Syntax in Chains

Marcus Kracht *
II. Mathematisches Institut
Freie Universität Berlin
Arnimallee 3
D–14195 Berlin
kracht@math.fu-berlin.de

December 1, 2000

Abstract

In transformational grammar the notion of a chain has been central ever since their introduction in the early 80's. However, an insightful theory of chains has hitherto been missing. This paper develops such a theory of chains. Though it is applicable to virtually all chains, we shall focus on movement induced chains. It will become apparent that chains are far from innocuous. A proper formulation of the structures and algorithms involved is quite a demanding task. Furthermore, we shall show that it is possible to define structures in which the notion of a chain coincides with that of a constituent, so that the notion of a chain becomes redundant, generally without making the theory more complicated. These structures avoid some of the problems that beset the standard structures (such as unbound traces created by remnant movement).

1 Introduction

This paper is the final part of a trilogy, which started with [13] and continued with [15]. This trilogy is concerned with three basic notions of GB [1]: nearness,

*This paper is the result of several long discussions with Hans–Martin Gärtner and Jens Michaelis. I am especially indebted to Hans–Martin for his broad knowledge of transformation grammar and general linguistics as well as for his never ending patience in ploughing through pages of mathematical detail. Furthermore, thanks to Tom Cornell, Fritz Hamm, Hans–Peter Kolb, Manfred Krifka, Uwe Mönnich and Markus Steinbach for useful discussions. Finally, thanks to the audience of the GLOW workshop 'Technical Aspects of Movement', in particular Mishy Brody and Ed Stabler, and an anonymous referee for L & P for critical assessment.

[1] The present paper deals with notions of transformational grammar since the 80's. It comes in various incarnations, the most important are *GB* (the theory of Govenment and Binding)

adjunction and chains. The first paper dealt with nearness, and the second with adjunction. Here we will be concerned with perhaps *the* most important concept of transformational grammar, namely *chains*.

We shall introduce the notion of *chains* and of *copy chain structures* (CCSs). The latter is a pair $\langle\mathfrak{T}, C\rangle$ where $\mathfrak{T}$ is a tree and C a set of copy chains over $\mathfrak{T}$ satisfying certain conditions. Furthermore, we shall show how to define transformations based on CCSs. Starting with these structures, we shall consider two alternative structures: *trace chain structures* (TCSs), the most commonly used ones, and *multidominance structures* (MDSs). Central to the theory of GB and MP is the idea that a chain is a unit. This idea can be found explicitly in [4], and again in [5]. However, in MP the device of copying has been (re-)introduced, which creates exponentially many copies of a single constituent. The ensuing so–called *explosion problem* is discussed in Gärtner [11]. (See also Gärtner [9] for a thorough discussion of the MP.) Basically, each copy introduces copies of unchecked features, which must somehow be taken care of once that feature has been checked. Even if feasible, the administration of this process needs enormous resources. It was proposed in [11] to introduce an equivalent of phrase–linking grammars to handle it. This motivates the introduction of multidominance structures (MDS). In an MDS, a chain is a single object, in fact, it is the same as a constituent. MDSs are not trees but directed acyclic graphs. The conditions under which an acyclic graph is an MDS are actually very simple. In fact, we shall establish that MDSs contain the same information and are easily translatable into ordinary trace chain structures used in GB. So, talk of trace chain structures is as good as talk of MDSs. The present results also show that many facts about the derivation can be read off very easily from the corresponding multidominance structures, for example, the length of the derivation (which equals the number of chain links).

The paper is organized as follows. In Section 2 we shall introduce the notion of a *chain*, more precisely the notion of a *copy chain* and a *trace chain*. Although trace chains are by far the most popular chains, the main part of the paper will study copy chains, since they are technically easier to handle. In Section 3 we shall define the notion of a CCS. In Section 4 it is shown that a pair $K = \langle\mathfrak{T}, C\rangle$ is obtained by means of movement transformations from $\langle\mathfrak{U}, D\rangle$, with D containing only trivial chains, iff K is a CCS. This result is proved by introducing a *blocking order*. Given the blocking order, all derivations of K can easily be found. Section 5 introduces the *pre–multidominance structures* (preMDSs). These are directed acyclic graphs satisfying a certain condition. The correspondence with preCCSs (these are CCSs in which the surface elements of the chains are not specified) and preMDSs are studied. For each preCCS there is exactly one preMDS, but several preCCSs can define the same preMDS. In Section 6 we shall discuss

and *MP* (the Minimalist Program. The differences between them are irrelevant for the purpose of this paper.

the effect of two specific derivational constraints, namely *Freeze* and *Shortest Steps*. In Section 7 the results are extended to CCSs and MDSs. This brings no further surprises and actually allows to simplify the Synchronisation condition on derivations. In Section 8 we will study the reduction of trace chains to copy chains. Finally, Sections 9 and 10 deal with order and adjunction in the presence of chains.

There has been some work in the past concerning the formalization and axiomatization of GB type theories (see for example Stabler [23] and Rogers [22]). However, there is so far little in the sense of a real theory of chains and chain structures that is more than a mere formal rendering of informally stated definitions and results. Rather, what is needed is a theory which presents new results and simplifications, and by which the subject matter can be understood more easily; and which provides directions into which it can be generalized fruitfully. We hope to present such a theory of chains. Moreover, we want to show that the structures of standard transformational grammar correspond quite closely to some alternative ideas about linguistic structures, which can be traced back to the phrase linking grammars by Peters and Ritchie (which remained samistad work, see [2]) and perhaps even earlier ideas by McCawley and the American structuralists. The main result in this direction is the abovementioned Proposition 78, showing that trace chain structures and multidominance structures contain *the same information*. The algorithms that mediate between them are very simple. In Blevins [2] one can actually read about various different proposals concerning the basic structures. For example, McCawley [17] has already proposed to allow for multidomination. Blevins himself argues for allowing discontinuity. A totally different approach to chains, coming from the rapprochement between categorial grammar, linear logic and the Minimalist Program (through the notion of resource sensitivity) can be found in the work of Cornell [8], [6], and [7], and Lecomte [16]. Moreover, the papers by Cornell also discuss the relation between derivational and representational aspects of the syntactic theory.

2 Chains

Let us fix some basic notation and terminology. In this section, proofs are generally a direct matter of verification and omitted. We shall also say that although in linguistics all structures are labelled, for the purposes of this paper, labelling is irrelevant and suppressed. It can of course be introduced if needed. If for example, subjacency and other nearness conditions come into play, labelling will be absolutely necessary.

We assume that the reader is acquainted with the usual notions of trees, and such linguistic entities as chains. Typically, a tree is defined as a pair $\langle T, < \rangle$ such that (a) $<$ is irreflexive and transitive, (b) there is an $r \in T$ (the **root**) such that $x < r$ for every $x \neq r$ and (c) if $y, z > x$ then either $y < z$ or $y = z$ or $y > z$. We write $x \prec y$ if $x < y$ and there is no z such that $x < z < y$. If $x > y$, we say that x **properly dominates** y, and if $x \succ y$ we say that x **immediately dominates** y. Moreover, x and y are called **comparable** if $x < y$, $x = y$ or $x > y$. For several reasons it is better in this context to define trees starting with $\prec$. (See the discussion in Section 5.) A pair $\mathfrak{T} = \langle T, \prec \rangle$, where T is a nonempty finite set and $\prec \subseteq T^2$, is a **tree** if (A) there is an $r \in T$ (the root) such that $r \prec x$ for no $x \in T$, (B) for every $x \neq r$ there is exactly one y such that $x \prec y$, (C) there is no sequence $x_0 \prec x_1 \prec \ldots \prec x_{n-1} \prec x_n = x_0$, $n > 0$. [2] Now put $<:=\prec^+$, the transitive closure of $\prec$. Then (C) can be replaced by (C$'$) $<$ is irreflexive. (This means that for *no* x: $x < x$.) We denote by $\mu(x)$, x not the root, the unique y such that $x \prec y$. $\mu(r)$ is undefined. We put $\downarrow x := \{y : y \leq x\}$. The **depth** of x, $dp(x)$, is defined as follows.

$$\begin{array}{rcl} dp(r) & := & 0 \\ dp(x) & := & dp(\mu(x)) + 1 \end{array}$$

(Here, $x \neq r$.) The following definition of c–command is standard in GB and MP and can be found in any textbook. (Notice that we do not yet consider adjunction.)

Definition 1 *x **c–commands** y if $x = r$ or $x \neq r$ and $\mu(x) \geq y$.*

$\mathfrak{C}$ is a **constituent of** $\mathfrak{T}$ if $\mathfrak{C} = \langle \downarrow x, \prec \cap (\downarrow x)^2 \rangle$ for some $x \in T$. We say that x **generates** $\mathfrak{C}$. We also write $m(\mathfrak{C}) := \downarrow x$. In general, we will not distinguish between $\downarrow x$ and the constituent $\mathfrak{C}$. For example, if $y \in m(\mathfrak{C})$, we write $y \in \mathfrak{C}$. We also write $\mathfrak{C} \subseteq \mathfrak{D}$ in place of $m(\mathfrak{C}) \subseteq m(\mathfrak{D})$, and say that $\mathfrak{D}$ **covers** $\mathfrak{C}$. If x generates $\mathfrak{C}$ and y generates $\mathfrak{D}$ then $\mathfrak{C} \subseteq \mathfrak{D}$ iff $x \leq y$. If $\downarrow x$ and $\downarrow y$ are constituents, then $\downarrow x$ **c–commands** $\downarrow y$ iff x c–commands y. We will standardly work with another basic relation, a relation which we call **ac–command** and which is used for example in Kayne [12].

Definition 2 *Let $\mathfrak{T}$ be a tree and x, y nodes. x **ac–commands** y if (1) x c–commands y but y does not c–command x and (2) x and y are incomparable. $\downarrow x$ ac–commands $\downarrow y$ iff x ac–commands y.*

Lemma 3 *x ac–commands y iff (a) x and y are incomparable and (b) $\mu(x) > \mu(y)$.*

Proposition 4 *The relation of ac–command is transitive.*

C–command is in general not transitive; however, note the following.

Lemma 5 *Suppose that x and y c–command z. Then either x c–commands y or y c–commands x.*

[2] We shall generally start counting with 0. Hence the locution '$i < n$' means '$i \in \{0, 1, \ldots, n-1\}$'.

In what is to follow we will present two definitions of chains, of which the last is the official one, but the other one is used whenever there is no risk of confusion.

Definition 6 *A* ***pre–chain in*** *(or* ***of****)* $\mathfrak{T}$ *is a nonempty set* $\mathbb{C}$ *of constituents of* $\mathfrak{T}$ *which is linearly ordered by ac–command. A* ***chain*** *is a pair* $\Delta = \langle \mathfrak{C}, \mathbb{C} \rangle$*, where* $\mathbb{C}$ *is a pre–chain and* $\mathfrak{C} \in \mathbb{C}$*.* $\mathfrak{C}$ *is* ***higher*** *than* $\mathfrak{D}$ *(and* $\mathfrak{D}$ ***lower*** *than* $\mathfrak{C}$*),* $\mathfrak{C}, \mathfrak{D} \in \mathbb{C}$*, if* $\mathfrak{C}$ *ac–commands* $\mathfrak{D}$*. A* ***member*** *of* Δ *is an element of* $\mathbb{C}$*. We also write* $\mathfrak{C} \in \Delta$ *to say that* $\mathfrak{C}$ *is a member of* Δ*.*

So, a chain is a pre–chain with a distinguished member. The difference between a pre–chain and a chain is in most cases marginal, and we will therefore speak of chains even when we mean pre–chains. Moreover, we will make the following convention. The pre–chain $\mathbb{C}$ will on occasion be identified with the chain $\langle \mathfrak{C}, \mathbb{C} \rangle$, where $\mathfrak{C}$ is the highest element of $\mathbb{C}$.

Before we discuss the implications of this definition, let us note that we do *not* require two members of a chain to be isomorphic. For example, the typical chains of linguistic literature are such that one element is empty, and the chain contains only one nonempty element. Moreover, in [20] chains may contain several overt elements. The linearity requirement is not as innocuous as it appears. It seems to exclude, for example, the analysis of parasitic gaps, where an antecedent is related to two empty elements (though it moved only from one position), so that these three elements are actually not linearly ordered by ac–command.

(1) *[Which books]$_1$ did you [[file t_1] [without reading e_1]]?*

For example, in (1), *which books* originates as the direct object of *file*. By moving to the specifier of CP it is in a position to bind the gap e_1, which is in the VP–adjunct *without reading* e_1. Now, *which books* c–commands both t_1 and e_1, but t_1 c–commands neither *which books* nor e_1, and e_1 c–commands neither *which boooks* nor t_1. We are faced with three possibilities.

1. Either we give up the requirement that chains are linearly ordered by ac–command or
2. we assume that either e_1 or t_1 is not contained in a chain with *which books* or
3. we assume that there are two chains, one formed by *which books* and t_1 and one by *which books* and e_1.

If we adhere to the metaphor of movement, we must discard the first alternative. (However, notice that across the board extraction is a serious problem, but also for the movement based approaches in general.) In fact, for the same reason we must discard the third alternative, too. However, we may say that *which books* does engage in two *types* of chains, namely a movement chain and a binding chain. As we are mainly concerned here with movement chains, we will have nothing to say about other kinds of chains in the sequel.

The ac–command relation on constituents is irreflexive but never linear, except in trivial cases. Therefore, the linearity requirement adds a constraint on the sets of constituents that can form a pre–chain. For example, a chain may not contain two constituents $\mathfrak{C}$ and $\mathfrak{D}$ which are sisters. For notice that in a linear order there are no two distinct elements x and y such that $x < y$ and $y < x$.

In a chain Δ, there exists a least element with respect to c–command, and this element is called the **foot** and denoted by Δ_f. Likewise, there exists a greatest element, and it is called the **head** and denoted by Δ_h. Finally, if $\Delta = \langle \mathfrak{C}, \mathbb{C} \rangle$, we put $\Delta_d := \mathfrak{C}$. Δ_d is the called the **designated element** of the chain. The surface elements [3] are designated elements, but the converse generally does not hold. A member which ac–commands the surface element is called an **LF–element** of that chain. A chain is **trivial** if it has only one element. In this case $\Delta_s = \Delta_d = \Delta_f$. A chain is an **S–chain** if $\Delta_d = \Delta_h$, otherwise it is an **LF–chain**.

A **copy–chain** is a chain in which all members are isomorphic as trees. A **trace–chain** is a chain in which only the head is not a trace. We will assume that a trace is simply a trivial tree, consisting of only one node. In labelled trees this node carries a nonterminal label (for example, an NP–trace is simply a one node tree whose only node has label NP). Generally, we will take it that the various empty elements give rise to distinct trees (for example, *pro*, *PRO*, *e* etc). Copy–chains encode the notion of movement as copying, while trace–chains encode the notion of movement as a sequence of copy–and–delete (where by deletion we mean not the marking by a special label as is nowadays fashionable, but the replacement by a trivial tree whose yield is a phonetically empty element).

A pre–chain can be put into the form $\{\mathfrak{C}_i : i < n\}$, where $\mathfrak{C}_i$ c–commands $\mathfrak{C}_j$ iff $i > j$. We call the pair $\langle \mathfrak{C}_{i+1}, \mathfrak{C}_i \rangle$ a **chain–link**. We may symbolize the chain also as follows. We list the elements in the order of c–command, putting an arrow under the distinguished element. Notice that by the numbering convention the numbers go down from left to right (= from greatest to lowest).

$$\begin{array}{cccccccc} \mathfrak{C}_{n-1} & \mathfrak{C}_{n-2} & \ldots & \mathfrak{C}_k & & \ldots & \mathfrak{C}_1 & \mathfrak{C}_0 \\ & & & \uparrow & & & & \end{array}$$

Given a subtree $\mathfrak{U}$ of $\mathfrak{T}$ and a (pre–)chain Δ, Δ gives rise to a (pre–)chain $\Delta \upharpoonright \mathfrak{U}$ in $\mathfrak{U}$. This we call the **residue** of the chain in $\mathfrak{U}$. As it turns out, however, this is not always well–defined.

Start with the following definition. Let $\Delta = \langle \mathfrak{C}_k, \{\mathfrak{C}_j : j < n\} \rangle$ be a chain. Put

$$m(\Delta) := \bigcup_{j<n} m(\mathfrak{C}_j) \ .$$

We will look at ways in which Δ intersects with $\mathfrak{U}$. $\mathfrak{U}$ can have nonempty intersection with exactly one or with several members of $m(\Delta)$. It is easy to see that

[3]These are to be defined later. Intuitively, they are the elements that we see at surface structure.

in the second case the intersection with $\mathfrak{C}_j$, $i < n$, is either empty or $\mathfrak{C}_j$. We investigate the case where there is exactly one $\mathfrak{C}_j$ such that $\Delta \cap \mathfrak{C}_j \neq \emptyset$. Then either (1) $\mathfrak{C}_j \cap \mathfrak{U} \subsetneq \mathfrak{C}_j$, or (2) $\mathfrak{C}_j = \mathfrak{U}$. In Case (1) the residue, if defined, would equal $\langle \mathfrak{U}, \{\mathfrak{U}\}\rangle$. It would be harmless to allow the residue to be defined even in this case, but we opt against that for conceptual clarity. Therefore, in Case (1), $\Delta \upharpoonright \mathfrak{U}$ is undefined. In Case (2), the residue is simply the one–membered (pre–)chain $\langle \mathfrak{C}_j, \{\mathfrak{C}_j\}\rangle$.

So, let us now assume $\mathfrak{U}$ properly contains a member of Δ. Then we may note the following.

$$\begin{array}{cccccccccc} \Delta & : & \mathfrak{C}_{n-1} & \mathfrak{C}_{n-2} & \ldots & \mathfrak{C}_k & \ldots & \mathfrak{C}_1 & \mathfrak{C}_0 \\ & & & & & \uparrow & & & \end{array}$$

If $\mathfrak{C}_j \cap \mathfrak{U} = \varnothing$ and $i > j$, then also $\mathfrak{C}_i \cap \mathfrak{U} = \varnothing$. The following may be also be noted. If $j > i$ and $\mathfrak{C}_j \cap \mathfrak{U} \neq \varnothing$ then not only is $\mathfrak{C}_i \cap \mathfrak{U} \neq \varnothing$, but also $\mathfrak{C}_j \cap \mathfrak{U}$ ac–commands $\mathfrak{C}_i \cap \mathfrak{U}$. This follows from the insensitivity of c–command with respect to embedding. Hence, the residue of a pre–chain is again a pre–chain. Let p be the largest number such that $\mathfrak{C}_p \cap \mathfrak{U} \neq \varnothing$. Then, by our assumptions, $\mathfrak{C}_q \cap \mathfrak{U} = \mathfrak{C}_q$, for all $q \leq p$. Put $\mathbb{C} := \{\mathfrak{C}_q : q < p + 1\}$. This defines the residue of the pre–chain $\mathbb{C}$. We may symbolize the residue of the pre–chain by inserting in brackets those members whose intersection with $\mathfrak{U}$ is nonempty. Then if the residue is defined and contains at least one chain member we get the following result.

$$\mathfrak{C}_{n-1} \quad \mathfrak{C}_{n-2} \quad \ldots \quad (\mathfrak{C}_p \quad \ldots \quad \mathfrak{C}_1 \quad \mathfrak{C}_0)$$

Now we turn to the definition of the residue of the chain Δ. Two cases need to be distinguished. The first is $p \geq k$. It looks as follows:

$$\begin{array}{ccccccc} \mathfrak{C}_{n-1} & \ldots & (\mathfrak{C}_p & \ldots & \mathfrak{C}_k & \ldots & \mathfrak{C}_1 \quad \mathfrak{C}_0) \\ & & & & \uparrow & & \end{array}$$

In this case $\Delta \upharpoonright \mathfrak{U} := \langle \mathfrak{C}_k, \mathbb{C} \upharpoonright \mathfrak{U}\rangle$. It may be that we have $p = k$.

$$\begin{array}{cccccc} \mathfrak{C}_{n-1} & \mathfrak{C}_{n-2} & \ldots & (\mathfrak{C}_k & \ldots & \mathfrak{C}_1 \quad \mathfrak{C}_0) \\ & & & \uparrow & & \end{array}$$

The second case is $p < k$.

$$\begin{array}{ccccccc} \mathfrak{C}_{n-1} & \mathfrak{C}_{n-2} & \ldots & \mathfrak{C}_k & \ldots & (\mathfrak{C}_p \ldots & \mathfrak{C}_0) \\ & & & \uparrow & & & \end{array}$$

Here, the intersection cuts out the distinguished element of Δ. In this case, our convention on pre–chains comes into play. Remember that if we like to turn a pre–chain into a chain then we take as default the highest member of the pre–chain as the distinguished element. So, here the result is

$$\begin{array}{ccccccc} \mathfrak{C}_{n-1} & \mathfrak{C}_{n-2} & \ldots & \mathfrak{C}_k & \ldots & (\mathfrak{C}_p & \ldots \quad \mathfrak{C}_0) \\ & & & & & \uparrow & \end{array}$$

Hence we put $\Delta \upharpoonright \mathfrak{U} := \langle \mathfrak{C}_p, \mathbb{C} \upharpoonright \mathfrak{U}\rangle$.

Definition 7 *Let $\mathbb{C}$ be a pre–chain over $\mathfrak{T}$ and $\mathfrak{U}$ a subconstituent of $\mathfrak{T}$. The* ***residue*** *of $\mathbb{C}$ in $\mathfrak{U}$, $\mathbb{C} \upharpoonright \mathfrak{U}$, is defined as follows. (1) $\mathbb{C} \upharpoonright \mathfrak{U}$ is undefined if $\mathfrak{U}$ is a proper subconstituent of some member of $\mathbb{C}$. (2) $\mathbb{C} \upharpoonright \mathfrak{U} := \{\mathfrak{C} \cap \mathfrak{U} : \mathfrak{C} \cap \mathfrak{U} \neq \varnothing\}$ else. Let $\Delta = \langle \mathfrak{C}, \mathbb{C} \rangle$ be a chain of $\mathfrak{T}$. The* ***residue*** *of Δ in $\mathfrak{U}$, $\Delta \upharpoonright \mathfrak{U}$, is defined only when $\mathbb{C} \upharpoonright \mathfrak{U}$ is defined and then $\Delta \upharpoonright \mathfrak{U} := \langle \mathfrak{X}, \mathbb{C} \upharpoonright \mathfrak{U} \rangle$, where $\mathfrak{X}$ is determined as follows. (1) If $\mathfrak{C} \subseteq \mathfrak{U}$ then $\mathfrak{X} := \mathfrak{C}$. (2) If $\mathfrak{C} \nsubseteq \mathfrak{U}$ then $\mathfrak{X}$ is the highest member of $\mathbb{C} \upharpoonright \mathfrak{U}$.*

In this way, chains can be relativized to subtrees. Now, unfortunately, this notion of a chain is not the one that is really needed. Rather it is the following.

