Semilinearity as a Syntactic Invariant*

Jens Michaelis[1] and Marcus Kracht[2]

[1] Universität Potsdam, Institut für Linguistik/Innovationskolleg, Postfach 60 15 53, D – 14415 Potsdam
michael@ling.uni-potsdam.de
[2] Freie Universität Berlin, II. Mathematisches Institut, Arnimallee 3, D – 14195 Berlin
kracht@math.fu-berlin.de

Abstract. Mildly context sensitive grammar formalisms such as multi-component TAGs and linear context free rewrite systems have been introduced to capture the full complexity of natural languages. We show that, in a formal sense, Old Georgian can be taken to provide an example of a non-semilinear language. This implies that none of the aforementioned grammar formalisms is strong enough to generate this language.

Introduction

What we have in mind when we use the term *syntactic invariant* is, roughly speaking, a property, valid within some (formal) grammar theory, which remains "robust under slight modifications" of this theory. In the following we direct our particular attention to one such property: *Semilinearity (of a language)*.

Introducing the definition of semilinearity, Parikh proved that any context free language (CFL) is semilinear (see e.g. [10]). It has been shown that there is a need to go beyond the class of all CFLs, if we want to define a formal language in terms of phrase structure grammar or some related formalism to capture the *complexity of natural language* (see e.g. [6], [15]). To cope with this problem, *mild context sensitivity* is intended to be one appropriate, but rather informally defined grammar type, determining a proper subclass of context sensitive grammars. Originally, a mildly context sensitive grammar (MCSG) was defined by three necessary properties (see e.g. [7]). *Constant growth* of the language produced by such a grammar is one of those, and a somewhat strengthened version of this property of a language is that of being *semilinear*. In fact, this latter property is common to all grammars within one of the "classical" formalism types where each of these types constitutes a subclass of all MCSGs. In particular if the grammar belongs to the class of all tree adjoining grammars (TAGs) or the

* This paper has been worked out at the University of Potsdam within the *Innovationskolleg 'Formale Modelle kognitiver Komplexität' (INK 12/A 1)* funded by the Deutsche Forschungsgemeinschaft (DFG). We have benefitted largely from discussions with Annius Groenink. Many thanks also to Winfried Boeder for his help concerning Old Georgian, and to two anonymous referees for fixing up several of the "very non-English" formulations within the preceding version of this paper.

class of all head grammars (HGs) as well as to their generalized extensions, to the class of all multicomponent TAGs (MCTAGs) or the class of all linear context free rewrite systems (LCFRSs), respectively. The class of TAGs is weakly equivalent to the one of HGs,[3] and the same holds for the class of MCTAGs and LCFRSs ([17]). See also [8] for a survey on MCSGs. Our special interest in the property of being semilinear is motivated by the question:

(Q) Is it reasonable to expect a grammar formalism to generate a semilinear language, if the formalism is intended to capture human language capacity?

We are going to argue that to answer (Q) is not a trivial matter. Not least, since due to the syntactic analysis of Boeder ([1]), the phenomenon of *Suffixaufnahme* of genitive suffixes in *Old Georgian* can be taken to provide a possible counterexample. But, before we consider this case (see Sect. 2), we want to give a formal definition of *semilinear sets* and *semilinear languages*, respectively. Then we state a (technical) proposition which will be used to show that, in a formal sense, Old Georgian is not semilinear. For these purposes we first mention some conventions applying to our notation. All these preliminaries are done in Sect. 1 of this paper, while a proof of the proposition is given in the Appendix. In Sect. 3 we consider two other languages, where each of them already has been introduced in the literature as an example of a language not derivable by any MCTAG. For each of these two languages we briefly check for the possibility of extending the corresponding result to the more general one that the language is non-semilinear at all. Section 4 is reserved for a discussion of our result presented in Sect. 2. In particular, our example is compared to the two examples mentioned in Sect. 3. Furthermore, we will show in outline how the technical tools we used in the Old Georgian case can be interpreted with regard to a general proof method within the framework of formal language theory. Some final remarks are given in Sect. 5.

1 Notations, Definitions, and a Proposition

We denote the set of real numbers and natural numbers (non negative integers) by $\mathbb{R}$ and $\mathbb{N}$, respectively. $\mathbb{N}_+$ is taken to be the set of all natural numbers $n > 0$. For any $n \in \mathbb{N}_+$ and any non-empty set M the set M^n is the set of all finite sequences of length n, or all n-tuples, $u = (u_0, \ldots, u_{n-1})$ in M, where $u_i \in M$ is the $(i+1)$–th component of u. In the case $M \subseteq \mathbb{R}$ and if $u_i = 1$ for some i while all other components of u are 0 we also write $e^{(i)}$ instead of u. $e^{(i)}$ is called the $(i+1)$–th unit tuple (unit vector). This is due to the fact that we will use $\mathbb{R}^n$ also as an abbrevation for the common n-dimensional (vector) space $(\mathbb{R}^n, +_{\mathbb{R}^n}, \cdot_{S_{\mathbb{R}}}, 0_{\mathbb{R}^n})$ over $\mathbb{R}$ defined in a canonical way. Here, $+_{\mathbb{R}^n}$ and $\cdot_{S_{\mathbb{R}}}$ denote the common addition and outer product (scalar product) on $\mathbb{R}^n$ defined

[3] More generally, these two classes fall into a broader range of weakly equivalent grammar types ([16]), where each of these has been proposed to capture natural language capacity in a formal way.

componentwise by means of the common addition $+_{\mathbb{R}}$ and multiplication $\cdot_{\mathbb{R}}$ on $\mathbb{R}$, respectively. Recall that $\cdot_{S_{\mathbb{R}}}$ is a function mapping $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$.[4] $0_{\mathbb{R}^n}$ denotes the neutral element with respect to $+_{\mathbb{R}^n}$ (the null-vector). Then, $\mathbb{N}^n$ can be considered as the substructure $(\mathbb{N}^n, +_{\mathbb{N}^n}, \cdot_{S_{\mathbb{N}}}, 0_{\mathbb{N}^n})$ of $\mathbb{R}^n$, where $+_{\mathbb{N}^n}$ and $\cdot_{S_{\mathbb{N}}}$ are the restrictions of $+_{\mathbb{R}^n}$ and $\cdot_{S_{\mathbb{R}}}$ to the domains $\mathbb{N}^n \times \mathbb{N}^n$ and $\mathbb{N} \times \mathbb{N}^n$, respectively, both taking values in $\mathbb{N}^n$. $0_{\mathbb{N}^n}$, the neutral element with respect to $+_{\mathbb{N}^n}$, thus is identical with $0_{\mathbb{R}^n}$. Furthermore, we sometimes think of $\mathbb{N}^n$ just as the "simple" monoid $(\mathbb{N}^n, +_{\mathbb{N}^n}, 0_{\mathbb{N}^n})$.

