Implications of a Revised Perspective on Minimalist Grammars

JENS MICHAELIS

Universität Potsdam, Postfach 601553, 14415 Potsdam, Germany

ABSTRACT: The type of a minimalist grammar (MG) *as introduced by Stabler [17] provides an attempt of a rigorous formalization of the perspectives adopted nowadays within the linguistic framework of transformational grammar. As shown in [11], MGs constitute a weakly equivalent subclass of* linear context–free rewriting systems (LCFRSs) *in the sense of Vijay–Shanker et al. [21]. Independent work of Harkema [5] and Michaelis [13] has proven the reverse to be true, as well. Hence, MGs as defined in [17] join to a series of formalism classes—among which there is e.g. the class of* multicomponent tree adjoining grammars (MCTAGs) *in their set–local variant of admitted adjunction (cf. [22])—all generating the same class of string languages. Inspired by current linguistic developments, a revised type of an MG as well as a certain type of a* strict MG (SMG) *has been proposed by Stabler [18]. Here we show that, in terms of derivable string languages, the revised MG–type as well as the SMG-type is not only subsumed by LCFRSs, but both also fall within a particular subclass of the latter: the righthand side of each rewriting rule of a corresponding LCFRS involves at most two nonterminals, and if two nonterminals appear on the righthand side then only simple strings of terminals are derivable from the first one. This result is in fact of specific interest, since conversely, in terms of weak equivalence, the corresponding LCFRS–subclass is provably subsumed by the class of revised MGs as well as the class of SMGs ([10]). Whether the inclusion of the respective classes of string languages derivable by the corresponding LCFRS–subclass and the class of all LCFRSs is proper or not seems to be an open problem. We briefly discuss what seems to constitute the crucial difference seen from the minimalist perspective at the end of this paper.*

KEYWORDS: *(revised) minimalist grammars, linear context–free rewriting systems, generative capacity*

1 Introduction

The type of a *minimalist grammar (MG)* as introduced in [17] provides an attempt of a rigorous algebraic formalization of the perspectives adopted nowadays within the linguistic framework of transformational grammar (see e.g. [2]). An MG, roughly speaking, is a formal device which specifies a countable set of finite, binary (ordered) trees each being equipped with a leaf–labeling function assigning a string of features to each leaf, and with an additional binary relation, the asymmetric relation of *(immediate) projection*, defined on the set of pairs of siblings. The base of an MG is formed by a *lexicon* (a finite set of single node trees in the above sense) and two structure building functions: *merge* (combining two trees) and *move* (transforming a given tree). Both functions build structure by canceling particular matching instances of features within the leaf–labels of the trees to which they are applied. The closure of the lexicon under these two functions is the set of trees characterized by the MG. As shown in [11], the MG–type introduced in [17] constitutes a weakly equivalent subclass of *linear context–free rewriting systems (LCFRSs)* [21, 22]. Independent work in [5] and [13] has proven the reverse to be true as well. Hence, MGs as defined in [17], beside LCFRSs, join to a series of formalism classes—among which there is e.g. the class of *multicomponent tree adjoining grammars* in their set–local variant of admitted adjunction (cf. [22])—all generating the same class of string languages. For a list of some further of such classes of generating devices see e.g. [15].

Mainly inspired by the linguistic work presented in [9], in [18] a revised type of an MG has been proposed whose departure from the version in [17] can be seen as twofold: the revised type of an MG neither employs any kind of *head movement* nor *covert phrasal movement*, and an additional restriction is imposed on the move–operator as to which maximal projection may move *overtly* into the highest specifier position. Deviating from the operator *move* as originally defined in [17], a constituent has necessarily to belong to the transitive complement closure of a given tree or to be a specifier of such a constituent in order to be movable. Closely in keeping with some further suggestions in [9], a certain type of a *strict minimalist grammar (SMG)* has been introduced in [18] as well. This MG–type allows only movement of constituents belonging to the transitive complement closure of a tree. But different from the just mentioned type, the triggering licensee feature may head the head–label of any constituent within the reflexive–transitive specifier closure of a moving constituent. Furthermore, due to the general definition of a lexical item of an SMG, an SMG does not permit the creation of multiple specifiers during the course of a derivation.

Employing and extending the methods developed in [11], this paper shows that, in terms of derivable string languages, the revised MG–type and the SMG–type are not only subsumed by LCFRSs, but both also fall within a particular subclass of the latter: the righthand side of each rewriting rule of a corresponding LCFRS involves at most two nonterminals, and if two nonterminals appear on the righthand side then only simple strings of terminals are derivable from the first one. The result is in fact of specific interest, since conversely, in terms of weak equivalence, the corresponding LCFRS–subclass is provably subsumed by the class of revised MGs as well as the class of SMGs ([10]). Consequently, the revised MG–type and the SMG–type are shown to determine the same class of derivable string languages, thereby confirming a conjecture explicitly stated in [18]. Whether the respective classes of string languages derivable by the corresponding LCFRS–subclass and the class of all LCFRSs—and thus the respective classes of string languages derivable by the class of MGs (or likewise SMGs) as defined in [18] and the class of MGs as defined in [17]—are identical seems to be an open problem. We briefly discuss what seems to constitute the crucial difference seen from the minimalist perspective at the end of this paper.

The result presented in this paper is also shown to hold in [12]. But in several respects, the proof given there is much more involved than the one given here. This mainly follows for the reason that, in [12], the corresponding inclusion in terms of derivable string languages within the particular LCFRS–subclass is proven for "re–extended" versions of both the revised MG–type and the SMG–type, namely, versions in which head movement and covert phrasal movement are "re–added," and w.r.t. the SMG–type the ban of multiple specifiers is revoked as well. Restricting our attention to both MG–types as defined in [18] allows us to significantly simplify the presentation of a proof yielding the intended result, and as we think, it becomes much more intelligible even at first glance. Thereby, the construction of a weakly equivalent LCFRS from a given (S)MG also becomes more directly accessible to further exploitation. The class of resulting LCFRSs may rather straightforwardly be interpreted as a succinct reformulation of the corresponding (S)MG–type comparable to the proposal made in [20] for a restricted version of the original MG–type which does not use any covert phrasal movement or head movement, but does not restrict the domain of the move–operator w.r.t. overt movement. Such a reformulation does not only open the possibility to adapt the polynomial time parsing methods developed in [4, 6] for the (restricted) original MG–type, but also opens, e.g., the field to a direct comparison with *mirror theoretic grammars* developed in [7] as a formalization of the syntactic theory proposed in [1].

The paper is structured as follows: the next section provides formal definitions of LCFRSs and the particular subtype mentioned above. In Section 3 we first define MGs and SMGs in the sense of [18], and then discuss in detail the specific properties which, in terms of weak equivalence, allow both types to be embedded into the LCFRS–subtype (cf. Section 3.1). In Section 4 methods of constructing a corresponding, weakly equivalent LCFRS from a given MG and SMG, respectively, are presented step by step, explicitly taking into account the discussion from Section 3.1. In Section 5 we sum up the immediate implications, and briefly compare them to the results which have been established before w.r.t. MGs as originally defined in [17].

Throughout the rest of the paper we refer to an MG of the type as originally defined in [17] as an *unrestricted MG (UMG)*. Attempting to avoid any confusion that might arise otherwise, this allows us to use the term *minimalist grammar* and its abbreviation *MG* exclusively in order to refer to an MG of the revised type as defined in [18].

2 Linear Context–Free Rewriting Systems

The class of *linear context–free rewriting systems (LCFRSs)* [21, 22] constitutes a proper subclass of *multiple context–free grammars (MCFGs)* [16], which in their turn form a subtype of *generalized context–free grammars* [14]. However, LCFRSs define the same class of derivable string languages as MCFGs.

DEFINITION 2.1 ([14])
A *generalized context–free grammar (GCFG)* is a five–tuple $G = \langle N, O, F, R, S\rangle$, where N is a finite non–empty set of *nonterminals*, and where O is a set of *(linguistic) objects*. F is a finite subset of $\bigcup_{n\in\mathbb{N}} F_n \setminus \{\emptyset\}$, F_n the set of partial functions from $\langle O\rangle^n$ into O.[1] R is a finite set of *(rewriting) rules*, i.e. $R \subseteq \bigcup_{n\in\mathbb{N}}(F\cap F_n)\times\langle N\rangle^{n+1}$. S is a distinguished symbol from N, the *start symbol*.

An $r = \langle f, \langle A_0, A_1, \dots, A_n\rangle\rangle \in (F\cap F_n)\times\langle N\rangle^{n+1}$ for some $n\in\mathbb{N}$ is written $A_0 \to f(A_1,\dots,A_n)$, and also $A_0 \to f(\emptyset)$ if $n = 0$. In case $n = 0$, i.e. if f is a constant in O, r is *terminating*, otherwise r is *nonterminating*.

For each $A \in N$ and $k \in \mathbb{N}$, $L^k_G(A) \subseteq O$ is given recursively by means of $\theta \in L^0_G(A)$ for each terminating $A \to \theta \in R$, and $\theta \in L^{k+1}_G(A)$ if $\theta \in L^k_G(A)$, or if there are $A \to f(A_1,\dots,A_n) \in R$ and $\theta_i \in L^k_G(A_i)$ for $1 \le i \le n$ such that $\langle\theta_1,\dots,\theta_n\rangle \in Dom(f)$ and $f(\theta_1,\dots,\theta_n) = \theta$.[2] The set $L_G(A) = \bigcup_{k\in\mathbb{N}} L^k_G(A)$ is the *language derivable from A (by G)*. $L_G(S)$, also denoted by $L(G)$, is the *language derivable by G*.

DEFINITION 2.2
A given GCFG G_1 and a given GCFG G_2 are *weakly equivalent* if $L(G_1) = L(G_2)$.

DEFINITION 2.3 ([16], [21, 22])
A *multiple context–free grammar (MCFG)* is a GCFG $G = \langle N, O, F, R, S\rangle$ with $O = \bigcup_{n\in\mathbb{N}}\langle\Sigma^*\rangle^{n+1}$ and satisfying (M1) and (M2), where Σ is a finite set of *terminals* with $\Sigma \cap N = \emptyset$.[3]

(M1) For each $f \in F$, some $n(f) \in \mathbb{N}$, $\varphi(f) \in \mathbb{N}\setminus\{0\}$ and $d_i(f) \in \mathbb{N}\setminus\{0\}$ for $1 \le i \le n(f)$ exist such that f is a (total) function from $\prod_{i=1}^{n(f)}\langle\Sigma^*\rangle^{d_i(f)}$ into $\langle\Sigma^*\rangle^{\varphi(f)}$ for which (f1) and (f2) hold.

(f1) Let $X_f = \{x_{ij} \mid 1 \le i \le n(f), 1 \le j \le d_i(f)\}$ be a set of pairwise distinct variables, for $1 \le i \le n(f)$ let $x_i = \langle x_{i1},\dots,x_{id_i(f)}\rangle$, and for $1 \le h \le \varphi(f)$ let f_h be the h–th component of f, i.e. the function from *Dom*(f) into Σ^* such that $f(\theta) = \langle f_1(\theta),\dots,f_{\varphi(f)}(\theta)\rangle$ for all $\theta \in$ *Dom*(f). Then, for each $1 \le h \le \varphi(f)$ there are an $l_h(f) \in \mathbb{N}$, a $\zeta(f_{hl}) \in \Sigma^*$ for $0 \le l \le l_h(f)$, and a $z(f_{hl}) \in X_f$ for $1 \le l \le l_h(f)$ such that f_h is represented by (c$_{f_h}$).

---

[1] $\mathbb{N}$ is the set of all non–negative integers. For $n \in \mathbb{N}$ and any sets $M_1,\dots,M_n$, $\prod_{i=1}^n M_i$ is the set of all n–tuples $\langle m_1,\dots,m_n\rangle$ with i–th component $m_i \in M_i$, where $\prod_{i=1}^n M_i := \{\emptyset\}$ for $n = 0$. We write $\langle M\rangle^n$ instead of $\prod_{i=1}^n M_i$ if for some set M, $M_i = M$ for $1 \le i \le n$.

[2] For each partial function g from a set M into a set M', *Dom*$(g) \subseteq M$ is the domain of g.

[3] For each set M, M^* is the Kleene closure of M, including ϵ, the empty string. M_ϵ denotes the set $M \cup \{\epsilon\}$.

(c_{f_h}) $\quad f_h(x_1, \ldots, x_{n(f)}) = \zeta(f_{h0})\, z(f_{h1})\, \zeta(f_{h1}) \cdots z(f_{hl_h(f)})\, \zeta(f_{hl_h(f)})$

(f2) Each $x \in X_f$ occurs at most once in all righthand sides of (c_{f_1})–($\mathrm{c}_{f_{\varphi(f)}}$), i.e. for the set $I_{\mathit{Dom}(f)} = \{\langle i,j\rangle \,|\, 1 \leq i \leq n(f), 1 \leq j \leq d_i(f)\}$ and for the set $I_{\mathit{Range}(f)} = \{\langle h,l\rangle \,|\, 1 \leq h \leq \varphi(f), 1 \leq l \leq l_h(f)\}$, the binary relation g_f on $I_{\mathit{Dom}(f)} \times I_{\mathit{Range}(f)}$ such that $\langle\langle i,j\rangle, \langle h,l\rangle\rangle \in g_f$ iff $x_{ij} = z(f_{hl})$ is an injective partial function onto $I_{\mathit{Range}(f)}$.