Definition 8 *Let $\mathfrak{T}$ be a tree. A* ***copy–chain link**** *of $\mathfrak{T}$ is a triple $\langle \mathfrak{C}, \varphi, \mathfrak{D} \rangle$ where $\mathfrak{C}$ and $\mathfrak{D}$ are constituents of $\mathfrak{T}$ such that $\mathfrak{C}$ ac–commands $\mathfrak{D}$ and φ an isomorphism from $\mathfrak{D}$ to $\mathfrak{C}$. A* ***pre–copy–chain**** *is a pair $\langle \Delta, \Phi \rangle$, where $\Delta = \{\mathfrak{C}_i : i < n\}$ is a pre–chain (in the sense of Definition 6) and $\Phi = \{\varphi_{ij} : i, j < n\}$ a set of isomorphisms $\varphi_{ij} : \mathfrak{C}_i \to \mathfrak{C}_j$ satisfying (a) φ_{ii} is the identity on $\mathfrak{C}_i$, (b) $\varphi_{ij} = \varphi_{ik} \circ \varphi_{kj}$ for all $i, j, k < n$. Finally, a* ***copy–chain**** *is a triple $\langle \Delta, \Phi, \mathfrak{C} \rangle$ such that $\langle \Delta, \Phi \rangle$ is a pre–copy–chain* and $\mathfrak{C}$ a member of Δ.* [4]

Alternatively, we may think of a pre–chain* as a sequence of chain links* such that if $\langle \mathfrak{F}, \psi, \mathfrak{E} \rangle$ follows $\langle \mathfrak{D}, \varphi, \mathfrak{C} \rangle$ in the sequence, then $\mathfrak{D} = \mathfrak{E}$. These notions are equivalent and used intrechangeably. If we have a pre–chain* $\langle \Delta, \Phi \rangle$ we can take $\Delta = \{\mathfrak{C}_i : i < n\}$, and the chain–links* are $\langle \mathfrak{C}_{i+1}, \varphi_{i,i+1}, \mathfrak{C}_i \rangle$. Conversely, given the chain–links*, we may enumerate the constituents as $\mathfrak{C}_i$, $i < n$, and choose the chain links* as above. Now put for $i > j$:

$$\varphi_{j,i} := \varphi_{j,j+1} \circ \varphi_{j+1,j+2} \circ \ldots \circ \varphi_{i-1,i}$$

Next $\varphi_{i,j} := \varphi_{i,j}^{-1}$, and $\varphi_{i,i}$ the identity on $\mathfrak{C}_i$. Now, put $\Phi := \{\varphi_{ij} : i, j < n\}$. Let Δ be a (pre–)chain*. We define $\approx_\Delta$ by $x \approx_\Delta y$ iff (ap1) $x, y \notin m(\Delta)$ and $x = y$ or (ap2) $x, y \in m(\Delta)$ and $x = \varphi(y)$ for some $\varphi \in \Phi$. This is an equivalence relation, and $x \approx_\Delta y$ iff Δ establishes a correspondence between $\downarrow x$ and $\downarrow y$.

The reader may wonder why it is necessary to have the isomorphisms in the definition of chains at all. For first of all we know that for two ordered binary branching trees there exists at most one isomorphism from one to the other. Second, for all that we required up to now, only the fact that the members of the chain are isomorphic is needed. Third, it is precisely the question whether in linguistic structures containing a chain $\Delta = \langle \mathfrak{C}, \mathbb{C} \rangle$ we know which nodes of the members of the chain are related to each other. As chains ideally form a

[4] To those readers for whom this is overly complicated we can only give the advice to check the section for the need of it. It order not to get confused with terminology, we will employ the following practice. For every technical term T there is a technical term $T^\star$ (typically involving talk of isomorphisms), which is in a sense more precise than T. Whenever we use $T^\star$ we will use the term T in its full technical detail and complexity. Otherwise, when this plays no essential role, we will use T rather than $T^\star$.

unit, with multiple 'stages' of the same constituent representing different copies of the same element (which gets spelled out only once) this should not be part of the representation. Moreover, knowing which node x_i in some member $\mathfrak{C}_i$ is the counterpart of which node x_j in $\mathfrak{C}_j$ is tantamount to having another chain, consisting of the constituents $\{\downarrow x_i : i < n\}$. (We will say that such a chain is a **parallel subchain** of Δ.) Parallel chains have never been assumed explicitly, even though implicitly some use has been made of them. A particular case is the work of Nunes (see [19]). For a clear discussion of the problems that this particular proposal raises see [10]. In a brief discussion on page 265 of [5], Chomsky considers the possibility that the computational system apart from looking at the proper chains also takes a glimpse at the resulting subchains. In MP the subchains are needed, for example, to handle pied–piping. However, as we shall see, the installment of such chains leads to a serious distortion of the theory. Movement cannot be formulated as an extension of single chains (and even less as the movement of a single element) but leads to a multitude of parallel 'movements' even in the most simple cases. So, not only do we not need parallel chains (in the ordered case), we also do not want them, for conceptual reasons.

However, there are reasons to assume the presence of the isomorphisms, a formal one and a linguistic one. The formal one is that we will note that the constructions of compression and decompression of Section 5 depend on the chosen isomorphisms. Second, there are processes that require the construction of such isomorphisms. We give an example. In [3] a new analysis of discontinuous constituents is given, in particular of split DPs in German. [5] The derivation of (2) proceeds as follows.

(2) *[Bücher gelesen] habe ich nicht viele.*
Books read have I not many.
I have not read many books.

The object *nicht viele Bücher* is generated inside the VP. The constituent part *nicht viele* carries a focus feature, which induces scrambling of *nicht viele Bücher*. This transformation is not realized as copy–and–delete, but as copy movement. Subsequently to this, a so–called partial deletion operation applies. It deletes parts of the upper copy and the complementary part of the lower copy. [6]

[5] This is by far not the only analysis of split–DPs (see Riemsdijk [21] against a derivational analysis), but we are not interested in the issue whether this analysis is correct. Rather, our question is: if it is correct, then what does that say about the theory.

[6] Deletion is marked by underlining. What is deleted depends on the focus structure. Ćavar and Fanselow, working within the Minimalist Program, assume that words are marked ±focus, and that the assignment of the focus feature steers the deletion. If this is so, there is an easy way to define the partial deletion: delete in the upper part everything that is −focus, and in the lower part everything that is +focus. In this way, talk of isomorphisms can be avoided.

(3a) *habe ich [[nicht viele Bücher] gelesen].*
(3b) *habe ich [nicht viele Bücher] [[nicht viele Bücher] gelesen.*
(3c) *habe ich [nicht viele Bücher] [[nicht viele Bücher] gelesen.*

The last step is the topicalization of the lower copy of the DP, and its complete deletion.

(3d) *[nicht viele Bücher gelesen] habe ich [nicht viele Bücher]*
[[nicht viele Bücher] gelesen.]

Interesting for us is the the step from (3b) to (3c). This is called *partial deletion*, since neither the upper part nor the lower part are completely deleted. Now, assuming that partial deletion is the result of applying a simple algorithm of PF, we need to ask what the input for this algorithm is. If it is just the structure without the marked deletions, then PF needs to compute for each of the copies which element of the upper copy is a counterpart of which element of the lower copy. This is tantamount to the construction of an isomorphism between the two constituents. With such an isomorphism given, however, partial deletion is easy to formulate. Namely, delete for each element x of the lower copy either x itself or $\varphi(x)$, where φ is the isomorphism.

3 Copy Chain Structures

In this section we will axiomatize what we call *Copy Chain Structures*. In the next section we will show that this is exactly the class of structures that can be obtained from trees by successive application of copy movement. The axiomatization of copy chain structures is clearly made with respect to capturing exactly this class of structures, since we are aiming — ultimately — at a purely representational theory of GB. Clearly, the structures of GB theory are derived from trees (D–structures) by successive application of Move–α. Before it is possible for us to approach the full theory, we will start with some simpler cases. First, we will deal with copy movement and later with the standard movement; second, we will first investigate the structures without looking at the designated elements.

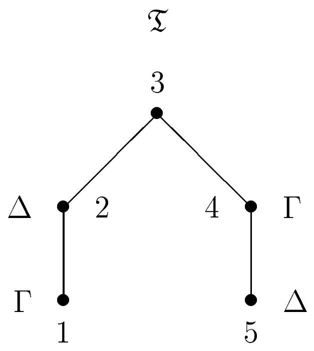
Definition 9 *A* ***(pre–)copy chain tree ((pre)CCT)*** *is a pair $\langle \mathfrak{T}, C\rangle$, where C is a set of (pre–)chains.*

If we look at the interaction of chains it is obvious that chains that share no node are chains which do not interact. Call Γ and Δ **disjoint** if $m(\Gamma) \cap m(\Delta) = \varnothing$.

Proposition 10 *If Γ and Δ are not disjoint, then there exist $\mathfrak{C} \in \Gamma$ and $\mathfrak{D} \in \Delta$ such that either $\mathfrak{C} \subseteq \mathfrak{D}$ or $\mathfrak{D} \subseteq \mathfrak{C}$.*

Proof. Let $x \in m(\Gamma) \cap m(\Delta)$. Then there exists a member of Γ, $\mathfrak{C}$, and a member of Δ, $\mathfrak{D}$, such that $x \in \mathfrak{C} \cap \mathfrak{D}$. It follows that $\mathfrak{C} \subseteq \mathfrak{D}$ or $\mathfrak{D} \subseteq \mathfrak{C}$. Q. E. D.

Figure 1: Violating *No Twisting*



Definition 11 *Let Γ and Δ be chains and $\mathfrak{X}$ a constituent. Δ* ***covers*** *$\mathfrak{X}$ if some member of Δ covers $\mathfrak{X}$. Δ* ***covers*** *Γ if Δ covers some member of Γ.*

We will formulate our first and main condition on structures with chains.

Constraint 1 (No Twisting) *In a chain structure, there are no two different chains Γ and Δ such that Γ covers Δ and Δ covers Γ.*

Figure 1 provides an example of two chains that violate the *No Twisting* constraint. (Each node carries a number for identification and a letter to denote the chain to which the constituent it heads belongs. Only nontrivial chains are annotated.) We take the pre–chains $\Gamma := \{\downarrow 1, \downarrow 4\}$ and $\Delta := \{\downarrow 2, \downarrow 5\}$. Then Γ covers Δ and Δ covers Γ.

We add another constraint, which falls into two parts. The first assures the existence of enough chains: every constituent is a member of some chain. Hence, movement will be done entirely by targeting chains. The other assumption is that a constituent is a member of only one chain.

Constraint 2 (Chain Existence) *Every constituent is a member of at least one chain.*

Constraint 3 (Chain Uniqueness) *Every constituent is a member of at most one chain.*

We require the existence of enough chains to make the algorithms smooth. Notice that it would be very awkward if the transformational component would have to target chains or constituents, depending on whether the constituents are part of a chain or not. Therefore, to assume that the transformations will always be able to manipulate chains we assume that in fact every constituent is contained in some chain. This however makes some cosmetics necessary in the formulation of movement (see the next section). That we require elements to be members of at most one chain excludes some types of movement, for example A–movement followed by $\overline{A}$–movement, since in GB the two chains are considered different. It is possible to generalize our theory so as to allow sequences of chains, but to allow for that right now would make matters complicated beyond need.

Chain Uniqueness is actually a consequence of *No Twisting*. For if $\mathfrak{C}$ is a member of both Γ and Δ then Δ and Γ cover each other. The pair $\langle \mathfrak{T}, \{\Gamma, \Delta\}\rangle$ does not satisfy neither *Chain Uniqueness* nor *Chain Existence*. If we put $\Theta := \{\downarrow 3\}$, then the structure $\langle \mathfrak{T}, \{\Gamma, \Delta, \Theta\}\rangle$ satisfies *Chain Existence* and *Chain Uniqueness* but not *No Twisting*. We will show that *Chain Uniqueness* implies *No Twisting* for copy–chains.

Proposition 12 *Let Γ and Δ be copy–chains. Then if Γ covers Δ and Δ covers Γ, all members of Γ and Δ are isomorphic. So if Γ and Δ are chains of a CCT satisfying Chain Uniqueness, Γ is identical to Δ.*

Proof. Let $\mathfrak{X}_i$ and $\mathfrak{X}_j$ be members of Γ, $\mathfrak{Y}_k$ and $\mathfrak{Y}_m$ members of Δ such that $\mathfrak{X}_i \supseteq \mathfrak{Y}_k$ and $\mathfrak{Y}_m \supseteq \mathfrak{X}_j$. Then $\mathfrak{X}_i$ has at least as many nodes as $\mathfrak{Y}_k$, and $\mathfrak{Y}_m$ has at least as many nodes as $\mathfrak{X}_j$. But $\mathfrak{X}_i$ and $\mathfrak{X}_j$ have the same number of nodes, and the same holds for $\mathfrak{Y}_k$ and $\mathfrak{Y}_m$. It follows that all four have the same number of nodes. Hence $\mathfrak{X}_i$ does not properly contain $\mathfrak{Y}_k$, and they are therefore identical. Hence, Γ and Δ consist of isomorphic members. This proves the first claim. The second is an immediate consequence. Q. E. D.

Chain Uniqueness implies that no constituent is moved (i. e. copied) twice. However, what may very well happen is that a chain member is moved as part of another constituent (piggy–backing). Depending on whether this element is the foot, an intermediate element or the head, different types of movement are generated.

Our next constraint concerns the case where we have two chains, Γ and Δ, where Γ covers Δ. In that case we have an element $\mathfrak{X}_i$ of Γ that covers an element $\mathfrak{Y}_j$ of Δ. It may also happen that $\mathfrak{X}_p$ of Γ covers some element $\mathfrak{Y}_q$ of Δ. If $i \neq p$ then also $j \neq q$ and we say that Δ is **doubly covered** by Γ and that Γ and Δ are (**partially**) **parallel**. (Here we assume that $\Gamma \neq \Delta$.)

Constraint 4 (No Parallelism) *No chain is doubly covered by another chain.*

We note that if we base our notion of c–command on trees rather than adjunction structures, chains cannot be doubly covered. For suppose that Γ covers Δ doubly, that is, $\mathfrak{X}_i$ and $\mathfrak{X}_j$ are members of Γ that cover $\mathfrak{Y}_p$ and $\mathfrak{Y}_q$, respectively. Then, by chain uniqueness, $\mathfrak{Y}_p \subsetneq \mathfrak{X}_i$. Suppose that x_i generates $\mathfrak{X}_i$ and y_p generates $\mathfrak{Y}_p$. Then $y_p < x_i$, and therefore the c–command domain of y_p is contained in $\mathfrak{X}_i$. Likewise, the c–command domain of y_q is contained in $\mathfrak{X}_j$. Since Δ is chain, y_p c–commands y_q or y_q c–commands y_p. From this it follows that $\mathfrak{X}_i = \mathfrak{X}_j$.

Proposition 13 *Let Γ and Δ be two distinct chains of a tree $\mathfrak{T}$. Then Δ does not cover Γ doubly.*

So, suppose that Γ covers Δ, and $\Gamma \neq \Delta$. Then there is exactly one member, say $\mathfrak{C}_i$, which is not disjoint with Δ. The set of points that are in $m(\Gamma) \cap m(\Delta)$ is therefore identical with $\mathfrak{C}_i \cap m(\Delta)$. Therefore, the intersection is nothing but the intersection with a constituent of the tree. We have seen earlier that this means that this set is a lower segment of Δ with respect to ac–command.

Proposition 14 *Let Γ and Δ be two distinct chains of $\mathfrak{T}$. Assume that Γ covers Δ. Then the members of Δ that are covered by Γ form a chain, which is extended by Δ. In particular, Γ covers Δ_f.*

This gives a good overview of the ways in which two chains can interact. Either Γ and Δ are disjoint, or Γ covers Δ or Δ covers Γ. If Γ covers Δ then there exists a unique member of Γ which is nondisjoint from $m(\Delta)$ and it covers a tail of Δ. There are no restrictions on the element of Γ which is nondisjoint with Δ.

The next postulate in this series is a little bit harder to motivate. It reflects an essential nature of the process of movement qua copying. Suppose namely that we have a member $\mathfrak{C}$ of a chain Γ that is not the foot. If we attempt a movement of a subconstituent $\mathfrak{D}$ of $\mathfrak{C}$ with landing site inside $\mathfrak{C}$, we will get a new copy of $\mathfrak{D}$ inside $\mathfrak{C}$ that has no equivalent or counterpart in the other chain members of Γ. Therefore, a non–foot may only engage in a chain if it covers the foot of that chain. This postulate is called

Constraint 5 (Liberation) *Let Γ and Δ be chains. Suppose that $\mathfrak{C}$ is a member of Γ that covers two members of Δ. Then $\mathfrak{C}$ is the foot of Γ.*

Look at the chain structure L in Figure 2. We have $\Gamma = \{\downarrow 5, \downarrow 11\}$ and $\Delta = \{\downarrow 8, \downarrow 10\}$. A structure containing Γ and Δ will violate *Liberation*, since Γ covers the entire chain Δ, but it is not the foot of Γ that covers Δ. Notice that Δ may consist of more elements, not visible in the figure. In the situation depicted here we have an occurrence of a movement inside a displaced constituent. It seems at first sight clear that we could have obtained the same tree (but with different chains) if we had first moved inside the foot of Γ and then raised it to 11 (see Figure 5). However, in the case of movement as copying we will not end up with the same structure (even though the difference is only with respect to empty material). In fact, the latter structure is the only one we can obtain at all, while the first one is illegitimate for the reason that it cannot be produced by copy–movement. To see this, look at the step prior to the formation of Δ and Γ. The chain structure, K_0, is shown in Figure 3. (Notice that substitution is defined a little bit differently, see Section 3.) There are no nontrivial chains. If we now move $\downarrow 5$ to form the chain Γ and then form the chain Δ in the head of Γ we get a structure almost like L, but the node 4 is missing (see Figure 4). This is so because 4 corresponds to the head of Δ, but the chain Θ has never been

Figure 2: The chain structure L

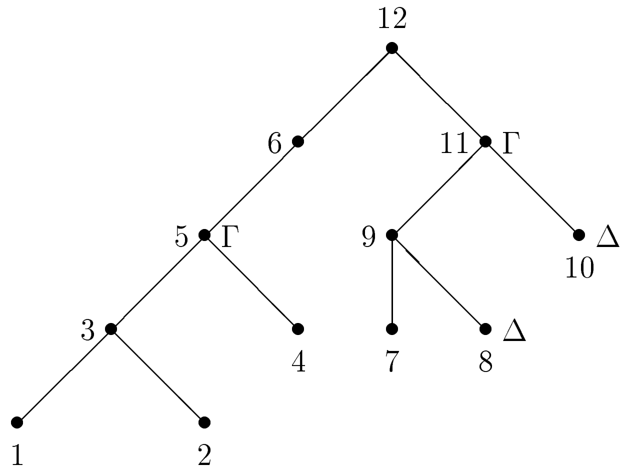


Figure 3: The chain structure K_0

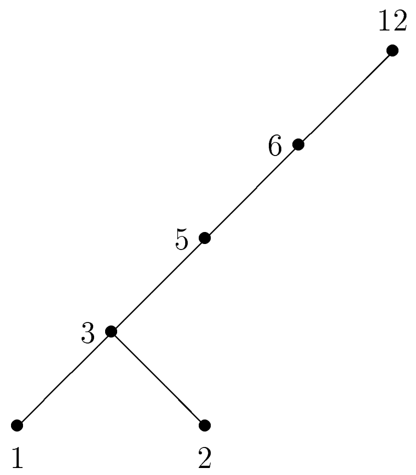
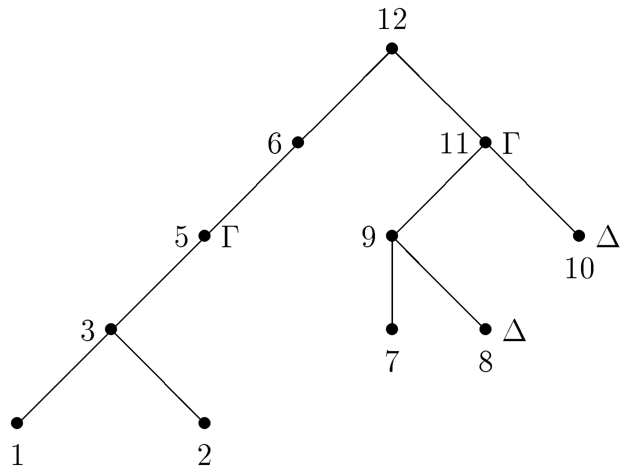


Figure 4: The derivation 1



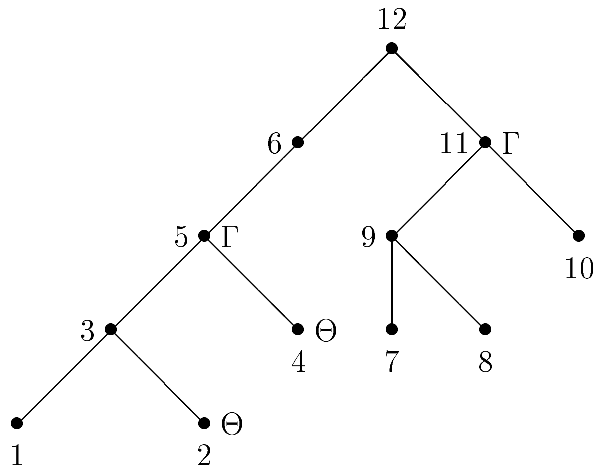
formed in $\downarrow 5$. On our view of chains, this is inadmissible, since the members of Γ are no longer isomorphic. [7]

Instead, we can form the chain $\{\downarrow 2, \downarrow 4\}$ and then move $\downarrow 5$ to form the chain Γ. Then we get the structure in Figure 5. It has the same tree but the chains are now $\{\downarrow 5, \downarrow 11\}$ and $\{\downarrow 2, \downarrow 4\}$. The chain Δ can simply not be formed, since the structure is already there. There is in this calculus no operation that allows to form a chain out of already existing constituents or chains: we can only form chains by extending a chain by *new material.* So, there is in particular also no operation of *chain composition*, which has sometimes been used in the GB–literature.

Recall that each chain* comes with a set of isomorphisms mapping the members onto each other. Given a chain* Γ, write $x \approx_\Gamma y$ if $x = y$ or there exists an isomorphism φ of Γ such that $\varphi(x) = y$. This is an equivalence relation. If $x < y$ and $\downarrow y$ is a member of Γ, then an isomorphism φ of Γ induces an isomorphism on the subconstituent $\downarrow x$ of $\downarrow y$, which we also denote by φ. Any composition of such maps is called a **composite isomorphism**. To distinguish the isomorphisms of chains* from the composite isomorphisms we call them **simple isomorphisms**. Furthermore, a **link map** is a simple isomorphism coming from a chain link*. So, if $\{\mathfrak{C}_i : i < n\}$ is the underlying pre–chain of the chain* such that $\mathfrak{C}_j$ ac–

[7]One may object, though, that our definition of chains is representational (and so nondynamic), while we are now concerned with the genesis of chains through copying. Hence we are mixing up two essentially different criteria: dynamic and static criteria. We have not much to say in defense, except that the theory we are developing here does not really allow to distinguish these viewpoints. The axiomatization of chain structures is clearly nondynamic, although developed with an eye to a derivational process.

