10 Limitations of TAG: Scrambling in German

Rambow (1994); Kulick (2000); Kallmeyer (2005)

10.1 Scrambling and LTAG

Scrambling is a movement of arguments and adjuncts within the Mittelfeld.

(50) a. [Peter]$_{VF}$ hat [seiner Schwester das Buch]$_{MF}$ gegeben

b. [Peter]$_{VF}$ hat [das Buch$_i$ seiner Schwester t_i]$_{MF}$ gegeben

A special case is *long-distance* scrambling where arguments or adjuncts of an embedded infinitive are 'moved' out of the embedded VP. This occurs for instance in languages such as German, Hindi, Japanese and Korean. These languages are therefore often said to have a free word order.

(51) ... dass [es]$_1$ der Mechaniker [t_1 zu reparieren] verspricht

In German there is

1. no bound on the depth of scrambling (i.e., in terms of movement, the number of VP borders crossed by the moved element),

   (52) a. ... dass [den Kühlschrank]$_1$ niemand [[t_1 zu reparieren] zu versuchen] versprochen hat

   b. ... dass [den Kühlschrank]$_1$ niemand [[[t_1 zu reparieren] zu versuchen] zu versprechen] bereit ist

2. and no bound on the number of scrambled elements.

   (53) ... dass [eine hiesige Firma] [meinem Onkel] [die Möbel] [vor drei Tagen] [ohne Voranmeldung] zugestellt hat

   Any permutation of the five elements in the Mittelfeld in (53) is grammatical.

Not only NPs and PPs, also VPs can be scrambled and scrambling out of scrambled VPs is possible as well:

(54) ... dass [den Kühlschrank]$_1$ ihm$_2$ [t_1 zu reparieren]$_3$ niemand [t_2 t_3 zu versprechen] bereit ist

As shown in Becker et al. (1991), TAG are not powerful enough to describe scrambling in German in an adequate way (if the above-mentioned unboundedness is assumed). I.e., an analysis with each argument attaching to the verb it depends on is not possible.[15]

Reason why scrambling is problematic for TAG: more than one occurrence of scrambling is possible (see (54)) ⇒ more than one VP can be separated into discontinuous parts. If only one was separated, one could adjoin the intervening material (as in the standard analysis of wh-movement).

⇒ possible solution: break the VPs into different pieces, i.e., use multicomponents.

First proposal in this direction: non-local MCTAG with dominance links (Becker et al. (1991)). Dominance links are additional constraints requiring dominance relations between a foot node and a root node of another elementary tree inside one elementary tree set. The dominance links must be satisfied in the derived tree.

[15]Note that scrambling up to depth 2 (2 crossed VP borders) actually can be described within TAG. I.e., for the examples above TAG analyses are possible. But from depth 3 on, only some word order variations can be accounted for within TAG, others not.