(M2) There exists a function d_G from N into $\mathbb{N}$ such that $d_G(S) = 1$, and such that, if $A_0 \to f(A_1, \ldots, A_n) \in R$ for some $n \in \mathbb{N}$ then $\varphi(f) = d_G(A_0)$ and $d_i(f) = d_G(A_i)$ for $1 \leq i \leq n$.

The *rank of G*, *rank*(G), is the number $\max\{n(f) \,|\, f \in F\}$. $L(G)$ is a *multiple context–free language (MCFL)*. $L(G) \subseteq \Sigma^*$, because $d_G(S) = 1$.

If for each $f \in F$ condition (f3) holds in addition to (f1) and (f2) then G is a *linear context–free rewriting system (LCFRS)*, and $L(G)$ is a *linear context–free rewriting language (LCFRL)*.

(f3) Each $x_{ij} \in X_f$ has to appear in one of the righthand sides of (c_{f_1})–($\mathrm{c}_{f_{\varphi(f)}}$), i.e. the function g_f from (f2) is total, and therefore, a bijection.

The class of all MCFLs and the class of all LCFRLs are known to be identical (cf. [16, Lemma 2.2]). Theorem 11 in [15], therefore, shows that for each MCFG G there is a weakly equivalent LCFRS G' with *rank*$(G') \leq 2$.

DEFINITION 2.4
An *MCFG*$_{1,2}$ (*LCFRS*$_{1,2}$) is an MCFG (LCFRS) G in the sense of Definition 2.3 such that *rank*$(G) \leq 2$, and such that $d_1(f) = 1$ for each $f \in F$ with $n(f) = 2$. In this case $L(G)$ is an *MCFL*$_{1,2}$ (*LCFRL*$_{1,2}$).

In Section 4.1 and 4.2, we shall in fact construct an LCFRS$_{1,2}$, “not just” an LCFRS of rank 2, that derives the same (string) language as a given minimalist grammar and strict minimalist grammar, respectively.

3 (Strict) Minimalist Grammars

Throughout we let ¬*Syn* and *Syn* be a finite set of *non–syntactic features* and a finite set of *syntactic features*, respectively, in accordance with (F1) and (F2) below. We take *Feat* to be the set ¬*Syn* ∪ *Syn*.

(F1) ¬*Syn* is disjoint from *Syn* and partitioned into a set *Phon* of *phonetic features* and a set *Sem* of *semantic features*.

(F2) *Syn* is partitioned into a set *Base* of *(basic) categories*, a set *Select* of *selectors*, a set *Licensees* of *licensees* and a set *Licensors* of *licensors*. For each $x \in \mathit{Base}$, usually typeset as `x`, the existence of a matching $x' \in \mathit{Select}$, denoted by `=x`, is possible. For each $x \in \mathit{Licensees}$, usually depicted as `-x`, the existence of a matching $x' \in \mathit{Licensors}$, denoted by `+X`, is possible. *Base* includes at least the category `c`.

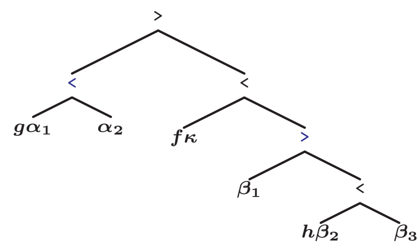

FIG. 1. A typical expression over *Feat*.

DEFINITION 3.1
An *expression (over Feat)* is a five–tuple $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau \rangle$ obeying (E1)–(E4).

(E1) $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ is a finite, binary (ordered) tree defined in the usual sense: N_τ is the finite, non–empty set of *nodes*, and $\triangleleft^*_\tau$ and $\prec_\tau$ are the respective binary relations of *dominance* and *precedence* on N_τ.[4]

(E2) $<_\tau \subseteq N_\tau \times N_\tau$ is the asymmetric relation of *(immediate) projection* that holds for any two siblings in $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$, i.e., for each $x \in N_\tau$ different from the root of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ either $x <_\tau$ *sibling*$_\tau(x)$ or *sibling*$_\tau(x) <_\tau x$ holds.[5]

(E3) *label*$_\tau$ is the *leaf–labeling function* from the set of leaves of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ into *Syn***Phon***Sem**.

(E4) $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ is a subtree of the natural interpretation of a tree domain.[6]

We take *Exp*(*Feat*) to denote the set of all expressions over *Feat*.

Let $\tau = \langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau \rangle \in \mathit{Exp}(\mathit{Feat})$.[7]

For each $x \in N_\tau$, the *head of x (in τ)*, denoted by *head*$_\tau(x)$, is the (unique) leaf of τ with $x \triangleleft^*_\tau$ *head*$_\tau(x)$ such that each $y \in N_\tau$ on the path from x to *head*$_\tau(x)$ with $y \neq x$ projects over its sibling, i.e. $y <_\tau$ *sibling*$_\tau(y)$. The *head of* τ is the head of τ's root. τ is said to be a *head* (or *simple*) if N_τ consists of exactly one node, otherwise τ is said to be a *non–head* (or *complex*).

A five–tuple $\upsilon = \langle N_\upsilon, \triangleleft^*_\upsilon, \prec_\upsilon, <_\upsilon, \mathit{label}_\upsilon \rangle$ is a called *subexpression of* τ in case $\langle N_\upsilon, \triangleleft^*_\upsilon, \prec_\upsilon \rangle$ is a subtree of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$, $<_\upsilon = <_\tau \restriction_{N_\upsilon \times N_\upsilon}$, and $\mathit{label}_\upsilon = \mathit{label}_\tau \restriction_{N_\upsilon}$.

[4]Thus, $\triangleleft^*_\tau$ is the reflexive–transitive closure of $\triangleleft_\tau \subseteq N_\tau \times N_\tau$, the relation of *immediate dominance* on N_τ

[5]*sibling*$_\tau(x)$ denotes the (unique) sibling of any given $x \in N_\tau$ different from the root of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$. If $x <_\tau y$ for some $x, y \in N_\tau$ then x is said to *(immediately) project over* y.

[6]A *tree domain* is a non–empty set $N_\upsilon \subseteq \mathbb{N}^*$ such that for all $\chi \in \mathbb{N}^*$ and $i \in \mathbb{N}$ it holds that $\chi \in N_\upsilon$ if $\chi\chi' \in N_\upsilon$ for some $\chi' \in \mathbb{N}^*$, and $\chi i \in N_\upsilon$ if $\chi j \in N_\upsilon$ for some $j \in \mathbb{N}$ with $i < j$. $\langle N_\upsilon, \triangleleft^*_\upsilon, \prec_\upsilon \rangle$ is the *natural (tree) interpretation* of N_υ in the case that for all $\chi, \psi \in N_\upsilon$ it holds that $\chi \triangleleft_\upsilon \psi$ iff $\psi = \chi i$ for some $i \in \mathbb{N}$, and $\chi \prec_\upsilon \psi$ iff $\chi = \omega i \chi'$ and $\psi = \omega j \psi'$ for some $\omega, \chi', \psi' \in \mathbb{N}^*$ and $i, j \in \mathbb{N}$ with $i < j$.

[7]Note that the leaf–labeling function *label*$_\tau$ can easily be extended to a total labeling function ℓ_τ from N_τ into *Feat*$^* \cup \{$< , >$\}$, where < and > are two new distinct symbols: to each non–leaf $x \in N_\tau$ we can assign a label from $\{$< , >$\}$ by ℓ_τ such that $\ell_\tau(x) =$ < iff $y <_\tau z$ for $y, z \in N_\tau$ with $x \triangleleft_\tau y, z$, and $y \prec_\tau z$. In this sense a concrete $\tau \in \mathit{Exp}(\mathit{Feat})$ is depictable in the way demonstrated in Fig. 1.

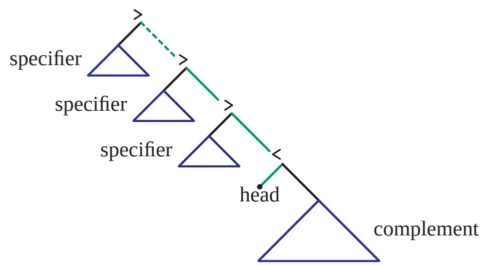

FIG. 2. The typical structure of a (minimalist) expression over *Feat*.

Thus, $\upsilon \in \mathit{Exp}(\mathit{Feat})$. Such an υ is a *maximal projection (in τ)* if υ's root is a node $x \in N_\tau$ such that x is the root of τ, or such that $\mathit{sibling}_\tau(x) <_\tau x$. $\mathit{MaxProj}(\tau)$ is the set of all maximal projections in τ.

$\mathit{comp}_\tau \subseteq \mathit{MaxProj}(\tau) \times \mathit{MaxProj}(\tau)$ is the binary relation defined such that for all $\upsilon, \phi \in \mathit{MaxProj}(\tau)$ it holds that $\upsilon\, \mathit{comp}_\tau\, \phi$ iff $\mathit{head}_\tau(r_\upsilon) <_\tau r_\phi$, where r_υ and r_ϕ are the roots of υ and ϕ, respectively. If $\upsilon\, \mathit{comp}_\tau\, \phi$ for some $\upsilon, \phi \in \mathit{MaxProj}(\tau)$ then ϕ is a *complement of υ (in τ)*. comp^+_τ and comp^*_τ are the transitive and the reflexive–transitive closure of comp_τ. $\mathit{Comp}^+(\tau)$ and $\mathit{Comp}^*(\tau)$ are the sets $\{\upsilon \mid \tau\, \mathit{comp}^+_\tau\, \upsilon\}$ and $\{\upsilon \mid \tau\, \mathit{comp}^*_\tau\, \upsilon\}$, respectively.

$\mathit{spec}_\tau \subseteq \mathit{MaxProj}(\tau) \times \mathit{MaxProj}(\tau)$ is the binary relation defined such that such that for all $\upsilon, \phi \in \mathit{MaxProj}(\tau)$ it holds that $\upsilon\, \mathit{spec}_\tau\, \phi$ iff $r_\phi = \mathit{sibling}_\tau(x)$ for some $x \in N_\tau$ with $r_\upsilon \triangleleft^+_\tau x \triangleleft^+_\tau \mathit{head}_\tau(r_\upsilon)$, where r_υ and r_ϕ are the roots of υ and ϕ, respectively. If $\upsilon\, \mathit{spec}_\tau\, \phi$ for some $\upsilon, \phi \in \mathit{MaxProj}(\tau)$ then ϕ is a *specifier of υ (in τ)*. spec^*_τ is the reflexive–transitive closure of spec_τ. $\mathit{Spec}(\tau)$ and $\mathit{Spec}^*(\tau)$ are the sets $\{\upsilon \mid \tau\, \mathit{spec}_\tau\, \upsilon\}$ and $\{\upsilon \mid \tau\, \mathit{spec}^*_\tau\, \upsilon\}$, respectively.

An $\upsilon \in \mathit{MaxProj}(\tau)$ is said to *have (open) feature f* if the label assigned to υ's head by label_τ is non–empty and starts with an instance of $f \in \mathit{Feat}$.[8]

τ is *complete* if its head–label is in $\{\mathtt{c}\}\mathit{Phon}^*\mathit{Sem}^*$ and each other of its leaf–labels in $\mathit{Phon}^*\mathit{Sem}^*$. Hence, a complete expression over *Feat* is an expression that has category c, and this instance of c is the only instance of a syntactic feature within all leaf–labels.

The *phonetic yield of τ*, denoted by $Y_{\mathit{Phon}}(\tau)$, is the string which results from concatenating in "left–to–right–manner" the labels assigned to the leaves of $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau \rangle$ via label_τ, and replacing all instances of non–phonetic features with the empty string, afterwards.

An $\upsilon = \langle N_\upsilon, \triangleleft^*_\upsilon, \prec_\upsilon, <_\upsilon, \mathit{label}_\upsilon \rangle \in \mathit{Feat}(\mathit{Exp})$ is *(label preserving) isomorphic to τ* if there is a bijective function i from N_τ onto N_υ with $x \triangleleft_\tau y$ iff $i(x) \triangleleft_\upsilon i(y)$, $x \prec_\tau y$ iff $i(x) \prec_\upsilon i(y)$, $x <_\tau y$ iff $i(x) <_\upsilon i(y)$, and with $\mathit{label}_\tau(x) = \mathit{label}_\upsilon(i(x))$ for $x, y \in N_\tau$. i is an *isomorphism (from τ to υ)*.

[8] Thus the expression depicted in Fig. 1 has feature f, while its specifier and its complement have feature g and h, respectively.

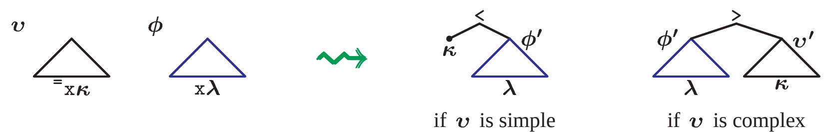

FIG. 3. *merge*(υ, ϕ) according to (me).