Figure 5: The derivation 2



commands $\mathfrak{C}_i$ iff $j > i$ then the link maps are the maps $\varphi_{i,i+1} : \mathfrak{C}_i \to \mathfrak{C}_{i+1}$. There are some noteworthy facts about composite maps. First, any composite map is a composition of links maps or their inverses. Second, the composition $\psi \circ \chi$ is defined iff $im(\chi) \cap dom(\psi) \neq \varnothing$. This intersection is a constituent and so $dom(\psi \circ \chi)$ is a constituent (and isomorphic to $im(\psi \circ \chi)$). Let C be a set of chains*. Then we let $\approx_C$ be the equivalence relation generated by all the $\approx_\Gamma$. This means that $x \approx_C y$ iff there is a sequence Γ_i, $i < k$, of chains* and elements x_i, $i < k+1$, such that

$$x = x_0 \approx_{\Gamma_0} x_1 \approx_{\Gamma_1} x_2 \ldots \approx_{\Gamma_{k-1}} x_k = y$$

We have $x \approx_C y$ iff x and y are multiple copies of the same element within the whole derivation. Let $[x]_K := \{y : x \approx_C y\}$. We often write $[x]_C$ or simply $[x]$ in place of $[x]_K$, depending on need of precision. If $\mathfrak{C} = \downarrow x$ and $\mathfrak{D} = \downarrow y$ we write $\mathfrak{C} \approx_C \mathfrak{D}$ iff $x \approx_C y$. Here we have an important

Lemma 15 *$x \in [y]_C$ iff there is an composite isomorphism ψ mapping $\downarrow x$ onto $\downarrow y$.*

Proof. Assume $x \in [y]_C$. Then $x \approx_C y$. By definition there exist chains Γ_i, $i < p$, such that

$$x = x_0 \approx_{\Gamma_0} x_1 \approx_{\Gamma_2} x_2 \approx \ldots \approx_{\Gamma_{p-1}} x_p = y$$

Now for each $i < p$ there exists an isomorphism φ_i such that $\varphi_i(x_{i-1}) = x_i$. Put $\psi := \varphi_{p-1} \circ \ldots \circ \varphi_1 \circ \varphi_0$. We then have $y = \psi(x)$. Furthermore, all maps are isomorphisms of the constituents headed by the respective nodes. Hence, ψ is an isomorphism from $\downarrow x$ onto $\downarrow y$. And conversely. Q. E. D.

Lemma 16 *Suppose that K satisfies Chain Uniqueness and Liberation. Let φ and φ' be link maps such that $im(\varphi) \cap im(\varphi') \neq \varnothing$. Then $\varphi = \varphi'$.*

Proof. Let $\varphi : \mathfrak{C} \to \mathfrak{D}$, $\varphi' : \mathfrak{C}' \to \mathfrak{D}'$ and $x = \varphi(u) = \varphi'(u')$. Then $x \in \mathfrak{D}$ and $x \in \mathfrak{D}'$. Then either $\mathfrak{D} \subseteq \mathfrak{D}'$ or $\mathfrak{D}' \subseteq \mathfrak{D}$. Without loss of generality we assume the first. Suppose $\mathfrak{D} \neq \mathfrak{D}'$. Then $\mathfrak{D}'$ properly covers $\mathfrak{D}$ and therefore also $\mathfrak{C}$. It follows by *Liberation* that $\mathfrak{D}'$ is the foot of its chain. Contradiction. Hence $\mathfrak{D} = \mathfrak{D}'$. By *Chain Uniqueness*, and the fact that φ and φ' are link maps, $\varphi = \varphi'$. Q. E. D.

Definition 17 *Let ψ be a composite isomorphism. Call ψ* ***ascending*** *if it is a composition of link isomorphisms.*

Lemma 18 *Let $\langle \mathfrak{C}', \varphi, \mathfrak{C} \rangle$ be a chain link. Then for all $x \in \mathfrak{C}$:*

$$dp(\varphi(x)) < dp(x).$$

Proof. Let x have depth k in $\mathfrak{C}$. Then $\varphi(x)$ has depth k in $\mathfrak{C}'$. Now, if $\mathfrak{C} = \downarrow y$ and $\mathfrak{C}' = \downarrow y'$ then $dp(y') < dp(y)$. So, $dp(\varphi(x)) = k + dp(y') < k + dp(y) = dp(x)$. Q. E. D.

Lemma 19 *Let K be a CCT satisfying Liberation and Chain Uniqueness. Let ψ and ψ' be two ascending maps with identical domain. Let $\psi = \varphi_{n-1} \circ \varphi_{n-2} \circ \ldots \circ \varphi_1 \circ \varphi_0$ and $\psi' = \varphi'_{m-1} \circ \varphi'_{m-2} \circ \ldots \circ \varphi'_1 \circ \varphi'_0$ be their decompositions into link isomorphisms. Suppose that $\psi = \psi'$. Then $m = n$ and $\varphi'_i = \varphi_i$ for all $i < n$.*

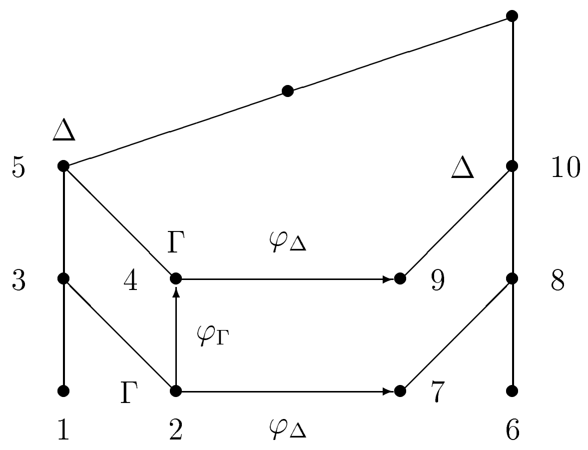
Proof. From Lemma 16 by induction on n. If $n = 0$, then ψ is the identity. Then ψ' also is the identity. But no nontrivial composition of link maps is the identity. Hence, $m = 0$ as well. Let $x \in im(\psi)$. Then $x \in im(\varphi_{n-1})$. Also, $x \in \psi'$ and so $x \in im(\varphi'_{m-1})$. Hence, $\varphi_{n-1} = \varphi'_{m-1}$ and so

$$\varphi_{n-2} \circ \ldots \circ \varphi_1 \circ \varphi_0 = \varphi'_{m-2} \circ \ldots \circ \varphi'_1 \circ \varphi'_0$$

By induction hypothesis, $m - 1 = n - 1$ and $\varphi'_i = \varphi_i$ for all $i < n - 1$. This shows the claim. Q. E. D.

Consider Figure 6. There are two nontrivial chains, Γ and Δ. Γ contains the constituents $\downarrow 2$ and $\downarrow 4$. The link map is $\varphi_\Gamma : 2 \mapsto 4$. Δ contains the constituents $\downarrow 5$ and $\downarrow 10$ and the link map is $\varphi_\Delta : 1 \mapsto 6, 2 \mapsto 7, 3 \mapsto 8, 4 \mapsto 9$ and $5 \mapsto 10$. (Recall that left–to–right order is at present irrelevant.) We have $[2]_C = \{2, 4, 7, 9\}$. We already have an ascending map from 2 to 4 and one from 2 to 7. There is an ascending map $\psi : 2 \mapsto 9$; it has a unique decomposition into $\varphi_\Delta \circ \varphi_\Gamma$. For notice that there is no link map from 7 to 9. Otherwise, *Liberation* would be violated. Hence, any ascending map ψ is uniquely decomposable into a product of link maps. We call this product the **canonical decomposition** of ψ. It follows that if there is a unique sequence of link maps that maps $\downarrow x_d$ to a given $\downarrow x$. Let it be $\varphi_{n-1} \ldots \circ \varphi_1 \circ \varphi_0$.

Figure 6: The Canonical Decomposition



Definition 20 *Let K satisfy Chain Uniqueness and Liberation. Then for each $x \in T$ there exists a unique x_r such that (r1) there exists an ascending map $\psi : x_r \to x$, (r2) x_r is not in the image of a link map. x_r is called the* ***root*** *of x. Furthermore, the unique ascending map sending x_r to x is called the* ***root map*** *of x. If $x \neq x_r$, that is, if x is not its own root, x is said to be in* ***derived position****.*

Similarly we speak of root constituents and constituents in derived position. Clearly, x is in derived position iff its root map is not the identity iff its canonical decomposition is nonempty.

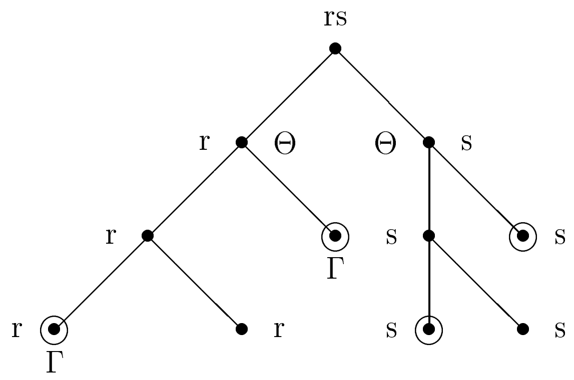
Lemma 21 *The following holds.*

1. $[x]_C = [y]_C$ *iff there exists two ascending maps χ and ψ such that* $y = \chi \circ \psi^{-1}(x)$.
2. $[x]_C = [y]_C$ *iff* $x_r = y_r$.

Proof. (1) From right to left follows from Lemma 15. Now let $[x]_C = [y]_C$. By Lemma 15 there exists a composite map σ such that $\sigma(x) = y$. σ is a composition $\tau_{n-1} \circ \tau_{n-2} \circ \ldots \circ \tau_0$, where τ_i is either a link map or its inverse. Suppose that n is chosen minimal. Now assume that τ_i is an inverse link map and τ_{i-1} is a link map. Then $(\tau_i)^{-1}$ is a link map, and by Lemma 16 we have $(\tau_i)^{-1} = \tau_{i-1}$. By minimality of n, this does not occur. Hence, there exists a j such that τ_i is an inverse link map iff $i < j$. So, put

$$\begin{array}{rcl} \chi & := & \tau_{i-1} \circ \tau_{i-2} \circ \ldots \circ \tau_j \\ \psi & := & (\tau_0^{-1} \circ \tau_1^{-1} \circ \ldots \circ \tau_{j-1}^{-1}) \end{array}$$

Figure 7: Too Many Surface Elements



Then $\sigma = \chi \circ \psi^{-1}$. Furhermore, χ and ψ are ascending. Now we show the second claim. Assume that $[x]_C = [y]_C$. Choose ascending maps ψ and χ such that $y = \chi \circ \psi^{-1}(x)$. Consider $\psi^{-1}(x)$. It is easy to see, using (1) and the definition of x_r, that there is an ascending map τ such that $\tau(x_r) = \psi^{-1}(x)$. Hence $y = \chi \circ \tau(x_r)$. Now, $\chi \circ \tau$ is ascending. x_r is a root. Hence $y_r = x_r$. Conversely, let $x_r = y_r$. Then there exist ascending maps ψ and χ such that $x = \chi(x_r)$ and $y = \psi(x_r)$. Hence $y = \psi \circ \chi^{-1}(x)$, and so $x \approx_C y$. Q. E. D.

Now, here is an alternative definition of the root elements.

Definition 22 *Let $K = \langle \mathfrak{T}, C \rangle$ be a CCT. $x \in T$ is called a* ***deep element*** *if either (d1) it is the root or (d2) it is the daughter of a deep element and $\downarrow x$ is a foot of some chain. We denote by $D^\star(K)$ the set of deep elements of K.*

Lemma 23 *An element is a root element iff it is in $D^\star(K)$.*

Proof. Suppose that x is not a root element. So $x = \varphi(y)$ for some link map, and so there exists some $u \geq x$ which is not the foot of some chain. Hence u and therefore x is not deep. Conversely, if x is not deep, there exists some $u \geq x$ that is a non–foot of some chain. So, there exists a link map φ and some y such that $\varphi(y) = x$. Hence x is not a root element. Q. E. D.

An analogous definition of surface or LF–element fails. Look at the structure of Figure 7; the elements enclosed in circles form an equivalence class. There are two members in one class that are labelled 's' if the definition is analogous to the definition of deep elements. This is an incorrect result. The same problems arise with LF–elements. Simply note that in the example given (using pre–chains) the surface elements and the LF–elements coincide. So, for the latter kind of elements we need a different definition. Finally, notice that there are two elements of the same, minimal depth in the class.

Lemma 24 $x \in D^\star(K)$ *iff* $\uparrow x \subseteq D^\star(K)$.

Proof. This follows from the definition. Q. E. D.

Lemma 25 *Suppose that K satisfies Liberation and Chain Uniqueness. Let $\mathfrak{C}$ be a non–foot of some chain. Then $\mathfrak{C}$ ac–commands its root.*

Proof. Let $\mathfrak{D}$ be the root of $\mathfrak{C}$ and $\mathfrak{C}_0$ the foot of the (unique) chain containing $\mathfrak{C}$. Suppose that $\mathfrak{C}$ does not c–command $\mathfrak{D}$. Then there is a constituent $\mathfrak{E}$ containing $\mathfrak{C}_0$ and $\mathfrak{C}$ but not $\mathfrak{D}$. (Namely, if $\mathfrak{C} = \downarrow x$ put $\mathfrak{E} := \downarrow \mu(x)$.) By *Liberation*, $\mathfrak{E}$ is not in derived position. Since $\mathfrak{C}_0 \neq \mathfrak{D}$, there must however be a chain Δ such that $\mathfrak{C}_0$ is covered by some non–foot $\mathfrak{G}$ of Δ. It is easy to see that $\mathfrak{G}$ covers $\mathfrak{C}$ as well, which is impossible. (It is easy to see also that the root cannot ac–command any constituent equivalent to it.) Q. E. D.

Lemma 26 *Let K be a CCT satisfying Liberation and Chain Uniqueness. Suppose that $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are distinct constituents such that $\mathfrak{C}_1 \approx_C \mathfrak{C}_2$ and that both ac–command their root. Then $\mathfrak{C}_1$ c–commands $\mathfrak{C}_2$ or $\mathfrak{C}_2$ c–commands $\mathfrak{C}_1$. If the tree is furthermore binary branching, $\mathfrak{C}_1$ ac–commands $\mathfrak{C}_2$ or $\mathfrak{C}_2$ ac–commands $\mathfrak{C}_1$.*

This follows from Lemma 25 and Lemma 5. Now define the **root–line** of $[x]_C$ or x_r, $r([x]_C)$, by

$$r([x]_C) := \{y : y \in [x]_C, y = x_r \text{ or } y \text{ ac–commands } x_r\}$$

Moreover, we also call the root line the set of all constituents $\downarrow y$ such that $y \in r([x]_C)$. The following is immediate to check.

Lemma 27 *x is a member of the root line of x_r iff $\downarrow x$ is the target of some link map iff $\downarrow x$ is the non–foot of some chain.*

Now let x be an element and $\varphi_{n-1} \circ \varphi_{n-2} \circ \ldots \circ \varphi_0$ the canonical decomposition of the root map of x. Then call the set

$$T(x) := \{\varphi_j \circ \varphi_{j-1} \circ \ldots \circ \varphi_0(x_r) : j < n\}$$

the **trajectory** of x. In a sense, the trajectory reflects the derivational history of x. However, as we shall see, there is only one trajectory for which this picture really fits. Define a binary relation $\lhd$ by $x \lhd y$ iff $dp(x) > dp(y)$.

Lemma 28 *$T(x)$ is linearly ordered by $\lhd$.*

Proof. Assume $T(x) = \{x_i : i < n+1\}$, where $x_{i+1} = \varphi_i(x_i)$, for some link map. By Lemma 18, $dp(x_{i+1}) < dp(x_i)$. Q. E. D.

Definition 29 *Let K be a CCT satisfying Chain Uniqueness and Liberation. Let x be an element. Call the highest member of the root line of $[x]_C$ the **peak** of $[x]_C$, and denote it by x_p. Further, let π_x be the (unique) product of chain links mapping x to x_p.*

Definition 30 *Let $\langle \mathfrak{T}, C\rangle$ be a CCT satisfying Chain Uniqueness and Liberation. The **zenith** of $[x]_C$, x_ζ, and the **zenith map**, ζ_x, are defined in tandem. If x is the root of $\mathfrak{T}$, x is the zenith of $[x]_C$, and $\zeta_x = id_T$. If x is distinct from the root of $\mathfrak{T}$, let y be the mother of x_p, the peak of $[x]_C$. Then the zenith of $[x]_C$ is $x_\zeta := \zeta_y(x_p)$, and the zenith map is $\zeta_x := \zeta_y \circ \pi_x$. The **orbit** of $[x]_C$, $O([x]_C)$, is the trajectory of the zenith of $[x]_C$. $Z^\star(K)$ denotes the set of zenith elements of K.*

This definition says the following. Suppose x is a given element, and assume that it is a root element. Then we follow the root line of x until its peak. If all $y > x$ are in trivial chains, then this position of x is already the zenith. If not, then some element $y > x$ is a member of a nontrivial chain. Then we must compute the zenith map of y first. Any element of $\downarrow y$ is moved up by the zenith map of y. Hence the zenith map of y moves $\downarrow x$ up. The definition requires specifically that y be the mother of x_p and that we take the zenith map of y. Should y be in a trivial chain, then the zenith map of y is the same as that of its mother, as is easily seen. It is crucial that one follows this procedure in the definition, otherwise the result is not unique. Notice that y in the definition is always in nonderived position, by *Liberation*. Hence, the zenith map is defined on y. The following is clear from the definitions.

Lemma 31 *$x \in Z^\star(K)$ iff $\uparrow x \subseteq Z^\star(K)$.*

Using the orbits we can define the last in our series of postulates. If we think of the orbit of $[x]_C$ as the series of intermediate positions of the root x_r then there should be no movements that make use of non–orbit positions.

Definition 32 *Let $\lambda = \langle \mathfrak{C}', \varphi, \mathfrak{C}\rangle$ be a chain link* and $\mathfrak{C} = \downarrow x$. λ is called **orbital** if φ is a member of the canonical decomposition of the zenith map of $[x]_C$.*

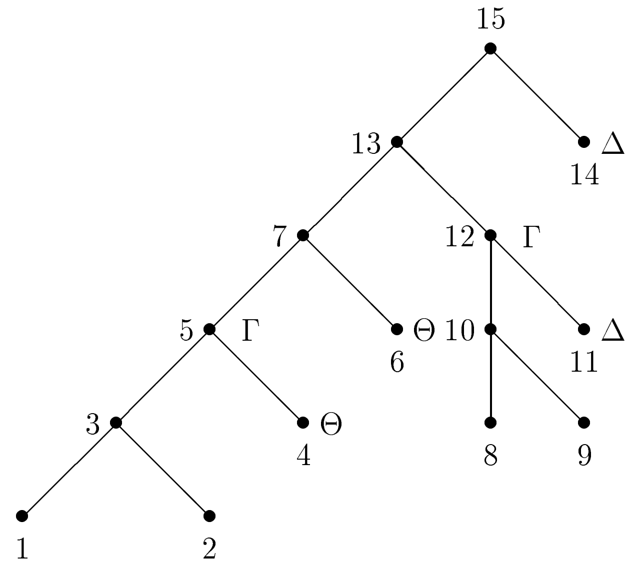
Constraint 6 (No Recycling) *All chain links must be orbital.*

A configuration that is excluded is shown in Figure 8. Here, the zenith of $[4]_C$ is 14. The zenith map ζ_4 has the unique decomposition

$$\zeta_4 = \varphi_\Delta \circ \varphi_\Gamma$$

where φ_Γ and φ_Δ are the unique link maps of the chains Γ and Δ. The trajectory of 4 is $\{4, 11, 14\}$. Now, the chain link of Θ is not orbital, as is clearly seen. Notice that the situation changes when we remove the chain Δ. For now 6 is the zenith of

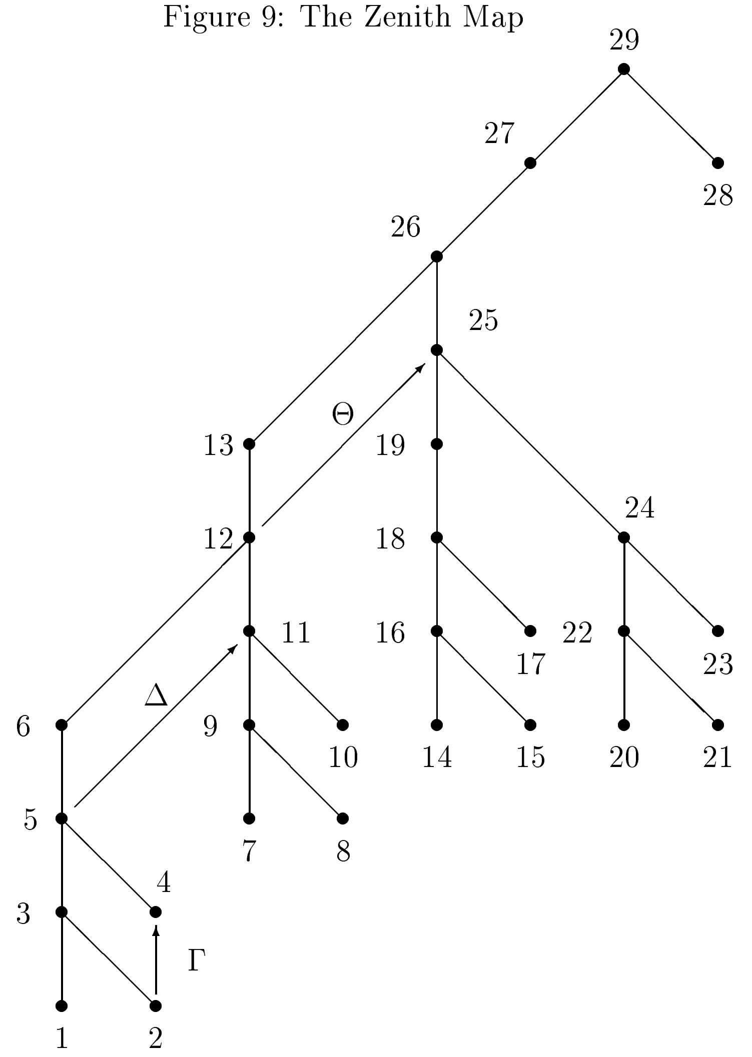
Figure 8: K



[4] and the link of Θ becomes orbital. So, the idea of the constraint *No Recycling* is that when a constituent is extracted, the lower copies of it may no longer become a member of some chain. For otherwise they would create links which are not orbital. To our knowledge, violations of *No Recycling* are nowhere attested in the theories. So, it might actually be considered to be trivially satisfied. [8] However, in the paper [3] split DPs are generated by scrambling the DP to VP via copy movement, subjecting the two copies to partial deletion, and then topicalizing the lower copy rather than the upper copy (see Section 2). A quick look at the structure (3d) shows that if the upper copy was topicalized instead, a different structure would have been generated. It is therefore precisely the mechanism of partial deletion (which we do not employ here) that makes the assumption of *No Recycling* a nonvoid requirement. If we assume *No Recycling* the analysis of Ćavar and Fanselow cannot be carried out. Since it is not clear what the status of the proposal is (in particular of partial deletion), we will leave it

[8] Indeed, as long as the theory had only trace–chains, *No Recycling* was excluded simply by the requirement that traces may not head chains. This is the principle *No Trace Recycling* discussed in Section 8. Chomsky [5] discusses the equivalent principle *Trace is immobile*, which he then adopts. (In MP, traces are actually copies, so there is a certain amount of terminological confusion.) The problem of mobile traces arises since he toys the idea that certain weak checking positions could be skipped overtly and then checked countercyclically (covertly) by the trace. This is discussed in detail in Gärtner [9]. These countercyclic movements are highly problematic in view of the resource problem and the explosion problem identified in Gärtner [11] and [9], because we do not have genuine traces but copies.