For any finite non-empty set Σ (a set of terminals), Σ^* is not only taken to denote the set of all finite sequences (of all strings) in Σ including the empty string ϵ, but also the monoid $(\Sigma^*, \cdot_{\Sigma^*}, \epsilon)$, where $\cdot_{\Sigma^*}$ is the usual concatenation operation of strings in Σ^*.

We will drop the subscripts of the just introduced operation symbols in all cases when it does not lead to misunderstandings, and furthermore usually drop the remaining dot "$\cdot$" after doing so. If, for some non-empty set M, we simply refer to a set M^n we will (tacitly) assume n to be in $\mathbb{N}_+$.

Definition 1. Let $M \subseteq \mathbb{N}^n$. Then

(a) we call M **linear**, if for some $k \in \mathbb{N}$ there are some $u^{(0)}, u^{(1)}, \ldots, u^{(k)} \in \mathbb{N}^n$, such that $M = \{u^{(0)} + \sum_{i=1}^{k} n_i\, u^{(i)} \mid n_i \in \mathbb{N} \text{ for } 1 \leq i \leq k\}$.
(b) we call M **semilinear**, if for some $k \in \mathbb{N}$ there are some linear $M_1, \ldots, M_k \subseteq \mathbb{N}^n$, such that $M = \bigcup_{i=1}^{k} M_i$.

It can be shown ([4]),

Lemma 2. *Let $M, N \subseteq \mathbb{N}^n$ be semilinear. Then, $M \cap N \subseteq \mathbb{N}^n$ is semilinear.*

Definition 3. For some $n \in \mathbb{N}_+$ let $\{w_i \mid 0 \leq i < n\}$ be an enumeration of Σ, a finite set of terminal symbols with cardinality n. The **Parikh mapping** $p_\Sigma : \Sigma^* \to \mathbb{N}^n$ is then defined inductively by means of $\cdot_{\Sigma^*}$ and $+_{\mathbb{N}^n}$ in the following way:

$$\begin{aligned} \epsilon &\mapsto 0_{\mathbb{N}^n} && \\ w_i &\mapsto e^{(i)} && \text{for } 0 \leq i < n \\ \alpha\beta &\mapsto p_\Sigma(\alpha) + p_\Sigma(\beta) && \text{for all } \alpha, \beta \in \Sigma^* \end{aligned}$$

The image $p_\Sigma[L] := \{p_\Sigma(\alpha) \mid \alpha \in L\} \subseteq \mathbb{N}^n$ of a language $L \subseteq \Sigma^*$ under p_Σ is called the **Parikh image** of L. If $p_\Sigma[L]$ is semilinear, L is called a **semilinear language**.

Due to its inductive definition, the aforementioned Parikh mapping p_Σ is a surjective homomorphism which maps the monoid $(\Sigma^*, \cdot_{\Sigma^*}, \epsilon)$ onto the monoid $(\mathbb{N}^n, +_{\mathbb{N}^n}, 0_{\mathbb{N}^n})$; and p_Σ is 1–1 (injective) iff Σ consists of exactly one element.

[4] For any two non-empty sets M_1 and M_2, $M_1 \times M_2$ denotes the set of all pairs, built up by the elements of M_1 and M_2.

For each string $\alpha \in \Sigma^*$ the $(i+1)$–th component of the Parikh image $p_\Sigma(\alpha)$ just provides the result of counting the number of appearances of the terminal symbol $w_i \in \Sigma$ within α. Thus, whether or not a language L is semilinear depends only on the strings belonging to L and not on the (derivational) structures (corresponding to a possibly underlying grammar) represented by those strings. Moreover, the order of the terminals that appear within those strings is disregarded.[5]

Definition 4. For any $m \in \mathbb{N}$ and any $\alpha = (a_0, \dots, a_m) \in \mathbb{R}^{m+1}$ with $a_m \neq 0$, we take P_α to denote the function $f : \mathbb{R} \to \mathbb{R}$ defined by $x \mapsto \sum_{i=0}^{m} a_i x^i$ for every $x \in \mathbb{R}$. That is, P_α is the **(real) polynomial (of degree m) corresponding to $\boldsymbol{\alpha}$.**

We are now able to state the proposition which we actually are interested in.

Proposition 5. *For some natural numbers $m, n \geq 2$ let $\alpha = (a_0, \dots, a_m)$ be a finite sequence in $\mathbb{R}^{m+1}$, where $a_m > 0$, and let a given $M \subseteq \mathbb{N}^n$ have the following properties:*

(i) For any $k \in \mathbb{N}_+$ there are some numbers $l_2^{(k)}, \dots, l_{n-1}^{(k)} \in \mathbb{N}$ for which the n-tuple $(k, P_\alpha(k), l_2^{(k)}, \dots, l_{n-1}^{(k)})$ belongs to M.

(ii) For any $k \in \mathbb{N}_+$ the value $P_\alpha(k)$ provides an upper (lower) bound for the second component l_1 of any n-tuple $(k, l_1, \dots, l_{n-1}) \in M$ (that is $l_1 \leq P_\alpha(k)$ ($l_1 \geq P_\alpha(k)$) for any such n-tuple).

Then M is not semilinear.

A proof of this more general proposition is given in the Appendix. Here, with regard to our further considerations, we just want to state a corollary, which emphasizes a special case of the last proposition; namely the case $m = 2$, $\alpha = (1/2, -1/2, 0)$, and P_α an upper bound in the sense of condition (ii). This corollary also completes the part of the rather technical preliminaries.

Corollary 6. *Let M be a subset of $\mathbb{N}^n$, where $n \geq 2$, which has the properties:*

(i) For any $k \in \mathbb{N}_+$ there are some numbers $l_2^{(k)}, \dots, l_{n-1}^{(k)} \in \mathbb{N}$ for which the n-tuple $\big(k, (k^2 - k)/2, l_2^{(k)}, \dots, l_{n-1}^{(k)}\big)$ belongs to M.