DEFINITION 3.2
For $\tau = \langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau\rangle \in \mathit{Exp}(\mathit{Feat})$ with $N_\tau = tN_\upsilon$ for some $t \in \mathbb{N}^*$ and some tree domain N_υ, and for $r \in \mathbb{N}^*$, $(\tau)_r$ denotes the *expression shifting* τ *to* r, i.e. the expression $\langle N_{\tau(r)}, \triangleleft^*_{\tau(r)}, \prec_{\tau(r)}, <_{\tau(r)}, \mathit{label}_{\tau(r)}\rangle$ over *Feat* with $N_{\tau(r)} = rN_\upsilon$ such that the function $i_{\tau(r)}$ from N_τ onto $N_{\tau(r)}$ with $i_{\tau(r)}(tx) = rx$ for all $x \in N_\upsilon$ is an isomorphism from τ to $(\tau)_r$.[9]

For $\upsilon, \phi \in \mathit{Exp}(\mathit{Feat})$ let $\chi = \langle N_\chi, \triangleleft^*_\chi, \prec_\chi, <_\chi, \mathit{label}_\chi\rangle$ be a complex expression over *Feat* with root ϵ such that $(\upsilon)_0$ and $(\phi)_1$ are the two subexpressions of χ whose roots are immediately dominated by ϵ. Then χ is of one of two forms: in order to refer to χ we write $[_<\upsilon, \phi]$ if $0 <_\chi 1$, and $[_>\upsilon, \phi]$ if $1 <_\chi 0$.

DEFINITION 3.3 ([18])
A *minimalist grammar (MG)* is a five–tuple $G = \langle \neg\mathit{Syn}, \mathit{Syn}, \mathit{Lex}, \Omega, \mathtt{c}\rangle$ with Ω being the operator set consisting of the structure building functions *merge* and *move* defined w.r.t. *Feat* as in (me) and (mo) below, respectively, and with *Lex* being a *lexicon (over Feat)*, i.e., *Lex* is a finite set of simple expressions over *Feat*, and each lexical item $\tau \in \mathit{Lex}$ is of the form $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau\rangle$ such that $N_\tau = \{\epsilon\}$, and such that $\mathit{label}_\tau(\epsilon) \in (\mathit{Select} \cup \mathit{Licensors})^* \mathit{Base}\, \mathit{Licensees}^* \mathit{Phon}^* \mathit{Sem}^*$.

(me) *merge* is a partial mapping from $\mathit{Exp}(\mathit{Feat}) \times \mathit{Exp}(\mathit{Feat})$ into $\mathit{Exp}(\mathit{Feat})$. A pair $\langle \upsilon, \phi\rangle$ with $\upsilon, \phi \in \mathit{Exp}(\mathit{Feat})$ belongs to $\mathit{Dom}(\mathit{merge})$ if for some $\mathtt{x} \in \mathit{Base}$ and $\kappa, \lambda \in \mathit{Feat}^*$, conditions (i) and (ii) are fulfilled:

(i) the head–label of υ is $^{=}\mathtt{x}\kappa$ (i.e. υ has selector $^{=}\mathtt{x}$), and
(ii) the head–label of ϕ is $\mathtt{x}\lambda$ (i.e. ϕ has category $\mathtt{x}$).

Then,

(me.1) *merge*$(\upsilon, \phi) = [_<\upsilon', \phi']$ if υ is simple, and
(me.2) *merge*$(\upsilon, \phi) = [_>\phi', \upsilon']$ if υ is complex,

where υ' and ϕ' result from υ and ϕ, respectively, just by deleting the instance of the feature that the respective head–label starts with (cf. Fig. 3).

(mo) *move* is a partial mapping from $\mathit{Exp}(\mathit{Feat})$ into $\mathit{Exp}(\mathit{Feat})$. An $\upsilon \in \mathit{Exp}(\mathit{Feat})$ is in $\mathit{Dom}(\mathit{move})$ if for some $\mathtt{-x} \in \mathit{Licensees}$ and $\kappa \in \mathit{Feat}^*$, (i)–(iii) are true:

---

[9] Note that for each $\tau = \langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau\rangle \in \mathit{Exp}(\mathit{Feat})$, a $t \in \mathbb{N}^*$ and tree domain N_υ with $N_\tau = tN_\upsilon$ exist by (E4).

(i) the head–label of υ is $\texttt{+X}\kappa$ (i.e. υ has licensor $\texttt{+X}$),

(ii) there is exactly one $\phi \in \mathit{MaxProj}(\upsilon)$ with head–label $\texttt{-x}\lambda$ for some $\lambda \in \mathit{Feat}^*$ (i.e. there is exactly one $\phi \in \mathit{MaxProj}(\upsilon)$ that has feature $\texttt{-x}$), and

(iii) there exists a $\chi \in \mathit{Comp}^+(\upsilon)$ with $\phi = \chi$ or $\phi \in \mathit{Spec}(\chi)$.

Then, $\mathit{move}(\upsilon) = [_> \phi', \upsilon']$, where $\upsilon' \in \mathit{Exp}(\mathit{Feat})$ results from υ by canceling the instance of $\texttt{+X}$ the head–label of υ starts with, while the subtree ϕ is replaced by a single node labeled ϵ. $\phi' \in \mathit{Exp}(\mathit{Feat})$ arises from ϕ by deleting the instance of $\texttt{-x}$ the head–label of ϕ starts with (cf. Fig. 4).

FIG. 4. *move*(υ) according to (mo).

DEFINITION 3.4 ([18])
A *strict minimalist grammar (SMG)* is a five–tuple of the form $\langle \neg\mathit{Syn}, \mathit{Syn}, \mathit{Lex}, \Omega, , \texttt{c}\rangle$ with Ω being the operator set consisting of the structure building functions *merge* and *move*$^{\mathrm{s}}$ defined w.r.t. *Feat* as in (me) above and (smo) below, respectively, and with *Lex* being a lexicon over *Feat* defined as in Definition 3.3 such that it additionally holds that $\mathit{label}_\tau(\epsilon) = \mathit{Select}_\epsilon(\mathit{Select} \cup \mathit{Licensors})_\epsilon \mathit{Base}\,\mathit{Licensees}^*\mathit{Phon}^*\mathit{Sem}^*$ for each $\langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau\rangle \in \mathit{Lex}$.

(smo) *move*$^{\mathrm{s}}$ is a partial mapping from $\mathit{Exp}(\mathit{Feat})$ into $\mathit{Exp}(\mathit{Feat})$. An $\upsilon \in \mathit{Exp}(\mathit{Feat})$ is in $\mathit{Dom}(\mathit{move})$ if for some $\texttt{-x} \in \mathit{Licensees}$ and $\kappa \in \mathit{Feat}^*$, (i)–(iii) are true:

(i) the head–label of υ is $\texttt{+X}\kappa$ (i.e. υ has licensor $\texttt{+X}$),

(ii) there is exactly one $\phi \in \mathit{MaxProj}(\upsilon)$ with head–label $\texttt{-x}\lambda$ for some $\lambda \in \mathit{Feat}^*$ (i.e. there is exactly one $\phi \in \mathit{MaxProj}(\upsilon)$ that has feature $\texttt{-x}$), and

(iii) there exists a $\chi \in \mathit{Comp}^+(\upsilon)$ with $\phi \in \mathit{Spec}^*(\chi)$.[10]

Then, $\mathit{move}^{\mathrm{s}}(\upsilon) = [_> \chi', \upsilon']$, canceling the instance of $\texttt{+X}$ the head–label of υ starts with, while the subtree χ is replaced by a single node labeled ϵ. $\chi' \in \mathit{Exp}(\mathit{Feat})$ arises from χ by deleting the instance of $\texttt{-x}$ the head–label of ϕ starts with (cf. Fig. 5).

For each (S)MG $G = \langle \neg\mathit{Syn}, \mathit{Syn}, \mathit{Lex}, \Omega, \texttt{c}\rangle$ the *closure of* G, $\mathit{CL}(G)$, is the set $\bigcup_{k \in \mathbb{N}} \mathit{CL}^k(G)$, where $\mathit{CL}^0(G) = \mathit{Lex}$, and for $k \in \mathbb{N}$, $\mathit{CL}^{k+1}(G) \subseteq \mathit{Exp}(\mathit{Feat})$ is recursively defined as the set

[10] Note that such a $\chi \in \mathit{Comp}^+(\upsilon)$ is unique.

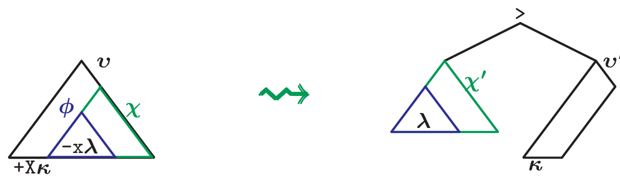

FIG. 5. *move*$^{\mathrm{s}}(\upsilon)$ according to (smo).

$$\begin{aligned}CL^k(G) \;\cup\; &\{merge(\upsilon,\phi) \mid \langle\upsilon,\phi\rangle \in Dom(merge) \cap CL^k(G) \times CL^k(G)\}\\ \cup\; &\{move'(\upsilon) \mid \upsilon \in Dom(move') \cap CL^k(G)\},\end{aligned}$$

where $move' \in \Omega \setminus \{merge\}$. The set $\{Y_{Phon}(\tau) \mid \tau \in CL(G)$ and τ complete$\}$, denoted by $L(G)$, is the *(string) language derivable by* G.

DEFINITION 3.5
A set L is a *(strict) minimalist language ((S)ML)* if $L = L(G)$ for some (S)MG G.

Just in order to complete the picture in terms of a formal definition we give

DEFINITION 3.6
An (S)MG G and an MCFG G' are *weakly equivalent* if they derive the same (string) language, i.e. if $L(G) = L(G')$.

3.1 Relevant Expressions of an (S)MG

Throughout the end of this section we assume $G = \langle\neg Syn, Syn, Lex, \Omega, \mathtt{c}\rangle$ to be an (S)MG and turn now to a "concept of relevance" being of central importance, when we examine the weak generative power of (S)MGs. We write *move*$'$ in order to refer to the corresponding move–operator, *move* or *move*$^{\mathrm{s}}$, respectively, belonging to Ω, i.e. we have $\Omega = \{merge, move'\}$.

We start with a brief motivation of the corresponding formal settings: apart from the head, each leaf of a complete expression of G is labeled by a string that does not contain any instance of some syntactic feature, whereas the head–label contains exactly one such instance, namely, an instance of the completeness category c. Thus, all instances of syntactic features within the leaf–labels of an expression τ of G that serves to derive a complete expression of G have to be canceled at some later stage of the derivation, except for the possible appearance of that instance of the completeness category which finally becomes the instance within the head–label of the complete expression. Each leaf of such an expression τ of G determines a maximal projection in τ, where the label of each leaf different from the head is a string that contains beside instances of non–syntactic features at most instances of syntactic features which belong to the set of licensees. These instances of licensee features can be checked only by a performance of the corresponding move–operator, *move*$'$. Since, in order to end up in a complete expression, each of these licensee instances is checked at some stage of the derivation by applying the move–operator either to τ or to an expression of G derived on the base of τ, τ has to fulfil a particular property, namely, that of *being relevant (to G)* in the sense of exactly one of the following two definitions, depending on whether G is an MG or an SMG, respectively.

DEFINITION 3.7
In case G is an MG, a given $\tau \in \mathit{CL}(G)$ is *relevant (to G)* if it has property (R_{MG}).

(R_{MG}) For each $\texttt{-x} \in \mathit{Licensees}$ there is at most one $\tau_{\texttt{-x}} \in \mathit{MaxProj}(\tau)$ that has feature $\texttt{-x}$, and if it exists then $\tau_{\texttt{-x}} \in \mathit{Comp}^{+}(\tau)$ or $\tau_{\texttt{-x}} \in \mathit{Spec}(\chi)$ for a $\chi \in \mathit{Comp}^{*}(\tau)$.

DEFINITION 3.8
In case G is an SMG, a given $\tau \in \mathit{CL}(G)$ is *relevant (to G)* if it has property ($\mathrm{R}_{\mathrm{SMG}}$).

($\mathrm{R}_{\mathrm{SMG}}$) For each $\texttt{-x} \in \mathit{Licensees}$ there is at most one $\tau_{\texttt{-x}} \in \mathit{MaxProj}(\tau)$ that has feature $\texttt{-x}$, and if it exists then $\tau_{\texttt{-x}} \in \mathit{Spec}^{*}(\chi)$ for a $\chi \in \mathit{Comp}^{*}(\tau)$.

DEFINITION 3.9
The *relevant closure of* G, denoted by $\mathit{RCL}(G)$, is the set of all relevant $\tau \in \mathit{CL}(G)$.

Consider some $\tau \in \mathit{CL}(G)$ that is relevant to G according to the respective definition, (R_{MG}) or ($\mathrm{R}_{\mathrm{SMG}}$). First of all τ must not contain two different maximal projections that have the same licensee $\texttt{-x}$. Furthermore, the uniqueness of a maximal projection in τ that has a particular licensee feature must be accompanied by a further property: the fulfillment of a particular condition on where this maximal projection is located in τ. This additional demand depends on whether G is an MG or an SMG as formulated in (R_{MG}) and ($\mathrm{R}_{\mathrm{SMG}}$), respectively. In fact, this kind of expression structure is typical of each $\tau \in \mathit{CL}(G)$ involved in creating a complete expression of G as will become clear immediately. This is not to say that each expression which is relevant to G by definition has to occur within some derivation of a complete expression of G. In this sense, *being relevant to* G rather means *being potentially relevant* in order to derive any complete expression of G.

