Local Tree Description Grammars

Laura Kallmeyer

Universität Tübingen
lk@sfs.nphil.uni-tuebingen.de

1 Introduction

A lot of interest has recently been paid to constraint-based definitions and extensions of Tree Adjoining Grammars (TAG). Examples are the so-called quasi-trees (see Vijay-Shanker (1992) and Rogers (1994)), D-Tree Grammars (see Rambow et al. (1995)) and Tree Description Grammars (TDG) (see Kallmeyer (1996a,b)). The latter are grammars consisting of a set of formulas denoting trees. TDGs are derivation-based where in each derivation step a conjunction is built of the old formula, a formula of the grammar and additional equivalences between node names of the two formulas. This formalism is more powerful than TAGs. TDGs offer the advantages of MC-TAG (see Joshi (1987a)) and D-Tree Grammars for natural languages, and they allow underspecification. However, the problem is that TDGs might be unnecessarily powerful for natural languages. To solve this problem, in this paper, I will propose local TDGs, a restricted version of TDGs. Local TDGs still have the advantages of TDGs but they are semilinear and therefore more appropriate for natural languages.

First, the notion of semilinearity is defined. Then local TDGs are introduced, and, finally, semilinearity of local Tree Description Languages is proven.

2 Semilinearity

Let N be the set of non-negative integers. For $(a_1, \cdots, a_n), (b_1, \cdots, b_n) \in N^n$ and $m \in N$ we define: $(a_1, \cdots, a_n) + (b_1, \cdots, b_n) := (a_1 + b_1, \cdots, a_n + b_n)$ and $m(a_1, \cdots, a_n) := (ma_1, \cdots, ma_n)$.

For some alphabet $X = \{a_1, \cdots, a_n\}$ with some (arbitrary) fixed order of the elements, a function $p : X^* \to N^n$ is called a Parikh-function, if:
For all $w \in X^*$: $p(w) := (|w|_{a_1}, |w|_{a_2}, \cdots, |w|_{a_n})$, where $|w|_{a_i}$ is the number of occurences of a_i in w. For all $L \subseteq X^*$: $p(L) := \{p(w) | w \in L\}$.

Two strings $x_1, x_2 \in X^*$ are letter equivalent if they contain equal number of occurences of each symbol, i.e. if $p(x_1) = p(x_2)$ for some Parikh-function p. Two languages $L_1, L_2 \subseteq X^*$ are letter equivalent if every element in L_1 is letter equivalent to an element in L_2 and vice-versa, i.e. if $p(L_1) = p(L_2)$ for some Parikh-function p.

Definition 1 *(Semilinearity)*

1. *Let $x_0, x_1, \cdots, x_m, 0 \leq m$ be in N^n. A linear subset of N^n is a set*
   $\{x_0 + n_1x_1 + \cdots + n_mx_m \quad | \quad n_i \in N \textit{ for } 1 \leq i \leq m\}$.

2. *The union of finitely many linear subsets of N^n is a semilinear subset of N^n.*
3. *A language $L \subseteq X^*$ is semilinear, if there is a Parikh-function p such that $p(L)$ is a semilinear subset of N^n.*

Proposition 1 *(Parikh-Theorem) Each context free language is semilinear.*

Clearly, each language that is letter equivalent to a semilinear language is semilinear as well. Because of the Parikh-Theorem (proven by Parikh (1966)), this means that for some language L, in order to prove the semilinearity of L, it is sufficient to show that L is letter equivalent to a context free language.

Semilinearity is an important language property because it seems plausible that natural languages are semilinear (see Joshi (1987b) and Vijay-Shanker et al. (1987)). As far as I know, the only example of a possibly non-semilinear phenomenon is case stacking in Old Georgian (see Michaelis and Kracht (1996)). Since it is not clear whether there is really a (theoretically) infinite progression of stacking possible, there is no reason to assume natural languages not to be semilinear, as long as these are the only examples of nonsemilinear phenomena. If natural languages are semilinear, then it is desirable that the languages generated by grammar formalisms intended to capture human language capacity are semilinear as well.