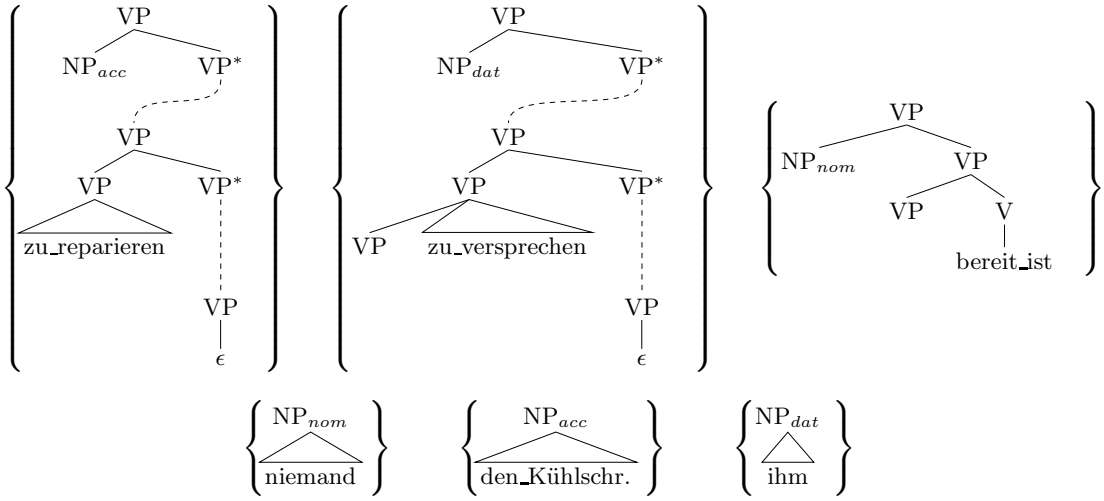

Figure 17: non-local MCTAG with dominance links: analysis of (54)

See analysis of (54) in Fig. 17: the arguments of the different verbs in the tree sets can be adjoined arbitrarily high as long as their foot node dominates the root of the lower VP node in the derived tree. Note that an argument must be adjoined at the same time as its verb because of the simultaneity constraint. Therefore, adjunction on foot nodes is necessary. Furthermore, trees form the same set cannot be adjoined into each other ⇒ extra sets needed for non-scrambling cases (arguments immediately preceding their verbs) and also for groups of arguments scrambled together.

Problem of non-local MCTAG: formal properties not well-studied, probably not mildly context-sensitive, in particular not polynomially parsable. Also, there is no locality constraint on the derivations. ⇒ locality constraints for movement (e.g., scrambling in German cannot proceed out of tensed clauses, see (55)) must be obtained by some additional mechanism.

(55) a. ... dass Peter sagt, dass er [das Auto zu reparieren] versucht

b. *... dass Peter [das Auto]$_1$ sagt, dass er [t_1 zu reparieren] versucht

Exercise 19 *What are the sentences generated by the non-local MCTAG with dominance links from Fig. 17?*

10.2 V-TAG

Rambow (1994)

V-TAG are like non-local MCTAG with dominance links except that the trees from one elementary tree set are *not* required to be added *simultaneously*.

A V-TAG without dominance links is called a Vector MCTAG (V-MCTAG).

Definition 19 (V-MCTAG and V-TAG)

1. *A Vector-MCTAG (V-MCTAG) is a tuple* $G = \langle N, T, I, A, O, C, \mathcal{I}, \mathcal{A} \rangle$ *such that:*
   - $G_{TAG} := \langle N, T, I, A, O, C \rangle$ *is a TAG with adjunction constraints, and*
   - $\mathcal{I} \subseteq P(I \cup A)$ *such that for all* $v \in \mathcal{I}$*:* $|v \cap I| = 1$.
   - $\mathcal{A} \subseteq P(A)$.

The $v \in \mathcal{I} \cup \mathcal{A}$ are called vectors.

A tree γ can be derived in G iff there is one vector $v_\mathcal{I} \in \mathcal{I}$ with initial tree α and there are n vectors $v_1, \ldots, v_n \in \mathcal{I} \cup \mathcal{A}$ such that there are node addresses $p_1, \ldots, p_r$ with $\gamma = \alpha[p_1, \gamma_1][p_2, \gamma_2], \ldots, [p_r, \gamma_r]$ and $\{\gamma_1, \ldots, \gamma_r\} = \bigcup_{i=1}^{n} v_i \cup v_\mathcal{I} \setminus \{\alpha\}$.

2. *A Vector-MCTAG with Dominance Links (V-TAG) is a tuple $G = \langle N, T, I, A, O, C, \mathcal{I}, \mathcal{A}\rangle$ such that:*

   - *$G_{TAG} := \langle N, T, I, A, O, C\rangle$ is a TAG with adjunction constraints, and*
   - *$\mathcal{I}$ and $\mathcal{A}$ are sets of vectors as in a V-MCTAG equipped with additional dominance links: For each $v \in \mathcal{I} \cup \mathcal{A}$, the dominance links dom_v form a binary relation over the set of nodes in the trees in v such that: if $dom_v(\mu_1, \mu_2)$, then μ_1 is the foot node of an auxiliary tree.*

   The derivation is defined as in V-MCTAG except that the derived tree must satisfy the dominance constraints coming from the vectors involved in the derivation.