PROPOSITION 3.10
Let $\tau \in \mathit{RCL}(G)$, and let $\upsilon, \phi \in \mathit{CL}(G)$. If $\tau = \mathit{merge}(\upsilon, \phi)$ then $\upsilon, \phi \in \mathit{RCL}(G)$, and if $\tau = \mathit{move}(\upsilon)$ then $\upsilon \in \mathit{RCL}(G)$.

Instead of providing a formal proof we want to address the important point underlying the last proposition in a descriptive way: recall that in any case an arbitrary expression $\tau' \in \mathit{CL}(G)$ belongs to the domain of the respective move–operator only if there is exactly one maximal projection in τ' that has a particular licensee feature allowing the projection's movement into a specifier position, i.e. only if (i) and (ii) of ((s)mo) are fulfilled by τ'. But in addition this maximal projection is subject to a condition concerning its structural position within τ', namely, the corresponding condition (iii) of (mo) or (smo). This implies that, whenever an expression $\tau' \in \mathit{CL}(G)$ is irrelevant to G, it contains a maximal projection that has some licensee feature `-x` which cannot be checked by applying the move–operator of G to τ'. Crucially, this "bad property" is inherited by any expression of G which is derived with the help of an instance such a τ'. That is to say, if τ' participates in a derivation of some expression τ of G, the corresponding instance of the licensee `-x` will still be unchecked within τ. Thus, because each complete $\tau \in \mathit{CL}(G)$ is relevant to G, we can fix

COROLLARY 3.11
Each $\tau' \in \mathit{CL}(G)$ that serves to derive a complete expression of G is also an element of $\mathit{RCL}(G)$, and we have $L(G) = \{Y_{\mathit{Phon}}(\tau) \,|\, \tau \in \mathit{RCL}(G) \text{ and } \tau \text{ complete}\}$.

The converse of Proposition 3.10 does not hold. Indeed, it may happen that from expressions in $\mathit{RCL}(G)$ irrelevant expressions are derivable.

PROPOSITION 3.12
$\mathit{RCL}(G)$ is generally closed neither under the merge–operator nor the move–operator.

The following paragraph gives an overview of the cases revealing the last proposition, where τ is supposed to be an element of $\mathit{CL}(G)$.

merge :

(ir.1) Assume that $\tau = \mathit{merge}(\upsilon, \phi)$ for some $\upsilon, \phi \in \mathit{RCL}(G)$.

Then $\tau \notin \mathit{RCL}(G)$ in each of the cases (ir.1.1)–(ir.1.4).

(ir.1.1) There are $\upsilon' \in \mathit{MaxProj}(\upsilon)$ and $\phi' \in \mathit{MaxProj}(\phi)$ that both have the same licensee `-y`.

(ir.1.2) There is a $\chi \in \mathit{MaxProj}(\upsilon) \cup \mathit{MaxProj}(\phi)$ that has licensee `-y`, and the head–label of ϕ is of the form $\mathtt{x\text{-}y}\lambda$ for some $\mathtt{x} \in \mathit{Base}$ and $\lambda \in \mathit{Feat}^*$, i.e. $\lambda \in \mathit{Licensees}^*\mathit{Phon}^*\mathit{Sem}^*$.

(ir.1.3) G is an MG, the expression υ is complex, and there is a $\phi' \in \mathit{MaxProj}(\phi)$ that has some licensee `-y`.

(ir.1.4) G is an SMG, the expression υ is complex, and there is a $\phi' \in \mathit{MaxProj}(\phi)$ that has some licensee `-y` such that $\phi' \notin \mathit{Spec}^+(\phi)$, hence $\phi' \notin \mathit{Spec}^*(\phi)$.

move' :

(ir.2) Assume that $\tau = \mathit{move}'(\upsilon)$ for some $\upsilon \in \mathit{RCL}(G)$.

Then $\tau \notin \mathit{RCL}(G)$ in each of the cases (ir.2.1)–(ir.2.3).

(ir.2.1) There is some $\phi \in \mathit{MaxProj}(\upsilon)$ whose head–label is of the form $\mathtt{-x\text{-}y}\lambda$ such that υ has licensor `+X`, and such that there is some $\phi' \in \mathit{MaxProj}(\upsilon)$ that has feature `-y`, where $\mathtt{-x}, \mathtt{-y} \in \mathit{Licensees}$ and $\lambda \in \mathit{Feat}^*$.

(ir.2.2) G is an MG, and there exists a $\phi \in \mathit{MaxProj}(\upsilon)$ that has some licensee `-x` such that υ has licensor `+X`, and such that there is some $\phi' \in \mathit{MaxProj}(\phi)$ different from ϕ that has some licensee `-y`.

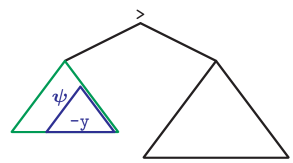

FIG. 6. A typical $\tau \in \mathit{CL}(G)$ in the sense of (ir.1.3), (ir.1.4), (ir.2.2) or (ir.2.3).

(ir.2.3) G is an SMG, and there exists a $\phi \in \mathit{MaxProj}(\upsilon)$ that has some licensee `-x` such that υ has licensor `+X`, and such that there is some $\phi' \in \mathit{MaxProj}(\phi)$ that has some licensee `-y`, but that does not belong to $\mathit{Spec}^*(\chi)$ for the existing $\chi \in \mathit{Comp}^+(\upsilon)$ with $\phi \in \mathit{Spec}^*(\chi)$.

We see that in cases (ir.1.1), (ir.1.2) and (ir.2.1) a corresponding application of *merge* or *move*′ produces an irrelevant expression that contradicts (R_{MG}) as well as (R_{SMG}), i.e. an expression of G that is not relevant to G independently of whether G is an MG an SMG, because the resulting expression τ contains two different maximal projections which have the same licensee, namely `-y`. The other cases arise specificly when dealing with an MG or SMG, respectively. We want to take up the latter two, (ir.2.2) and (ir.2.3), in somewhat more descriptive terms here: assume that we have $\upsilon \in \mathit{Dom}(\mathit{move})$, respectively $\upsilon \in \mathit{Dom}(\mathit{move}^{\mathrm{s}})$, for some $\upsilon \in \mathit{RCL}(G)$. Thus υ has licensor `+X` for some licensee `-x`. Let ϕ be the maximal projection in υ that has feature `-x`. In case G is an MG, $\mathit{move}(\upsilon)$ belongs to $\mathit{RCL}(G)$ only if ϕ does not properly include a maximal projection ψ that has some licensee feature `-y`. In case G is an SMG, such a ψ may exist. But it has to belong to the same set $\mathit{Spec}^*(\chi)$ for some $\chi \in \mathit{Comp}^+(\upsilon)$ as ϕ does in order to let $\mathit{move}^{\mathrm{s}}(\upsilon)$ become an element of $\mathit{RCL}(G)$. Otherwise, applying *move*, respectively *move*$^{\mathrm{s}}$, results in an expression that does not fulfil property (R_{MG}), respectively (R_{SMG}), since ϕ, respectively χ, is moved into a specifier position (cf. Fig. 6). Similar considerations arise if two expressions are merged in case the selecting tree is complex. This is due to the fact that, if we have $\upsilon, \phi \in \mathit{RCL}(G)$ such that υ is complex and $\mathit{merge}(\upsilon, \phi)$ is defined, ϕ is selected as a specifier by υ (cf. Fig. 6).

Indeed, the possibility to construct an $\mathrm{LCFRS}_{1,2}$, "not just" an LCFRS of rank 2, which is weakly equivalent to the G results essentially from G's property that no proper subconstituent is extractable out of some specifier in order to move into another specifier position, i.e. the phonetic yield of a specifier cannot be "devided" into proper substrings by any application of some structure building function.[11]

Note that (R_{MG}) and (R_{SMG}) "deviate" from (mo) and (smo), respectively, since they allow, for each $\upsilon \in \mathit{RCL}(G)$, a maximal projection $\phi \in \mathit{MaxProj}(\upsilon)$ having some licensee `-x` to belong to $\mathit{Spec}(\upsilon)$ or $\mathit{Spec}^*(\upsilon)$, respectively. This does not preclude the head–label of such an υ from starting with an instance of a matching licensor feature,

[11] Note that, in this respect, both the MG–type and the SMG–type differ from the UMG–type, i.e. MGs as originally defined in [17].

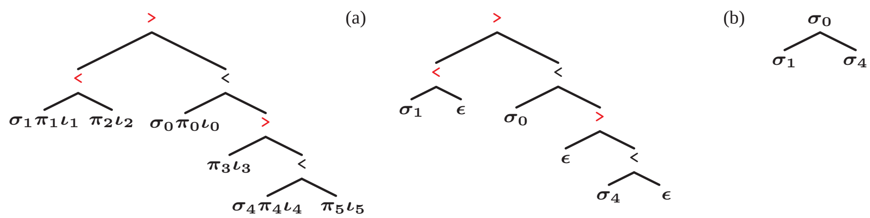

FIG. 7. Basic idea to define a finite partition of $RCL(\widetilde{G})$.

"although" υ does not belong to the domain of the respective move–operator in this case. Nevertheless, we generally have to be aware of the fact that the corresponding instance of `-x` would determine a potentially movable maximal projection as soon as υ was selected as a (right) complement.

4 Transforming (S)MGs into weakly equivalent MCFGs

Throughout this section we let $\widetilde{G} = \langle \neg\mathit{Syn}, \mathit{Syn}, \mathit{Lex}, \Omega, \mathtt{c} \rangle$ be an MG, respectively an SMG, in the sense of Definition 3.3, respectively Definition 3.4. For the appropriate $m \in \mathbb{N}$ we take $\langle \mathtt{-l}_i \rangle_{1 \leq i \leq m}$ to be an enumeration of *Licensees* $\subseteq$ *Syn*. Below, we construct a weakly equivalent $\text{LCFRS}_{1,2}$ $G = \langle N, O, F, R, S \rangle$, i.e., $\mathit{rank}(G) = 2$ and, if $A \to f(B, C) \in R$ then $d_G(B) = 1$ (cf. Definition 2.4). In order to achieve this, G will operate at least w.r.t. syntactic features, on equivalence classes of a finite partition of $RCL(\widetilde{G})$ rather than on single expressions.

The basic idea in order to define the corresponding equivalence class of a given $\tau \in RCL(\widetilde{G})$ is the following: (a) delete all non–syntactic features within the leaf-labels of τ, and (b) reduce τ to those nodes which are the root of some maximal projection with an open (i.e. unchecked) syntactic feature, while in parallel, the head–label of such a maximal projection becomes the label of its root (cf. Fig. 7). In this sense $RCL(\widetilde{G})$ is partitioned into a finite number of equivalence classes: the tree resulting from τ has at most $m + 1$ different nodes, since τ is relevant to $\widetilde{G}$. Furthermore, each node label of the resulting tree is the suffix of the syntactic prefix of the label of a lexical item, because of the particular feature consuming character of the structure building operators. Regarding this partition, applications of the structure building functions do crucially not depend on the chosen representatives. Our weakly equivalent LCFRS G will actually operate w.r.t. a somewhat finer, but still finite partition, since some more structural information about each relevant τ is necessary in both cases, $\widetilde{G}$ being an MG or an SMG, in order to be able at all to construct such an LCFRS. The tree resulting from the corresponding reduction of a $\tau \in RCL(\widetilde{G})$, at least virtually, becomes a nonterminal $T \in N \setminus \{S\}$ representing τ in G. τ's phonetic yield will be separately coded by some $p_T \in O_\Sigma$, a finite tuple of strings of phonetic features, that takes into account the structural information stored in T (cf. Definition 4.1, respectively 4.5). p_T will be derivable from T in G as a finite recursion on functions in F, since for each particular application of *merge* or *move*, respectively *move*$^{\mathrm{s}}$, in $\widetilde{G}$ there will be some nonterminating rule in R simulating the corresponding structure building step in $\widetilde{G}$ (Proposition 4.2, respectively 4.6).[12] Vice versa it will hold that, whenever some $p_T \in O_\Sigma$ is derivable in G from some $T \in N \setminus \{S\}$, there is some $\tau \in \mathit{RCL}(\widetilde{G})$ to which T and p_T correspond as outlined above (Proposition 4.3, respectively 4.7).