3 Local TDGs

The tree logic used for local TDGs is the same as for TDGs (see Kallmeyer (1996b)). It is similar to the logic proposed by Rogers (1994) for TAGs. The logic is a quantifier-free first order logic with variables K (node names), binary relations $\lhd$ (parent or immediate dominance), $\lhd^*$ (dominance), $\prec$ (linear precedence) and $\approx$ (equality), a symbol δ for the labelling function, sets of constants N and T for the nonterminal and terminal labels, and logical connectives $\neg$, $\wedge$ and $\vee$. Satisfaction is defined with respect to special models (finite labelled trees) and variable assignments. ϕ_1 entails ϕ_2 ($\phi_1 \models \phi_2$) for two formulas ϕ_1, ϕ_2 iff all finite labelled trees satisfying ϕ_1 with respect to an assignment g also satisfy ϕ_2 with respect to g. A sound, complete and decidable notion of syntactic consequence, $\phi_1 \vdash \phi_2$, can be defined for this logic.

In the formulas in TDGs (descriptions) certain subtrees are uniquely described together with dominance relations between these trees. A negation free, disjunction free satisfiable formula ϕ is a *description* if there is at least one $k \in node(\phi)$ ($k \in K$ occuring in ϕ) such that $\phi \vdash k \lhd^* k'$ for all $k' \in node(\phi)$ (k is called *minimal* in ϕ), and if for all k_1, k_2, k_3:

- If $\phi \vdash k_1 \lhd k_2 \wedge k_1 \lhd^* k_3$, then either $\phi \vdash k_1 \approx k_3$ or there is a k_4 with $\phi \vdash k_1 \lhd k_4 \wedge k_4 \lhd^* k_3$.
- If $\phi \vdash k_1 \lhd k_2 \wedge k_1 \lhd k_3$, then either $\phi \vdash k_2 \prec k_3$ or $\phi \vdash k_2 \approx k_3$ or $\phi \vdash k_3 \prec k_2$.

To guarantee that in each derivation step, descriptions with disjoint sets of node names can be chosen, an equivalence relation on $\{(\phi, K_\phi); \phi$ is a description and

$K_\phi \subseteq node(\phi)\}$ is needed: $(\psi_1, K_{\psi_1}) \approx_K (\psi_2, K_{\psi_2})$ iff ψ_1 and ψ_2 only differ in a bijection (variable renaming) $f_K : K \to K$ with $K_{\psi_2} = f_K(K_{\psi_1})$.

A *TDG* is a tuple $G = (N, T, D, \phi_S)$, such that:

1. N and T are pairwise disjoint finite sets, the nonterminals and terminals
2. D is a finite set of equivalence classes $\overline{(\psi, K_\psi)}$ (wrt $\approx_K$), such that for all $(\psi, K_\psi) \in \overline{(\psi, K_\psi)}$, ψ is a description with constants N and T. ψ is called an *elementary description* of G, and each $k \in K_\psi$ is called *marked* in ψ.
3. ϕ_S is a description (with constants N and T), the *start description*.

In a derivation step $\phi_1 \overset{\psi}{\Rightarrow} \phi_2$, the result ϕ_2 is the conjunction of ϕ_1, an elementary ψ and equivalences $k_1 \approx k_2$ with $k_1 \in node(\phi_1)$ and $k_2 \in \{k; k$ minimal in ψ or $k \in K_\psi\}$. The main idea of local TDGs is to restrict the derivation mode such that all $k_1 \in node(\phi_1)$ used for new equivalences occur in one single elementary ψ_d that was added before. Furthermore, each $k_1 \in node(\phi_1)$ can be used but once to introduce a new equivalence. Then the derivation step only depends on ψ_d, and the derivation process can be described by a context-free grammar. Doing this, letter equivalence of local TDLs (the string languages of local TDGs) and context-free languages can be shown, and, consequently, local TDLs are semilinear.

