free grammars (e.g. Standard German) and those where there is such evidence (e.g. Swiss German). I first argue that this case has just as much legitimacy as the one of a regular restriction even though proponents of the complexity parameter view have generally not addressed this difference. However, I argue that what would be parametrized is the method of memory access within syntax, which seems highly implausible. In the second section, I consider the claim that there are natural languages parametrized to be restricted to regular grammars. I first consider what kind of view of syntactic memory this type of parametrization seems to be assume: Namely, one where different aspects of syntactic memory are recruited on demand by the acquisition process. Then, I argue that a lan-

guage faculty with such limited memory could not predict the right interpretations for the regular fragment of English. More generally, I then argue that the standard semantics of scope requires at least a memory system that supports context-free grammars. Truly regular languages should display a very different semantics. Since no convincing evidence for such a difference has been shown, the complexity parameter view has no empirical support.

2.2 Context-free vs. context-sensitive languages?

The complexity parameter view as far as I know has not paid attention to the difference between languages that require a (mildly) context-sensitive analysis and those for which a context-free analysis might be available at the level of weak generative capacity. If such a division was contemplated, the dividing line might fall between Dutch and Swiss German on one side and English and Standard German on the other for all that is presently known about the formal complexity of these languages.[7] While my goal is not to advocate the complexity parameter view for the context-free/context-sensitive division either, it is instructive to consider why it is generally agreed that in this case a complexity parameter is not the right way to capture the cross-linguistic variation within West Germanic. Since the conclusion in this case seems to be universally held up, we can then compare this case to the controversial case of the context-free vs. regular language division.

Before the counter-arguments to the context-free/context-sensitive parameter, consider one argument why this might be a strong case for the complexity parameter view: In this case, the languages involved are well-studied and therefore the evidence basis is solid. English is in all likelihood the best studied language and German is also well studied, so if nobody has found strong evidence that a context-free analysis of English or German is impossible, we can be quite sure that it is indeed possible. The amount of evidence available for English contrasts strongly with the languages hypothesized to be of regular complexity which are either extinct (e.g. Old Babylonian) or spoken in remote locations (e.g. Pirahã). For these languages, the amount of evidence available is much more limited and there are no native speakers active in the field of linguistics. So in these cases, there is a non-negligible likelihood that evidence against a regular analysis was overlooked by the existing work on these languages or may simply be missing from the limited historical records.

In the following sections, I discuss five potential arguments that militate against a direct parametrization of the context-free/context-sensitive distinction, of which I

7 With respect to English, there has been quite some discussion of the question of whether it requires a complexity greater than context-free. Other candidates of languages that require a context-sensitive analysis are Bambara and Mohawk (see Manaster-Ramer 1986 and references therein).

consider the last three to be convincing. First, I discuss the absence of a clear direction of historical progress in the context-free/context-sensitive division. Second, I address language acquisition including artificial grammar learning. Third, I consider plausible cognitive implementations of a context-free/context-sensitive parameter. The fourth and fifth argument both relate to grammatical evidence that underlies existing analysis of the languages involved. On the one hand, actual grammars of both types of languages both involve mechanisms like movement that go beyond a context-free analysis. On the other hand, existing proposals for parametrization contain parameters that account for the domain of variation that would be addressed by a context-free/context-sensitive complexity parameter.

Probably the most important reason the context-free/context-sensitive division has not been viewed from the complexity-parameter perspective is that there is no clear historical development. The four example languages, Dutch, English, Standard German, and Swiss German make this point very well. All four languages are closely related, and only diverged from each other quite recently in their history. In fact, Swiss German, Standard German and Dutch have been viewed as points of a dialect continuum that also includes many Low German dialects that are quite close to Dutch (though I don't know about their verb clusters) and Upper German varieties of Southern Germany and Switzerland. Furthermore, the level of cultural development across all four languages is not perceived to be substantially different at present time. However, the absence of historical development alone is not a strong reason to reject the complexity parameter view in this case. For example, cultural evolution may still require some time for the greater expressivity of context-sensitive languages to take effect. Or it may be that other technological developments compensate for the expressive handicap that context-free languages carry with them. So, while the lack of evidence for a historical development has sociological importance, it alone should not discourage the complexity parameter view.