Since each $\tau \in \mathit{Lex}$ is simple, we identify τ with its head–label. Also for the sake of convenience, we can w.l.o.g. assume that for each $\mathtt{x} \in \mathit{Base}$ a corresponding feature $\mathtt{=x}$ is present within the set *Select*, and for each $\mathtt{-x} \in \mathit{Licensees}$ a corresponding licensor $\mathtt{+X}$ belongs to the set *Licensors*. For technical reasons we define sets $\mathrm{suf}(\mathit{Syn})$, $\mathrm{suf}(\mathtt{-l}_i)$ for $1 \leq i \leq m$, and I_m:

- $\mathrm{suf}(\mathit{Syn}) = \{\kappa \in \mathit{Syn}^* \,|\, \text{there are a } \kappa' \in \mathit{Syn}^* \text{ and } \zeta \in \neg\mathit{Syn}^* \text{ with } \kappa'\kappa\zeta \in \mathit{Lex}\}$
- $\mathrm{suf}(\mathtt{-l}_i) = \{\kappa \in \mathrm{suf}(\mathit{Syn}) \,|\, \kappa = \mathtt{-l}_i\lambda \text{ for some } \lambda \in \mathit{Syn}^*\} \cup \{\epsilon\}$
- $I_m = \{i_1 \cdots i_n \,|\, n \in \mathbb{N}, i_1, \ldots, i_n \in \{1, \ldots, m\} \text{ with } i_j \neq i_k \text{ if } j \neq k\}$

The set I_m is finite, because in particular $|\iota| \leq m$ for each $\iota \in I_m$. By (N1), each $\mathrm{suf}(\mathtt{-l}_i)$ as well as $\mathrm{suf}(\mathit{Syn})$ is finite, and $\mathrm{suf}(\mathtt{-l}_i) \subseteq \mathit{Licensees}^*$.

Our next goal is to show how particular $m+1$–tuples can be recruited to code a finite (ordered) tree whose set of nodes is a subset of $\{0, \ldots, m\}$ such that 0 belongs to this subset and represents the root of the tree. For this purpose consider any $m+1$–tuple $\alpha = \langle \alpha_0, \ldots, \alpha_m \rangle$ such that $\alpha_i \in \{1, \ldots, m\}^*$ for each $0 \leq i \leq m$, and such that the requirements (a1) and (a2) are met.

(a1) $\alpha_0 \cdots \alpha_m \in I_m$.

(a2) If $\alpha_i = \beta j \gamma$ for some $i, j \in \{1, \ldots, m\}$ and $\beta, \gamma \in \{1, \ldots, m\}^*$ then

(i) $i \neq j$, and

(ii) $\alpha_k = \beta' i \gamma'$ for some $k \in \{0, \ldots, m\}$ and $\beta', \gamma' \in \{1, \ldots, m\}^*$.

Let $\lhd_\alpha$ be the binary relation on $\{0, \ldots, m\}$ such that $i \lhd_\alpha j$ iff

($\lhd_\alpha$) $\alpha_i = \beta j \gamma$ for some $\beta, \gamma \in \{1, \ldots, m\}^*$.

Let $\lhd_\alpha^+$ and $\lhd_\alpha^*$ be the transitive and the reflexive–transitive closure of $\lhd_\alpha$, respectively, and assume $\prec_\alpha$ to be the binary relation on $\{0, \ldots, m\}$ such that $i \prec_\alpha j$ iff

($\prec_\alpha$) $\alpha_k = \beta i' \gamma j' \delta$ for some $\beta, \gamma, \delta \in \{1, \ldots, m\}^*$
and some $i', j', k \in \{0, \ldots, m\}$ such that $i' \lhd_\alpha^* i$ and $j' \lhd_\alpha^* j$.

[12] For each $\tau \in \mathit{Lex}$ there will be a terminating rule $T \to p_T \in R$ with $T \in N$ and $p_T \in O_\Sigma$ coding τ in the sense just mentioned.

Then $\Delta_\alpha = \{i \in \{0,\dots,m\} \,|\, 0 \lhd_\alpha^* i\}$ consists of $|\alpha_0 \cdots \alpha_m| + 1$ elements, and $\langle \Delta_\alpha, \lhd_\alpha^* \restriction_{\Delta_\alpha \times \Delta_\alpha}, \prec_\alpha \restriction_{\Delta_\alpha \times \Delta_\alpha} \rangle$ is a finite ordered tree with root 0.

We take *Treecodings*(m) to denote the set of all $m+1$–tuples $\langle \alpha_0, \dots, \alpha_m \rangle$, where $\alpha_i \in \{1,\dots,m\}^*$ for $0 \leq i \leq m$, that are in accordance with (a1) and (a2). *Treecodings*(m) is finite, since I_m is.

4.1 MLs as MCFLs

Assume $\widetilde{G}$ to be an MG. We write G_{MG} instead of $\widetilde{G}$. In order to construct the weakly equivalent LCFRS $G = \langle N, O, F, R, S \rangle$ as desired we take `sim`, `com`, `true` and `false` to be pairwise distinct new symbols and next define the sets N and O, the set of nonterminals and the set of tuples of terminal strings, respectively.

• • • Each nonterminal $T \in N$ is either the start symbol S or an $m+2$–tuple of the form $\langle \widehat{\mu}_0, \widehat{\mu}_1, \dots, \widehat{\mu}_m, t \rangle$ with $t \in \{\mathtt{sim}, \mathtt{com}\}$ and $\widehat{\mu}_i$ a quadruple $\langle \mu_i, \alpha_i, f_i, t_i \rangle$ for $0 \leq i \leq m$, where

(n1) $\mu_0 \in \text{suf}(\textit{Syn}) \setminus \bigcup_{i=1}^m \text{suf}(\mathtt{-l}_i)$ and $\mu_i \in \text{suf}(\mathtt{-l}_i)$ for $1 \leq i \leq m$,

(n2) $\alpha_i \in \{1,\dots,m\}^*$ for $0 \leq i \leq m$ with $\langle \alpha_0, \dots, \alpha_m \rangle \in$ *Treecodings*(m), and

(n3) $f_0 = 0$ and $f_i \in \{-1\} \cup \{0,\dots,m\}$ for $1 \leq i \leq m$

such that for $1 \leq j \leq m$ it holds that[13]

(n4) $\mu_j \neq \epsilon$ iff $i \lhd_T j$ for some $0 \leq i \leq m$ iff $f_j \geq 0$.

Extending the catalogue of requirements, the following ones are met:

for $f_{\max} = \max\{f_i \,|\, 0 \leq i \leq m\}$ it holds that

(n5) $\{f_i \,|\, 0 \leq i \leq m\} \cap \{0,\dots,m\} = \{0,\dots,f_{\max}\}$,

and for $1 \leq j \leq m$, (n6) is true.

(n6) If $\mu_j \neq \epsilon$ then $i = j$ for each $0 \leq i \leq m$ with $f_i = f_j$.[14]

Finally, we demand

(n7) $t_0 = \mathtt{true}$ and $t_i \in \{\mathtt{true}, \mathtt{false}\}$ for $1 \leq i \leq m$

such that for $1 \leq j \leq m$ it holds that

(n8) if $\mu_j = \epsilon$ then $t_j = \mathtt{true}$.

[13] In the following, for each corresponding $T \in N \setminus \{S\}$ we write $\lhd_T$ and $\prec_T$ instead of $\lhd_\alpha$ and $\prec_\alpha$, respectively, to denote the relations as they result from the "tree–coding" tuple $\alpha = \langle \alpha_0, \dots, \alpha_m \rangle$ according to ($\lhd_\alpha$) and ($\prec_\alpha$).

[14] Hence, we may conclude from (n3)–(n6) that the function $\widetilde{f}$ from $\{0\} \cup \{1 \leq j \leq m \,|\, \mu_j \neq \epsilon\}$ into $\{0,\dots,f_{\max}\}$ defined by $\widetilde{f}(i) = f_i$ is a bijection with $\widetilde{f}(0) = 0$.

• • • $O = \bigcup_{n \in \mathbb{N}} \langle \Sigma^* \rangle^{n+1}$, $\Sigma = \mathit{Phon}$ the set of phonetic features in G_{MG}.

It is straightforward that the set N is in fact finite, since, up to the start symbol S, N constitutes a subset of a finite product of finite products of finite sets. Disregarding non–syntactic features, we can use N to characterize the relevant expressions of G_{MG}, the set $\mathit{RCL}(G_{\text{MG}})$. This set itself is generally non–finite. The phonetic yield of an expression $\tau \in \mathit{RCL}(G_{\text{MG}})$ can be characterized then as a particular tuple from O depending on a corresponding nonterminal from N which depicts τ up to non–syntactic features. Definition 4.1 below aims to motivate the definitions of N and O in some more detail: for this purpose let $\tau = \langle N_\tau, \triangleleft_\tau^*, \prec_\tau, <_\tau, \mathit{label}_\tau \rangle \in \mathit{RCL}(G_{\text{MG}})$, and for $1 \leq i \leq m$ take, if there is any $\tau' \in \mathit{MaxProj}(\tau)$ that has licensee $\texttt{-l}_i$, τ_i to be such a τ'.[15] Otherwise, take τ_i to be a single node labeled ϵ. Furthermore, set $\tau_0 = \tau$, and for $0 \leq i \leq m$ let r_i be the root of τ_i. Also, let $T = \langle \widehat{\mu}_0, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, t \rangle \in N$ with $t \in \{\texttt{sim}, \texttt{com}\}$, and with $\widehat{\mu}_i = \langle \mu_i, \alpha_i, f_i, t_i \rangle$ for $0 \leq i \leq m$ in accordance with (n1)–(n5). Then define $f_{\max} = \max\{f_i \,|\, 0 \leq i \leq m\}$ and finally choose some $p_T = \langle \pi_0, \pi_1, \ldots, \pi_{f_{\max}} \rangle$ from $\langle \mathit{Phon}^* \rangle^{f_{\max}+1}$.

DEFINITION 4.1
The pair $\langle T, p_T \rangle$ *corresponds to* τ if (D1)–(D4) are true.

(D1) For $0 \leq i \leq m$, μ_i is the prefix of τ_i's head-label which consists of just the syntactic features, and $t = \texttt{sim}$ iff τ is simple.

(D2) For $0 \leq i, j \leq m$ with $\mu_i, \mu_j \neq \epsilon$, $i \triangleleft_T^+ j$ iff $r_i \triangleleft_\tau^+ r_j$, and $i \prec_T j$ iff $r_i \prec_\tau r_j$.

(D3) If $\mu_i \neq \epsilon$ for some $0 \leq i \leq m$ then π_{f_i} is the phonetic yield of τ_i', where τ_i' results from τ_i by replacing for each $1 \leq j \leq m$ with $i \triangleleft_T^+ j$ the (proper) subtree τ_j of τ_i by a single node labeled ϵ in case there is no $1 \leq k \leq m$ with $i \triangleleft_T^+ k \triangleleft_T^+ j$.[16]

(D4) If $t_i = \texttt{false}$ for some $1 \leq i \leq m$ then $\tau_i \in \mathit{Spec}(\tau)$. If $t_i = \texttt{true}$ and $\mu_i \neq \epsilon$ for some $1 \leq i \leq m$ then $\tau_i \in \mathit{Comp}^+(\tau)$ or $\tau_i \in \mathit{Spec}(\tau')$ for some $\tau' \in \mathit{Comp}^+(\tau)$.

Note that (D1) determines a particular, finite partition $\mathcal{P}$ on $\mathit{RCL}(G_{\text{MG}})$: in the corresponding manner, to each $\tau \in \mathit{RCL}(G_{\text{MG}})$ exactly one element belonging to the product $\text{suf}(\mathit{Syn}) \times \text{suf}(\texttt{-l}_1) \times \ldots \times \text{suf}(\texttt{-l}_m) \times \{\texttt{sim}, \texttt{com}\}$ can be assigned.[17]

(D2) can be seen then as introducing a refinement $\mathcal{P}_{\mathit{ref}}$ of $\mathcal{P}$: expressions τ from one equivalence class are distinguished w.r.t. proper dominance, $\triangleleft_\tau^+$, and precedence, $\prec_\tau$, as it holds between each two distinct maximal projections τ_i and τ_j whose head–labels start with some licensee $\texttt{-l}_i$ and $\texttt{-l}_j$, respectively. This distinction can be manifested by assigning to each $\tau \in \mathit{RCL}(G_{\text{MG}})$ a particular $m+1$–tuple $\langle \alpha_0, \ldots, \alpha_m \rangle$ from *Treecodings*(m) by means of (D2).

---

[15] Then τ_i is unique, because τ fulfills property (R_{MG}).

[16] Recall that for $1 \leq i \leq m$, $\mu_i \neq \epsilon$ iff $f_i > 0$. Hence, if $\mu_i \neq \epsilon$ for some $0 \leq i \leq m$ then the corresponding π_{f_i} in fact exists.

[17] As a finite product of finite sets this product is also a finite set.

Again let $T = \langle \widehat{\mu}_0, \ldots, \widehat{\mu}_m, t\rangle \in N$ with $t \in \{\mathtt{sim}, \mathtt{com}\}$ and $\widehat{\mu}_i = \langle \mu_i, \alpha_i, f_i, t_i\rangle$ for $0 \leq i \leq m$ as in (n1)–(n8), define $f_{\max} = \max\{f_i \,|\, 0 \leq i \leq m\}$, and choose some $p_T = \langle \pi_0, \pi_1, \ldots, \pi_{f_{\max}}\rangle$ from $\langle \mathit{Phon}^*\rangle^{f_{\max}+1}$. Assume that $\langle T, p_T\rangle$ corresponds to some $\tau \in \mathit{RCL}(G_{\text{MG}})$ as in Definition 4.1.