Figure 9: The Zenith Map



out of discussion here. *No Recycling* may therefore be safely assumed. We note here that if we assume binary branching, a derivation violating *No Recycling* can be 'rectified' in the following way. If Θ and Θ' are two chains recycling each other's feet, then the non–feet are ac–comparable and therefore form a pre–chain, Δ. Furthermore, one of the feet is ac–commanded by all non–feet of Θ_1 and Θ_2. So, we add it to the pre–chain and get a new chain which does not contain the offending instance.

To make the reader acquainted with the concepts, we have drawn in Figure 9 a tree with some set D of chains. Trivial chains are not shown. We have $\varphi_\Gamma : 2 \mapsto 4$, $\varphi_\Delta : k \mapsto k+6$ $(1 \leq k \leq 5)$, $\varphi_\Theta : k \mapsto k+13$ $(1 \leq k \leq 12)$. [9] Now, it is computed

[9] The reader is asked to check what would happen if we had chosen $\varphi'_\Delta : 1 \mapsto 8, 2 \mapsto 7, 3 \mapsto 9, 4 \mapsto 10, 5 \mapsto 11$. This will help to explain why the choice of isomorphisms is essential.

that $[1]_D = \{1, 7, 14, 20\}$, $[2]_D = \{2, 4, 8, 10, 15, 17, 21, 23\}$, $[3]_D = \{3, 9, 16, 22\}$, $[5]_D = \{5, 11, 18, 24\}$, $[12] = \{12, 25\}$, and so on. The roots are 1, 2, 3, 5, 6, 12, 13, 26, 27, 28 and 29. The root line of $[12]_D$ is $\{12, 25\}$, the root line of $[5]_D$ is $\{5, 11\}$ and the root line of $[2]_D$ is $\{2, 4\}$. We compute the following peak and zenith maps:

$$\begin{array}{lcl lcl} \pi_{12} & = & \varphi_\Theta & \zeta_{12} & = & \varphi_\Theta \\ \pi_5 & = & \varphi_\Delta & \zeta_5 & = & \varphi_\Theta \circ \varphi_\Delta \\ \pi_2 & = & \varphi_\Gamma & \zeta_2 & = & \varphi_\Theta \circ \varphi_\Delta \circ \varphi_\Gamma \end{array}$$

Clearly, $\pi_{29} = \zeta_{29} = id$, $\zeta_{28} = \zeta_{29}$, $\zeta_1 = \zeta_5$, $\zeta_3 = \zeta_5$, and so on for the other elements.

Now suppose we add a chain link involving 28, so that it becomes part of the root line of 2. Then we have eight choices for a link. However, not all of them are eligible. Suppose, for example, we add a chain link from 21 to 28. The trajectory of 28 will not contain 4. However, the map φ_Γ must be an orbital map, by *No Recycling*. Likewise, 2, 8, 15 and 21 may not enter a link with 28. This leaves us with three choices: we can link 28 with 4. In this case the trajectory of 28 is $\{2, 4, 28\}$. This derivation satisfies *Freeze* (which will be defined later). We can also link 28 with 10. Then the trajectory of 28 is $\{2, 4, 10, 28\}$. And finally we can link 28 with 23. This derivation satisfies *Shortest Steps*, which will also be defined later. The trajectory of 28 is now $\{2, 4, 10, 23, 28\}$. It is to be noted that the trajectories get longer the shorter the link is. This is no coincidence.

We will now present the definition of a copy chain structure, but at first only for the case of pre–chains.

Definition 33 *A* ***pre–copy chain structure*** *(a* ***preCCS****) is a preCCT satisfying Chain Existence, Chain Uniqueness, No Twisting, Liberation and No Recycling.*

Chains can interact in various ways with each other, as we have seen. We will study here how to classify these interactions. The basic setup is this. We have a CCS $K = \langle \mathfrak{T}, C \rangle$ and we want to add a chain link. Then we obtain a CCS $K' = \langle \mathfrak{T}', C' \rangle$ in which the new chain link is, say, $\lambda = \langle \mathfrak{D}, \mathfrak{D}' \rangle$. Let us take a chain Γ and see how it lies with respect to λ. There are two trivial cases. (A) $\mathfrak{D}$ is identical to the head of Γ. Then λ **extends** Γ. (B) $\mathfrak{D} \cap m(\Gamma) = \varnothing$. Then we say that λ **commutes** with Γ. Let us exclude these cases in the sequel.

Assume that $\mathfrak{D}$ is contained in a member of Γ, $\mathfrak{C}$. Then it is properly contained in $\mathfrak{C}$. Two cases arise. (C) $\mathfrak{C}$ is the foot of Γ. Then this represents the typical case of **extraction**. (D) $\mathfrak{C}$ is not the foot of Γ. Then we have what is known as **subextraction**.

Next suppose that $\mathfrak{D}$ is not contained in a member of Γ. Then the intersection of $\mathfrak{D}$ and Γ is a chain, the residue $\Gamma \upharpoonright \mathfrak{D}$. Now take a link $\mu = \langle \mathfrak{C}_i, \mathfrak{C}_{i+1} \rangle$. (E) Both members are covered by $\mathfrak{D}$. Then we say that λ **piggy–backs** μ. (F) Only one member is covered by $\mathfrak{D}$ (which is then $\mathfrak{C}_i$). In this case we say that $\mathfrak{D}$ is a **remnant** of μ and that λ is a case of **remnant movement**.

4 Copy Derivations

In this section we shall define a derivation by means of copy–movement and we shall show that chain structures defined earlier capture exactly the class of structures that can be obtained from trees by means of copy–movement. Modulo a property, Synchronization, which is still to be defined, we say the following.

Definition 34 *A* ***copy chain structure (CCS)*** *is a CCT which is a preCCS satisfying Synchronization. K is a* ***tree*** *if all chains are trivial.*

First we define the notion of a 1–step chain extension.

Definition 35 *Let $\Delta = \langle \mathfrak{D}, \mathbb{D} \rangle$ and $\Gamma = \langle \mathfrak{C}, \mathbb{C} \rangle$ be chains. Δ is a* ***1–step extension*** *of Γ if $\mathbb{D} = \{\mathfrak{X}_i : i < n+1\}$, $\mathbb{C} = \{\mathfrak{X}_i : i < n\}$ for a certain $n \in \omega$ and constituents $\mathfrak{X}_i$, and furthermore if $\mathfrak{D} \neq \mathfrak{X}_n$ then $\mathfrak{D} = \mathfrak{C}$. In case $\mathfrak{D} \neq \mathfrak{C}$ we call Δ an* ***S–extension*** *of Γ, otherwise an* ***LF–extension***. *Δ is an* ***extension*** *of Γ if it can be obtained from Γ through a series of 1–step extensions. Δ is a* ***proper extension of*** *Γ if Δ is an extension of Γ but $\Delta \neq \Gamma$.*

So, a 1–step extension is a chain with one more chain link. For chains* we also have to define a new chain link* $\langle \mathfrak{C}_n, \varphi_{n-1,n}, \mathfrak{C}_{n-1} \rangle$, with the new isomorphism $\varphi_{n-1,n}$. The full set of isomorphisms of the new chain* is obtained by closing the old set plus $\varphi_{n-1,n}$ under composition and inverse.

In contrast to the popular picture in transformational grammar, transformations will not be operations on trees or phrase markers but rather on chain structures. Therefore, a **CCS–derivation** is a sequence $K_0, K_1, \ldots, K_p$, where (i) K_0 is a tree, (ii) for each $i < p+1$, K_i is a copy chain structure, and (iii) for each $i < p$ the structure K_{i+1} is obtained from K_i by a single transformation. Of course, these transformations will be induced by simple operations on single chains, but there is a certain amount of cosmetics that needs to be done.

So, a derivation by movement proceeds by manipulation of chain structures. If we look at copy–movement, matters are relatively straightforward. Given a particular chain structure, *Move–α* will target a chain $\Gamma = \langle \mathfrak{C}, \mathbb{C} \rangle$, where $\mathbb{C} = \{\mathfrak{X}_i : i < n\}$ and produce a new chain, which in the case of copy chains is an extension of Γ. Namely, it will add a new constituent $\mathfrak{X}_n$ to $\mathfrak{T}$ that c–commands $\mathfrak{X}_{n-1}$. Then it will go over to either the 1–step extension $\langle \mathfrak{X}_n, \mathbb{C} \cup \{\mathfrak{X}_n\} \rangle$ or $\langle \mathfrak{C}, \mathbb{C} \cup \{\mathfrak{X}_n\} \rangle$. In the first case we speak of **S–movement**, in the second case of **LF–movement**. Moreover, for theoretical reasons it is advisable to add all one–membered chains for those constituents that are not members of a chain already. [10] This is the general construction. [11]

Suppose that Δ is an extension of Γ. If Γ is an LF–chain, the surface elements must remain the same, so Δ is an LF–chain, too. If Γ is an S–chain then Δ may also be an S–chain, in which case the surface element must change, or Δ will be an LF–chain, in which case the surface elements of Δ and Γ coincide.

Definition 36 *Let $K_1 = \langle \mathfrak{T}_1, C_1 \rangle$ and $K_2 = \langle \mathfrak{T}_2, C_2 \rangle$ be CCSs. K_2 is obtained from K_1 by **(1–step) copy–movement** if the following holds.*

1. *$\mathfrak{T}_1 \subseteq \mathfrak{T}_2$.*
2. *$\mathfrak{C} := \mathfrak{T}_2 - \mathfrak{T}_1$ is a constituent of $\mathfrak{T}_2$.*
3. *$\mathfrak{C}$ is in zenith position in K_2.*
4. *Every chain of K_1 is the residue of a chain of K_2.*
5. *There is exactly one nontrivial chain Γ such that $m(\Gamma) \cap \mathfrak{T}_1 \neq \varnothing$ and $m(\Gamma) \cap \mathfrak{C} \neq \varnothing$. Moreover:*
   (a) *$m(\Gamma) \cap \mathfrak{C} = \mathfrak{C}$.*
   (b) *$\mathfrak{C}$ is the highest element of Γ.*
   (c) *For the second highest element $\mathfrak{D}$ of Γ: $\mathfrak{D}$ is in zenith position of K_1.*
   (d) *If $\mathfrak{C}$ is the distinguished element of Γ, all chains of K_2 are S–chains.*

*L is obtained from K by **copy–movement** iff there exists a sequence K_i, $i < n+1$, such that $K = K_0$, $L = K_n$ and K_{i+1} is obtained from K_i by 1–step copy–movement. The sequence $\langle K_i : i < n+1 \rangle$ is also called a **derivation** of K. This derivation has **length** n.*

Some facts are worth noting. Since in going back from K_2 to K_1 we are removing one link only and this link is to a zenith element, the trajectories of all elements distinct from $\mathfrak{C}$ remain the same. This amounts to the same whether we compute the trajectory of $\mathfrak{D} \neq \mathfrak{C}$ in K_2 or in K_1. We require that the element that is being copied is in zenith position in $\mathfrak{T}_1$ and the copy is in zenith position in $\mathfrak{T}_2$. Therefore, orbits are monotone in the derivation. This definition does not show how to obtain K_2 from K_1, only how the two are related. Clearly, they are

[10] This is one, easy, example of the cosmetics to be done when one works with chain structures. The reader may wonder why such details are actually necessary. However, we believe that in general the technical frameworks are hardly worked out in such detail as is done here, and hence one may get the impression that they are essentially simpler. But this is not so, as far as we can see. When the details are worked out, as in the book by Stabler [23], one clearly sees how much there needs to be done. With respect to the problems induced by trivial chains see also the discussion in Gärtner [9].

[11] However, when we look at copy–and–delete, matters are more delicate.

related in such a way that the residue of Γ in K_2 is extended to Γ. Below we will show how K_2 is constructed from K_1, given the residue of Γ and the point that it targets by movement.

We note here that for technical convenience, movement does not consist in replacement of an empty node, but rather in the addition of an entire constituent. So, rather than substituting for a node that has been there before, it creates all the nodes afresh. The reader is assured that this is just a technical simplification which is of no theoretical significance. First of all we will show that there is always a way to move a constituent:

Proposition 37 *Let $K = \langle \mathfrak{T}, C\rangle$ be a CCS, $\Gamma \in C$ and x a node of $\mathfrak{T}$ in zenith position that dominates all nodes of $m(\Gamma)$, but which does not dominate any node of $m(\Gamma)$ immediately. Further, assume that every $y \geq x$ is in a trivial chain. Let $\mathfrak{D}$ be a tree disjoint from $\mathfrak{T}$ and isomorphic to the head of Γ. Then form a tree $\mathfrak{U}$ by adding $\mathfrak{D}$ as a new daughter constituent to x. Now, for any $y \in \mathfrak{D}$ which is not the root r' of $\mathfrak{D}$, let Δ_y be the trivial chain containing $\downarrow y$. For a chain $\Delta \in C$, $\Delta \neq \Gamma$, let $\Delta^\dagger := \Delta$ if no point of $m(\Delta)$ dominates x; else $\Delta = \langle \downarrow y, \{\downarrow y\}\rangle$ (in $\mathfrak{T}$) for some $y \geq x$, and then $\Delta^\dagger := \langle \downarrow y, \{\downarrow y\}\rangle$ (in $\mathfrak{U}$). Finally, $\Gamma^\dagger := \langle \mathfrak{C}, \mathbb{C} \cup \{\mathfrak{D}\}\rangle$ or $\Gamma^\dagger := \langle \mathfrak{D}, \mathbb{C} \cup \{\mathfrak{D}\}\rangle$, but the latter only if C does not contain any LF–chains. Finally, let $E := \{\Delta^\dagger : \Delta \in C\} \cup \{\Delta_x : x \in D - \{r'\}\}$ and $L := \langle \mathfrak{U}, E\rangle$. Then L is obtained from K by copying the head of Γ to x.*

It is easily verified that the proposed construction indeed produces a CCS and that it is obtained from K by movement. We note that we require that all $y \geq x$ are in trivial chains to satisfy *Liberation* and that x is zenith position in order to guarantee *No Recycling*. Hence we can produce nontrivial structures from trees by movement. We wish to show that any CCS can be obtained in this way. The following is straightforwardly verified.

Lemma 38 *Suppose that K_2 is obtained from K_1 by copy–movement. Then if K_1 is a CCS so is K_2.*

Lemma 39 *Suppose that $K_2 = \langle \mathfrak{T}_2, C_2\rangle$ is obtained from $K_1 = \langle \mathfrak{T}_1, C_1\rangle$ by copying $\downarrow x$ to $\downarrow y$. Let $\langle \downarrow y, \varphi, \downarrow x\rangle$ be the additional link. Then for all $z \in T_1$: $T_{K_2}(z) = T_{K_1}(z)$. Furthermore, if $z \leq y$ there is a u such that $\varphi(u) = z$ and $T_{K_2}(z) = T_{K_1}(u) \cup \{z\}$. Further: if $z \not\leq x$ is in zenith position in K_1 then it is in zenith position in K_2, and if $u \leq x$ is in zenith position in K_1 then $\varphi(u)$ is in zenith position in K_2.*

Now we come to an important notion, that of *blockage*. By means of blockage we will show that every CCS is derivable by means of copy–movement from a tree.

Definition 40 *Let $\Gamma = \langle \mathfrak{C}_m, \{\mathfrak{C}_j : j < p\}\rangle$ and $\Delta = \langle \mathfrak{D}_n, \{\mathfrak{D}_j : j < q\}\rangle$ be two different nontrivial chains. $\mathfrak{C}_j$ **blocks** $\mathfrak{D}_i$ where $i > 0$ if either (a) $j > 0$ and $\mathfrak{C}_{j-1}$ covers $\mathfrak{D}_{i-1}$ or (b) some $\mathfrak{D}_{i'}$ covers $\mathfrak{C}_j$ for some $i' \geq i$ or (c) $i \leq n$ and $j > m$. In a single chain, $\mathfrak{C}_j$ **blocks** $\mathfrak{C}_i$ if $0 < i < j$.*

Figure 10: Blocking Configurations

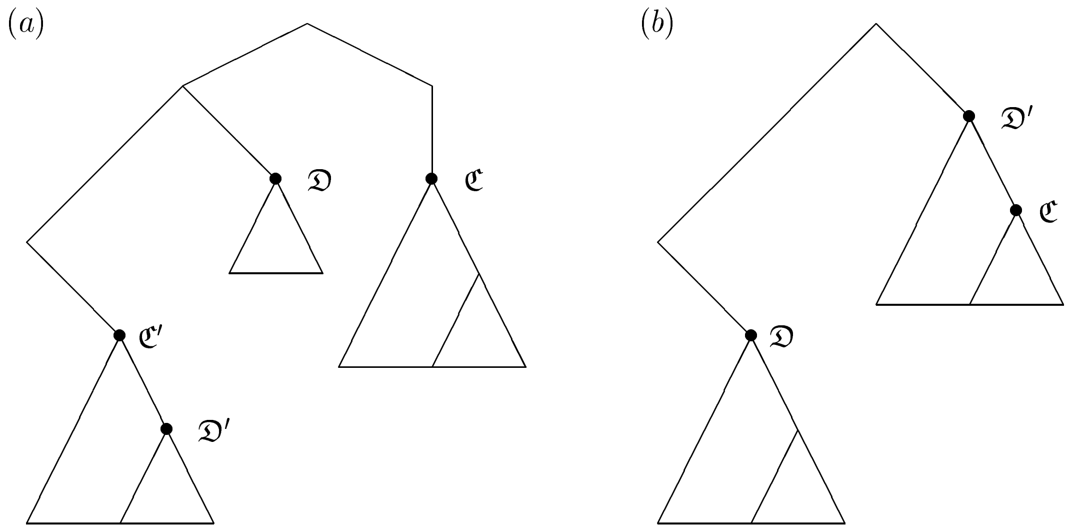


Figure 10 shows the blocking configurations of Clauses (a) and (b). In (a), both $\mathfrak{C}$ and $\mathfrak{D}$ are the non–foot members of some chains and the predecessor of $\mathfrak{C}$, $\mathfrak{C}'$, properly covers the predecessor $\mathfrak{D}'$ of $\mathfrak{D}$. In (b), $\mathfrak{C}$ is contained in $\mathfrak{D}'$, which is higher in the chain than $\mathfrak{D}$. Again, $\mathfrak{C}$ blocks $\mathfrak{D}$.

Blockage is a relation between constituents that are members of some nontrivial chains. We will show that the notion of blockage characterizes exactly which elements have been added prior to which others. That is to say, if $\mathfrak{C}$ blocks $\mathfrak{D}$ then all derivations introduce $\mathfrak{C}$ after $\mathfrak{D}$. For example, consider the state of the derivation of (a) before $\mathfrak{C}$ and $\mathfrak{D}$ have been added. Suppose that $\mathfrak{D}$ is added after $\mathfrak{C}$ (so that it would be removed before $\mathfrak{C}$). Then after adding $\mathfrak{C}$, $\mathfrak{D}'$ is actually no longer in zenith position, and so it may not enter a link with $\mathfrak{D}$. A different situation is captured with (b). Here, the chain containing $\mathfrak{C}$ must be resolved first before we can attempt to remove $\mathfrak{D}'$. So, $\mathfrak{C}$ blocks $\mathfrak{D}$ and $\mathfrak{D}'$ as long as it is a member of a nontrivial chain.

Lemma 41 *Suppose that L is obtained from K by 1–step copy–movement. Let $\mathfrak{C}$ be the constituent that is added to K to obtain L. Then $\mathfrak{C}$ is not blocked.*

Proof. Clearly, $\mathfrak{C}$ is not blocked inside its own chain since it is the head. Let $\mathfrak{D}$ be a non–foot of some chain. If $\mathfrak{D}$ blocks $\mathfrak{C}$, then either (a) the predecessor $\mathfrak{D}'$ of $\mathfrak{D}$ covers the predecessor $\mathfrak{C}'$ of $\mathfrak{C}$ or (b) $\mathfrak{C}$ covers $\mathfrak{D}$, or (c) $\mathfrak{D}$ is an LF–element but $\mathfrak{C}$ is not. (a) cannot hold. For $\mathfrak{C}$ is not in the zenith of K since there is a link map moving $\mathfrak{D}'$ to $\mathfrak{D}$. Hence, $\mathfrak{C}'$ was not not eligible for copying. Contradiction.

(b) does not arise, since $\mathfrak{C}$ properly covers $\mathfrak{D}$, and so $\mathfrak{D}$ is in a trivial chain. (c) Suppose that $\mathfrak{C}$ is not an LF–member of its chain. Then $\mathfrak{D}$ is also not an LF–member in its own chain, by construction. Q. E. D.

We can now formulate the condition on synchronization. The motivation comes from the fact that derivations in GB proceed in two stages: first, the surface structure is created. At this stage, no LF–movement is permitted, any (copy–)movement consists in a 'visible' displacement. After the surface structure is established, no more visible displacements are possible. The surface chain members are frozen to their place and every movement is an LF–movement. We will see that the following condition does the job as intended.