Consider now evidence from language acquisition and artificial grammar learning. To my knowledge, there has been little effort to investigate whether this domain provides evidence for a context-free/context-sensitive parameter. Such evidence might come from a comparison of Standard German speakers' acquisition of Swiss German or Dutch with the reverse. If parameter resetting was to require greater effort, we would expect Swiss German and Dutch to be the more difficult to acquire for Standard German speakers than vice versa. As far as I know, no such difference in learning difficulty has been reported, so at least we can conclude that the comparative difficulty is not dramatically different. But we cannot rule out that there may be a small difference at this point – we simply lack the relevant evidence. Furthermore, a similar contrast would be expected with Artificial Grammar Learning: for English or Standard German speakers, the copy grammar ought to be harder to learn than for Dutch and Swiss German speakers. As far as I know, no results on such comparisons are out in print. Tecumseh Fitch (p.c.) has mentioned to me unpublished experiments on English, where speakers found the copy grammar not harder to learn than a center-

embedding context-free grammar. Beyond these unpublished data that would argue against a complexity parameter, though, there is just little evidence from acquisition at present bearing on the complexity parameter view.

For the first real argument against the complexity parameter view, consider how the context-free/context-sensitive parameter would be implemented. From the production perspective, it seems natural to simply propose that only languages with the context-sensitive parameter setting are permitted to have context-sensitive rules. However, if we took this view, we might as well not have the complexity parameter, but have for each potential context-sensitive rule a parameter as to whether that rule is part of the grammar or not. From a parsing perspective, though, there may be a more natural implementation of the complexity parameter. Specifically, the Chomsky hierarchy has a fairly direct interpretation in terms of a memory structure: finite state automata correspond to regular grammars, push-down automata correspond to context-free grammars, and context-sensitive grammars correspond to linear-bounded automata. Therefore, one might assume that the context-free parametrization corresponds to a mental state where the syntactic parser can only access memory in the manner of a push-down stack. The context-sensitive setting correspondingly would be captured by a freer access to memory within syntax. But, such a view strikes me as implausible, since the method of access to memory would need to be part of the cognitive implementation of memory, and therefore it is plausible that also English and German speakers should be able to access memory freely if other humans are capable of a non-stack like access to memory in syntax.

A second real argument against the complexity parameter view comes from the consideration of actual proposals for the grammar of the languages involved. As for example Chomsky (1965) discusses in connection with the distinction between weak and strong generative capacity, the evidence for a specific syntactic structure comes mostly from other sources than the set of grammatical strings. Though the syntactic analysis of English is highly debated, there is general agreement that purely context-free analyses have been on the wrong track, even though they may be technically feasible. All current analyses of English include mechanisms that go beyond those of context-free grammars. One such mechanism is the movement transformation or its equivalents in other grammatical formalisms. In sum, all current work on English grammar argues against the complexity parameter view.

The third real argument against the complexity parameter view is that existing parameters make a complexity parameter superfluous. Consider for example the difference between Standard German and Swiss German with infinitival verb clusters illustrated above. For example, Schmid (2005) proposes an optimality theoretic analysis of word-order variation in such clusters. According to this analysis, actually all speakers generate word orders on both sides of the complexity divide as candidates for an optimality theoretic competition and then select one such analysis as determined by other parameters. In general, syntacticians working on infinitival clusters and word-order

have argued that the extent of typological variation in this domain is better accounted for by a set of word-order parameters than by a complexity parameter.

In sum, a context-free/context-sensitive parameter to account for the differences between English/Standard German on the one hand and Dutch/Swiss German on the other is indeed implausible. In this way, I concur with a belief generally held among typologists. For this reason, I assume that the reasons I presented in this section are tacitly held by many in the field, since, as far as I know, there are no advocates of the context-free/context-sensitive parameter. Now it is interesting to compare this case with the one of context-free vs. regular languages.