To understand the intuitions behind the definition of local TDGs, it is helpful to have an idea of the semilinearity proof. In this proof, for a given local TDG G_T a letter equivalent context-free grammar G_{CF} is obtained as follows: The nonterminals in G_{CF} describe "states" of elementary descriptions used in the course of a derivation. For a derived description ϕ in the corresponding derivation in G_{CF} there is one nonterminal Z_{ψ_d} for each start or elementary description ψ_d added in the course of the derivation of ϕ. Z_{ψ_d} specifies in which way the names of ψ_d can be used in a new derivation step. For each $k \in node(\psi_d)$, Z_{ψ_d} gives information about whether k has a parent or daughter in ϕ, whether k is minimal or does not dominate any other name in ϕ and whether k is strongly dominated by a name k' such that $\phi \vdash \delta(k') \approx X$ for some label X. (A strong dominance in ϕ is a conjunct $k_1 \lhd^* k_2$ in ϕ that is not entailed by the rest of ϕ, i.e. ϕ without this conjunct. Notation: $\phi \vdash_s k_1 \lhd^* k_2$.)

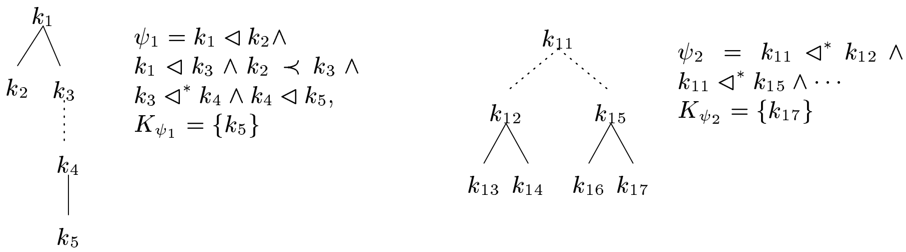
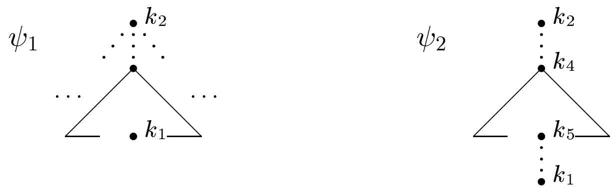


Figure 1: non-local elementary descriptions

For the old description ϕ in a derivation step the following should hold: Only for the elementary ψ_d (in ϕ) used in this derivation step may the state change. Therefore "subtree descriptions" (e.g. the part with k_{12}, k_{13}, k_{14} in ψ_2 in Fig. 1) must not be inserted into strong dominances $\phi \vdash_s k \lhd^* k'$ with $k' \notin node(\psi_d)$. To guarantee this the form of the descriptions is restricted by defining local descriptions. The descriptions of Fig. 1 for example are not local. If k_{13} or k_{14} was marked, then ψ_2 would be local.

Definition 2 *(Local description) An elementary description ψ in a TDG G is local, if for all $k_1, k_2, k_3 \in node(\psi)$:*

1. *If $\psi \vdash k_1 \approx k_2$, then $k_1 = k_2$.*
2. *If $\psi \vdash_s k_2 \lhd^* k_1$ and $\psi \vdash_s k_3 \lhd^* k_1$, then $k_2 = k_3$.*
3. *If $\psi \vdash_s k_1 \lhd^* k_2$ and $\psi \vdash_s k_1 \lhd^* k_3$, then either $k_2 = k_3$ or: k_1 is minimal or marked in ψ and there are $k_4, k_5 \in K_\psi$ with $\psi \vdash k_2 \lhd^* k_4$ and $\psi \vdash k_3 \lhd^* k_5$.*
4. *If $k_1 \in K(\psi)$ and k_2 is marked or minimal in ψ with $k_1 \neq k_2$ and $\psi \vdash k_2 \lhd^* k_1$, such that there is no further marked name between k_1 and k_2, then:*
   - *There is a $k \in node(\psi)$ with $\psi \vdash_s k_2 \lhd^* k$ and $\psi \vdash k \lhd^* k_1$, and for all $k_3 \in K_\psi$: if $\psi \vdash k \lhd^* k_3$, then $\psi \vdash k_1 \lhd^* k_3$.*
   - *If there are k_4, k_5 with $\psi \vdash k_4 \lhd^* k_5$, $\psi \vdash_s k_2 \lhd^* k_4$ and $\psi \vdash_s k_5 \lhd^* k_1$, then: there is an $X \in N$ with $\psi \vdash \delta(k_i) \approx X$ for all $i \in \{1, 2, 4, 5\}$, and if there is a k with $\psi \vdash_s k_2 \lhd^* k$, then $k = k_4$ holds.*



By this definition two kinds of marked names k_1 with k_2 as next marked or minimal name dominating k_1 are allowed: first (see ψ_1) names k_1 with no $k \neq k_2$ strongly dominating k_1. The second type (see ψ_2) are marked names where underspecification can occur. This is the case, if k_2 strongly dominates some k_4 and k_1 is strongly dominated by some k_5. k_4, k_5, k_1 and k_2 then have the same labels, and there are no other names strongly dominated by k_2. Generally, names k that are not marked or minimal do not strongly dominate more than one name.