Constraint 7 (Synchronization) *No LF–member of a chain is blocked by a non–LF–member of some chain.*

Lemma 42 *Let K be a CCS. Assume that K has an unblocked constituent in zenith position. Then there exists a CCS L from which K is obtained by 1–step copy–movement.*

Proof. Let $K = \langle \mathfrak{T}, C\rangle$, $\Gamma \in C$ a chain. Assume that the head of Γ, $\mathfrak{C}$, is unblocked and in zenith position. We claim first that every nontrivial chain of K different from Γ is disjoint from $\mathfrak{C}$. So, let Δ be a nontrivial chain and $\Delta \neq \Gamma$. By definition of blocking (clause (b)), $\mathfrak{C}$ does not cover any member of Δ. Since $\mathfrak{C}$ is in zenith position, no member of Δ covers $\mathfrak{C}$. Hence, Δ is disjoint from $\mathfrak{C}$. Therefore, let $\mathfrak{U}$ be the tree defined from removing from $\mathfrak{T}$ the constituent $\mathfrak{C}$, and let $L := \langle \mathfrak{U}, D\rangle$, where D consists of the residues of all chains of K. We check that L is a CCS. First, all residues of chains are chains again. For let $\mathfrak{D} \restriction \mathfrak{U}$ and $\mathfrak{D}' \restriction \mathfrak{U}$ be distinct members of some chain, $\Theta \restriction \mathfrak{U}$. Then one ac–commands the other. Further, they are isomorphic. For Θ is either disjoint from Γ or $\mathfrak{D}$ and $\mathfrak{D}'$ are constituents of Γ different from $\mathfrak{C}$. In both cases, $\mathfrak{D} \restriction \mathfrak{U} \cong \mathfrak{D}' \restriction \mathfrak{U}$. It is easy to see that *Chain Uniqueness* and *Chain Existence* are satisfied. Now *Liberation*. Suppose that L contains a chain with a non–foot covering two members of some other chain. Then it is clear that K must contain these chains as well, contrary to our assumption. We postpone the verification of *No Recycling*. $\mathfrak{C}$ is in zenith position in K. Let $T(\mathfrak{C})$ be the orbit of $\mathfrak{C}$ in K and let $\mathfrak{D}$ be the highest element of $T(\mathfrak{C}) - \{\mathfrak{C}\}$ with respect to $\lhd$. There is a link map $\varphi : \mathfrak{D} \to \mathfrak{C}$ in K. We claim that $\mathfrak{D}$ is in zenith position in L. For if not, there is a link–map $\varphi' : \mathfrak{H} \to \mathfrak{K}$ such that $\mathfrak{D} \subsetneq \mathfrak{H}$. But then $\mathfrak{C}$ is actually blocked by $\mathfrak{K}$ by Clause (a) of blocking. Now we verify *No Recycling*. Take a link map φ of L. We have to show that it is orbital. To that effect, it is actually enough to show that the orbits of L are just the orbits of K restricted to $\mathfrak{U}$. Let $x \in T$. The canonical decomposition of $\psi : x_r \to x$ either contains no occurrence of φ (and hence the trajectory of x is entirely inside L) or $\psi = \varphi \circ \chi$ and χ contains no occurrence of φ. So, in the second case $x = \varphi(y)$ for some $y \notin \mathfrak{C}$. This shows that trajectories of K are extensions of trajectories of L by at most one element. Unless $x \in \mathfrak{D} \cup \mathfrak{C}$,

x is in zenith position in K iff it is in zenith position in L. If $x \in \mathfrak{C}$, x is in zenith position in K iff $\varphi^{-1}(x)$ is in zenith position in L. So, the claim for orbits follows, and *No Recycling* is shown. Finally, *Synchronization*. Suppose that $\mathfrak{E}$ is an LF–member and is blocked by $\mathfrak{F}$, a non–LF–member. Since both are in L, then already K violates *Synchronization*. Contradiction. So, L satisfies *Synchronization*. It follows that K is obtained from L by 1–step copy–movement. Q. E. D.

Lemma 43 *Let K be a CCS which is not a tree. Then K has an unblocked constituent $\mathfrak{C}$ which is in a nontrivial chain and in zenith position.*

Proof. Suppose first that K contains no LF–chains. Let M be the set of all constituents that are not members of a trivial chain. Since K is not a tree, $M \neq \varnothing$. Now choose a $\mathfrak{C} \in M$ such that for every $\mathfrak{D} \in M$ that c–commands $\mathfrak{C}$, also $\mathfrak{C}$ c–commands $\mathfrak{D}$. In other words, $\mathfrak{C}$ is not ac–commanded by any member of M. It follows that $\mathfrak{C}$ is the head of some nontrivial chain. We show that $\mathfrak{C}$ is not blocked. Suppose therefore that $\mathfrak{C}$ is properly contained in some $\mathfrak{D}$. Then, since $\mathfrak{C}$ is a member of a nontrivial chain, $\mathfrak{D}$ covers a nontrivial chain. Hence $\mathfrak{D}$ is the foot of a chain Δ. If Δ is nontrivial, it has a member which ac–commands $\mathfrak{C}$. Contradiction. So, Δ is trivial and therefore its member cannot block $\mathfrak{C}$. Now assume that $\mathfrak{C}$ properly contains some $\mathfrak{D}$. Suppose that $\mathfrak{D}$ is a member of a nontrivial chain, Δ. Since $\mathfrak{C}$ is also a member of a nontrivial chain, by *Liberation*, $\mathfrak{D}$ is the foot of Δ. Now, the head of Δ ac–commands $\mathfrak{C}$, since it ac–commands $\mathfrak{D}$ but is disjoint with $\mathfrak{C}$. Contradiction to the choice of $\mathfrak{C}$. Since there are no LF–chains, (c) does not arise. Hence, $\mathfrak{C}$ is unblocked. $\mathfrak{C}$ is also in zenith position. For, by *Liberation*, since $\mathfrak{C}$ is in a nontrivial chain, $\mathfrak{C}$ is in peak position. If $\mathfrak{C}$ is properly contained in $\mathfrak{D}$, $\mathfrak{D}$ is in root position, by *Liberation*. If $\mathfrak{D}$ is not in zenith position, then there exists some $\mathfrak{D}'$ which ac–commands $\mathfrak{D}$. Furthermore, it can be arranged that both $\mathfrak{D}$ and $\mathfrak{D}'$ are members of the same chain. Hence $\mathfrak{D}'$ is a non–foot. So, $\mathfrak{D}' \in M$ and $\mathfrak{D}'$ ac–commands $\mathfrak{C}$, contrary to our choice. This finishes the proof if no LF–chains exist. Now assume that there exist LF–chains. Let M be the set of constituents that are LF–members of some chain. $M \neq \varnothing$, by assumption. Let $\mathfrak{C}$ be an element which is not ac–commanded by a member of M. Then assume that $\mathfrak{D}$ blocks $\mathfrak{C}$. By Synchronization, $\mathfrak{D}$ is an LF–element, and therefore $\mathfrak{D} \in M$. Now reason as before. Q. E. D.

The following is now obvious.

Theorem 44 *Let K be a CCS. Then there exists a tree L such that K is derived from K by copy–movement.*

We can be a little bit more precise than that. First, look at the restriction of K to the set $D^*(K)$. It turns out that it is a tree, since the residue on any chain (if defined) is single membered. K is obtained by adding more and more constituents to L. Let H be the set of constituents of K which are not the foot of some chain.

Choose a transitive linear order $\ll$ on H such that if $\mathfrak{C} \ll \mathfrak{D}$ then (a) $\mathfrak{C}$ does not block $\mathfrak{D}$ and (b) if $[\mathfrak{C}]_K = [\mathfrak{D}]_K$ then $\mathfrak{C} \in T(\mathfrak{D})$. Enumerate $H = \{\mathfrak{C}_i : i < q\}$ such that $\mathfrak{C}_i \ll \mathfrak{C}_j$ iff $i < j$. Let L_j be the restriction of K to the set of nodes which are in U or in one of the $\mathfrak{C}_i$ for $i < j$. L_j is a chain structure. $L_0 = L$, $K = L_q$. And L_{j+1} is obtained from L_j by movement. So, any enumeration of the non–foot constituents that is compatible with the blocking order and respects the trajectories gives rise to a derivation. Conversely, of course, every derivation defines a set of non–foot constituents which are linearly ordered such that later elements are not blocked by earlier elements.

It follows that there exist derivations that can be factored in an S–structure derivation, where all chains are S–chains, and a subsequent LF–structure derivation, where all new chains are LF–chains. This is a direct consequence of the Synchronization condition. However, all derivations can be factored in this way, since the blocking order bans any surface movement that follows an LF–movement. Hence, we have a complete correspondence between structures that can be established derivationally from trees and structures that satisfy certain constraints.

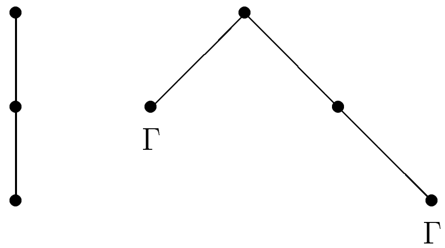
5 Multidominance Structures

Ideally, we wish to look at the chains as single entities and replace talk of members of chains by talk of chains simpliciter. This is possible, and it gives rise to structures that we will call *multidominance structures*. Their main difference with trees is that an element can have several mothers. However, in passing from chain structures to multidominance structures some information is lost. This means that there are nonisomorphic chain structures which have isomorphic multidominance structures. The idea behind the construction is quite simple. We will take the members of one chain as one single constituent of some structure that has various mothers. However, the details of this construction are quite delicate. We will therefore work first with *pre–copy chain structures* or *preCCSs*. This allows to skip the discussion of the surface element so that we can concentrate on the mechanics of the pre–chains. The position of the surface element will be discussed in the next section.

Definition 45 *Let G be a set and $\prec$ a binary relation. Suppose that the transitive closure of $\prec$, denoted by $<$, is irreflexive. Then $\langle G, \prec\rangle$ is called an* ***acyclic graph****. A pair $\langle x, y\rangle \in \prec$ is called a* ***link (of*** *x).*

(Usually, graphs are symmetric, therefore one usually calls these structures *directed acyclic graphs*. However, acyclicity only makes sense in the context of directed graphs. Notice that if $<$ is irreflexive and transitive it contains no cycles.) The acyclic graphs will replace the trees. A node in an acyclic graph may have several mother nodes. It is worthwhile pointing out why we have

Figure 11: Chain structures of $\mathfrak{G}_1$) (left) and $\mathfrak{G}_2$ (right)



defined structures using the immediate dominance relation rather than its transitive closure. Consider the structures $\mathfrak{G}_1 := \langle \{1, 2, 3\}, \{\langle 1, 2\rangle, \langle 2, 3\rangle\}\rangle$ and $\mathfrak{G}_2 := \langle \{1, 2, 3\}, \{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 2, 3\rangle\}\rangle$. Clearly, the transitive closure of the two relations of $\mathfrak{G}_1$ and $\mathfrak{G}_2$ are the same. As will become clear below, the two structures correspond to two different chain structures, shown in Figure 11. This is so because the node 3 has two immediate mothers in $\mathfrak{A}$ and so it will be split into two copies. We will begin by showing how to construct a tree out of an acyclic graph, and how to define an acyclic graph from a preCCS (Lemma 47 and Lemma 48). After that we are in a position to characterize intrinsically those acyclic graphs that arise from preCCSs via this construction (Definition 49). The exactness of this characterization is shown in Lemma 50.

Definition 46 *Let $\mathfrak{G} = \langle G, \prec\rangle$ be an acyclic graph. It is called **rooted** if there exists an $r \in G$ such that for all x we have $x \leq r$. An **identifier** is a sequence $I = \langle x_i : i < n + 1\rangle$ such that $x_0 = r$ and $x_{i+1} \prec x_i$, $i < n$. I has **length** n and **identifies** x_n.*

Identifiers are sequences. We write $I; J$ for the concatenation of sequences, and also $I; x$ for appending x to the sequence I. We say that J **extends** I (**properly**) if $J = I; K$ for some K (for some K which is not empty). We can easily define a tree based on the set of identifiers by putting $I \prec J$ iff $I = J; x$ for some x. Then $I \leq J$ iff I extends J and $I < J$ iff I properly extends J. We say that I **precedes** J if $I = K; x$ and $J = K; L$ for some nonempty sequence L not starting with x. The following is easily established.

Lemma 47 *Let $\mathfrak{G}$ be a rooted acyclic graph, $T(\mathfrak{G})$ the set of identifiers of $\mathfrak{G}$ and $I \prec J$ iff $I = J; x$ for some x. Then $\langle T(\mathfrak{G}), \prec\rangle$ is a tree. Moreover, I precedes J iff I ac–commands J in the tree $\langle T(\mathfrak{G}), \prec\rangle$.*

For example, the acyclic graph $\mathfrak{A}$ shown in Figure 12 is converted into the tree shown in Figure 13. (Or look again at Figure 11.) Recall the definition of

Figure 12: The acyclic graph $\mathfrak{A}$

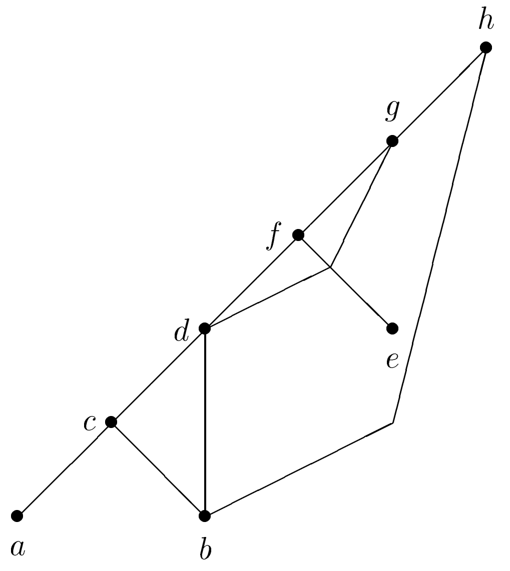
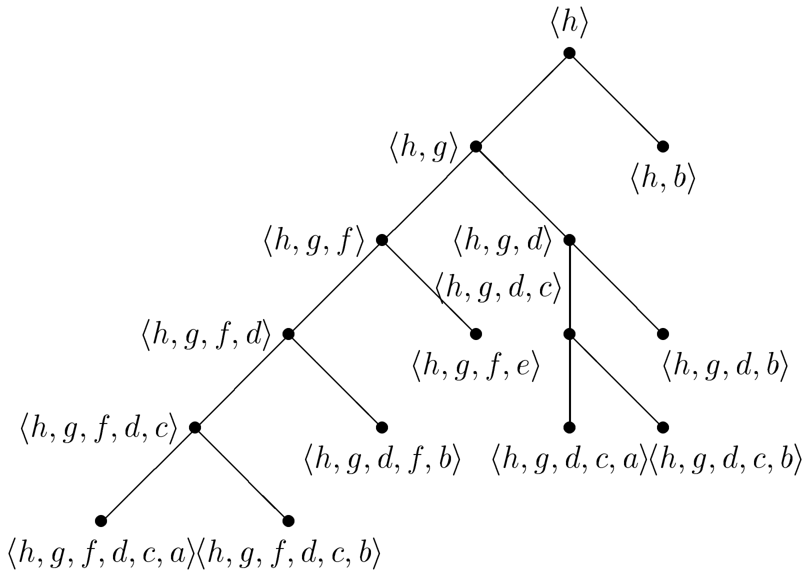


Figure 13: The tree corresponding to $\mathfrak{A}$



the relation $\approx_K$ from Section 2. We define $G(K) := \{[x]_K : x \in \mathfrak{T}\}$. We put $[x]_C \prec [y]_C$ iff there exists $x' \in [x]_K$ and $y' \in [y]_K$ such that $x' \prec y'$. Finally,

$$M(K) := \langle G(K), \prec \rangle$$

The following is easy to verify.

Lemma 48 *Let K be a CCS. Then $M(K)$ is a rooted acyclic graph.*

Instead of defining $M(K)$ on the set of the $[x]_K$, we could alternatively take the set $D^*(K)$ of all x_r. Our main aim is now to define what kind of structures are of the form $M(K)$, K a preCCS.

Definition 49 *Let M be a set and $\prec$ a binary relation on M. $\langle M, \prec \rangle$ is a **pre–multidominance structure (preMDS)** if (md1) $\langle M, \prec \rangle$ is a rooted acyclic graph and (md2) for every $x \in M$, the set $M(x) := \{y : x \prec y\}$ is linearly ordered by $<$.*

Lemma 50 *Let K be a preCCS. Then $M(K)$ is a preMDS.*

Proof. $M(K)$ is a rooted acyclic graph by Lemma 48. We have to show (md2). Suppose that $[x]_C \prec [y]_C$ and $[x]_C \prec [z]_C$. Then there are $x', x'' \in [x]_C$ and $y' \in [y]_C$, $z' \in [z]_C$ such that $x' \prec y'$ and $x'' \prec z'$. In fact, we may assume that x', x'' are in the root line of x_r. (Namely, take x' of maximal depth in $[x]_C$ such that $x' \prec u$ for some $u \in [y]_C$. Assume that it is not in root line. Then take the unique link map φ such that $x' \in im(\varphi)$. Then $y' \in im(\varphi)$, for otherwise x' is in root line. Now put $y'' := \varphi^{-1}(y')$ and $x''' := \varphi^{-1}(x')$. Then $x''' \prec y''$. Furthermore, x''' has larger depth than x'. Contradiction.) Hence, y' and z' both dominate x_r. It follows that $y' < z'$, $y' = z'$ or $z' < y'$. Hence either $[y]_C = [y']_C < [z']_C = [z]_C$, or $[y]_C = [z]_C$ or $[y]_C > [z]_C$. Q. E. D.

After we have isolated the preMDSs, we will give algorithms for extracting all possible preCCSs out of a preMDS. We observe first that there exist nonisomorphic preCCSs which define the same preMDS (for example the structure in Figure 12). Here is a linguistic example.

(4) ***Diese Bücher*** *hatte ich* ***alle*** *damals meinem Chef* ***t*** *gegeben.*
These books had I all then to my boss given.
I had then given all of these books to my boss.

Here, the direct object has been scrambled and its NP subpart has been topicalized, leaving behind the quantifier *alle* (Subextraction). However, there is an alternative, countercyclic analysis, namely where the noun phrase moves to its final position in one fell swoop, and the remnant consisting of the quantifier and the NP–trace is scrambled (Remnant Movement). Both lead to the same preMDS.

We shall describe an algorithm that extracts all possible preCCSs from a given preMDS. It will be apparent that there is at least one. First, we settle the existence question. Call an identifier I a **root identifier** if for all prefixes $J; y; x$, y is the smallest element in $M(x)$. Clearly, each y has exactly one root identifier, and all its prefixes are root identifiers. Moreover, if I is a root identifier for x, and J identifies x, then I is longer than J. Having defined a root identifier, we can also define the **root line**. It consists of all identifiers preceding a root identifier identifying the same element.

Theorem 51 *Let $\mathfrak{M}$ be a preMDS. Then there exists a preCCS K such that $M(K) \cong \mathfrak{A}$.*

Proof. Let $\mathfrak{M} = \langle M, \prec \rangle$ be a preMDS. Then by Lemma 47, $\langle T(\mathfrak{M}), \prec \rangle$ is a tree. Given $I, J \in T(\mathfrak{M})$, put $I \approx J$ iff I and J identify the same element. This is an equivalence relation. Our job is to find the chain links. Simply take as chain links all $\langle \downarrow I', \varphi_{II'}, \downarrow I \rangle$ such that (a) I' and I are in the root line of some element J, (b) no element in the root line of J is ac–commanded by I' and ac–commands I, and (c) if $I' = K'; y'$, $I = K'; L; y$, then $\varphi_{II'}(K'; L; y; M) := K'; y'; M$. Clearly, by construction $J \approx \varphi_{II'}(J)$ for all $J \leq I$. This defines the chain links* and also the chains*. Now put $K := \langle T(\mathfrak{M}), \prec, C \rangle$. K is a preCCS, as is easily verified. Finally, we have to show that $M(K) \cong \mathfrak{A}$. It is clear by construction that $\approx_C \subseteq \approx$. The converse needs to be shown. First, I is its own root line iff $I = H; u$ for some root identifier H. Call such an I a **near root**. Put $M(u) = \{v_i : i < m\}$ such that $v_i < v_{i+1}$ in $\mathfrak{A}$. The elements of the root line of I are of the form $L_i; v_i; u$, $i < m$. Furthermore, $L_i = L_{i+1}; P_i$ for some P_i. The link maps are of the form $L_{i+1}; P_i; v_i; u; L \mapsto L_{i+1}; v_{i+1}; u; L$. Hence, every near root identifier is the image of a root identifier under a composition of these maps. We show that any given identifier is the image of a root identifier under a suitable composition of these maps. Let P be an identifier. If P is a root identifier, we are done. Otherwise, let H be the largest prefix that is a near root identifier. H is not a root. $P = H; Q$ for some Q. H is the image of a root identifier H' under a composite map χ. So, $P = \chi(H'; Q)$. Now it is enough to show the claim for $P' = H'; Q$. And so on. To see that this reduction comes to an end, notice the following. Let H'' be the largest near root identifier in P'. Then $P' = H''; Q'$ for some Q'. It is easy to see that Q' is a proper prefix of Q. Hence, the 'defect' of the identifier gets smaller in each step. Finally, it disappears. Q. E. D.

By construction, the root line of each element is a pre–chain. For the links are members of the same root line, by (a), and for each $\downarrow I$, $\downarrow I'$ is the next higher element of the root line, by (b).

Lemma 52 *Let $K = \langle \mathfrak{T}, C \rangle$ be a preCCS. Suppose that D is a set of pre–chains such that for any x in the root line of $[x]_C$ we have $x \approx_D x_r$. Then we have $\approx_C \subseteq \approx_D$.*

Proof. By induction on the depth of $[x]_C$ in $M(K)$ we show that $x \approx_C y$ implies $x \approx_D y$ for all $y \in T$. Suppose the claim is proved for all nodes of depth $< dp([x]_C)$. Let $y \approx_C x$. Pick an ascending map $\chi : x \to y$ in C. Its links are either maps from $\downarrow u$ to $\downarrow v$ where $dp([x]_C) > dp([u]_C)$ or link maps of some $u \approx_C x$. In the first case, we have by assumption $u \approx_D v$. In the second case v is in the root line of x and so $x \approx_D v$, by assumption on D. So, $x \approx_D y$. Q. E. D.

We describe now a general procedure for generating all possible preCCSs given a preMDS. In the light of the previous it is enough if we construct the preCCS given some $\mathfrak{T}$ and $\approx$ such that $\approx = \approx_D$ for some preCCS $\langle \mathfrak{T}, D\rangle$. We construct a sequence C_i in the following way. First, $C_0 := \varnothing$. Assume that C_i is constructed. If $\approx_{C_i} = \approx$, we stop. Otherwise, we choose x such that $dp([x]_D)$ is minimal such that $[x]_D \neq [x]_{C_i}$. Let the root line of x be $\{x_j : j < n\}$, where x_j ac–commands x_{j-1}, $0 < j < n$. (Then x_0 is the root.) For each x_j we shall find a y_j and a link map $\tau_j : \downarrow y_j \to \downarrow x_j$. Now, suppose y_j is found. Then fix a map $\tau_j : \downarrow y_j \to \downarrow x_j$ such that τ_j is an isomorphism of trees and $\tau_j(z) \approx z$ for all $z \leq y_j$. τ_j exists but is not unique. Simply proceed from y downwards in the tree, choosing for each member of $\downarrow y$ an appropriate image. (The isomorphism in case we deal with identifier trees may be of the form $\tau(J, K) : J; L \mapsto K; L$.) So, we only need to find y_j. For each $j < n$: choose y_j freely such that (i) $y_j \approx x_j$, (ii) x_j ac–commands y_j, (iii) x_{j-1} does not c–command y_j, (iv) there is an ascending map in C_i from x_{j-1} to y_j, and (v) there is no ascending map in C_i having y_j in its domain. (Notice that the identity is an ascending map. Hence x_{j-1} is a valid choice for y_j. Actually, (iv) makes (iii) redundant.) Then for all $0 < j < n$, $\langle \downarrow x_j, \tau_i, \downarrow y_j\rangle$ is added as a link. This adds one or several nontrivial chains. This defines C_{i+1}. Clearly, $C_i \subsetneq C_{i+1}$. Hence $\approx_{C_i} \subsetneq \approx_{C_{i+1}}$. Suppose that the construction ends with C_n. Finally, we add to C_n all trivial chains for elements that are not yet in a chain. Let the result be C. Then we have $\approx_D \subseteq \approx_C$ by Lemma 52. Next we show that $\langle \mathfrak{T}, C\rangle$ is a preCCS. We shall show *Liberation* and *No Recycling*, the other postulates being easy. Suppose that there is a chain containing two constituents $\downarrow y$ and $\downarrow y'$, and that $y, y' < x$. We show that $\downarrow x$ is in a trivial chain. It is clear that we may assume that there is a chain link to at least one element, say $\downarrow y$. Then y is in the root line, at this in turn means that x is in root position. (Otherwise, y cannot c–command its own root.) Now for *No Recycling*. The chain links are orbital. The proof is by induction on the construction. Suppose that C_i satisfies *No Recycling*. Let the new links be for the root line of x. Consider the peak map of x. It consists of a composition of the following maps:

$$x_0 \mapsto y_1 \mapsto x_1 \mapsto y_2 \mapsto \ldots y_{n-1} \mapsto x_{n-1}$$

The maps from x_{j-1} to y_j are compositions of link maps of C_i, and the maps from y_j to x_j are links maps of C_{i+1} by construction. Hence all added chain links are orbital, being members of the canonical decomposition of the zenith map. So we have CCS. Invoking Lemma 52 again, we get that $\approx_C \subseteq \approx_D$, whence the two are equal.