Then t as well as each μ_i and each α_i for $0 \leq i \leq m$ is unique, because (D1) and (D2) hold. For each possible combination consisting of f_0, f_1, $\ldots$, f_m there is exactly one p_T that satisfies the requirement (D3). The μ_i's, the α_i's and t determine the equivalence class of τ w.r.t. the refined partition $\mathcal{P}_{\mathit{ref}}$ on $\mathit{RCL}(G_{\text{MG}})$. If $f_i = -1$ for some $1 \leq i \leq m$ then $\mu_i = \epsilon$. Recall that $f_0 = 0$. If $f_i \geq 0$ for some $0 \leq i \leq m$ then $\mu_i \neq \epsilon$. In this case, according to (D3), the component π_{f_i} of p_T specifies that part of the phonetic yield of τ_i that is definitely “non–extractable.” That is to say, no movement can apply to τ such that a proper subexpression of τ_i is extracted pied–piping some (proper) subpart of π_{f_i}.

The requirement (D4) finally equips T, in terms of correspondence, with some further structural information about τ as to where the maximal projections in τ that have a licensee feature are located. To put it differently, we are provided by means of T with some further knowledge of τ as it is necessary concerning the potential application of the move–operator in G_{MG}. If $t_i = \mathtt{false}$ for some $1 \leq i \leq m$ then T mirrors the fact that, within the configuration of the corresponding τ, the attraction of τ_i by the head of τ into a specifier position is not allowed under any circumstances, i.e. an application of *move* would be blocked even if the head–label of τ started with an instance of the corresponding licensor $\mathtt{+L}_i$. Having this information accessible in T will be one important prerequisite in order to successfully define the LCFRS $G = \langle N, O, F, R, S\rangle$ such that it will be weakly equivalent to G_{MG}.

••• The set F of functions and the set R of rewriting rules are simultaneously defined w.r.t. the occurrence of an $f \in F$ within an $r \in R$.

Nonterminating rules: First of all we define two initial rules by

(r0) $S \to id_{\mathit{Phon}^*}(T) \in R$ for $T = \langle \widehat{\mu}_0, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, t\rangle \in N$

with $\widehat{\mu}_0 = \langle \mathtt{c}, \epsilon, 0, \mathtt{true}\rangle$, with $\widehat{\mu}_i = \langle \epsilon, \epsilon, -1, \mathtt{true}\rangle$ for $1 \leq i \leq m$, and with $t \in \{\mathtt{sim}, \mathtt{com}\}$. id_{Phon^*} is the identity function on *Phon**.

Next, for $\mathtt{x} \in \mathit{Base}$ suppose that

$\mathtt{x}\lambda \in \text{suf}(\mathit{Syn})$ with $\lambda \in \mathit{Syn}^*$, i.e., $\lambda \in \mathit{Licensees}^*$, /cf. (ii) of (me)/
$s\kappa \in \text{suf}(\mathit{Syn})$ with $s = \mathtt{=x}$ and $\kappa \in \mathit{Syn}^*$, /cf. (i) of (me)/
$\nu_i, \xi_i \in \text{suf}(\mathtt{-l}_i)$ for $1 \leq i \leq m$ /cf. (R_{MG})/

such that for $1 \leq i \leq m$ it holds that

$\nu_i = \epsilon$ or $\xi_i = \epsilon$, /cf. (ir.1.1)/
$\nu_i = \xi_i = \epsilon$ if $\lambda = \mathtt{-l}_i\lambda'$ with $\lambda' \in \mathit{Syn}^*$. /cf. (ir.1.2)/

Choose g_i, $h_i \in \{-1\} \cup \{0, \dots, m\}$, u_i, $v_i \in \{\texttt{true}, \texttt{false}\}$ for $1 \le i \le m$, β_i, $\gamma_i \in \{1, \dots, m\}^*$ for $0 \le i \le m$, and $u, v \in \{\texttt{sim}, \texttt{com}\}$ such that

$$U = \langle \langle s\kappa, \beta_0, 0, \texttt{true} \rangle, \langle \nu_1, \beta_1, g_1, u_1 \rangle, \dots, \langle \nu_m, \beta_m, g_m, u_m \rangle, u \rangle \in N,$$
$$V = \langle \langle \texttt{x}\lambda, \gamma_0, 0, \texttt{true} \rangle, \langle \xi_1, \gamma_1, h_1, v_1 \rangle, \dots, \langle \xi_m, \gamma_m, h_m, v_m \rangle, v \rangle \in N,$$

and such that, additionally,

if $u = \texttt{com}$ then $\xi_i = \epsilon$ for each $1 \le i \le m$, /cf. (ir.1.3)/
if $u = \texttt{sim}$ then $\nu_i = \epsilon$ for each $1 \le i \le m$.[18] /cf. (D1)/

Proceeding, let $g_{\max} = \max\{g_i \mid 0 \le i \le m\}$, and $h_{\max} = \max\{h_i \mid 0 \le i \le m\}$.

In case $\lambda = \epsilon$ we set $j = 0$ and take

$$T = \langle \langle \kappa, \gamma_0\beta_0, 0, \texttt{true} \rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \texttt{com} \rangle \in N \text{ , and}$$

in case $\lambda \neq \epsilon$ we choose the uniquely existing $1 \le j \le m$ such that $\lambda = \texttt{-l}_j\lambda'$ for some $\lambda' \in \text{suf}(\mathit{Syn}^*)$ and we take

$$T = \langle \langle \kappa, j\beta_0, 0, \texttt{true} \rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \texttt{com} \rangle \in N,$$

where in both cases for $1 \le i \le m$ we have

$$\widehat{\mu}_i = \begin{cases} \langle \nu_i, \beta_i, g_i, u_i \rangle & \text{if } i \neq j \text{ and } \xi_i = \epsilon \\ \langle \xi_i, \gamma_i, g_{\max} + h_i, \texttt{true} \rangle & \text{if } i \neq j \text{ and } \xi_i \neq \epsilon \\ \langle \lambda, \gamma_0, g_{\max} + h_{\max} + 1, \texttt{ true } \rangle & \text{if } i = j \text{ and } u = \texttt{sim} \\ \langle \lambda, \gamma_0, g_{\max} + h_{\max} + 1, \texttt{false} \rangle & \text{if } i = j \text{ and } u = \texttt{com} \end{cases}$$

For the function $\mathit{merge}_{U,V} \in F$ as defined below, we finally let

(r1) $T \to \mathit{merge}_{U,V}\big(U, V\big) \in R$ if $u = \texttt{sim}$, and
$T \to \mathit{merge}_{U,V}\big(V, U\big) \in R$ if $u = \texttt{com}$.

Assume x_0, $x_1, \dots, x_m$, y_0, $y_1, \dots, y_m$ to be pairwise distinct variables, and define $x = \langle x_0, x_1 \dots, x_{g_{\max}} \rangle$ and $y = \langle y_0, y_1, \dots, y_{h_{\max}} \rangle$. Note that $g_{\max} = 0$ if $u = \texttt{sim}$, and that $h_{\max} = 0$ if $u = \texttt{com}$. Moreover, let $\widehat{\jmath} \in \{0, 1\}$ be such that $\widehat{\jmath} = 0$ iff $j = 0$.

In case $u = \texttt{sim}$, $\mathit{merge}_{U,V}$ is the function from $\mathit{Phon}^* \times \langle \mathit{Phon}^* \rangle^{h_{\max}+1}$ into the set $\langle \mathit{Phon}^* \rangle^{h_{\max}+\widehat{\jmath}+1}$ defined by

$$\langle x, y \rangle \mapsto \begin{cases} \langle x_0 y_0, y_1, \dots, y_{h_{\max}} \rangle & \text{if } \widehat{\jmath} = 0 \\ \langle \ x_0 \ \ , y_1, \dots, y_{h_{\max}}, y_0 \rangle & \text{if } \widehat{\jmath} = 1 \end{cases}$$

In case $u = \texttt{com}$, $\mathit{merge}_{U,V}$ is the function from $\mathit{Phon}^* \times \langle \mathit{Phon}^* \rangle^{g_{\max}+1}$ into the set $\langle \mathit{Phon}^* \rangle^{g_{\max}+\widehat{\jmath}+1}$ defined by

[18] According to (n2) and (n4) the latter two restrictions are equivalent to the respective following two: if $u = \texttt{com}$ then for each $0 \le i \le m$ with $0 \lhd_V^* i$ it holds that $i = 0$; if $u = \texttt{sim}$ then for each $0 \le i \le m$ with $0 \lhd_U^* i$ it holds that $i = 0$.

$$\langle y, x\rangle \mapsto \begin{cases} \langle y_0 x_0, x_1, \ldots, x_{g_{\max}}\rangle & \text{if } \widehat{\jmath} = 0 \\ \langle\ x_0\ \ , x_1, \ldots, x_{g_{\max}}, y_0\rangle & \text{if } \widehat{\jmath} = 1 \end{cases}$$

Now, for some $1 \leq j \leq m$, suppose that

$\nu_j \in \text{suf}(\texttt{-l}_j)$ with $\nu_j = \texttt{-l}_j\lambda$ for a $\lambda \in$ *Licensees**, /cf. (ii) of (mo)/
$l\kappa \in \text{suf}(\textit{Syn})$ with $l = \texttt{+L}_j$ and $\kappa \in$ *Syn**, /cf. (i) of (mo)/
$\nu_i \in \text{suf}(\texttt{-l}_i)$ for $1 \leq i \leq m$ with $i \neq j$ /cf. (R_{MG})/

such that for $1 \leq i \leq m$ with $i \neq j$ it holds that

$\nu_k = \epsilon$ if $\lambda = \texttt{-l}_k\lambda'$ with $\lambda' \in$ *Syn**. /cf. (ir.2.1)/

Next choose elements $g_i \in \{-1\} \cup \{0, \ldots, m\}$, $u_i \in \{\texttt{true}, \texttt{false}\}$ for $1 \leq i \leq m$, and $\beta_i \in \{1, \ldots, m\}^*$ for $0 \leq i \leq m$ such that

$$U = \langle\langle l\kappa, \beta_0, 0, \texttt{true}\rangle, \langle\nu_1, \beta_1, g_1, u_1\rangle, \ldots, \langle\nu_m, \beta_m, g_m, u_m\rangle, \texttt{com}\rangle \in N,$$

and such that, additionally,

$u_j = \texttt{true}$, /cf. (iii) of (mo)/

and furthermore,

$\beta_i = \epsilon$ for each $1 \leq i \leq m$ with $j \lhd_U^+ i$. /cf. (ir.2.2)/

Let $j' \in \{0, \ldots, m\}$ such that $\beta_{j'} = \zeta_{j'} j \eta_{j'}$ for some $\zeta_{j'}, \eta_{j'} \in \{1, \ldots, m\}^*$.[19]

In case $\lambda = \epsilon$ we set $k = 0$ and take

$$T = \langle\langle \kappa, b_0, \beta_j\beta, 0, \texttt{true}\rangle, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, \texttt{com}\rangle \in N.$$

In case $\lambda \neq \epsilon$ we choose the uniquely existing $1 \leq k \leq m$ such that $\lambda = \texttt{-l}_k\lambda'$ for some $\lambda' \in$ *Syn** and we take

$$T = \langle\langle \kappa, b_0, k\beta, 0, \texttt{true}\rangle, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, \texttt{com}\rangle \in N,$$

where in both cases β is defined by $\beta = \zeta_{j'}\eta_{j'}$ if $j' = 0$, and $\beta = \beta_0$ otherwise. Furthermore, for each $1 \leq i \leq m$ we have

$$\widehat{\mu}_i = \begin{cases} \langle\lambda, \beta_j, g_j, \texttt{false}\rangle & \text{if } i = k \\ \langle\epsilon, \epsilon, -1, \texttt{true}\rangle & \text{if } i \neq k \text{ and } i = j \\ \langle\nu_i, \zeta_i\eta_i, \widetilde{g}_i, u_i\rangle & \text{if } i \neq k \text{ and } i = j' \\ \langle\nu_i,\ \beta_i\ , \widetilde{g}_i, u_i\rangle & \text{if } i \neq k,\ i \neq j \text{ and } i \neq j' \end{cases}$$

where $\widetilde{g}_i = g_i - 1$ if $k = 0$ and $g_i > g_j$, and $\widetilde{g}_i = g_i$ otherwise.

Now, for the function *move*$_U \in F$ as defined below we let

---

[19] Thus, $j' \lhd_U j$ and therefore $j' \neq j$. Such a j' exists and is unique by (n2) and (n4).

(r2) $T \to \mathit{move}_U\big(U\big) \in R$.