A *local TDG* is a TDG $G = (N, T, D, \phi_S)$ where ϕ_S and all elementary descriptions are local. As already mentioned, the main idea of local derivation is to use for new equivalences only names from one elementary ψ_d in the old description ϕ_1, and to use each $k \in node(\phi_1)$ at most once.

Definition 3 *(Local derivation) Let G be a local TDG. For an elementary ψ in G and descriptions ϕ_1, ϕ_2 with $\phi_S \overset{*}{\Rightarrow}_l \phi_1$ and $node(\psi) \cap node(\phi_1) = \emptyset$: $\phi_1 \overset{\psi}{\Rightarrow}_l \phi_2$ holds (ϕ_2 is locally derived from ϕ_1 in one step), if there is a ψ_d with $\psi_d = \phi_S$ or $\phi_S \overset{*}{\Rightarrow}_l \phi \overset{\psi_d}{\Rightarrow}_l \phi' \overset{*}{\Rightarrow}_l \phi_1$, such that:*

1. *$\phi_2 \vdash \phi_1 \wedge \psi$.*
2. *For all $k_1 \in node(\phi_1), k_2 \in node(\psi)$ such that $\phi_2 \vdash k_1 \approx k_2$, there is a k_1' with $\phi_1 \vdash k_1 \approx k_1'$ and*

(i) $k'_1 \in node(\psi_d)$*, and* k_2 *is marked or minimal in* ψ*.*

(ii) For all k *with* $\phi_1 \vdash k'_1 \approx k$*: either* $k'_1 = k$*, or* $\psi_d \neq \phi_S$ *and* $\phi' \vdash k'_1 \approx k$*.*

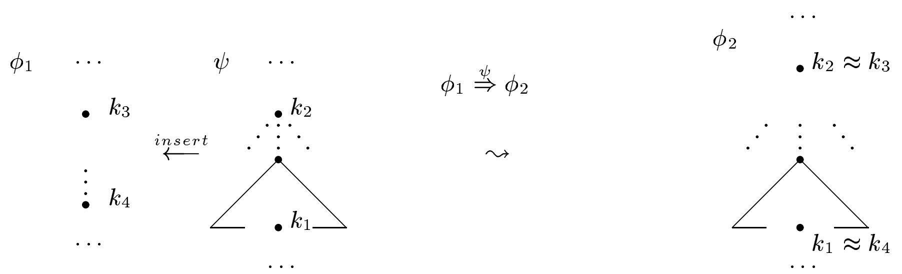
(iii) If k_m *is the next marked or minimal name dominating* k_2 *and there are* k'_m, k'_2 *with* $\psi \vdash_s k_m \lhd^* k'_m$ *and* $\psi \vdash k'_m \lhd^* k'_2 \wedge k'_2 \lhd k_2$*, then: There is a* k *with* $\phi_1 \vdash_s k \lhd^* k'_1$ *such that for all* k'*: if* $\psi \vdash k'_m \lhd^* k'$*, then* $\phi_2 \vdash k \lhd^* k'$*.*

(iv) If there is no $k_3 \in K_\psi$*,* $k_2 \neq k_3$*, such that* $\psi \vdash k_2 \lhd^* k_3$*, then either* k'_1 *is a leaf name in* ϕ_1 *or* k_2 *is a leaf name in* ψ*. (*k *is a leaf name in* ϕ *iff for all* k'*: If* $\phi \vdash k \lhd^* k'$*, then* $\phi \vdash k \approx k'$*.)*