6 Constraints on Derivations

We have seen that there exist nonisomorphic CCSs with isomorphic MDSs. In fact, one can construct MDSs having exponentially many (depending on the number of nodes) nonisomorphic CCSs. However, in linguistic theory a number of constraints have been imposed on derivations that are reflected in the structure of CCSs. We discuss mainly two, namely *Freeze* (see [18] and references therein for a discussion of this principle) and *Shortest Steps* (see [5]).

Principle 1 (Freeze) *Suppose that $x < y$ and that $\downarrow y$ is a non–foot of some chain. Then $\downarrow x$ forms a trivial chain.*

There is another way to describe *Freeze*. Consider the canonical decomposition of an ascending map. If the derivation satisfies *Freeze*, a link map cannot follow a non–link–map. Since the targets of the link maps are exactly the members of the root line by Lemma 27, the following is clear.

Lemma 53 *A preCCS satisfies* Freeze *iff for all roots $\mathfrak{C}$, the root line of $\mathfrak{C}$ forms a pre–chain.*

The preCCS we have constructed to prove Theorem 51 therefore was a *Freeze*–preCCS. The next principle to be considered is the following.

> *Shortest Steps.* If two competing derivations for the same structure exist, then the one using the shortest steps is the only legitimate derivation.

The principle of *Shortest Steps* is not so easily made rigorous. There is a global variant (the sum of all steps is minimal) and a local variant (each step should be minimal). In the present context, we may simply take as a measure of complexity the following. Take a pair $\delta = \langle \downarrow y, \downarrow x \rangle$. Let z be the least common ancestor of x and y. The *length* of δ, $\lambda(\delta)$, is defined to be the sum of the sizes of the open intervals $]x, z[:= \{u : x < u < z\}$ and $]y, z[:= \{u : y < u < z\}$. For a chain Γ, let $\lambda(\Gamma)$ be the sum of all lengths of its chain links. Finally, let $\lambda(K)$ be the sum of all $\lambda(\Gamma)$, Γ a chain of K. Now, *Shortest Steps* can be reformulated as follows.

> *Shortest Steps (Global Version). Suppose K and K' are preCCS with the same preMDS. Suppose that $\lambda(K) < \lambda(K')$. Then $\lambda(K')$ represents an illegitimate derivation.*

However, this definition can be replaced by another (and slightly stronger) one, which is more concise and more handy in proofs. In fact, it characterizes not only that the sum of all lengths of links is minimal, but that each link is actually minimal.

Definition 54 *Let K be a CCS and $\lambda = \langle \mathfrak{C}, \varphi, \mathfrak{D} \rangle$ a link of K. λ is called **minimal** if no $\mathfrak{E}$ exists such that*

1. *$\mathfrak{C}$ ac–commands $\mathfrak{E}$,*
2. *$\mathfrak{E}$ ac–commands $\mathfrak{D}$ and*
3. *$\mathfrak{E}$ contains some $\mathfrak{D}'$ and there is an ascending map $\psi : \mathfrak{D} \to \mathfrak{D}'$.*

Principle 2 (Shortest Steps) *All links are minimal.*

We note that the definition of minimality coincides with another one that requires that the links should have shortest length. Notice the following. If a link is not minimal, then there exists an ascending map $\psi : \mathfrak{D} \to \mathfrak{D}'$. The following lemma asserts that under this condition the link does not have the least length of all competing links.

Lemma 55 *Let $\downarrow x, \downarrow y_1$, $\downarrow y_2$ be constituents and $\downarrow y_1 \neq \downarrow y_2$. Suppose that φ is a link map and ψ an ascending map such that $\varphi(\downarrow y_1) = \psi(\downarrow y_2) = \downarrow x$. Then $\lambda(\langle y_1, x\rangle) < \lambda(\langle y_2, x\rangle)$.*

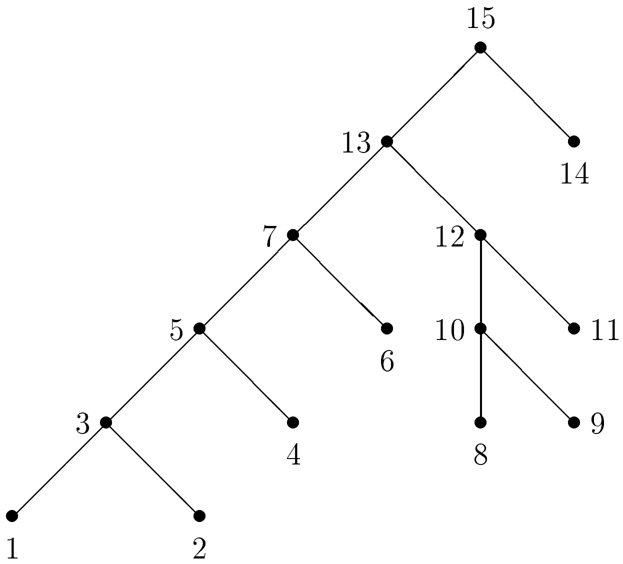
Proof. In the canonical decomposition, $\psi = \varphi \circ \chi$ for some ascending map χ which is not the identity. Hence, $\chi(y_2) = y_1$. Let z be the common ancestor of y_2 and y_1 and z' the common ancestor of y_1 and x. Now, since $z, z' \geq x$, either $z < z'$ or $z = z'$ or $z > z'$. But $z \geq z'$ cannot hold. For y_1 is in derived position. Hence it is contained in uniquely defined derived constituent $\mathfrak{C}$. $\mathfrak{C}$ ac–commands (and hence does not contain) $\downarrow y_2$. (This follows from *Liberation*. For inside $\mathfrak{C}$ there exist only trivial chains, so there exist no link maps. Hence no ascending map can send y_2 to y_1 if they would both be in $\mathfrak{C}$. To establish ac–command between $\mathfrak{C}$ and $\downarrow y_2$ observe that they cannot be sisters, since y_2 must be in some member of the chain containing $\mathfrak{C}$.) Likewise there is a uniquely defined constituent $\mathfrak{E}$ ac–commanding $\downarrow y_1$ and containing x. Then z' is the mother of the generator of $\mathfrak{E}$. $\mathfrak{E}$ also ac–commands $\downarrow y_2$ and so $\mathfrak{D}$. Let z be the mother of the generator of $\mathfrak{C}$. Then $z' > z$. Hence, z' is also the common ancestor of x and y_2. Further, $card(]y_1, z[) < card(]y_2, z[)$, since χ is ascending, and so $card(]y_1, z'[) < card(]y_2, z'[)$. Now

$$\begin{aligned} \lambda(\langle y_2, x\rangle) &= card(]y_2, z'[) + card(]x, z'[) \\ &> card(]y_1, z'[) + card(]x, z'[) \\ &= \lambda(\langle y_1, x\rangle) \end{aligned}$$

This shows the claim. Q. E. D.

The principle *Shortest Steps* favours subextraction. So remnant movement is assumed only when the remnant moves out of the c–command domain of the antecedent. Otherwise, subextraction is favoured. Therefore, call a preCCS **cyclic** if for all Γ and Δ such that $\mathfrak{C}_i$ covers $\mathfrak{D}_j$, $\mathfrak{D}_{i-1}$ does not c–command $\mathfrak{C}_{j-1}$.

Figure 14: $T(\mathfrak{A})$



Theorem 56 *For every preMDS* $\mathfrak{A}$ *there exists up to isomorphism exactly one preCCS* K *satisfying Shortest Steps such that* $M(K) \cong \mathfrak{A}$. K *is also cyclic.*

Theorem 57 *For every preMDS* $\mathfrak{A}$ *there exists up to isomorphism exactly one preCCS* K *satisfying Freeze such that* $M(K) \cong \mathfrak{A}$.

Now, to construct a preCCS satisfying *Shortest Steps*, we must define the links in such a way that a member of the root line is connected with the 'highest' possible member. We have seen in Lemma 55 that there is a unique element in the derivation that minimizes the length of the link.

The following is clear. Given $K = \langle \mathfrak{T}, C \rangle$, there is a natural isomorphism S from $\mathfrak{T}$ onto the tree $T(M(K))$. It is defined by induction on the depth of x. Namely, if x is the root then $S(x) := [x]_C$. If $x \prec y$ then $S(x) := S(y); [x]_C$. Now, given $M(K)$, we will construct a CCS over $T(M(K))$. Using the isomorphism $T := S^{-1}$, this then gives back the original structure K.

Lemma 58 *Let* K *be a preCCS and let* I, J *be identifiers of* $M(K)$. *Put* $I \approx J$ *iff* I *and* J *identify the same element. Then* $I \approx J$ *iff* $T(I) \approx T(J)$.

Let us illustrate the construction with our preMDS $\mathfrak{A}$. We reproduce Figure 13 in Figure 14, using numbers instead of identifiers. The equivalence relation contains as non–singleton sets only $\{2, 4, 9, 11, 14\}$ and $\{5, 12\}$. We have two nontrivial root lines, namely $\{5, 12\}$ and $\{2, 4, 14\}$. We start with the largest constituent, $\downarrow 5$. Here, we have no choice but to link 12 to 5. Next, we proceed to 2. We must link 4 and 14 to some member of the equivalence class of 2, $\{2, 4, 11, 14, 9\}$. 4 can only be linked to 2, since this is the only element that it ac–commands. Now, the link $\langle 4, 2\rangle$ must be an orbital link, and so 14 may not be linked to 9 or to 2. In the procedure that we have defined this is taken care of by the requirement that there must be an ascending map from 4 to the link partner of 14. There is no ascending map from 4 to 9. Hence the only choices are 4 and 11. The results are shown in Figure 15 and 16. The derivations satisfy *Freeze* and *Shortest Steps*, respectively.

We can now define the notion of a derivation on preMDSs. Say that y is **not derived** if either y is the root or $M(y) = \{z\}$, where z is not derived. (So, above y the relation $<$ is linear.)

Definition 59 *Let $\mathfrak{A} = \langle A, \prec_A\rangle$ and $\mathfrak{B} = \langle B, \prec_B\rangle$ be preMDSs. $\mathfrak{B}$ is a* ***1–step link extension*** *of $\mathfrak{A}$ if $B = A$ and $\prec_B = \prec_A \cup\{\langle x, y\rangle\}$, where $\langle x, y\rangle \notin \prec_A$ and y is not derived. $\mathfrak{B}$ is a* ***link extension*** *of $\mathfrak{A}$ if there exists a finite sequence $\langle \mathfrak{B}_i : i < n + 1\rangle$ such that $\mathfrak{A}_0 = \mathfrak{A}$, $\mathfrak{B}_n = \mathfrak{B}$ and $\mathfrak{B}_{i+1}$ is a 1–step link extension of $\mathfrak{B}_i$ for $i < n$.*

Theorem 60 *Every preMDS is a link extension of a tree. More precisely, the following holds.*

1. *If L is isomorphic to a 1–step extension of K then $M(L)$ is isomorphic to a 1–step link extension of $M(K)$.*
2. *If $\mathfrak{A} = M(K)$ and $\mathfrak{B}$ is a 1–step link extension of $\mathfrak{A}$ then there exists a 1–step extension L of K such that $M(L) \cong \mathfrak{B}$. L is unique up to isomorphism.*

Proof. Suppose that $M(K) = \mathfrak{A}$, $K = \langle \mathfrak{T}, C\rangle$. Inductively, we may assume that $\mathfrak{A}$ consists of the root elements of K. Then suppose that $L = \langle \mathfrak{U}, D\rangle$ is obtained from K by a 1–step extension, adding the link $\langle \mathfrak{D}, \varphi, \mathfrak{C}\rangle$. Suppose that $\mathfrak{C} = {\downarrow}x$, $\mathfrak{D} = {\downarrow}z$ and that $y \succ z$. Then add the link $\langle x, y\rangle$ to $\mathfrak{A}$. This defines $\mathfrak{B}$. By definition of 1–step extensions, y is in root and in zenith position, and so is every $u \geq y$. So ${\uparrow}y$ is linear in $\mathfrak{A}$. Hence, $\mathfrak{B}$ is a 1–step link extension of $\mathfrak{A}$. Moreover, $u \approx_D v$ iff (1) $u, v \in T$ and $u \approx_C v$ or (2) $u \notin T$ and $v \in T$ and $v \approx_C \varphi^{-1}(u)$ or (3) $u \in T$, $v \notin T$ and $u \approx_C \varphi^{-1}(v)$ or (4) $u, v \notin T$ and $\varphi^{-1}(u) \approx_C \varphi^{-1}(v)$. Hence the map $\{y \in T : y \approx_K x\} \mapsto \{y \in U : y \approx_L x\}$ is a bijection between the equivalence classes of T and the equivalence classes of U. Furthermore, $[x]_D \prec [y]_D$ iff for some $x' \approx_D x$ and some $y' \approx_D y$ we have $x' \prec_U y'$. Now, $x' \prec_U y'$ iff $x \prec_U y$ iff $[x]_C \prec_T [y]_C$ for $x', y' \in T$. Careful examination of the other cases yields the result that $M(L) \cong \mathfrak{B}$. Now, assume that K, $\mathfrak{A}$ and $\mathfrak{B}$ are given, and $\mathfrak{A} = M(K)$ and $\mathfrak{B}$ is a 1–step link extension of $\mathfrak{A}$. Then there is a $y \in A$ such that $M(y) = \{z\}$, and an $x \in A$ such that $\prec_B = \prec_A \cup\{\langle x, y\rangle\}$, and $\langle x, y\rangle \notin\prec_A$. We may actually identify the set of nodes of K with the set of identifiers of K. Now, choose an identifier $I; x$ of x in zenith position. (This identifier is actually unique.) There is a unique identifier J of y by assumption that y is not derived. Now, put $U := T \cup \{J;x;K : I;x;K \in T\}$. $\prec_U$ is canonically defined. Let the chain links of L be all chain links of K plus $\langle \downarrow J;x, \varphi, \downarrow I \rangle$, where $\varphi : I;x;K \to J;x;K$. Then L is obtained from K by 1–step extension. Q. E. D.

We remark that isomorphism can be replaced by identity if the structure $M(K)$ is constructed on the root elements of K rather than their equivalence classes. The upshot of the theorem is as follows. Look at the diagram below

$$\begin{tikzcd}
K \arrow[r, "E"] \arrow[d, "M"'] & L \arrow[d, "M"] \\
\mathfrak{A} \arrow[r, "E"] & \mathfrak{B}
\end{tikzcd}$$

We have denoted by E the relation of 1–step extension and 1–step link extension, respectively. The idea is that whenever three of the four objects of this diagram are given then the fourth one and the maps to it are determined by the others up to isomorphism. For example, if $\mathfrak{A}$, K and L are given, $\mathfrak{B}$ is determined up to isomorphism by $\mathfrak{B} \cong M(L)$. If $\mathfrak{A}$, K and $\mathfrak{A}$ are given, L is determined up to isomorphism, and if $\mathfrak{A}$, $\mathfrak{B}$ and L are given, then K is unique up to isomorphism. It is relatively easy to see why this is so. Let us be given $\mathfrak{A}$ and $\mathfrak{B}$. If K is known, we construct L by adding the link in form of a copy that is added to $\mathfrak{A}$ to obtain $\mathfrak{B}$. We know where to put the chain link since we must choose in K a zenith element. Conversely, if L is known, we can construct K by removing the constituent that is highest and corresponds to x, where the link $\langle x, y\rangle$ has been retracted.

We close this section with some remarks on alternative optimality criteria on derivations. A *Freeze*–derivation satisfies a minimality principle other than *Shortest Steps*, namely *Fewest Chains*. The reader may verify in the given example that a derivation satisfying *Freeze* has the least number of all chains. This is no accident.

Principle 3 (Fewest Chains) *Suppose that K and K' are preCCSs such that $M(K) \cong M(K')$ and that K has fewer nontrivial chains than K'. Then K' is illegitimate.*

Theorem 61 *If a preCCS satisfies **Freeze** then it satisfies **Fewest Chains**. Furthermore, the number of chains can be computed on the corresponding preMDS by*

$$\gamma(\mathfrak{A}) := \mathit{card}\{x : \mathit{card}(M(x)) > 1\}\ .$$

$\gamma(\mathfrak{A})$ *is called the chain number of* $\mathfrak{A}$.

Proof. Given $\mathfrak{A}$, there exists up to isomorphim a unique K satisfying *Freeze* such that $M(K) \cong \mathfrak{A}$. The number of chains of K equals the number of nontrivial root lines. This is exactly the number $\gamma(M(K))$. Since $M(K) \cong \mathfrak{A}$, $\gamma(M(K)) = \gamma(\mathfrak{A})$. It is clear that $\gamma(\mathfrak{A})$ is the minimum number of chains in K, since we know that for each element with $card(M(x)) \neq 1$ there must be a chain, and if $x \neq y$ these chains must be different. Q. E. D.

What happens if we try to prove the converse? Consider the following strategy. Suppose, *Freeze* is violated. Then there is a chain–link $\langle \downarrow J, \downarrow I\rangle$ such that $\downarrow I$ is in derived position. That is, there is a $\downarrow K$ properly covering $\downarrow I$, and $\downarrow K$ is in a chain with foot $\downarrow K'$, $K' \neq K$. Now $\downarrow K'$ contains some I' such that $I' \approx I$. Remove the chain link and add instead the link $\langle \downarrow J, \downarrow I'\rangle$. Continue this procedure until $\downarrow I'$ becomes a member of some chain Γ (this must happen at some point). At this point, the chain Δ containing $\downarrow J$ and Γ become fused, since $\downarrow I$ is replaced by $\downarrow I'$. So, eventually some two chains become one, and therefore the original derivation did not satisfy *Fewest Chains*. The success of this argumentation rests on the condition that Γ is a nontrivial chain. In that case the argument carries through. Otherwise not. An examples is given in the Figures 17 and 18. The first derivation does not satisfy *Freeze*, the second does. Yet both have two nontrivial chains, each with two members, and 5 trivial chains. In the first derivation we have the offending chain link $\langle \downarrow 8, \downarrow 6\rangle$, since there is the chain $\Theta = \{\downarrow 7, \downarrow 3\}$. If we use the proposed procedure and exchange $\downarrow 2$ for $\downarrow 6$ in Δ, Δ is fused with the chain containing $\downarrow 2$. But here $\Gamma = \{\downarrow 2\}$ is a singleton chain, and we end up with a *Freeze*–derivation that has no less chains, shown in Figure 18. We remark here finally the following. Define the **link number** of $\mathfrak{A}$ as

$$\nu(\mathfrak{A}) := \sum_{x \in A - \{r\}} (card M(x) - 1)$$

where r is the root of $\mathfrak{A}$. We notice that this number counts how many non–surface mothers an element has. This is exactly the number of chain links that x must enter in a derivation, no matter in what ways the derivation is composed. For the following can be established.

Theorem 62 *Let K be a preCCS. Then the number of chains links in K is exactly $\nu(M(K))$. More precisely, if x is not a root, then there are exactly $card(M(x)) - 1$ many chain links involving some $\downarrow I$, where I identifies x.*

It follows that if one tries to optimize the number of links, one does not reduce the set of competing derivations. An alternative proof is as follows. Observe that if only $\mathfrak{T}$ and $\approx$ are given, we can compute the number of links as follows. For every class $[I]$, let $\rho([I])$ be the size of the root line of $\downarrow I$ minus 1. Notice namely

Figure 15: A *Freeze* Derivation

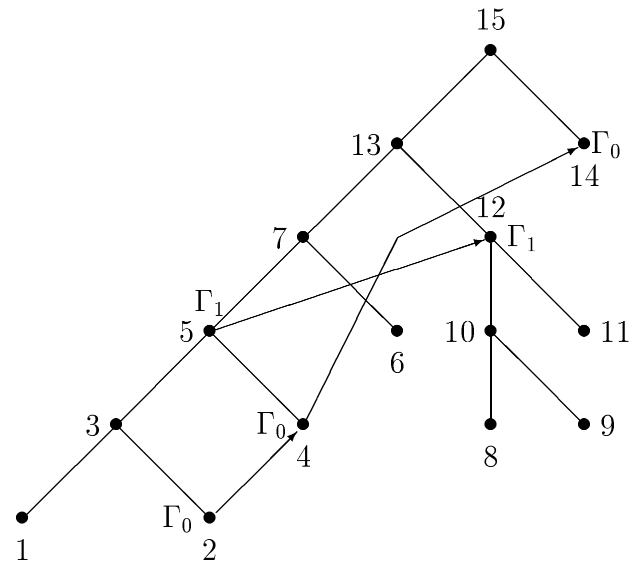


Figure 16: A *Shortest Step* Derivation

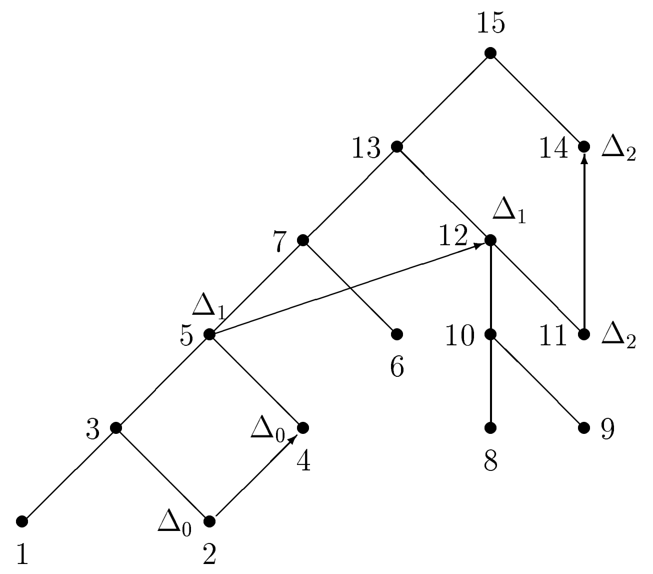


Figure 17: *Freeze* Contra *Fewest Steps* I

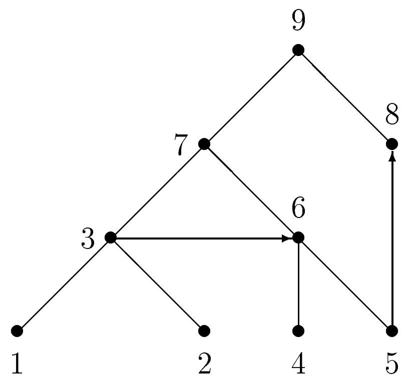
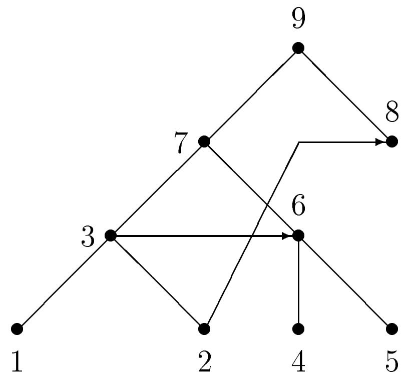


Figure 18: *Freeze* Contra *Fewest Steps* II



that $\downarrow J$ is in the root line of $[I]$ if $J \in [I]$ and for the unique root (= longest member) I' of $[I]$, J ac–commands I. Now, put

$$\rho(\mathfrak{T}, \approx) := \sum_{[I]} \rho([I])$$

This number is exactly the number of chain links. For by *Liberation*, any chain link adds some member to the root line of some element. So, each step in the derivation increases the number ρ by 1.