Let $g_{\max} = \max\{g_i \,|\, 0 \le i \le m\}$, let $x_0, x_1, \dots, x_m$ be pairwise distinct variables, define $x = \langle x_0, \dots, x_{g_{\max}} \rangle$, and let $\widehat{k} \in \{0,1\}$ be such that $\widehat{k} = 0$ iff $k = 0$.

move_U is the function from $\langle \mathit{Phon}^* \rangle^{g_{\max}+1}$ into $\langle \mathit{Phon}^* \rangle^{g_{\max}+\widehat{k}}$ defined by

$$x \mapsto \begin{cases} \langle x_{g_j} x_0, x_1, \dots, x_{g_j - 1}, x_{g_j + 1}, \dots, x_{g_{\max}} \rangle & \text{if } k = 0 \\ \langle x_0, x_1, \dots, x_{g_{\max}} \rangle & \text{if } k \neq 0 \end{cases}$$

Terminating rules: For each $\kappa\pi\iota \in \mathit{Lex}$ with $\kappa \in \mathit{Syn}^*$, $\pi \in \mathit{Phon}^*$ and $\iota \in \mathit{Sem}^*$ let

(r3) $T \to \pi \in R$, where $T = \langle\langle \kappa, \epsilon, 0, \texttt{true}\rangle, \widehat{\nu}_1, \dots, \widehat{\nu}_m, \texttt{sim}\rangle \in N \setminus \{S\}$

with $\widehat{\nu}_i = \langle \epsilon, \epsilon, -1, \texttt{true}\rangle$ for $1 \le i \le m$.

The weak equivalence of the given MG G_{MG} and the LCFRS G as constructed above from G_{MG} is fixed in Corollary 4.4, which can be seen as a consequence of two propositions, namely, Proposition 4.2 and 4.3.

PROPOSITION 4.2
If $\tau \in \mathit{RCL}(G_{\text{MG}})$ then a $T \in N \setminus \{S\}$ and a $p_T \in L_G(T)$ exist such that $\langle T, p_T \rangle$ corresponds to τ according to Definition 4.1.

PROOF (SKETCH). The proof follows from an induction verifying (4.2$_k$) for $k \in \mathbb{N}$.

(4.2$_k$) If $\tau \in \mathit{CL}^k(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$, there are some $T = \langle \widehat{\mu}_0, \dots, \widehat{\mu}_m, t \rangle \in N$ and $p_T \in L^k_G(T)$ such that $\langle T, p_T \rangle$ corresponds to τ in the sense of Definition 4.1.

Because $\mathit{Lex} = \mathit{CL}^0(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$, (4.2$_0$) holds according to (r3). As far as the induction step is concerned we will skip specific details, emphasizing the crucial line of argumentation: whenever for some $k \in \mathbb{N}$, some $\tau \in \mathit{CL}^{k+1}(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$ results from an application of *merge* to a pair $\langle \upsilon, \phi \rangle$ for some $\upsilon, \phi \in \mathit{CL}^k(G_{\text{MG}})$, υ and ϕ must not be in line with (ir.1.1)–(ir.1.3), since τ is relevant to G_{MG}. Note that υ and ϕ are from $\mathit{RCL}(G_{\text{MG}})$ by Proposition 3.10. Hence, for any $U, V \in N$ and $p_U, p_V \in O_\Sigma$ such that $\langle U, p_U \rangle$ and $\langle V, p_V \rangle$ correspond to υ and ϕ, respectively, U and V obey the restrictions demanded in order to appear on the righthand side of some rule of the form (r1). By hypothesis such $U, V \in N$ and $p_U, p_V \in O_\Sigma$ exist such that $p_U \in L^k_G(U)$ and $p_V \in L^k_G(V)$. By definition, for the corresponding nonterminal T from (r2) and the tuple $p_T = \mathit{merge}_{U,V}(p_U, p_V) \in L_G(T)$, the pair $\langle T, p_T \rangle$ corresponds to τ.

If for some $k \in \mathbb{N}$, some $\tau \in \mathit{CL}^{k+1}(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$ is of the form $\mathit{move}(\upsilon)$ for some $\upsilon \in \mathit{CL}^k(G_{\text{MG}})$, υ cannot be subject to (ir.2.1) and (ir.2.2). $\upsilon \in \mathit{RCL}(G_{\text{MG}})$ by Proposition 3.10. Hence, for any $U \in N$ and $p_U \in O_\Sigma$ such that $\langle U, p_U \rangle$ corresponds to υ, U provides the case of an U as it appears in (r2). Here by hypothesis, a corresponding U and p_U exist such that $p_U \in L^k_G(U)$, and by definition, for the corresponding nonterminal T from (r2) and the tuple $p_T = \mathit{move}_U(p_U) \in L_G(T)$, the pair $\langle T, p_T \rangle$ corresponds to τ. $\square$

PROPOSITION 4.3
If $p_T \in L_G(T)$ for some $T \in N \setminus \{S\}$ and $p_T \in O$, then there is a $\tau \in \mathit{RCL}(G_{\text{MG}})$ that corresponds to $\langle T, p_T \rangle$ in the sense of Definition 4.1.

PROOF (SKETCH). Here, an induction proving (4.3$_k$) for $k \in \mathbb{N}$ yields the result.

(4.3$_k$) If $p_T \in L_G^k(T)$ for some $T \in N \setminus \{S\}$ then $\langle T, p_T \rangle$ corresponds to some $\tau \in \mathit{CL}^k(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$.

Again because $\mathit{Lex} = \mathit{CL}^0(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$, (4.3$_0$) holds according to (r3). Regarding the induction step, we summarize the decisive points: consider first the case that for some $k \in \mathbb{N}$, there are some $U, V \in N$, and some $p_U \in L_G^k(U)$ and $p_V \in L_G^k(V)$ such that U and V fulfil the requirements to occur on the righthand side of some rule of the form (r1). Then for the corresponding T from (r1) and $p_T = \mathit{merge}_{U,V}(p_U, p_V)$, we have $p_T \in L_G^{k+1}(T)$. By hypothesis there are not only $\upsilon, \phi \in \mathit{CL}^k(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$ such that $\langle U, p_U \rangle$ and $\langle V, p_V \rangle$ correspond to υ and ϕ, respectively. The restrictions applying to U and V also ensure that $\tau = \mathit{merge}(\upsilon, \phi)$ is defined, and they ensure that υ and ϕ do not chime in with one of the cases (ir.1.1)–(ir.1.3). Thus, τ is relevant to G_{MG}. Moreover, $\langle T, p_T \rangle$ corresponds to τ.

Now, assume that for some $k \in \mathbb{N}$ there are any $U \in N$ and $p_U \in L_G^k(U)$ such that U fulfills the requirements to occur on the righthand side of some rule of the form (r2). For the corresponding T from (r2) and $p_T = \mathit{move}_U(U)$, we have $p_T \in L_G^{k+1}(T)$. By hypothesis there is an $\upsilon \in \mathit{CL}^k(G_{\text{MG}}) \cap \mathit{RCL}(G_{\text{MG}})$ such that $\langle U, p_U \rangle$ corresponds to υ. The restrictions applying to U imply that $\upsilon \in \mathit{Dom}(\mathit{move})$, and they prevent υ from constituting a case in the sense of (ir.2.1) or (ir.2.2). Therefore, $\tau = \mathit{move}(\upsilon)$ is defined and relevant to G_{MG}. The pair $\langle T, p_T \rangle$ corresponds to τ. $\square$

COROLLARY 4.4
$\pi \in L(G)$ iff $\pi \in L(G_{\text{MG}})$ for each $\pi \in \mathit{Phon}^*$.

PROOF. To show that the "if–part" holds, choose complete $\tau \in \mathit{CL}(G_{\text{MG}})$ with phonetic yield $\pi \in \mathit{Phon}^*$. Let $T = \langle \widehat{\mu}_0, \ldots, \widehat{\mu}_m, t \rangle \in N$ with $t \in \{\texttt{sim}, \texttt{com}\}$, and with $\widehat{\mu}_i = \langle \mu_i, \alpha_i, f_i, t_i \rangle$ for $0 \leq i \leq m$ as in (n1)–(n8), and let $p_T \in \langle \mathit{Phon}^* \rangle^{f_{\max}+1}$, where $f_{\max} = \max\{f_i \,|\, 0 \leq i \leq m\}$. Assume that $\langle T, p_T \rangle$ corresponds to τ in the sense of Definition 4.1. By Proposition 4.2, T and p_T exist such that $p_T \in L_G(T)$. Because τ is complete, we have $\widehat{\mu}_0 = \langle \texttt{c}, \epsilon, 0, \texttt{true} \rangle$ and $\widehat{\mu}_i = \langle \epsilon, \epsilon, -1, \texttt{true} \rangle$ for $1 \leq i \leq m$. From (r0) we therefore conclude that $\pi \in L_G(S) = L(G)$.

To prove the "only if"–part, take some $\pi \in L(G) = L_G(S)$. By definition of R each rule applicable to S is of the form (r0). Thus, there is a $p_T \in L_G(T) \subseteq \mathit{Phon}^*$ such that $\pi = id_{\mathit{Phon}^*}(p_T) = p_T$ for $T \in N$ as in (r0). $\langle T, p_T \rangle$ corresponds to some $\tau \in \mathit{RCL}(G_{\text{MG}})$ by Proposition 4.3. This τ is complete by (D1), and $\pi = p_T$ is the yield of τ by (D3). $\square$

4.2 SMLs as MCFLs

We now suppose $\widetilde{G}$ to be an SMG, and we write G_{SMG} instead of $\widetilde{G}$. Constructing the weakly equivalent LCFRS$_{1,2}$ $G = \langle N, O, F, R, S\rangle$, we again take $\mathtt{sim}$ and $\mathtt{com}$ to be pairwise distinct new symbols and continue by providing the definitions of N and O, the set of nonterminals and the set of tuples of terminal strings, respectively.

• • • Each nonterminal $T \in N$ is either the start symbol S or an $m + 2$–tuple of the form $\langle \widehat{\mu}_0, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, t\rangle$ with $t \in \{\mathtt{sim}, \mathtt{com}\}$, and with $\widehat{\mu}_i$ a triple $\langle \mu_i, \alpha_i, f_i\rangle$ for $0 \leq i \leq m$, where

(n1) $\mu_0 \in \mathrm{suf}(\mathit{Syn}) \setminus \bigcup_{i=1}^{m} \mathrm{suf}(\mathtt{-l}_i)$ and $\mu_i \in \mathrm{suf}(\mathtt{-l}_i)$ for $1 \leq i \leq m$,

(n2) $\alpha_i \in \{1, \ldots, m\}^*$ for $0 \leq i \leq m$ with $\langle \alpha_0, \ldots, \alpha_m\rangle \in \mathit{Treecodings}(m)$, and

(n3) $f_0 = 0$ and $f_i \in \{-1\} \cup \{0, \ldots, m\}$ for $1 \leq i \leq m$

such that for $1 \leq j \leq m$ it holds that[20]

(n4) $\mu_j \neq \epsilon$ iff $i \lhd_T j$ for some $0 \leq i \leq m$ iff $f_j \geq 0$,

and such that for $f_{\max} = \max\{f_i \,|\, 0 \leq i \leq m\}$ it holds that

(n5) $\{f_i \,|\, 0 \leq i \leq m\} \cap \{0, \ldots, m\} = \{0, \ldots, f_{\max}\}$.[21]

Note that, by (n1)–(n4), it holds that

- ○ for each $0 \leq j \leq m$ with $\mu_j \neq \epsilon$ there is exactly one $0 \leq i \leq m$ with $f_i = f_j$ such that $i \lhd_T^* k$ or $i \prec_T k$ for each $0 \leq k \leq m$ with $f_i = f_k$.

• • • We let $O = \bigcup_{n \in \mathbb{N}} \langle \Sigma^* \rangle^{n+1}$, $\Sigma = \mathit{Phon}$ the set of phonetic features in G_{SMG}.

Consider $T = \langle \widehat{\mu}_0, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, t\rangle \in N$ with $t \in \{\mathtt{sim}, \mathtt{com}\}$ and $\widehat{\mu}_i = \langle \mu_i, \alpha_i, f_i\rangle$ for $0 \leq i \leq m$ in accordance with (n1)–(n5). From one perspective, each such nonterminal T could be seen as being "internally" less complex than a corresponding nonterminal of the "MG–case." Here, each of the first $m + 1$ components, i.e. each $\widehat{\mu}_i$ with $0 \leq i \leq m$, is a triple instead of a quadruple that has just lost its fourth component $t_i \in \{\mathtt{true}, \mathtt{false}\}$. But at the same time, the relation between those first $m + 1$–components has become more involved. To put it differently, where condition (n6) in conjunction with (n3) formerly ensured the uniqueness of the third component f_i of each $\widehat{\mu}_i$ with first component $\mu_i \neq \epsilon$ for $0 \leq i \leq m$, the omission of condition (n6) now allows such an f_i to be identical to another f_j for some $0 \leq j \leq m$ with $j \neq i$ under particular circumstances. Of course, up to the start symbol S, N is

[20] In the following, adhering to the notational conventions of the last subsection, for each corresponding $T \in N \setminus \{S\}$ we write $\lhd_T$ and $\prec_T$ instead of $\lhd_\alpha$ and $\prec_\alpha$, respectively, to denote the relations as they result from the "tree–coding" tuple $\alpha = \langle \alpha_0, \ldots, \alpha_m\rangle$ according to ($\lhd_\alpha$) and ($\prec_\alpha$).

[21] Hence, we may conclude from (n3)–(n5) that the function $\widetilde{f}$ from $\{0\} \cup \{1 \leq j \leq m \,|\, \mu_j \neq \epsilon\}$ into $\{0, \ldots, f_{\max}\}$ defined by $\widetilde{f}(i) = f_i$ is a surjection with $\widetilde{f}(0) = 0$.

nevertheless a subset of a finite product of finite products of finite sets and, for this reason, a finite set itself.