(ii) For any $k \in \mathbb{N}_+$ the value $P(k) := (k^2 - k)/2$ provides an upper bound for the second component l_1 of any n-tuple $(k, l_1, \dots, l_{n-1}) \in M$ (that means $l_1 \leq P(k)$ for any such n-tuple).

Then M is not semilinear.

[5] Notice that a mapping $p_\Sigma : \Sigma^* \to \mathbb{N}^n$ according to Definition 3 for some given set of terminals Σ depends on the enumeration chosen for Σ. But, independently of such a choice, the corresponding mapping is unique up to an isomorphism on $\mathbb{N}^n$ induced by a permutation π on $\{0, \dots, n-1\}$. In this sense it is reasonable to speak of *the* Parikh mapping p_Σ with respect to a given set of terminals Σ.

2 Old Georgian

In this section, first we briefly introduce the phenomenon, shown by Old Georgian, to which our further argumention is related. Then, we are going to determine a sublanguage (fragment) L_M of Old Georgian of which the Parikh image $p_\Sigma[L_M]$ is a set $M \subseteq \mathbb{N}^n$ according to Corollary 6 above. This fragment is not only non-semilinear itself, but also prevents the entire Old Georgian language from being semilinear.

Old Georgian is one of those languages which show the phenomenon called *Suffixaufnahme* (literal: *taking up of suffixes*[6]). In particular, the Old Georgian grammar allows for multiple case(-number)-marking of nouns by adding "extra" case suffixes to the "inner" case suffix, where the (possibly empty) inner case suffix is obligatory and signals case in the "usual" sense, that is to say, the case immediately assigned to the smallest NP to which a noun belongs (which is the only marked case, for example, in the Indo-European language family). Each additional case-marker is the result of some indirect case assignment, and in this sense an explicit "reference" to the syntactical function of a noun(phrase) as a part of some more comprehensive constituent. Taking the case-suffixes of a noun in their "left-to-right-order" they can be described as the reflex of a left recursive, increasing structural embedding of NPs. Such a description is possible at least with regard to an underlying basic structure as it is assumed in [1]. More concretely, in Old Georgian, complex NPs, including stacked genitive NPs, cannot only be of the following basic form.[7]

Davit-is	*galob-isa*	*muql-ta*	*ama-t*	(1)
[[[[David-Gen]$_{\mathrm{NP}_4}$	singing-Gen]$_{\mathrm{NP}_3}$	verse-Pl(Gen)	Art-Pl(Gen)]$_{\mathrm{NP}_2}$	

c̣artkuma-j
recitation-Nom]$_{\mathrm{NP}_1}$

'the recitation of the verses of the song of David'

The "sub–NPs" can also appear in reversed order. But then corresponding to the stacking of the NPs, a stacking of case suffixes appears as well.

tkuenda	*micemul*	*ars*	*cnob-ad*	*saidumlo-j*	*igi*	*sasupevel-isa*	(2)
to=you	given	is	knowing-Adv	mystery-Nom	Art=Nom	kingdom-Gen	

m-is	*γmrt-isa-jsa-j*
Art-Gen	god-Gen-Gen-Nom

'Unto You it is given to know the mystery of the kingdom of God'

---

[6] But, see [11] for the (non-)possibility of an appropriate translation of the term *Suffixaufnahme*.

[7] Actually, the structure of (1) is only nearly basic, since Boeder assumes nouns to be phrase-final within the basic structure, while in (1) the article *ama-t* has shifted to the right of its noun *muql-ta* by undergoing clitic movement.

qovel-i igi sisxl-i saxl-isa-j m-is Saul-is-isa-j (3)
all-Nom Art=Nom blood-Nom house-Gen-Nom Art-Gen Saul-Gen-Gen-Nom

'all the blood of the house of Saul'

According to Boeder ([1]), the examples (2) and (3) are a result of ordinary agreement, instantiated by an optional application of some recursive (transformational) rules to the underlying basic structure as given with the example (1). If we apply Boeder's analysis to the general case, especially the analysis of constructions like (3), then (abstracting from determiners and other lexical categories) complex nominative NPs consisting of k stacked NPs, $k \in \mathbb{N}_+$, may have the form

$$\mathrm{N}_1\text{-}\mathrm{Nom}\ \ \mathrm{N}_2\text{-}\mathrm{Gen}\text{-}\mathrm{Nom}\ \ \mathrm{N}_3\text{-}\mathrm{Gen}^2\text{-}\mathrm{Nom}\ \ \ldots\ \ \mathrm{N}_k\text{-}\ \mathrm{Gen}^{k-1}\text{-}\mathrm{Nom} \tag{4}$$

Here, for $1 \leq i \leq k$, N_i denotes some noun(stem). That is to say, for any such NP, $\sum_{i=0}^{k-1} i = k^2/2 - k/2$ is the number of genitive suffixes appearing in it. Moreover,

the number of all genitive suffixes of all nouns within a complex NP consisting of k stacked NPs, where $k \in \mathbb{N}_+$, is bounded by $k^2/2 - k/2$. (5)

This is due to the restrictions on the applicability of the rules, proposed by Boeder. These restrictions come up with the given basic structure

$$[\,[\,[\ldots[\,\mathrm{N}_k\text{-}\mathrm{Gen}\,]_{\mathrm{NP}_k}\ \ldots\ \mathrm{N}_3\text{-}\mathrm{Gen}\,]_{\mathrm{NP}_3}\ \ \mathrm{N}_2\text{-}\mathrm{Gen}\,]_{\mathrm{NP}_2}\ \ \mathrm{N}_1\text{-}\mathrm{Nom}\,]_{\mathrm{NP}_1} \tag{6}$$

Roughly speaking, an indirect case assignment to N_i-Gen is only possible "upwards" with respect to the stacked NPs. That is to say, such an assignment to N_i-Gen is only possible with respect to the NPs in which NP_i is properly embedded, and possible only one time for each NP_j with $1 \leq j < i$. Thus, the maximal number of suffixes that can be taken up by N_i-Gen is $i - 1$.

Take the set Σ to be the Old Georgian lexicon, and $L_G \subseteq \Sigma^*$ to be the language of Old Georgian. Before we proceed, we just want to emphasize one observation which can be made so far. If we want to deal with a finite lexicon (a set of terminal symbols or atomic elements), (4) demands that at least some kind of "global" genitive suffix Gen has to be a lexical entry (terminal symbol) on its own. Notice, due to (4) there are uncountable many (pairwise distinct) possibilities of multiple-case marking of a noun: there is no upper bound on the number of genitive suffixes that "can be added" because of the fact that the number $k \in \mathbb{N}_+$ of instances of some nouns within a complex NP is not (finitely) limited.