As an example consider again the analysis using non-local MCTAG with dominance links in Fig. 17. In V-TAG this analysis is also possible, even with NA constraints on all foot nodes. In V-TAG these trees account even for all scrambling permutations of (56) since simultaneity is not required. I.e., the auxiliary trees can be adjoined according to the order of the words from right to left.

(56) ... dass niemand [ihm [den Kühlschrank zu reparieren] zu versprechen] bereit ist

Derivations in V-TAG are non-local. In order to account for locality constraints for movement, Rambow introduces *integrity constraints*: A node in an elementary tree can have an integrity constraint. In this case, in the derived tree, this node must not be between the two nodes of a dominance link. On CP nodes there are for example integrity constraints to prevent a scrambled element from moving into a higher tensed clause.

Rambow has shown the following:

Proposition 26 *Lexicalized V-TAGs with integrity constraints are polynomially parsable.*

Exercise 20 *What are the sentences generated by the MCTAG with dominance links from Fig. 17 if this MCTAG is considered being a V-TAG and if, additionally, all foot nodes carry NA constraints?*

10.3 SN-MCTAG

Tree-local MCTAG with shared nodes (SN-MCTAG), Kallmeyer (2005).

Aim: define a local TAG variant for scrambling, such that locality constraints for movement need not be obtained by some extra mechanism.

Idea: MCTAG with tree-local derivations. But: assume a node μ in some elementary tree γ such that β adjoins to μ. In *standard TAG*, in the derived tree, the root node of β belongs *only to* β, i.e., further adjunctions at that node are adjunctions to β. In contrast to this, in SN-MCTAG, the node in question is considered as *being shared by γ and β* since it is a merging or unification of μ (belonging to γ) and of the root of β. Further adjunctions at that node can be considered being either adjunctions to γ or adjunctions to β.

Definition of SN-MCTAG via properties of underlying TAG derivation tree.

See for example the derivation of (51) in Fig. 18. After the first derivation step, the root node in the derived tree belongs to both, *reparieren* and *verspricht*. Consequently, arguments of both can be adjoined here. Similar, after the second derivation step.

Consider the SN-MCTAG from Fig. 19. This grammar generates all scrambling variations of (56).

Restricted SN-MCTAG (RSN-MCTAG): at least one tree from each elementary tree set must be attached to a node where nothing else has been attached so far.