7 The Surface Element

At this point it becomes essential to remove a gap in our definitions, namely that of the surface elements. We have already defined the root elements and the zenith elements. Furthermore, both the set of root elements and the set of zenith elements are upward closed in the tree. This is essential. We want that at any stage of the derivation the zenith elements form a tree. Since the root elements together and the surface elements together represent stages of the derivation, we want that the set of surface elements is upwards closed as well. The definition of surface element is as follows.

Definition 63 *Let K be a CCS and Γ a chain. A link $\lambda = \langle \mathfrak{D}, \varphi, \mathfrak{C} \rangle$ of Γ is called **visible** if $\mathfrak{D}$ does not ac–command the designated element of Γ. A link map is **visible** iff its associated link is visible.*

Definition 64 *Let K be a CCS. x is called a **visible element** of K (and $\downarrow x$ a **visible constituent**) if the canonical decomposition of the root map of x contains only visible link maps. x is called a **surface element** (and $\downarrow x$ a **surface constituent**) if x is the highest visible element in its orbit with respect to $\lhd$. The set of surface elements of K is denoted by $S^\star(K)$.*

Equivalently, since the next higher element in an orbit is reachable by means of a link map, x is a surface element iff it is orbital and the canonical decomposition of its root map is maximal with respect to using only visible links.

Lemma 65 *For every x, $[x]_C \cap S^\star(K)$ contains exactly one member. This is called the **surface element** of $[x]_C$, and denoted by x_s.*

Lemma 66 *Let $x < y$. Then the canonical decomposition of the root map of y is a prefix of the canonical decomposition of the root map of x.*

Proof. It is enough to show the theorem for the case $x \prec y$. Let ψ be the root map of y. Put $\widehat{x} := \psi^{-1}(x)$. Then $\psi(\widehat{x}) = x \prec y = \psi(y_r)$, and so $\widehat{x} \prec y_r$. Now, let χ be the root map of $\widehat{x}$. Since $\widehat{x} \approx x$, x_r is the root of $\widehat{x}$. Hence $\chi(x_r) = \widehat{x}$ and so $x = \psi \circ \chi(x_r)$. The product of the decompositions for χ and ψ is a decomposition of the root map of x, and by uniqueness of the latter, it is the canonical decomposition. This shows the claim. Q. E. D.

Lemma 67 *If x is visible and $x < y$ then also y is visible. Further, if $x \in S^\star(K)$ then also $y \in S^\star(K)$.*

Proof. Suppose that x is visible and $x < y$. Then the root map of x is a composition of visible link maps, and so the root map of y is a composition of visible link maps, by the previous theorem. Now assume that $x \in S^\star(K)$. Let χ be the root map of x. Let $x \prec y$ and ψ be the root map of y. Then y is visible. Suppose that there is a visible link map φ such that $\varphi(y)$ is also visible. We must show that φ is not orbital for y. Consider the link $\langle \mathfrak{D}, \varphi, \mathfrak{C}\rangle$. It is visible. Therefore $\varphi(x)$ is visible. Clearly, if φ is orbital for x, we would have a contradiction to the choice of x. So, the link is not orbital for x. This means, however, that $x \in \mathfrak{C}'$ where we have another link, $\langle \mathfrak{D}', \varphi', \mathfrak{C}'\rangle$. This link is orbital. By choice of x, φ' is not visible. Hence, $\mathfrak{C}'$ is the designated element of its chain and $\mathfrak{D}'$ an LF–member. Clearly, $\mathfrak{C}' \neq \mathfrak{C}$. Suppose that $\mathfrak{C}' \subsetneq \mathfrak{C}$. Then we have a violation of *Synchronization*: $\mathfrak{D}$, a non–LF–member, blocks $\mathfrak{D}'$, an LF–member. Suppose now $\mathfrak{C} \subsetneq \mathfrak{C}'$. Then the link map φ is not orbital for y. This had to be shown. Q. E. D.

Naively, one would expect that the surface elements of K show up as subsets of the underlying set of $M(K)$. However, as each equivalence class $[x]_K$ intersects nontrivially with $S^\star(K)$, every element of $M(K)$ would thus be a surface element. Therefore, we resort to a different solution, which is generic for the way in which the tree based notions are transformed into (equivalent) notions about multidominance structures. Instead of S we take the relation $\Sigma := \{\langle x, y\rangle : x \prec y, x \in S^\star(K)\}$. (We may require also that $y \in S^\star(K)$, but this amounts to the same, according to Lemma 67.) We define

$$\propto := \{\langle [x]_K, [y]_K\rangle : x \,\Sigma\, y\}$$

We call $\propto$ the **surface relation**. A member of $\propto$ is also called a **surface link**. We notice the following fact, which basically follows from Lemma 67.

Lemma 68

1. *$[x]_K \propto [y]_K$ iff $x_s \prec y_s$.*
2. *If $[x]_K \propto [y]_K$ then also $[x]_K \prec [y]_K$.*
3. *Let $[x]_K$ be distinct from the root. Then there exists exactly one $[y]_K$ such that $[x]_K \propto [y]_K$.*

This shows some necessary conditions on $\propto$. There is an additional condition, which enshrines the condition of *Synchronization*.

Definition 69 *Let $\mathfrak{A}$ be a preMDS. The link $\langle x, y\rangle \in\prec$ is **higher** than the link $\langle x', y'\rangle \in\prec$ if $x = x'$ and $y > y'$. A link is called a **root link** (of x) if it is a lowest link (of x). A link (of x) is an **invisible link** (of x) if it is higher than a surface link (of x).*

Definition 70 *A* ***multidominance structure (MDS)*** *is a triple* $\langle G, \prec, \propto \rangle$ *such that*

1. $\langle G, \prec \rangle$ *is a preMDS.*
2. $\propto \subseteq \prec$.
3. *For every non–root* x *there exists exactly one* y *such that* $x \propto y$.
4. *If* $\langle x, y \rangle$ *is an invisible link, then all non–root links of* y *are invisible as well.*

Definition 71 *Let* $\mathfrak{A} = \langle A, \prec, \propto \rangle$ *and* $\mathfrak{B} = \langle B, \prec', \propto' \rangle$ *be MDSs. Then* $\mathfrak{B}$ *is a* ***1–step link extension*** *of* $\mathfrak{A}$ *if*

1. $\langle B, \prec' \rangle$ *is a 1–step link extension of* $\langle A, \prec \rangle$.
2. *If the newly added link* $\langle x, y \rangle$ *is visible,* $\mathfrak{A}$ *contains no invisible links and* $\propto' = (\propto - \{\langle x, z \rangle\}) \cup \{\langle x, y \rangle\}$, *where* $\langle x, z \rangle$ *is the highest link of* x *in* $\mathfrak{B}$.
3. *If the newly added link is invisible,* $\propto' = \propto$.

Locutions such as **is a link extension of** etc. are straightforwardly defined. We extend our theorems on precCCSs and preMDSs to the full case.

We will first show that an analogous theorem of Theorem 44 can be proved for MDSs. A MDS $\langle A, \prec, \propto \rangle$ is a **tree** if $\langle A, \prec \rangle$ is a tree (and so $\propto = \prec$).

Lemma 72 *Let* $\langle A, \prec, \propto \rangle$ *be a structure such that* $\prec, \propto \subseteq A \times A$. $\mathfrak{A}$ *is an extension of a tree iff it is an MDS.*

Proof. The claims are proved by induction. First, if $\mathfrak{A}$ is a 1–step extension of $\mathfrak{B}$ and $\mathfrak{B}$ is a MDS then so is $\mathfrak{A}$. A tree is an MDS. This shows the left–to–right direction. From right to left, let $\mathfrak{A}$ be an MDS. Suppose it contains an invisible link. Then let x be such that there is an invisible link $\langle x, y \rangle$ but no $z \geq y$ has invisible links. Then all $z \geq y$ have only a root link. Define $\mathfrak{B} := \langle A, \prec', \propto \rangle$, where $\prec' := \prec - \{\langle x, y \rangle\}$. It is easy to see that $\mathfrak{B}$ is an MDS and that $\mathfrak{A}$ is a 1–step link extension of $\mathfrak{B}$. Now suppose that $\mathfrak{A}$ contains only visible links. Then there is a link $\langle x, y \rangle$ such that all $z \geq y$ have only root links. Let $\langle x, z \rangle$ be the next lower link of x. Define $\mathfrak{B} := \langle A, \prec', \propto' \rangle$ where $\prec' := \prec - \{\langle x, y \rangle\}$ and $\propto' := (\propto - \{\langle x, y \rangle\}) \cup \{\langle x, z \rangle\}$. Then $\mathfrak{B}$ is an MDS and $\mathfrak{A}$ is obtained by a 1–step extension. Clearly, this procedure comes to a halt when there are no non root links. But then we have a tree. Q. E. D.

The theorem can be made more precise than that. We will show that there is a bijective map between the derivations of CCSs and derivations of MDSs (modulo isomorphism of structures). That means, given a derivation of a CCS there is a unique derivation on MDS matching it (via $M(-)$) and given a derivation on the MDSs there is a unique derivation on the CCSs. Hence, on the derivational side there is a complete match. Nevertheless, as we have seen earlier, more than one CCS may correspond to single MDS. So, the MDSs actually forget part of the derivational history.

Theorem 73 *The following holds.*

1. *If L is isomorphic to a 1–step extension of K then $M(L)$ is isomorphic to a 1–step link extension of $M(K)$.*
2. *If $\mathfrak{A} = M(K)$ and $\mathfrak{B}$ is a 1–step link extension of $\mathfrak{A}$ then there exists a 1–step extension L of K such that $M(L) \cong \mathfrak{B}$. L is unique up to isomorphism.*

Proof. We follow the proof of Theorem 60. Suppose that $M(K)$ and L is isomorphic to a 1–step extension of K. Then we have to show that if $M(K)$ is an MDS, so is $M(L)$, and it is isomorphic to a link extension of $M(K)$. Now $M(L)$ is obtained from $M(K)$ by adding a new link. Case I. Suppose that this link is a visible link. Then we have added a visible chain link $\langle \mathfrak{D}, \varphi, \mathfrak{C} \rangle$ to K. $\mathfrak{C}$ is in zenith in K, and $\mathfrak{D}$ in zenith in L. So, $\mathfrak{D}$ is the new surface constituent of $[\mathfrak{C}]$. Let y be the generator of $\mathfrak{D}$ and $z \succ y$. Now take an element x. We wish to show that if $[x]_L \neq [y]_L$ then $[x]_L \propto [u]_L$ iff $[x]_K \propto [u]_K$ and that $[y]_L \propto [z]_L$. (Here, x_s^K denotes the surface element of x in K.) Assume first $[x]_L \neq [y]_L$. Then $[x]_L \propto [u]_L$ iff $x_s^L \prec u_s^L$. Two cases arise. (I.a) φ is orbital for x, (I.b) is not orbital for x. In Case (I.a), $x_s^L = \varphi(x_s^K)$. Since $[x]_K \neq [y]_K$, φ is defined on u_s^K. Now $u_s^L = \varphi(u_s^K)$ since the link must also be orbital for u. Now we have

$$\begin{aligned} [x]_L \propto [u]_L \quad & \text{iff } x_s^L \prec u_s^L \\ & \text{iff } \varphi(x_s^K) \prec \varphi(u_s^K) \\ & \text{iff } x_s^K \prec u_s^K \\ & \text{iff } [x]_K \propto [u]_K \end{aligned}$$

Case (I.b). Then $x_s^L = x_s^K$ and so $u_s^L = u_s^L$. The claim now follows. Finally, assume that $[x]_L = [y]_L$. The link φ is orbital, hence $[y]_L \propto [u]_L$ iff $[u]_L = [z]_L$. Case II. An invisible link has been added. Then $x_s^L = x_s^K$ for all $x \in T$. It follows that for all $x, u \in T$: $[x]_L \propto [u]_L$ iff $x_s^L \prec u_s^L$ iff $x_s^K \prec u_s^L$ iff $[x]_K \propto [u]_K$. Since for every class $[v]_L$ there exists an $x \in T$ such that $[x]_L = [v]_L$, the claim is shown. The first claim is now entirely proved. Now we assume that $\mathfrak{B}$ is a 1–step link extension of $M(K)$. We need to find an L such that $M(L) \cong \mathfrak{B}$. Let the new link be $\langle x, y \rangle$. Case I. This link is visible. Then all links of $M(K)$ are visible. We add to K the corresponding link. This is unique, by Theorem 60. Make the newly added constituent the designated member of its chain. It is routine to check that $M(L) \cong \mathfrak{B}$. Now assume that an invisible link has been added. Again, L is canonically defined. We have to check that $M(L)$ is an MDS. But this is clear from Theorem 60 and the fact that $\propto' = \propto$ (under identification). Q. E. D.

8 Trace Chains

By far the most popular chain is the *trace chain.* In a trace chain, there are two types of elements. At most one element is not a trace. Proposition 78 shows us that trace chain structures are naturally isomorphic to MDSs, so that working with either of them is a matter of convenience, not of empirical of theoretical substance. It turns out that the *No Recycling* condition reduces to the requirement that there always exists a constituent in a given chain that is not a trace. This member is usually taken to be the head of the chain, but that need not be the case. One may decide, for example, that it be the surface member. (That would save marking the surface member of the chain separately, given that *No Trace Recycling* holds.) However, we will follow here the conventional approach and take non–trace to be the head of the chain. Typically, a trace is a two–node tree consisting of a nonterminal label dominating a terminal node with label t. (In earlier literature, the symbol for the empty string, e, was used, since the theory was mainly formulated in terms of string rewriting.) By our definition of chains, no indexation is required. Further, it is rather unnecessary (and technically inconvenient) to have the node carrying the trace–label t. It is enough to keep the mother. So, an NP–trace is simply a one–node tree carrying the label NP. However, labels are irrelevant in the present context.

Definition 74 *A* ***trace*** *is a one–node tree. Let $\mathfrak{T}$ be a tree. A* ***trace pre–chain*** *over $\mathfrak{T}$ is a pre–chain over $\mathfrak{T}$ such that all non–heads are traces. A* ***trace chain*** *is a chain $\langle \mathfrak{C}, \mathbb{C} \rangle$, where $\mathbb{C}$ is a trace pre–chain.*

Definition 75 *A* ***trace chain tree*** *is a pair $\langle \mathfrak{T}, C \rangle$, where $\mathfrak{T}$ is a tree and C a set of trace chains over $\mathfrak{T}$.*

It is clear how movement shall be formulated using trace chains. Given a trace chain, $\Gamma = \langle \mathfrak{C}, \mathbb{C} \rangle$, 1–step extension is a chain $\Delta = \langle \mathfrak{D}, \mathbb{D} \rangle$, where all non–head members of $\mathbb{C}$ are in $\mathbb{D}$, the head member is replaced by a trace, and a copy is placed into the head of $\mathbb{D}$. The problem with trace chains as opposed to copy chains is that the trees and the chains do not grow, but change in a non–increasing way. Rather than giving more details we will show how one can pass directly from a CCS to a trace chain. In doing so, we avoid having to spell out exactly what sort of structures we obtain from trees by using copy–and–delete creating trace chains. Using the zenith–elements we define the trace chain structures. The basic idea is to eliminate all nodes that are below an empty node, since we do not allow traces inside traces. However, by eliminating these nodes we may actually kill some traces. For example, if $\mathfrak{D}$ covers the chain Γ entirely, and $\mathfrak{D}$ is the foot of some chain Δ, then $\mathfrak{D}$ is reduced to a trace, and with it the chain Γ. For although the head of Δ contains a record of the members, the chain is lost. To circumvent this problem, we first pass to the corresponding MDS and define the trace chain structure from the MDS. For the purpose of the next definition, let us say that for x and y members of an MDS, x is in **lf–relation** with y if $\langle x, y\rangle$ is the highest link of x.

Definition 76 *Let $\mathfrak{M}$ be an MDS. An identifier I of $\mathfrak{M}$ is an **LF–identifier** if either (a) $I = x$, where x is the root, or (b) $I = J; y; x$, where $J; y$ is an LF–identifier and x is in lf–relation with y. Call an identifier I a **trace identifier** if it is not an LF–identifier and of the form $I = J; x$, where J is an LF–identifier. Call I **essential** if it is either an LF–identifier or a trace. $E(\mathfrak{M})$ denotes the set of all essential identifiers of $\mathfrak{M}$.*

Definition 77 *Let $I \in E(\mathfrak{M})$ be an essential identifier identifying x. Then there is a unique $J^\star$ such that $J^\star$ identifies y, and y is the surface element relation of $[x]$. Then let*

$$\Gamma(I) := \langle \{\downarrow J \cap E(\mathfrak{M}) : \downarrow J \approx \downarrow I\}, \downarrow J^\star \cap E(\mathfrak{M})\}$$

*and call $\Gamma(I)$ the **chain associated with** I. Finally, put*

$$TC(\mathfrak{M}) := \langle E(\mathfrak{M}), < \cap E(\mathfrak{M})^2, \{\Gamma(I) : I \in E(\mathfrak{M})\rangle$$

*$TC(\mathfrak{M})$ is called the **trace chain structure (TCS)** associated with $\mathfrak{M}$.*

Let K be a TCS. We define as before the relation $\approx_K$. This allows to derive an MDS in just the same way as with a CCS. For — as the reader may notice — the definition of $M(K)$ did not depend on the fact that it consisted of copy chains. The result is now the following.

Proposition 78 *Let $\mathfrak{M}$ be an MDS and K a TCS. Then*

1. *$\mathfrak{M} \cong M(TC(\mathfrak{M}))$.*
2. *$K \cong TC(M(K))$.*

Proof. First, notice that if $I \in E(\mathfrak{M})$ and $I \approx_K J$ then I and J are in fact members of the same chain. Hence, $I \approx_K J$ iff I and J identify the same element in $\mathfrak{M}$. Hence we see that $\mathfrak{M} \cong M(TC(\mathfrak{M}))$. Now consider a TCS K. By definition there is an $\mathfrak{M}$ such that $K = TC(\mathfrak{M})$. Hence $M(K) = M(TC(\mathfrak{M})) \cong \mathfrak{M}$ and so $TC(M(K)) \cong TC(\mathfrak{M}) = K$. Q. E. D.

Now, consider a derivation using copy–and–delete in the typical sense. We want to show that the structures we get through this process are exactly the TCSs. To see this, we need to do only an induction on the derivation and monitor the effect on both sides. Consider a step of extending a copy chain Γ, forming K_2 from K_1. It corresponds to adding a link in the MDS $M(K_2)$. (For simplicity we are dealing in fact only with the relevant pre–chains.) Constructing the trace chains we see that it corresponds indeed to copying the head of the chain Γ under the node where the highest link has been added, and reducing the previous head of Γ to a trace. This process can be reversed. If we make a copy–and–delete step in $TC(M(K_2))$, it amounts to adding a new link in $M(K_1)$, and this in turn corresponds to copying the head of Γ to the new position.

$$\begin{tikzcd}
K_1 \arrow[r] \arrow[d] & M(K_1) \arrow[r] \arrow[d] & TC(M(K_1)) \arrow[d] \\
K_2 \arrow[r] & M(K_2) \arrow[r] & TC(M(K_2))
\end{tikzcd}$$

Hence, the definition of the trace chain structures is complete. These are exactly the structures that can be obtained from trees using copy–and–delete.

Now, how are TCSs defined directly? Why do we need such a roundabout way of defining these structures? The answer is that trace chains do not even satisfy the simplest requirement: that the trace is ac–commanded by its antecedent. Consider for example the structure in Figure 19. Suppose 2 moves to 4 and then to 6. Suppose next that 5 moves to 12. The corresponding trace chain structure is shown in Figure 20 (with circles around the traces). We see that the first chain link ($\langle\downarrow 4, \downarrow 2\rangle$) is copied onto the chain link $\langle\downarrow 11, \downarrow 9\rangle$, where ac–command holds. [12] However, the second chain link $\langle\downarrow 4, \downarrow 6\rangle$ is now transformed into the chain link $\langle\downarrow 11, \downarrow 6\rangle$. No ac–command relation holds. This problem appears for example in Topicalization in German. Assuming that the arguments of the verb are in canonical order in base position, alternate serializations can only be realized by moving the objects out of the verb phrase. Also, all the objects can be moved out of the verb phrase and the verb therefore forms a verb phrase by itself. Suppose this happens and that the verb phrase is subsequently topicalized, as in example (5).

(5) *Verkaufen wollte er das Auto nie.*
Sell wanted he the car never.
He never wanted to sell the car.

The full derivation is the following.

(6) *e e e e nie [er das Auto verkaufen] wollte.*
e e e [das Auto]$_1$ *nie [er* t_1 *verkaufen] wollte.*
e e er$_2$ *[das Auto]*$_1$ *nie [*t_2 t_1 *verkaufen] wollte.*
e wollte$_3$ *er*$_2$ *[das Auto]*$_1$ *nie [*t_2 t_1 *verkaufen]* t_3.
[t_2 t_1 *verkaufen]*$_4$ *wollte*$_3$ *er*$_2$ *[das Auto]*$_1$ *nie* t_4 t_3.

Several transformations have to apply. The direct object has to move out of the verb phrase; the subject moves out of the verb phrase. The inflected auxiliary moves to INFL/COMP and finally, the verb phrase is topicalized. Most chains are

[12] Indeed, in contrast to CCSs, we must assume that when an element moves it takes along all chains that have been defined on elements below it. Otherwise, practically all information concerning the structure is lost.