As announced at the begining of this section, we now determine a non-semilinear fragment $L_M \subseteq L_G$. For any $k \in \mathbb{N}_+$ this fragment includes at least one sentence in which a complex NP corresponding to (4) can be found. But, in addition, we try to keep such a sentence "as simple as possible" with respect to its structural complexity. The idea is to fix an intransitive verb and to define the members of L_M in such a way that each member is a grammatical sentence which consists of exactly one instance of the fixed verb and an NP, which is therefore necessarily the subject. For some $n \in \mathbb{N}$ let us now take $\{w_i \mid 0 \leq i < n\}$ to be an enumeration of Σ. For simplicity, not effecting our considerations in general, we assume

$$\begin{aligned}&w_0 \text{ to be some fixed noun(stem)} ,\\&w_1 \text{ the genitive suffix } \mathsf{Gen} ,\\&w_2 \text{ a nominative suffix} ,\\&w_3 \text{ a genitive article} ,\\&w_4 \text{ a nominative article} ,\\&w_5 \text{ a fixed intransitiv verb} .\end{aligned} \tag{7}$$

Now consider the linear, and hence semilinear, set

$$R \;=\; \Big\{ e^{(4)} + e^{(5)} + \sum_{i=0}^{3} n_i e^{(i)} \mid n_i \in \mathbb{N} \Big\} \;\subseteq\; \mathbb{N}^n \tag{8}$$

Then the full pre-image of R with respect to the Parikh mapping p_Σ is the language

$$L_R \;:=\; p_\Sigma^{-1}[R] \;=\; \{\alpha \in \Sigma^* \mid \text{there is an } u \in R \text{ with } p_\Sigma(\alpha) = u\} \;\subseteq\; \Sigma^* \tag{9}$$

The language L_R includes all strings of Σ^* which consists of exactly one appearance of w_4 and w_5 and an arbitrary number of appearances of $w_0, \ldots, w_3$. We now define the language L_M as the set of all strings belonging not only to L_R but also to L_G, that is

$$L_M \;:=\; L_G \cap L_R \;\subseteq\; \Sigma^* \tag{10}$$

Then, the Parikh image $p_\Sigma[L_M]$ of L_M clearly is a subset of the intersection of the Parikh images of L_G and L_R. But $p_\Sigma[L_M]$ even is identical with this intersection, since L_R is defined as the full pre-image of R with respect to p_Σ; that is

$$M \;:=\; p_\Sigma[L_M] \;=\; p_\Sigma[L_G] \cap p_\Sigma[L_R] \;=\; p_\Sigma[L_G] \cap R \tag{11}$$

Recall that L_M only includes strings of $L_R \subseteq \Sigma^*$. But all strings within L_M are grammatical in the sense that they belong to L_G as well. Since we can assume that the intransitive verb needs to be combined with one nominative subject-NP to produce a well-formed sentence, it is reasonable to conclude that each string of L_M consists of one (possibly complex) nominative subject-NP besides the single instantiation of w_5, the fixed intransitive verb. Moreover, since an NP cannot assign nominative case to its complement but only the oblique case genitive, at least the underlying basic structure of this subject-NP can only be of the form (6). That is to say, the subject-NP itself can only consist of k stacked NPs, $k \in \mathbb{N}_+$, where each of these stacked NPs has genitive as its immediate case, except for the "highest" one which must have nominative case. But then, because of (4) and (5) the set M can not be semilinear according to Corollary 6. Now, we recall that the set R is in particular semilinear, and that the intersection of two semilinear sets is semilinear itself (Lemma 2). Then, due to (11), we may finally conclude that $p_\Sigma[L_G]$ can not be semilinear, and thus Old Georgian is a non-semilinear language.

3 Semilinearity and MCTAGs

Although a "weak" property on its own,[8] semilinearity seems to be a "strong" property with respect to natural language formalisms, as we have just seen. Remember that (as is shown in [17]) semilinearity is a necessary property of e.g. any multicomponent TAL (MCTAL), that is, any language generated by an MCTAG. Due to other examples, it has already been doubted whether TAGs and/or even MCTAGs, and thus LCFRSs, are appropriate tools for natural language. E.g. Rambow ([14]) states the inadequacy of MCTAGs to capture the phenomenon of Scrambling as it appears in German. Nevertheless, also his proposal for a revised grammar formalism only gives rise to grammars that produce a semilinear language.

Manaster-Ramer ([9]) argues that Dutch is no tree adjoining language (TAL) referring to a certain subset of coordination phrases which is no TAL. He informally outlines that, on the basis of a generalized extension of this subset, the same kind of argumentation is even possible if the term *TAL* is replaced by *MCTAL*. A formal elaboration of this is given by Groenink ([5]). Radzinski ([13]) proves that the language of all Chinese number-names is no MCTAL. Like Groenink, Radzinski refers to a certain sublanguage which is a non-MCTAL, and he also uses a pumping lemma to show this. In both cases, Dutch and Chinese number-names, the corresponding sublanguage can be shown to be non-semilinear in general by means of Proposition 5, or at least a similar one in the Dutch case.

The final conclusion of Radzinski, as well as that of Groenink, is based on the argument that the considered sublanguage is the intersection of the entire language with a certain regular language, since both authors refer to the property that MCTALs are closed under intersection with regular sets. We implicitly used a similar, but strictly weaker property of semilinear languages in the Old Georgian case. It is related to the fact that a language is semilinear iff it is *letter equivalent* to a regular language (see Sect. 4 for more details). And, taking for granted Radzinski's general assumptions about the Chinese number-names, in a nearly identical way as in the Old Georgian case, the corresponding sublanguage can be used to show that the entire Chinese number-names system is non-semilinear as well, and thus not only a non–MCTAL.

3.1 Chinese Number-Names

The sublanguage of the Chinese number-names, considered by Radzinski ([13]), includes all well-formed strings composable only of instances of *wu (five)* and *zhao (*am. *trillion)* and consisting of at least one instance of each. This set is the following one:

$$L_M := \{ wu\ zhao^{k_1} \ldots\ wu\ zhao^{k_m}\ \mid m, k_i \in \mathbb{N}_+\ ,\ k_1 > k_2 > \ldots > k_m \} \quad (12)$$

[8] E.g., if Σ is a finite set with at least two elements there are uncountable many semilinear languages $L \subseteq \Sigma^*$.