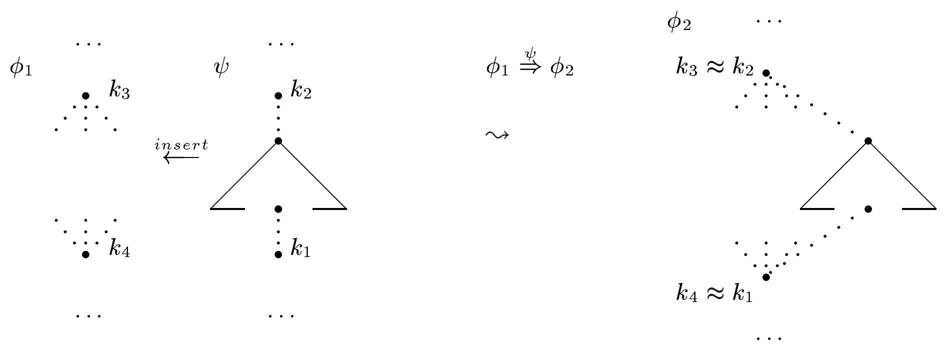
(v) If there is a $k_3 \in K_\psi$ *with* $\psi \vdash k_2 \lhd^* k_3$ *and* $k_2 \neq k_3$*, if there is no marked name between* k_2 *and* k_3*, and if there are* k'_2, k'_3 *with* $\psi \vdash_s k_2 \lhd^* k'_2$ *and* $\psi \vdash_s k'_3 \lhd^* k_3$ *and* $\psi \vdash k'_2 \lhd^* k'_3$*, then: If* $k_4 \in node(\psi_d)$ *with* $\phi_2 \vdash k_4 \approx k_3$*, then for all* $k \in node(\phi_1)$*:* $\phi_1 \not\vdash k'_1 \lhd k \vee k \lhd k_4$*.*

3. For all ϕ_3 *such that 1. and 2. hold for* ϕ_3*: If* $\phi_2 \vdash \phi_3$*, then* $\phi_3 \vdash \phi_2$*.*

(i) makes sure that all $k \in node(\phi_1)$ used in one derivation step are from one elementary ψ_d. (ii) says that each name can be used only once for a derivation step. Because of (iii), parent relations in ϕ_2 come from exactly one of the descriptions ϕ_1 or ψ, and everything between two marked or minimal names in ψ must be inserted into one single strong dominance. With (iv) a $k \in K_\psi$ not dominating any other $k' \in K_\psi$ either is a leaf name or it is identified with a leaf name in ϕ_1. Because of *1.* and *3.*, ϕ_2 must entail ψ and ϕ_1, and ϕ_2 must be maximally underspecified.



For $k_1, k_2 \in node(\psi)$ either marked or minimal with no marked names in between and with $\phi_2 \vdash k_4 \approx k_1 \wedge k_3 \approx k_2$ for $k_3, k_4 \in node(\phi_1)$: Either there is no $k \neq k_2$ with $\psi \vdash_s k \lhd^* k_1$. Then the derivation step is as in the preceding figure. Or, if there is such a k, (see *(v)*) the derivation step has the form:



In a local TDG G, $L^l_D(G)$ is the set of descriptions that can be locally derived from ϕ_S. The tree language is the set of *minimal trees* of these descriptions. A minimal tree of ϕ is a tree that satisfies ϕ such that all parent relations in the tree are already described in ϕ. The set of strings yielded by these trees is the string language.

Local TDGs are still powerful enough to describe $\{a_1^n a_2^n \cdots a_k^n\}$ and copy languages. Local Tree Description Languages (TDL) are a true superset of Tree Adjoining Languages. With local TDGs, as with MC-TAGs, several subtree descriptions can be added simultaneously, and subsertion-like derivation steps as in D-Tree Grammars are possible. Furthermore, in cases of scope ambiguities, underspecified representations can be derived (see Kallmeyer (1996b, 1997)).