Figure 19: The Copy Chain Structure K

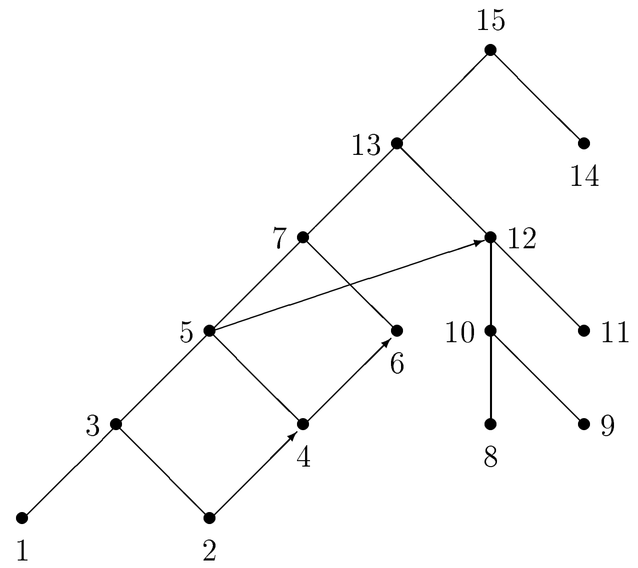
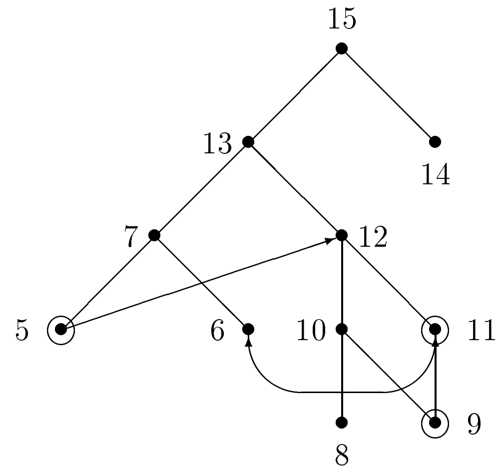


Figure 20: The corresponding Trace Chain Structure



unproblematic, but the last step raises a number of problems, which are adressed, for example, in Müller [18]. For it makes the subject trace (t_2) and the object trace (t_1) unbound.

Under a strict reading of GB, these structures would have to be excluded. For traces have to be bound at all levels, which means here in particular that in the trace chain structure at the end of the derivation all chains must be linearly ordered by ac–command. There is no way to reconcile these requirements without making nontrivial changes to the theory. We notice here only that if we use MDSs, no such problem arises.

We conclude with some remarks on *No Recycling*. The idea behind *No Recycling* is that we never attempt to copy a constituent that is actually empty. By empty we mean here that it is not an LF–member (before copying). *No Recycling* ensures that this never happens. For trace chains, this corresponds to moving a non–head of some chain to some other place (and leaving a trace for it). However, non–heads are always traces. Additionally, if non–heads are not eligible for movement, heads cannot be traces (there is no way to produce traces in head position of a chain). Hence, *No Recycling* reduces to the following requirement on TCCs.

Constraint 8 (No Trace Recycling) *No trace is the head of a chain.*

9 Ordered Trees

Up to now we have only dealt with unordered trees. Now we shall show how to introduce an ordering. For trees there exist several choices. However, not all of them transfer equally well to MDSs. The first solution, favoured in [14], is to add to the tree a relation $\sqsubset$ such that $x \sqsubset y$ iff x and y are sisters and x is immediately left of y. Another solution was proposed in [1]. We split $\prec$ into several relations, $\prec_i$, where i is a natural number, such that $x \prec_i y$ iff x is the ith daughter of y, counting from the left. It is the latter solution that turns out to be more practical here. Since we have to specify the number of daughters in advance, let us from now on work with binary branching trees (this is anyway standard). It goes without saying that the definition of chains* must now be revised; the isomorphisms of a chain* must be isomorphisms of the ordered constituents involved.

Definition 79 *An* ***ordered (binary branching) multidominance structure (OMDS)*** *is a quadruple* $\mathfrak{M} = \langle M, \prec_0, \prec_1, \propto \rangle$ *such that*

(a) $\langle M, \prec_0 \cup \prec_1, \propto \rangle$ *is an MDS,*

(b) $\prec_0 \cap \prec_1 = \varnothing$*,*

(c) $(\forall x)((\exists y) y \prec_1 x \to (\exists z) z \prec_0 x)$*,*

(d) $(\forall xyz)(y \prec_0 x \wedge z \prec_0 x \rightarrow y = z)$,

(e) $(\forall xyz)(y \prec_1 x \wedge z \prec_1 x \rightarrow y = z)$.

An ***ordered (binary branching) tree*** *is an ordered, binary branching MDS* $\langle A, \prec_0, \prec_1, \propto\rangle$ *such that* $\langle A, \prec_0 \cup \prec_1\rangle$ *is a tree.*

We speak of $\langle A, \prec_0, \prec_1\rangle$ as an **ordering** of $\langle A, \prec_0 \cup \prec_1\rangle$. The definition states in informal terms that $\prec_0$ ('first daughter of') and $\prec_1$ ('second daughter of') are mutually exclusive relations whose inverses are partial functions. Moreover, if a second daughter exists, a first daughter must exist as well. Finally, any daughter is a first daughter or a second daughter.

Before we extend the results of the previous sections to the ordered case we show two easy facts about the relationship between ordered and unordered trees. Let $\mathfrak{T} = \langle T, \prec^T\rangle$ and $\mathfrak{U} = \langle U, \prec^U\rangle$ be binary unordered branching trees and $j : \mathfrak{T} \rightarrow \mathfrak{U}$ an isomorphism. If $\langle T, \prec_0^T, \prec_1^T\rangle$ is an ordering of $\mathfrak{T}$, then $\langle U, j(\prec_0^T), j(\prec_1^T)\rangle$ is an ordering of U. (Here, $j(\prec_i^T) := \{\langle j(x), j(y)\rangle : \langle x, y\rangle \in \prec_i^T\}$.) The following is also clear.

Lemma 80 *Let* $\mathfrak{T}$ *and* $\mathfrak{U}$ *be ordered binary branching trees and* $j, j' : \mathfrak{T} \rightarrow \mathfrak{U}$ *isomorphisms. Then* $j = j'$.

This fact is important insofar as in the ordered case we can dispense with keeping track of the isomorphisms between members of a chain (which are essential in the unordered case).

We can now define **ordered CCS** and so on in just the same way. The notation remains more or less the same, for convenience. In defining $M(K)$ in the ordered case we need to define the ordering. We put

$$\begin{array}{lll} \prec_0^M & := & \{\langle [x], [y]\rangle : x \prec_0 y\} \\ \prec_1^M & := & \{\langle [x], [y]\rangle : x \prec_1 y\} \end{array}$$

Recall that each chain comes with a set of isomorphisms mapping the members onto each other. If $x < y$ and $\downarrow y$ is a member of Γ, then the isomorphisms of Γ induce isomorphisms on the subconstituents $\downarrow \varphi(x)$. We use this for the proofs of the following

Theorem 81 *The following holds.*

1. *If* K *is an ordered CCS then* $M(K)$ *is an ordered MDS.*

2. *If* $\mathfrak{M}$ *is an ordered MDS and* $K = \langle \mathfrak{T}, C\rangle$ *is a CCS such that* $M(K)$ *is isomorphic to the unordered reduct of* $\mathfrak{M}$*, then there is exactly one ordering of* $\mathfrak{T}$ *that turns* K *into an ordered CCS* L *with* $M(L) \cong \mathfrak{M}$.

Proof. For (a) we must show that $\prec\ =\ \prec_0 \cup \prec_1$. Clearly, either $[x] \prec_0^M [y]$ or $[x] \prec_1^M [y]$ holds. Since $M(K)$ is binary branching, we have $\prec_0^M \cup \prec_1^M = \prec$. Now (b). Assume that $u \prec_0 x$, $v \prec y$ and $x \approx_K y$. By Lemma 15 there is a composite isomorphism ψ mapping x to y. Now, $\psi(u) = v$ iff $v \prec_0 y$. Hence, if $u \approx_K v$ then $v \prec_1 y$ cannot hold. Likewise for $u \prec_1 x$. This establishes that we do not have both $[u] \prec_0^M [x]$ and $[u] \prec_1^M [x]$. Next, to show (c), suppose that there is a $[y]$ such that $[y] \prec_1^M [x]$. Then for some $x' \approx_K x$ and some $y' \approx_K y$ we have $y' \prec_1 x'$. So there exists a z such that $z \prec_0 x'$, from which follows $[z] \prec_0^M [x'] = [x]$. For (d), assume that $[y], [z] \prec_0^M [x]$. Then there are $y' \in [y]$, $z' \in [z]$ and $x', x'' \in [x]$ such that $y' \prec_0 x'$ and $z' \prec_0 x''$. By Lemma 15 there is a composite isomorphism ψ such that $\psi(x') = x''$. Put $z'' := \psi(y')$. Then, by uniqueness of the daughters, $z'' = z'$. Hence $y' \approx z'$, and so $[y] = [z]$, as required. Likewise for (e).

To show the second claim, assume that K is an unordered CCS, and that $\prec_0$ and $\prec_1$ introduce an ordering on $M(K)$. We need to define an ordering on K that makes it into an ordered CCS. We may assume for simplicity that K consists of identifiers of $M(K)$. Now put $I; x \prec_i^K I$ iff I identifies y, and $x \prec_i y$. This is unique. Furthermore, $\prec_0^K \cap \prec_1^K = \varnothing$ and $\prec_0^K \cup \prec_1^K = \prec^K$. The other postulates are verified similarly. Now, finally, we need to show that the chain isomorphisms $\phi_{i,j}$ are isomorphisms of the ordered constituents. Since they are isomorphisms of the unordered constituents, we need to show that they respect the ordering. Hence, let $I \prec_0^K J$ and let $\varphi_{i,j}$ be an isomorphism of a chain. Then $I' := \varphi_{i,j}(I) \prec^K J' := \varphi_{i,j}(J)$, hence either $I' \prec_0^K J'$ or $I' \prec_1^K J'$. However, $I' = J'; \varphi_{i,j}(x)$. Since $x \approx \varphi_{i,j}(x)$, we have $I' \prec_0 J'$ as well. Likewise if $I \prec_1^K J$. Q. E. D.

10 Adjunction

In [15] we have studied the notion of an adjunction structure and the ensuing complications for the theory of grammar. We will incorporate the insights of that paper with minimal disturbance to the previous constructions. This means that we will make certain adjustments. Some complications need to be left unadressed, for example the problem of the constituents (strong versus weak) and the question of legitimate objects that may move. Here, we will take a neutral stance, allowing as many options as we possibly can. Previously, we have defined an adjunction structure as a triple $\langle T, \prec, \Pi \rangle$, where Π is a partition of T into sets which are linear and convex. As with order, there are problems in lifting this approach to multidominance structures. Adjunction is therefore marked by means of a binary relation $\rtimes$. Intuitively, if $x \rtimes y$ then x and y are segments of the same category and y immediately dominates x.

Definition 82 *An* ***adjunction structure*** *is a triple $\langle T, \prec, \rtimes \rangle$, where $\langle T, \prec \rangle$ is a tree and $\rtimes \subseteq \prec$ is such that for each $x \in T$ there is at most one y with $y \rtimes x$. A* ***multidominant adjunction structure (MAS)*** *is a triple $\langle A, \prec, \rtimes \rangle$, where*

$\langle A, \prec \rangle$ is an MDS and $\rtimes \subseteq \prec$ is such that for each $x \in A$ there is at most one y with $y \rtimes x$.

Given $\rtimes$, a set which is maximally connected via $\rtimes$ is called a **block**. We write $x_{\rtimes}$ to denote the block of x. A block $\mathfrak{b}$ has a unique maximal member, $\mathfrak{b}^{\circ}$ and a unique minimal member, $\mathfrak{b}_{\circ}$. We write $x_{\circ}$ for $(x_{\rtimes})_{\circ}$ and x° for $(x_{\rtimes})^{\circ}$.

It is immediate how to define ordered adjunction structures, ordered multidominance structures etc. Much of what we have said about incorporating order carries over to adjunction structures. In the definition of chains, for example, we need to require that the chain members are isomorphic as ordered adjunction structures. [13] Now, the crucial aspect of adjunction structures is the new notion of c–command that arises with it. Define $\nu(x)$ as follows. (A) If $x_{\rtimes}$ is the root block, $\nu(x)$ is undefined. (B) If $x_{\rtimes}$ is not the root block, let y the least element $\geq x$ which is not in $x_{\rtimes}$. Then put $\nu(x) := y^{\circ}$. We call y° also the **strong mother** of x. We construe the new c–command relation as a relation between nodes.

Definition 83 *Let $\langle T, \prec, \rtimes \rangle$ be an adjunction structure. x **c–commands** y iff $x_{\rtimes}$ is the root block or else $\nu(x) \geq y$. x **ac–commands** y if (1) x and y are incomparable and (2) x c–commands y but y does not c–command x.*

Lemma 84 *x ac–commands y iff (a) x and y are incomporable and (b) $\nu(x) > \nu(y)$.*

Proof. Suppose that x ac–commands y. Then $\nu(x) \geq y$, by definition of c–command. It is easy to see that $\nu(x) \geq y^{\circ}$. Moreover, we must have $\nu(x) > y^{\circ}$, since they belong to different blocks. Let $z := \mu(y^{\circ})$ (sic!). If $z \in x_{\rtimes}$, then y c–commands x. Hence $z \notin x_{\rtimes}$ and so $\nu(x) > z$. Therefore $\nu(x) > z^{\circ} = \nu(y)$. Now assume conversely that (a) and (b) hold. Then x c–commands y. Suppose that y also c–commands x. $\nu(x)$ and $\nu(y)$ are defined and $\nu(y) \geq x$. Hence, $\nu(y)$ and $\nu(x)$ are comparable, and so it easily seen that they are equal. This cannot be, however. Hence $\nu(y) < \nu(x)$, as required. Q. E. D.

Proposition 85 *The relation of ac–command is transitive.*

Virtually all definitions remain intact. We note only that the postulates of chain structures are not in the same way independent as they were before. It does not follow, for example, that chains cannot be doubly covered. It must be excluded by a separate postulate. This is why we have chosen to introduce it into the definition of a chain structure. We recall here the discussion of [15] concerning the notion of a constituent. Here, we will take the widest possible option: a constituent is any substructure of the form $\langle U, \prec\restriction U, \rtimes\ \restriction U \rangle$ where $U \subseteq T$ is a downward closed set of the form $\downarrow x$. Now, we denote such a structure also by

[13] In contrast to [15] we will not conflate the segments of a category into a single object.

$\downarrow x$, and we write $\prec$ and $\rtimes$ rather than $\prec\restriction U$ and $\rtimes \restriction U$. If matters are defined in this way, it is allowed to split up a block and raise some parts of the material. An example is excorporation, where a complex head is broken into two parts. The existence of such movement is largely a theory internal problem as well as an empirical one; both are of no concern to us.

Now, define as before the transition from a CCS to a multidominance structure. The CCS is now defined over an adjunction structure. We will put $[x] \rtimes [y]$ iff for some $x' \in [x]$ and some $y' \in [y]$ we have $x' \rtimes y'$. To show that this allows to recover the original relation between the nodes from the multidominance relation we need the following theorem.

Lemma 86 *Suppose that $x \approx_K x'$ and $y \approx_K y'$. Then $x \rtimes y$ iff $x' \rtimes y'$.*

Proof. Since $y \approx_K y'$ there is a composite isomorphism ψ such that $y' = \psi(y)$. Since $\psi : \downarrow y \to \downarrow y'$, we have $\psi(x) \rtimes \psi(y) = y'$. However, $\psi(x) = x'$, since $x \approx_K x'$, and there is exactly one $z \prec y'$ such that $x \approx_K z$. Since $\psi(x) \prec y'$, we have $\psi(x) = x'$. Q. E. D.

As a consequence, if $[x] \rtimes [y]$ and $x \prec y$, then also $x \rtimes y$. So, when we are given an (ordered) MAS, we reconstruct as before the (ordered) CCS. Now, take two identifiers I and J, identifying x and y, respectively. We put $I \rtimes J$ iff $I = J; x$ and $x \rtimes y$.

The adjunction structures allow for a new type of movement. When $\langle \mathfrak{C}, \mathfrak{C}' \rangle$ and $\langle \mathfrak{D}, \mathfrak{D}' \rangle$ are chain links and $\mathfrak{C}' \subsetneq \mathfrak{D}$, then it is not necessarily the case that $\mathfrak{D}$ covers $\mathfrak{C}$. Movement that produces such links we call **accumulative movement**. The typical case is iterated head movement. For example, when the verb moves to INFL (to pick up the tense features) and then to COMP, each of the movements end in an adjunction to the relevant head. So, when V moves to INFL it adjoins to it. Subsequently, the complex head consisting of V+INFL moves, not V alone. Hence, we have two chains of length 2 rather than one chain of length 3. There is also a number of movement types that are admitted by the system though rarely considered an option. For example, if a DP scrambles to VP then it can be adjoined to by any other object that scrambles to VP. So, rather than (7) we also have (8). (Here we assume that scrambling is adjunction to VP or some higher functional projection.)

(7) *[Das neue Auto]$_1$ wollte [Alfred$_2$ [seiner Freundin$_3$ [denn doch nicht [t_2 t_3 t_1 geben]]]].*

(8) *[Das neue Auto]$_1$ wollte [[Alfred$_2$ [seiner Freundin$_3$]] [denn doch nicht [t_2 t_3 t_1 geben]]].*

It might be that adjunction to DP is limited for some reasons to special categories. However, we note here that from a structural point of view, no diagnostic can distinguish (7) from (8). All relevant syntactic domains are identical (if defined along [13] and [15], as most of them are). For example, binding theory cannot distinguish the two unless some stipulations are being made. The c–command domains cover the same elements in both cases.

11 Conclusion

In this paper we have studied the notion of a chain and identified three different ways to represent syntactic structures involving chains: either as copy chain structures (CCSs), as multidominance structures (MDSs) or as trace chain structures (TCSs). Although not exactly identical in content (the CCSs contain more information) the three can be said to be identical for all linguistic purposes. This is proved in Proposition 78 and Theorem 81. They say effectively that there is a one–to–one and onto map from derivations of CCSs to derivations of MDSs and back, and from derivations of MDSs to derivations of TCSs and back. There is a natural isomorphism between MDSs and TCSs. For an MDS there exist in the worst case exponentially many nonisomorphic CCSs, and for every CCS in the worst case exponentially many derivations.

Especially interesting is the idea — which has surfaced in the literature from time to time — that we could dispense with talk of several copies (or for that matter, talk of traces) if we endorse the view that an element can be attached to several nodes in a structure. This has advantages also for formalizing these structures. While we would otherwise need to talk of isomorphisms of constituents using MDSs this is actually needless, since isomorphy is replaced by identity. Moreover, MDSs are far more compact representations than are CCSs. In the worst case, the size of a CCS is exponential in the size of its MDS, as is easy to show. Hence, abundant copying leads to low performance of the human computation system. [14]

In retrospect one may wonder whether the introduction of chains has been sufficiently motivated. We think the present paper does not allow to construct an argument against using chains per se. However, it shows that linguists should be very careful not to reject formal investigations on the grounds that the notions keep changing anyway or — even worse — that we understand what this is all about in enough detail. The problem is that a deeper understanding of the mechanics of chains allows to characterize the options that we have in using chains, and what the consequences are. In this way it can be foreseen far better whether we want such a theory or not.

[14] In fact, fast copying of data files in computers is done by adding only a new link from the second (virtual) copy to the first file. If one of the copies of a file is actually needed, for example for editing it, it is copied anyway into the buffer. Only when it is stored after editing, the link is replaced by a distinct file, since the second copy may (but need not) be different from the first file.

References

[1] Patrick Blackburn, Wilfried Meyer-Viol, and Maarten de Rijke. A Proof System for Finite Trees. In H. Kleine Büning, editor, *Computer Science Logic '95*, number 1092 in Lecture Notes in Computer Science, pages 86 – 105. Springer, 1996.

[2] James Blevins. *Syntactic Complexity.* PhD thesis, University of Massachusetts at Amherst, 1990.

[3] Damir Ćavar and Gisbert Fanselow. Discontinuous constituents in Slavic and Germanic languages. manuscript, 1998.

[4] Noam Chomsky. *Lectures on Government and Binding.* Foris, Dordrecht, 1981.

[5] Noam Chomsky. A minimalist program for linguistic theory. In K. Hale and Keyser S. J., editors, *The View from Building 20: Essays in Honour of Sylvain Bromberger*, pages 1 – 52. MIT Press, 1993.

[6] Thomas L. Cornell. Derivational and representational views of minimalist syntactic calculi. In *Proceedings of LACL '97*, Nancy, France, 1997. To appear in the Lecture Notes in Artificial Intelligence series by Springer.

[7] Thomas L. Cornell. A type–logical perspective on minimalist derivations. In *Proceedings of Formal Grammar '97*, Aix–en–Provence, France, August 1997.

[8] Thomas L. Cornell. Representational minimalism. In Hans-Peter Kolb and Uwe Mönnich, editors, *The Mathematics of Sentence Structure*, pages 301 – 339. de Gruyter, Berlin, 1998.

[9] Hans-Martin Gärtner. *Generalized Transformations and Beyond.* Akademie Verlag. to appear.

[10] Hans-Martin Gärtner. Review of 'the Copy Theory of Movement and Linearization of Chains in the Minimalist Program'. by Jairo Nunes. *GLOT International 3.8*, pages 16 – 19, 1998.

[11] Hans-Martin Gärtner. Phrase–linking Meets Minimalist Syntax. In Sonya Bird, Andrew Carnie, Jason Haugen, and Peter Norquest, editors, *Proceedings of WCCFL 18*, pages 159 – 169, 1999.

[12] Richard S. Kayne. *The Antisymmetry of Syntax.* Number 25 in Linguistic Inquiry Monographs. MIT Press, 1994.

[13] Marcus Kracht. Mathematical aspects of command relations. In *Proceedings of the EACL 93*, pages 240 – 249, 1993.

[14] Marcus Kracht. Syntactic Codes and Grammar Refinement. *Journal of Logic, Language and Information*, pages 41 – 60, 1995.

[15] Marcus Kracht. Adjunction structures and syntactic domains. In Hans-Peter Kolb and Uwe Mönnich, editors, *The Mathematics of Sentence Structure*, pages 259 – 299. Mouton–de Gruyter, Berlin, 1998.

[16] Alain Lecomte. POM–nets and minimalism. In Claudia Casadio, editor, *Proceedings of the IV Roma Workshop: Dynamic Perspectives in Logic and Linguistics*. SILPS Group in Logic and Natural Languages, 1997.

[17] James McCawley. Concerning the base component of a transformational grammar. *Foundations of Language*, 4:243 – 269, 1968.

[18] Gereon Müller. *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German.* Kluwer Academic Publishers, Dordrecht, 1998.

[19] Jairo Nunes. *The Copy Theory of Movement and Linearization of Chains in the Minimalist Program.* PhD thesis, University of Maryland, 1995.

[20] Tanya Reinhart and Eric Reuland. Reflexivity. *Linguistic Inquiry*, 24:657 – 720, 1993.

[21] Henk C. van Riemsdijk. The Case of German adjectives. In F. Heny and B. Richards, editors, *Linguistic Categories: Auxiliaries and Related Puzzles I*, Foris. Reidel, Dordrecht, 1983.

[22] James Rogers. *Studies in the Logic of Trees with Applications to Grammar Formalisms.* PhD thesis, Department of Computer and Information Sciences, University of Delaware, 1994.

[23] Edward P. Stabler. *The Logical Approach to Syntax. Foundation, Specification and Implementation of Theories of Government and Binding.* ACL–MIT Press Series in Natural Language Processing. MIT Press, Cambridge (Mass.), 1992.