Let $\{w_i \mid 0 \leq i < n\}$ be an enumeration of the Chinese number-names lexicon Σ, where $w_0 = \mathsf{wu}$ and $w_1 = \mathsf{zhao}$, and let $L_{CN} \subseteq \Sigma^*$ be the language of all such number names. Then, by Proposition 5, the Parikh image $M := p_\Sigma[L_M]$ is a non-semilinear subset of $\mathbb{N}^n$. The appropriate polynomial P_α, providing a lower bound in the sense of (ii) of Proposition 5, is of degree 2 and looks similar to that of the Old Georgian case, namely $P_\alpha(x) = x^2/2 + x/2$. Due to the definition of L_M and (12), we may conclude that

$$M = p_\Sigma[L_{CN}] \cap \{e^{(0)} + e^{(1)} + n_0 e^{(0)} + n_1 e^{(1)} \mid n_0, n_1 \in \mathbb{N}\} \tag{13}$$

Thus, if L_{CN} were semilinear, then M, as the intersection of two semilinear sets, would be as well. Hence, L_{CN} cannot be semilinear in general.

3.2 Dutch Coordination Phrases

Groenink ([5]) considers a fragment L_M of the language Dutch, which can be written as the infinite union $\bigcup_{k \in \mathbb{N}} L_k$, where for each $k \in \mathbb{N}$, L_k is the following sublanguage of Dutch

$$\begin{aligned} L_k \;=\; & \textit{dat Jan Piet Marie Fred}^k \\ & (\textit{hoorde leren}^k \textit{ uitnodigen})^+ \textit{ en } (\textit{zag leren}^k \textit{ omhelzen}) \end{aligned} \tag{14}$$

That is, e.g. for $k = 1$ the shortest string belonging to L_k is

dat	*Jan*	*Piet*	*Marie*	*Fred*	*hoorde*	*leren*	*uitnodigen*	*en*	(14′)
that	Jan	Piet$_1$	Marie$_2$	Fred$_3$	hear-past$_1$	teach-inf$_2$	invite-inf$_3$	and	

zag	*leren*	*omhelzen*
saw-past$_1$	teach-inf$_2$	embrace-inf$_3$

‘that Jan heard Piet teach Marie to invite Fred and saw him teach her to embrace him’

Now, let be $w_0 = \mathsf{Fred}$ and $w_1 = \mathsf{leren}$ in an enumeration $\{w_i \mid 0 \leq i < n\}$ of the Dutch lexicon Σ. Then the Parikh image $M := p_\Sigma[L_M]$ is of the form

$$M = \{(k, k \cdot m + k, l_2^{(k \cdot m)}, \ldots, l_{n-1}^{(k \cdot m)}) \mid k \in \mathbb{N}, m \in \mathbb{N}_+\} \subseteq \mathbb{N}^n \tag{15}$$

for some appropriate $l_i^{(k \cdot m)} \in \mathbb{N}$, $2 \leq i < n$, for each $k \in \mathbb{N}$ and $m \in \mathbb{N}_+$. The set M can be shown to be non-semilinear.[9] Then it is an immediate consequence that L_M must be a non-MCTAL, as is shown in [5].

[9] We omit a proof here, since it differs from that of Proposition 5 in certain details, though it can be done in a similar manner.

4 Discussion

Radzinski is aware of the fact that it is not quite clear whether the Chinese number-names system can be considered as properly related to an I-language (in the sense of Chomsky) or rather as related to an interface between a natural language component of the human cognitive endowment and a "mathematical" one. Therefore, Radzinski admits that the consequences of his example for the natural language Chinese do not become obvious at once. But also, if the first possibility is supposed to be the right one, it does not immediately become decidable whether the Chinese language "in its entirety" is a non–semilinear language.[10]

In contrast to the case of Chinese number-names, it does not become clear as quickly that the entire Dutch language is non-semilinear. This is due to reasons we will not go into in detail here. They are related to the fact that coordination phenomena in general present rather problematic data for an appropriate linguistical analysis. This is already mentioned by Manaster-Ramer ([9]), when he discusses his argumentation. He notices that his observation strongly depends on what is assumed to belong to the Dutch language in general by means of a competence–performance distinction. Looking at some examples, Manaster-Ramer on the one hand points out the possibility of "dropping" some of the NPs explicitly appearing in coordination phrases like (14). On the other hand, he shows that to "pump in" some new NPs, without simultaneously pumping in new verbs, is possible as well. Roughly speaking, the latter possibility is to coordinate some of the NPs as a complex NP which itself is part of an NP coordination. As far as we see, both possibilities (dropping and pumping in) together, even make validity of the argumentation that Dutch is no (simple) TAL very uncertain, at least in terms of weak equivalence.

However, it is not our intention to judge the conclusions of Radzinski and Manaster-Ramer, respectively Groenink, rigorously. We just wanted to emphasize some possible objections raised by the authors themselves, since the Old Georgian example seems to avoid the corresponding difficulties. First of all, there should be no doubt that this example falls into the domain of natural language capacity. Secondly, compared to the Dutch coordination-phenomenon the complex NP constructions of Old Georgian, considered above, have the advantage that they represent a strict morpho-syntactic phenomenon. That is to say, with regard to the case suffixes the underlying (syntactic) structure is "directly visible"; which is not necessarily the case e.g. within a coordination of coordinated NPs in Dutch. This kind of "visibility" turns out to be possible at least in the case of Suffixaufnahme as it is given with (4). Recall that, in accordance with (5), (4) is the "most complex" case by means of the total number of genitive suffixes

[10] Such a decision may be advanced, if (as is proposed in [9]) a certain kind of constraint on "classificatory capacity" is assumed to come up with an appropriate grammar for natural language. According to Radzinski, the validity of such a constraint at least implies that Chinese turns out to be a non-MCTAL in case of "proper inclusion" of the Chinese number-names within the Chinese language.

which may appear as suffixes of the nouns within a complex NP consisting of k stacked NPs for some $k \in \mathbb{N}_+$. And in fact, our argumention is in some sense strictly concentrated on this case. We do not have to take care of all possibilities of Suffixaufnahme coming up with the application of some (recursive) rules to a given basic structure of a complex NP like (6) as long as the total number of "addable" genitive suffixes is bounded as stated in (5). The definition of the Old Georgian fragment L_M is only an implicit one, also allowing for complex NPs in which the "full range of possible suffixes" is only partially realized like in (2). Moreover, since L_M actually is defined by means of its Parikh image $p_\Sigma[L_M]$, a priori there is no restriction on the order in which the stacked NPs have to appear within the subject-NP of a member of L_M. This is reasonable since there are also examples of preposed genitive NPs showing agreement by Suffixaufnahme.