4 Semilinearity of local TDLs

Proposition 2 *Local TDLs are letter equivalent to context-free languages.*

Proof (outline): Let $G_T = (N_T, T, D, \phi_S)$ be a local TDG such that without loss of generality for all elementary or start descriptions ϕ and all $k \in node(\phi)$ there is a $X \in N_T \cup T \cup \{\epsilon\}$ with $\phi \vdash \delta(k) \approx X$.
Construction of a letter equivalent context-free grammar $G_{CF} := (N_C, T, P, S)$: The nonterminals are states Z of the form $Z = \phi_Z \wedge \xi_Z$ with: $\phi_Z = \phi_S$ or ϕ_Z elementary in G_T (one representative for each class in D is chosen). ξ_Z is a conjunction of formulas $parent(k)$, $child(k)$, $leaf(k)$, $minimal(k)$, $dom_\uparrow(k, X)$ or $derive(k)$ or their negations with $k \in node(\phi_Z)$ and $X \in N_T$. For each state $Z = \phi_Z \wedge \xi_Z$ for all $k \in node(\phi_Z)$ and all such formulas $\psi = parent(k), \cdots$ either ψ or $\neg\psi$ must occur in ξ_Z.
Additionally N_C contains a start symbol S different from all other nonterminals. Let $Z^\sim = \phi_Z^\sim \wedge \xi_Z^\sim$ be equivalent to one $Z \in N$ ("equivalent" means that Z and $Z^\sim$ only differ in a bijection K). We define: A description ϕ with $\phi_S \stackrel{*}{\Rightarrow}_l \phi$ entails $Z^\sim$, $\phi \models Z^\sim$, as follows:

1. $\phi \models parent(k)$ iff there is a k' such that $\phi \models k' \lhd k$.
2. $\phi \models child(k)$ iff there is a k' such that $\phi \models k \lhd k'$.
3. $\phi \models leaf(k)$ iff k is a leaf name in ϕ.
4. $\phi \models minimal(k)$ iff k is minimal in ϕ.
5. $\phi \models derive(k)$ iff there are ϕ_1, ϕ_2 such that $\phi_S \stackrel{*}{\Rightarrow}_l \phi_1 \Rightarrow_l \phi_2 \stackrel{*}{\Rightarrow}_l \phi$, $k \in node(\phi_1)$ and $\phi_2 \models k \approx k'$ for one $k' \notin node(\phi_1)$.
6. $\phi \models dom_\uparrow(k, X)$ iff there is a k' with $\phi \vdash_s k' \lhd^* k \wedge \delta(k') \approx X$.
7. Apart from this, $\phi_1 \models \phi_2$ is defined as before.

Productions P:

1. If $Z_S \in N$ with $Z_S = \phi_S \wedge \xi_S$ and $\phi_S \models \xi_S$ and if $t_1, \cdots, t_n$ are all occurences of terminals in ϕ_S, then $S \rightarrow t_1 \cdots t_n Z_S \in P$.
2. Let Z and Z' be states for the same elementary or start description, Z_{new} a state for some elementary ψ, and $t_1, \cdots, t_n$ all occurences of terminals in ψ. $Z \rightarrow t_1 \cdots t_n Z' Z_{new} \in P$ iff the following holds:

For all ϕ, $\phi_S \stackrel{*}{\Rightarrow}_l \phi$ entailing a $Z^{\sim} = \phi^{\sim} \wedge \xi^{\sim}$ equivalent to Z: There is a ϕ' with $\phi \stackrel{\psi}{\Rightarrow}_l \phi'$ and $Z'^{\sim} = \phi'^{\sim} \wedge \xi'^{\sim}$ and $Z^{\sim}_{new} = \phi^{\sim}_{new} \wedge \xi^{\sim}_{new}$ equivalent to Z' and Z_{new} such that $\phi' \models Z'^{\sim} \wedge Z^{\sim}_{new}$. Furthermore $\phi^{\sim} = \phi'^{\sim}$ and $\psi = \phi^{\sim}_{new}$ hold and $\phi^{\sim}$ is the elementary ψ_d (see Def. 3) used in this derivation step.

3. For all $Z \in N$, $Z = \phi_Z \wedge \xi_Z$:
   $Z \to \epsilon \in P$ iff for all k in ϕ_Z: if X is the label of k, then either $parent(k)$ or $dom_{\uparrow}(k, X)$ or $derive(k)$ or $minimal(k)$ is in ξ_Z.