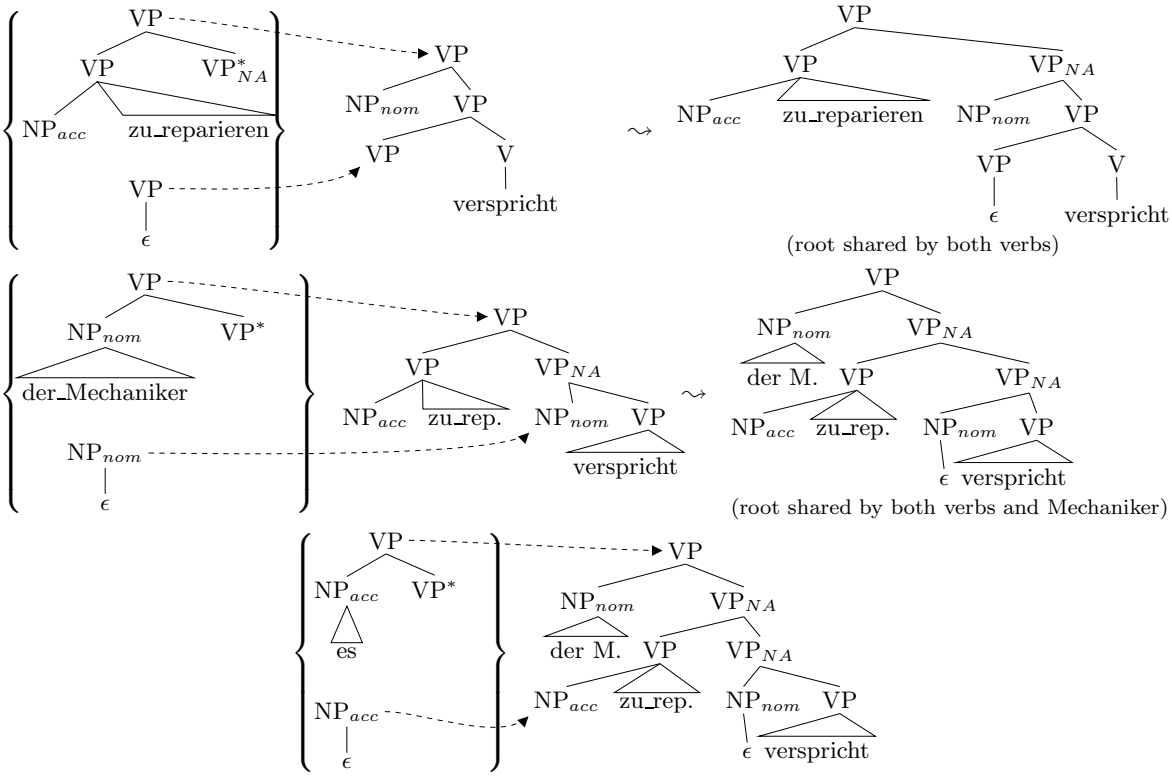

Figure 18: SN-MCTAG derivation of (51) *dass es der Mechaniker zu reparieren verspricht*

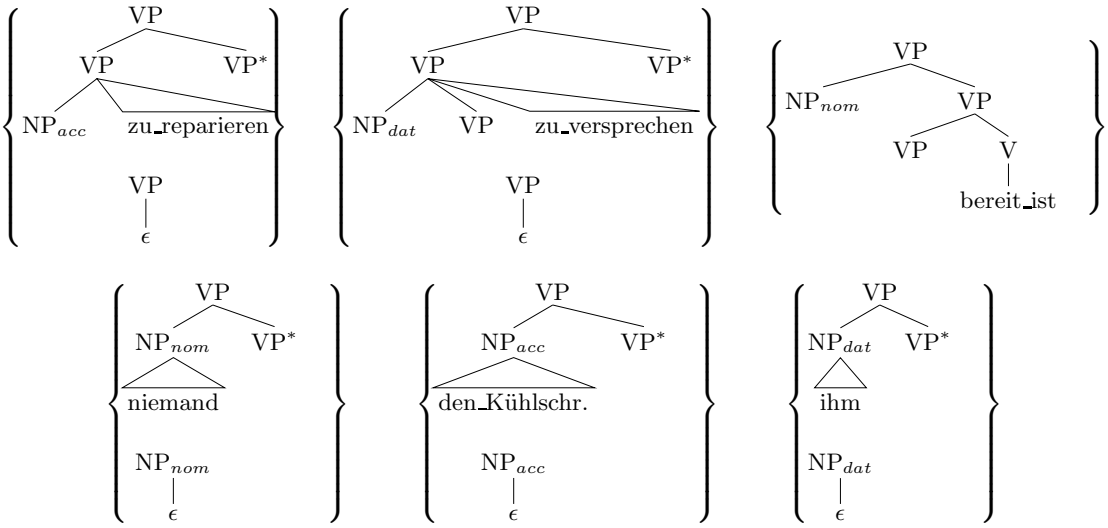

Figure 19: SN-MCTAG: analysis of (56) and scrambling variations

RSN-MCTAG of arity k: the contribution of an elementary γ (including everything attached to γ cannot be separated into more than k discontinuous parts.

Proposition 27 *(Kallmeyer 2005)*

For a given k, the string languages generated by RSN-MCTAGs of arity k are mildly context-sensitive.

Crucial: arity is variable, i.e., can be fixed arbitrarily large.