P̣avle-js-i *tav-isa* *mok̇ueta-j* (16)
Paul-Gen-Nom head-Gen cutting-Nom

'the decapitation of the head of Paul'

Although Boeder ([1]) argues that "leftward" Suffixaufnahme is non-recursive and even limited to the possibility of taking up one single suffix like in (16), the contrary would not effect our argumentation. This also holds if it turned out that sentences including NPs like (2.1) or (2.2) belong to Old Georgian. According to Boeder, such NPs are in all probability ungrammatical.

γmrt-isa-jsa sasupevel-isa saidumlo-j (2.1)

sasupevel-isa γmrt-isa-jsa saidumlo-j (2.2)

More generally, all our considerations are independent of the (generative) grammar formalism which we prefer to investigate the underlying structures of complex NPs in Old Georgian.[11] Our argumentation stays valid as long as, on the one hand, complex NPs are in general allowed to have a surface structure like (4), and on the other hand, a kind of "upper bound-restriction" on the appearing suffixes like (5) holds, as we assume to be the case due to the analysis of Boeder.

Besides the particular result concerning Old Georgian our approach offers a more general perspective on a formal approach to languages or on dealing with formal languages. Let us recall first the definition of *constant growth* as it is proposed in [7] as a necessary property of a mildly context sensitive language (MCSL), a language derivable by an MCSG. A language L has constant growth iff there is some constant $c_0 \in \mathbb{N}$ and some finite set of constants $C \subseteq \mathbb{N}$ such that for any $\alpha \in L$ with $|\alpha| \geq c_0$ there is some $\alpha' \in L$ and some $c \in C$ for which the equation $|\alpha| = |\alpha'| + c$ holds. Here, for any $\beta \in L$, $|\beta|$ denotes its length. Indeed, this definition can be thought of as a kind of a weak pumpability condition, which

[11] E.g., a grammar in terms of a GB- or minimalist framework in which we want (in particular and in contrast to Boeder's approach) some kind of Freezing principle (in the sense of [18]) to be valid, or at least descriptively adequate (see e.g. [2]).

also is true for the somewhat strengthend version, the definition of semilinearity of a language. (It should be clear that each semilinear language a fortiori has the constant growth property.) And, in this sense our Proposition 5 provides a kind of weak pumping lemma. We already mentioned in Sect. 3 that our formal argumentation applying to Old Georgian is similar to a standard proof method in the framework of formal languages: first, to figure out a fragment of a given language L, where the fragment does not fall into a class of formal languages $\mathcal{L}$ which is closed under intersection with regular languages. Second, to show that the fragment is the intersection of L and a regular language, that is to show that L cannot belong to $\mathcal{L}$. In our considerations we first fixed the non-semilinear fragment L_M of Old Georgian L_G, then we argued that the Parikh image $p_\Sigma[L_M]$ is the intersection of the Parikh images of L_G and of the semilinear language $L_R \subseteq \Sigma^*$. Finally, we concluded that $p_\Sigma[L_G]$ cannot be semilinear referring to the property that the set of semilinear subsets of $\mathbb{N}^n$ is closed under intersection.

Actually, we do not need the last mentioned property in this general form, since L_R is a context free language and we have $p_\Sigma[L_R] = R$. L_R even is a regular language. Assuming L_G to be semilinear there would be a regular language L'_G which is *letter equivalent* to L_G, since any language is semilinear iff it is letter equivalent to a regular language.[12] L_R is the full pre-image of R with respect to p_Σ, thus the following equation can be given.

$$p_\Sigma[L_G \cap L_R] \;=\; p_\Sigma[L_G] \cap p_\Sigma[L_R] \;=\; p_\Sigma[L'_G] \cap p_\Sigma[L_R] \;=\; p_\Sigma[L'_G \cap L_R] \quad (17)$$

But L_M is the intersection of L_G and L_R. According to (17), and because $L'_G \cap L_R$ is context free, even regular, the Parikh image $p_\Sigma[L_M]$ would be semilinear. This would contradict the result shown in Sect. 2.

5 Some Final Remarks

Old Georgian is only one of several languages which show the phenomenon of Suffixaufnahme.[13] But, as far as we know, there is no other such language which allows for an unbounded iteration of Suffixaufnahme as is possible in Old Georgian, at least with regard to genitive case suffixes. In most languages this process is strictly limited to the possibility of taking up just one additional case suffix. However, there are e.g. some Australian language families in which a limited stacking of up to four case suffixes can be found (see [3]). Thus Old Georgian does not only show a phenomenon which is not widespread among natural languages in general, it seems to be even a rather bizzare language among those which allow for multiple case-marking. And, it should be remarked here that, nowadays, Suffixaufnahme in Modern Georgian is at best only marginally possible and in any case non recursive.

[12] Two languages L_1, $L_2 \subseteq \Sigma^*$ are said to be *letter equivalent* if $p_\Sigma[L_1] = p_\Sigma[L_2]$. Although not explicitly stated, the mentioned equivalence is already proven in [10]. The “only if” is even shown for context free languages.

[13] [12] provides a collection of papers on this topic.