G_{CF} is unique and it is a context-free grammar.
By induction on the length n of the derivation the following can be shown:

$S \stackrel{n+1}{\Rightarrow} w_n$ wrt G_{CF} without applying ϵ-productions, and $Z_1, \cdots Z_n$ are all occurences of nonterminals in w_n
iff there is a derivation $\phi_S \stackrel{n}{\Rightarrow}_l \phi_n$ wrt G_T such that there are pairwise different $Z_1^{\sim}, \cdots, Z_n^{\sim}$ with $Z_i^{\sim} = \phi_i^{\sim} \wedge \xi_i^{\sim}$ equivalent to Z_i, with:

- The elementary or start descriptions that have been used in course of the derivation of ϕ_n, are exactly $\phi_1^{\sim}, \cdots, \phi_n^{\sim}$.
- $\phi_n \models Z_i^{\sim}$ for all $1 \leq i \leq n$.

With the ϵ-productions the following holds for w_n, ϕ_n as above: $w_n \stackrel{*}{\Rightarrow} w'_n$ can be derived by applying only ϵ-productions and $w'_n \in T^*$ iff ϕ_n has a minimal tree.
In general: $\phi_S \stackrel{*}{\Rightarrow}_l \phi$ wrt G_T, ϕ has a minimal tree yielding the string w iff there is a w' letter equivalent to w such that $S \stackrel{*}{\Rightarrow} w'$ wrt G_{CF}.

□

As a corollary local TDLs are semilinear.

5 Conclusion

TDGs have been developed to give a constraint-based TAG-extension that offers the advantages of MC-TAGs and D-Tree Grammars, and to introduce underspecification to TAGs. However, TDGs seem to be unnecessarily powerful for natural languages. For this reason I have presented local TDGs in this paper, a restriction of TDGs that is still much more powerful than TAGs. Local TDGs also have the advantages of MC-TAGs and D-Tree Grammars, and even underspecified representations are still possible in local TDGs (see Kallmeyer (1996b, 1997)). By describing the derivation process by a context-free grammar, I have proven that local TDGs are semilinear, which indicates that they really are an interesting alternative to other formalisms developed for natural language processing.

Acknowledgments

For critical discussions and helpful comments, I would like to thank Tom Cornell and Frank Morawietz.

References

Joshi, A. K.: 1987a, An introduction to Tree Adjoining Grammars, *in* A. Manaster-Ramer (ed.), *Mathematics of Language*, John Benjamins, Amsterdam.

Joshi, A. K.: 1987b, Tree adjoining grammars: How much contextsensitivity is required ro provide reasonable structural descriptions?, *in* D. Dowty, L. Karttunen and A. Zwicky (eds), *Narural Language Parsing*, Cambridge University Press.

Kallmeyer, L.: 1996a, Tree Description Grammars, *in* D. Gibbon (ed.), *Natural Language Processing and Speech Technology. Results of the 3rd KONVENS Conference*, Mouton de Gruyter, Berlin.

Kallmeyer, L.: 1996b, Underspecification in Tree Description Grammars, *Arbeitspapiere des SFB 340 81*, Universität Tübingen. To appear in: Hans Peter Kolb and Uwe Mönnich (eds.) *The Mathematics of Syntactic Structures.*

Kallmeyer, L.: 1997, A syntax-semantics interface with Synchronous Tree Description Grammars, *Proceedings of Formal Grammar*, Aix en Provence. to appear.

Michaelis, J. and Kracht, M.: 1996, Semilinearity as a syntactic variant. Abstract, *Logical Aspects of Computational Linguistics*, Nancy.

Parikh, R.: 1966, On context free languages, *Jounal of the ACM* **13**, 570–581.

Rambow, O., Vijay-Shanker, K. and Weir, D.: 1995, D-Tree Grammars, *Proceedings of ACL.*

Rogers, J.: 1994, *Studies in the Logic of Trees with Applications to Grammar Formalisms*, PhD thesis, University of Delaware.

Vijay-Shanker, K.: 1992, Using descriptions of trees in a tree adjoining grammar, *Computational Linguistics* **18**(4), 481–517.

Vijay-Shanker, K., Weir, D. J. and Joshi, A. K.: 1987, Characterizing structural descriptions produced by various grammatical formalisms, *Proceedings of ACL*, Stanford.