2.3 Context-free vs. regular natural languages?

In this section, I consider the proposal that languages are parametrized for the context-free vs. regular grammar distinction – the other major distinction of the Chomsky hierarchy relevant to our concerns. I argue first that such a parametric claim would only make sense at the level of cognitive mechanisms for similar reasons as those discussed in the previous section. But, from that it follows that not only purely syntactic evidence is relevant, but also semantic evidence for non-regular structures.

Throughout I will consider the regular fragment of English within the context of this argument. Recall from above that I assume that this fragment contains only sequences of nominal phrases and verbs such as those in (5), but does not contain the well-known types of structures that establish solely on the basis of weak generative capacity that English requires a context-free grammar.

(5) a. We let the children help John paint the house.
b. We helped John to paint the house.
c. John painted the house.

To generate the strings, assume a recursive, right-regular grammar[8] consisting of rules obeying the schema in (6a), where NP can be any of the terminals *we*, *John*, *the house*, and *the children*, and V can be any of the terminals *let*, *help*, and *paint*.[9]

(6) a. S → NP V S
b. S → NP V NP

While the facts about the language Pirahã are hard to ascertain presently (Everett 2005; Nevins et al. 2009), one possible description of the available data (and indeed a

8 As one of the reviewers points out, the term *right-linear* is also used instead of right-regular.

9 For expository reasons, I abstract away from the difference between inflected and uninflected verbs in the following.

plausible one) is that complement clauses in Pirahã have essentially the same grammar as the one in (6). Such a language – essentially English without relative clauses and some complex coordinations such as *either … or …* – is quite plausible on the basis of what we know about existing parametric variation. Specifically, the language Teiwa, for which Klamer (2010) provides a detailed description, seems to be a language of this type according to her.

Now consider how a language like the regular fragment of English would be captured within an approach assuming a complexity parameter. Recall that regular grammars can be characterized in at least two ways: One characterization is as set of strings that can be generated using only right- or only left-recursive phrase structure rules. The other is as strings that can be parsed by finite state automata. If the complexity parameter was to apply at the level of phrase structure rules, the parameter would govern the linear order of the possible recursion within such rules. However, it would need to apply across all rules of the grammar, while typically word-order parameters can vary from one phrase structure rule to another – for example, German and Japanese are both verb-final, but in German the noun precedes its arguments *der Brief an Maria* ('the letter to Mary'), while it follows them in Japanese *Mary-e-no tegami*. If each phrase structure rule has its own word-order parameters, the complexity parameter would be redundant at this level.

A more promising interpretation is to assume that the complexity parameter is implemented as part of the parser. According to this interpretation, the initial state of language would only allow parsing using finitely many memory states like a finite state automaton. But after being triggered by experience in the form of center-embedded strings, the language faculty would make available memory structures that at least allow in principle the processing of context-free grammar structures. Note that this interpretation of the complexity parameter is at odds with the assumption that human memory is limited in general. But, possibly for the present purposes it is sufficient to assume that the two parametrizations link to memory resources of a qualitatively different nature: one kind of memory that is self-contained and thereby finite, while the other can in principle recruit other memory resources without limit. The fact that these other resources are actually limited is not part of the system, but rather a performance limit imposed from the outside. This kind of difference is familiar from other domains: For example, it is well known that individuals' general working memory is limited (cf. Miller and Chomsky 1963; King and Just 1991). But nevertheless models of memory don't built in a hard limit. Indeed, if there was a hard limit of *n*-items to working memory, one would expect that subjects were aware of this boundary and that performance on tasks like a digit span would show a sharp drop-off exactly between remembering 7 vs. 8 digits. As things stand, however, the limit of short-term memory does not seem to be wired in like a computer's memory limit, but instead the structure of memory is in principle unlimited, but the accuracy of recall drops gradually around