Nevertheless, according to our argumentation Old Georgian provides a natural languge which, in a formal sense, is not semilinear in general, and thus no MCTAL in particular. Now, the intention underlying the definition of *mild context sensitivity* is to add a sufficient amount to the capacity of the context free grammar-type, which is just the necessary amount to capture natural language complexity. Thus, the question arises which kind of mildly context sensitive grammar formalism could be appropriate to deal with Old Georgian. The Simple Literal Movement Grammars (simple LMGs) of Groenink ([5]) seem to provide such a searched for tool. They can be thought of as a generalization of LCFRSs where a restricted non-linear rewriting is allowed. However, the class of all simple LMLs, the class of all languages derivable by some simple LMG, is known to be identical with PTIME ([5]). Thus for example also the language $\{a^{(2^n)} \mid n \in \mathbb{N}\}$ is a simple LML. Therefore, the class of simple LMLs includes languages which do not have constant growth. Groenink gives the simple LMG-type definition to cope with languages like the Dutch fragment to which we referred in Sect. 3 and which is no MCTAL. To restrict the class of simple LMGs in a reasonable way with respect to natural languages, Groenink proposes a revised definition of mild context sensitivity. In particular this definition contains a strengthend version of the constant growth property which he calls *finite pumpability* and which allows for several non-semilinear languages like the Dutch fragment. Whether this finite pumpability property also is fullfilled by the Old Georgian fragment L_M, which we have taken into account, depends on what is "really" contained in it. That is to say, it depends on an explicit definition of L_M which has not been necessary to give within our argumentation. Without looking at further details we want to finish our considerations here, just by claiming that, as far as we see, at least along the lines of Boeder's paper there does not seem to be a reason to exclude the Old Georgian fragment to be finite pumpable in the sense of Groenink.

Appendix

In order to prove Proposition 5 we first recall a classical result in Analysis as

Lemma 7. *For some $m \in \mathbb{N}$ let $\alpha = (a_0, \dots, a_m) \in \mathbb{R}^{m+1}$ be a finite sequence where $a_m > 0$. Then, there is some $r_0 \in \mathbb{R}$ such that $P_\alpha(x) = \sum_{i=0}^{m} a_i x^i > 0$ for all $x \in \mathbb{R}$ for which $x > r_0$.*

Proof of Proposition 5. First, we define a set $\overline{M}$ as a subset of M. For this purpose we choose fixed numbers $l_2^{(k)}, \dots, l_{n-1}^{(k)} \in \mathbb{N}$ for any $k \in \mathbb{N}_+$ according to property (i) and set

$$\overline{M} := \left\{ \left(k, P_\alpha(k), l_2^{(k)}, \dots, l_{n-1}^{(k)}\right) \mid k \in \mathbb{N}_+ \right\} \tag{18}$$

The proof will be done by contradiction. Assume M is semilinear and notice that M is not empty. Then there exists a $\mu \in \mathbb{N}_+$ and linear sets $N_1, \dots, N_\mu \subseteq \mathbb{N}^n$ for which we have

$$M = \bigcup_{j=1}^{\mu} N_j \tag{19}$$

Since $\overline{M}$ is an infinite subset of M there must be an N_j, $1 \leq j \leq \mu$, which contains an infinite subset of $\overline{M}$. We take N to be such an N_j and choose $u^{(0)}, u^{(1)}, \ldots, u^{(\nu)} \in \mathbb{N}^n$ for some appropriate $\nu \in \mathbb{N}$ such that

$$N = \left\{ u^{(0)} + \sum_{j=1}^{\nu} n_j u^{(j)} \mid n_j \in \mathbb{N} \text{ for } 1 \leq j \leq \nu \right\} \tag{20}$$

For $0 \leq i < n$ and for $0 \leq j \leq \nu$ let $u_i^{(j)}$ be the $(i+1)$–th component of $u^{(j)}$. In the special case $j = 0$ we will simply write u instead of $u^{(0)}$, and u_i instead of $u_i^{(0)}$, to denote the corresponding n–tuple and its components. Notice that $\nu \geq 1$ because N is infinite. In particular the set N contains an infinite subset of $\overline{M}$. Hence, there must be an unbounded (strongly) increasing (infinite) sequence $(k_i)_{i \in \mathbb{N}}$ in $\mathbb{N}_+$ for which we have

$$\left(k_i, P_\alpha(k_i), l_2^{(k_i)}, \ldots, l_{n-1}^{(k_i)} \right) \in N \tag{21}$$

Using (20) and (21) we now fix some $n_1^{(k_i)}, \ldots, n_\nu^{(k_i)} \in \mathbb{N}$ for every $i \in \mathbb{N}$, or every element of the sequence $(k_i)_{i \in \mathbb{N}}$, such that the following equation holds

$$u + \sum_{j=1}^{\nu} n_j^{(k_i)} u^{(j)} = \left(k_i, P_\alpha(k_i), l_2^{(k_i)}, \ldots, l_{n-1}^{(k_i)} \right) \tag{22}$$

That is to say, for all $i \in \mathbb{N}$ we have especially

$$u_0 + \sum_{j=1}^{\nu} n_j^{(k_i)} u_0^{(j)} = k_i \tag{22.1}$$

$$u_1 + \sum_{j=1}^{\nu} n_j^{(k_i)} u_1^{(j)} = P_\alpha(k_i) \tag{22.2}$$

Since not only N is infinite, but also the projection of N to the first component, there must be a natural number ω with $1 \leq \omega \leq \nu$ for which the first component $u_0^{(\omega)}$ of $u^{(\omega)}$ is different from zero. We may assume w.l.o.g. that $u_0^{(1)}, \ldots, u_0^{(\omega)}$ all are different from 0 and that $u_0^{(\omega+1)}, \ldots, u_0^{(\nu)}$ all are identical with 0. Then, for some fixed element K of the sequence $(k_i)_{i \in \mathbb{N}}$ we can distinguish two cases.

<u>1. case</u>: $n_j^{(K)} u_1^{(j)} = 0$ for any natural number j with $\omega < j \leq \nu$.

We first notice, because of (22.1) it follows immediately that

$$n_j^{(K)} \leq K \text{ for } 1 \leq j \leq \omega \text{ , since } u_0^{(j)} > 0 \text{ in such a case.} \tag{23}$$

Carrying on we define $m_1 := \max\{u_1^{(j)} \mid 1 \leq j \leq \omega\}$ and let β be the finite sequence $(a_0 - u_1, a_1 - \omega m_1, a_2, \ldots, a_m) \in \mathbb{R}^{m+1}$. Then, by Lemma 7 applied to the polynomial P_β and due to the fact that $(k_i)_{i\in\mathbb{N}}$ is an unbounded increasing sequence, we may suppose K to be great enough to fulfill

$$0 \;<\; (a_0 - u_1) + (a_1 - \omega m_1)K + \sum_{i=2}^{m} a_i K^i \;=\; P_\beta(K) \tag{24}$$

Thus, the following inequality can be given

$$\begin{array}{lll} & u_1 + \sum\limits_{j=1}^{\nu} n_j^{(K)} u_1^{(j)} & \\ = & u_1 + \sum\limits_{j=1}^{\omega} n_j^{(K)} u_1^{(j)} & \text{, since } n_j^{(K)} u_1^{(j)} = 0 \text{ for } \omega+1 \le j \le \nu \\ \le & u_1 + \omega m_1 K & \text{, by (23) and according to the definition of } m_1 \\ < & a_0 + a_1 K + \sum\limits_{l=2}^{m} a_i K^i & \text{, by (24)} \\ = & P_\alpha(K) & \end{array}$$

But, this provides a contradiction, since K is an element of the sequence $(k_i)_{i\in\mathbb{N}}$ and so (22.2) holds especially for K. (1. case)

2. case: $n_j^{(K)} u_1^{(j)} \neq 0$ for some natural number j with $\omega < j \le \nu$.

We may assume w.l.o.g. that $j = \nu$. Then $n_\nu^{(K)} > 0$ as well as $u_1^{(\nu)} > 0$. And so, we can continue by concluding

$$(n_\nu^{(K)} - 1)\, u_1^{(\nu)} \;<\; n_\nu^{(K)}\, u_1^{(\nu)} \;<\; (n_\nu^{(K)} + 1)\, u_1^{(\nu)} \qquad \text{, since } u_1^{(\nu)} > 0 \tag{25}$$

Hence, on the one hand we have

$$\begin{array}{ll} u_0 + \sum\limits_{j=1}^{\nu-1} n_j^{(K)} u_0^{(j)} + (n_\nu^{(K)} + 1)\, u_0^{(\nu)} = K & \text{, due to (22.1) since } u_0^{(\nu)} = 0 \\ u_1 + \sum\limits_{j=1}^{\nu-1} n_j^{(K)} u_1^{(j)} + (n_\nu^{(K)} + 1)\, u_1^{(\nu)} > P_\alpha(K) & \text{, by (22.2) and (25)} \end{array} \tag{26.1}$$

and on the other hand we have

$$\begin{array}{ll} u_0 + \sum\limits_{j=1}^{\nu} n_j^{(K)} u_0^{(j)} + (n_\nu^{(K)} - 1) u_0^{(\nu)} = K & \text{, due to (22.1) since } u_0^{(\nu)} = 0 \\ u_1 + \sum\limits_{j=1}^{\nu-1} n_j^{(K)} u_1^{(j)} + (n_\nu^{(K)} - 1) u_1^{(\nu)} < P_\alpha(K) & \text{, by (22.2) and (25)} \end{array} \tag{26.2}$$

Notice that the elements

$$z^{>} := u + \sum_{j=1}^{\nu-1} n_j^{(K)} u^{(j)} + (n_\nu^{(K)} + 1) u^{(\nu)}$$

$$z^{<} := u + \sum_{j=1}^{\nu-1} n_j^{(K)} u^{(j)} + (n_\nu^{(K)} - 1) u^{(\nu)}$$

both belong to N since $n_\nu^{(K)} > 0$, and therefore both belong to M too. But (26.1) and (26.2) point out that this is a contradiction to the property (ii) of the set M in any case. Assuming $P_\alpha(K)$ to be the corresponding upper bound for the second component of any n-tuple which belongs to M and which first component is K, we can refer to $z^{>}$, providing a contradiction by (26.1). Assuming $P_\alpha(K)$ to be the corresponding lower bound presupposed by property (ii) we can refer to $z^{<}$ and (26.2). (2. case)

The two considered cases are the only ones that can appear. So, we have just shown that the set M cannot be semilinear, that is to have proven the proposition. □

References

1. Winfried Boeder. Suffixaufnahme in Kartvelian. In [12], pages 151–215.
2. M. A. Browning. Bounding conditions on representations. *Linguistic Inquiry* 22:541–562, 1991.
3. Nick Evans. Multiple-case in Kayardild: Anti-iconic suffix ordering and the diachronic filter. In [12], pages 396–428.
4. Seymour Ginsburg and Edwin H. Spanier. Bounded ALGOL-like languages. *Transactions of the American Mathematical Society*, 113:333–368, 1964.
5. Annius V. Groenink. Mild context-sensitivity and tuple-based generalizations of context-free grammar. In D. E. Johnson and L. S. Moss, editors, *Special Issue on the Mathematics of Language, Linguistics and Philosophy*, forthcoming. Also available as Report CS-R9634, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 1996.
6. Riny Huybregts. The weak inadequacy of context-free phrase structure grammars. In G. de Haan, M. Trommelen, and W. Zonneveld, editors, *Van Periferie naar Kern*, pages 81–99. Foris Publications, Dordrecht, 1984.
7. Aravind K. Joshi. Tree adjoinig grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In D. Dowty, L. Karttunen, and A. Zwicky, editors, *Natural Language Parsing. Theoretical, Computational and Psychological Perspective*, pages 206–250. Cambridge University Press, New York, 1985.
8. Aravind K. Joshi, K. Vijay-Shanker, and David J. Weir. The convergence of mildly context-sensitive grammar formalisms. In P. Sells, S. Shieber, and T. Wasow, editors, *Foundational Issues in Natural Language Processing*, pages 31–81. MIT Press, Cambridge (Mass.), 1991.
9. Alexis Manaster-Ramer. Dutch as a formal language. *Linguistics and Philosophy*, 10:221–246, 1987.

10. Rohit J. Parikh. On context-free languages. *Journal of the Association for Computing Machinery*, 13:570–581, 1966.
11. Frans Plank. (Re-)introducing Suffixaufnahme. In [12], pages 3–110.
12. Frans Plank, editor. *Double Case. Agreement by Suffixaufnahme.* Oxford University Press, New York, 1995.
13. Daniel Radzinski. Chinese number-names, tree adjoining languages, and mild context-sensitivity. *Computational Linguistics*, 17:277–299, 1991.
14. Owen Rambow. *Formal and Computational Aspects of Natural Language Syntax.* PhD thesis, University of Pennsylvania, Philadelphia, PA, 1994.
15. Stuart M. Shieber. Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8:333–343, 1985.
16. K. Vijay-Shanker and David J. Weir. The equivalence of four extensions of context-free grammars. *Mathematical Systems Theory*, 27:511–546, 1994.
17. David J. Weir. *Characterizing Mildly Context-Sensitive Grammar Formalisms.* PhD thesis, University of Pennsylvania, Philadelphia, PA, 1988.
18. Kenneth Wexler and Peter W. Culicover. *Formal Principles of Language Aquisition.* MIT Press, Cambridge (Mass.), 1980.