Now, let $\tau = \langle N_\tau, \triangleleft^*_\tau, \prec_\tau, <_\tau, \mathit{label}_\tau \rangle \in \mathit{RCL}(G_{\text{SMG}})$. For $1 \le i \le m$ we choose, if possible at all, some $\tau_i \in \mathit{MaxProj}(\tau)$ that has licensee $\texttt{-l}_i$.[22] Otherwise, we take τ_i to be a single node labeled ϵ. Further, we set $\tau_0 = \tau$, and for $0 \le i \le m$ we let r_i denote the root of τ_i. We also take some $T = \langle \widehat{\mu}_0, \widehat{\mu}_1, \dots, \widehat{\mu}_m, t \rangle \in N$ with $t \in \{\texttt{sim}, \texttt{com}\}$ and $\widehat{\mu}_i = \langle \mu_i, \alpha_i, f_i \rangle$ for $0 \le i \le m$ in accordance with (n1)–(n5), and we then define $f_{\max} = \max\{f_i \,|\, 0 \le i \le m\}$. Depending on T we let $p_T = \langle \pi_0, \pi_1, \dots, \pi_{f_{\max}} \rangle$ be an $f_{\max} + 1$–tuple from O.

DEFINITION 4.5
The pair $\langle T, p_T \rangle$ *corresponds to* τ if (SD1)–(SD4) are true.

(SD1) For $0 \le i \le m$, μ_i is the prefix of τ_i's head-label which consists of just the syntactic features, and $t = \texttt{sim}$ iff τ is simple.

(SD2) For $0 \le i, j \le m$ with $\mu_i, \mu_j \ne \epsilon$, $i \triangleleft^+_T j$ iff $r_i \triangleleft^+_\tau r_j$, and $i \prec_T j$ iff $r_i \prec_\tau r_j$.

(SD3) If $\mu_i \ne \epsilon$ for some $0 \le i \le m$ then π_{f_i} is the phonetic yield of υ', where υ' results from the uniquely existing $\upsilon \in \mathit{Comp}^*(\tau)$ with $\tau_i \in \mathit{Spec}^*(\upsilon)$ in the following way: we replace by a single node labeled ϵ each $\phi \in \mathit{Comp}^+(\upsilon)$ for which there is some $1 \le j \le m$ with $\mu_j \ne \epsilon$ such that $\tau_j \in \mathit{Spec}^*(\phi)$, and such that there is no $1 \le k \le m$ with $\mu_k \ne \epsilon$ such that $\tau_k \in \mathit{Spec}^*(\chi)$ for some $\chi \in \mathit{Comp}^+(\tau)$ with $\phi \in \mathit{Comp}^+(\chi)$.[23]

(SD4) If there is some $\upsilon \in \mathit{Comp}^*(\tau)$ for which there is some $1 \le i \le m$ with $\mu_i \ne \epsilon$ such that $\tau_i \in \mathit{Spec}^*(\upsilon)$ then for each $0 \le j \le m$ with $\mu_j \ne \epsilon$ it holds that $\tau_j \in \mathit{Spec}^*(\upsilon)$ iff $f_j = f_i$.

We see that, fitting in with (D1) and (D2) of Definition 4.1, the corresponding *corresponds to*–definition in the "MG–case," (SD1) and (SD2) mirror the possibility of defining a particular finite partition $\mathcal{P}$ on $\mathit{RCL}(G_{\text{SMG}})$ by means of the product $\text{suf}(\mathit{Syn}) \times \text{suf}(\texttt{-l}_1) \times \dots \times \text{suf}(\texttt{-l}_m) \times \{\texttt{sim}, \texttt{com}\}$, and a particular refinement $\mathcal{P}_{\mathit{ref}}$ of $\mathcal{P}$ by means of *Treecodings*(m). In other words, the equivalence class in $\mathcal{P}_{\mathit{ref}}$ to which τ belongs is characterized in terms of $\langle \mu_0, \mu_1, \dots, \mu_m, t \rangle$ and $\langle \alpha_0, \alpha_1, \dots, \alpha_m \rangle$. Recall that the relevance condition (R_{SMG}) fulfilled by τ modifies the relevance condition (R_{MG}) for expressions of an MG without changing the requirement that for each licensee $\texttt{-x}$ there is at most one maximal projection in τ that has this licensee.

(SD3) departs from (D3) in a significant way. This is due to the fact that the move–operator move^{s} of the SMG G_{SMG} is not simply a restriction of the corresponding structure building function *move* from $\mathit{Exp}(\neg \mathit{Syn} \cup \mathit{Syn})$ into $\mathit{Exp}(\neg \mathit{Syn} \cup \mathit{Syn})$: whenever a maximal projection τ' in τ has some licensee $\texttt{-x}$ that triggers an application of $\mathit{move}^{\text{sr}}$,

[22] If such a τ_i exists, it is unique. Recall that τ has property (R_{SMG}).

[23] If a corresponding ϕ exists then it is unique, because $\mathit{Comp}^+(\upsilon)$ is totally ordered by comp^*_τ. Furthermore, according to (n4), for each $1 \le i \le m$ we have $\mu_i = \epsilon$ iff $f_i = -1$. Hence, if $\mu_i \ne \epsilon$ for some $1 \le i \le m$ then the corresponding π_{f_i} in fact exists.

it will be the maximal projection $\upsilon \in \mathit{Comp}^+(\tau)$ with $\tau' \in \mathit{Spec}^*(\upsilon)$ which will move into a specifier position. If the corresponding instance of the licensee $\texttt{-x}$ were to trigger an application of *move*, it would be τ' itself that would move. We clarify the intentions underlying (SD3) in somewhat more detail having a look at (SD4).

(SD4) is actually rather self–explanatory, taking into account the relevance condition (R_{SMG}) as it differs from the corresponding condition (R_{MG}) for MGs: because (R_{SMG}) is obeyed by τ, for each $1 \leq i \leq m$ with $\mu_i \neq \epsilon$ there is some expression $\upsilon \in \mathit{Comp}^*(\tau)$ such that $\tau_i \in \mathit{Spec}^*(\upsilon)$. Such an υ is unique, because $\mathit{Comp}^*(\tau)$ is totally ordered by the binary relation comp^*_τ. (SD4) now states that, whenever for some $\upsilon \in \mathit{Comp}^*(\tau)$ there is some $1 \leq i \leq m$ with $\mu_i \neq \epsilon$ such that $\tau_i \in \mathit{Spec}^*(\upsilon)$, then for all and only those $0 \leq j \leq m$ with $f_j = f_i$ the maximal projection τ_j also belongs to $\mathit{Spec}^*(\upsilon)$. In particular, this ensures a specific kind of unique relation between such an $\upsilon \in \mathit{Comp}^*(\tau)$ and some component of p_T, namely, π_{f_i}:[24] according to (SD3), this component π_{f_i} specifies that part of the phonetic yield of υ which is not "extractable" by any means. That is, the part that will not become subject to pied–piping if any movement of some maximal projection $\upsilon' \in \mathit{Comp}^+(\tau)$ takes place that has been potentially licensed by means of $\mu_j \neq \epsilon$ for some $1 \leq j \leq m$ with $\tau_j \in \mathit{Spec}^*(\upsilon')$.

• • • The set F of functions and the set R of rewriting rules are simultaneously defined w.r.t. the occurrence of an $f \in F$ within an $r \in R$.

<u>*Nonterminating rules*</u>: First of all we define two initial rules by

(r0) $S \longrightarrow id_{\mathit{Phon}^*}(T) \in R$ for $T = \langle \widehat{\mu}_0, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, t \rangle \in N$

with $\widehat{\mu}_0 = \langle \texttt{c}, \epsilon, 0 \rangle$, $\widehat{\mu}_i = \langle \epsilon, \epsilon, -1 \rangle$ for $1 \leq i \leq m$ and $t \in \{\texttt{sim}, \texttt{com}\}$. The function id_{Phon^*} from Phon^* onto Phon^* is the identity function represented by $x \mapsto x$ for some variable x.

For $\texttt{x} \in \mathit{Base}$ suppose that

$\texttt{x}\lambda \in \mathrm{suf}(\mathit{Syn})$ with $\lambda \in \mathit{Syn}^*$, i.e. $\lambda \in \mathit{Licensees}^*$, /cf. (ii) of (me)/
$s\kappa \in \mathrm{suf}(\mathit{Syn})$ with $s = \texttt{=x}$ and $\kappa \in \mathit{Syn}^*$, /cf. (i) of (me)/
$\nu_i, \xi_i \in \mathrm{suf}(\texttt{-l}_i)$ for $1 \leq i \leq m$ /cf. (R_{SMG})/

such that for $1 \leq i \leq m$ it holds

$\nu_i = \epsilon$ or $\xi_i = \epsilon$, /cf. (ir.1.1)/
$\nu_j = \xi_j = \epsilon$ if $\lambda = \texttt{-l}_j\lambda'$ with $\lambda' \in \mathit{Syn}^*$. /cf. (ir.1.2)/

We choose elements $g_i, h_i \in \{-1\} \cup \{0, \ldots, m\}$ for $1 \leq i \leq m$, $\beta_i, \gamma_i \in \{1, \ldots, m\}^*$ for $0 \leq i \leq m$, and $u, v \in \{\texttt{sim}, \texttt{com}\}$ such that

[24] Note also that, since $f_0 = 0$, and since $\tau = \tau_0$ is the only $\upsilon \in \mathit{Comp}^*(\tau)$ with $\tau_0 \in \mathit{Spec}^*(\upsilon)$, it is a consequence of (SD4) that for each $1 \leq j \leq m$ with $\mu_j \neq \epsilon$ we have $\tau_j \in \mathit{Spec}^*(\tau)$ iff $f_j = 0$.

$$U = \langle\langle s\kappa, \beta_0, 0\rangle, \langle\nu_1, \beta_1, g_1\rangle, \dots, \langle\nu_m, \beta_m, g_m\rangle, u\rangle \in N,$$
$$V = \langle\langle \mathtt{x}\lambda, \gamma_0, 0\rangle, \langle\xi_1, \gamma_1, h_1\rangle, \dots, \langle\xi_m, \gamma_m, h_m\rangle, v\rangle \in N,$$

and such that, additionally,

if $u = \mathtt{com}$ then $\xi = \epsilon$ for each $1 \leq i \leq m$ with $h_i \neq 0$,[25] /cf. (ir.1.4)/
if $u = \mathtt{sim}$ then $\nu_i = \epsilon$ for each $1 \leq i \leq m$. /cf. (SD1)/

Proceeding, we let $g_{\max} = \max\{g_i \,|\, 0 \leq i \leq m\}$, and $h_{\max} = \max\{h_i \,|\, 0 \leq i \leq m\}$.
In case $u = \mathtt{sim}$, $\lambda = \epsilon$ and $h_i \neq 0$ for each $1 \leq i \leq m$, we set $j = 0$, and we take

$$T = \langle\langle \kappa, \gamma_0\beta_0, 0\rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \mathtt{com}\rangle \in N.$$

In case $u = \mathtt{sim}$, $\lambda = \epsilon$ and $h_i = 0$ for some $1 \leq i \leq m$, we choose the uniquely existing $1 \leq j \leq m$ with $h_j = 0$ such that $j \triangleleft_V^* i$ or $j \prec_V i$ for each $1 \leq i \leq m$ with $h_i = 0$, and we take

$$T = \langle\langle \kappa, \gamma_0\beta_0, 0\rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \mathtt{com}\rangle \in N.$$

In case $u = \mathtt{sim}$ and $\lambda \neq \epsilon$, we choose the uniquely existing $1 \leq j \leq m$ such that $\lambda = \mathtt{-l}_j\lambda'$ for some $\lambda' \in \mathrm{suf}(\mathit{Syn}^*)$, and we take

$$T = \langle\langle \kappa, j\beta_0, 0\rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \mathtt{com}\rangle \in N,$$

where for all subcases of $u = \mathtt{sim}$, for $1 \leq i \leq m$ we have

$$\widehat{\mu}_i = \begin{cases} \langle\nu_i, \beta_i, g_i\rangle & \text{if } i \neq j \text{ and } \xi_i = \epsilon \\ \langle\xi_i, \gamma_i, h_i\rangle & \text{if } i \neq j \text{ and } \xi_i \neq \epsilon\text{, and if } h_i \neq 0 \\ \langle\xi_i, \gamma_i, h_{\max}+1\rangle & \text{if } i \neq j \text{ and } \xi_i \neq \epsilon\text{, and if } h_i = 0 \\ \langle\xi_i, \gamma_i, h_{\max}+1\rangle & \text{if } i = j \text{ and } \lambda = \epsilon \text{ (i.e. } h_i = 0) \\ \langle\lambda, \gamma_0, h_{\max}+1\rangle & \text{if } i = j \text{ and } \lambda \neq \epsilon \text{ (i.e. } h_i = -1) \end{cases}$$

In case $u = \mathtt{com}$ and $\lambda = \epsilon$, we set $j = 0$, and we let

$$T = \langle\langle \kappa, \gamma_0\beta_0, 0\rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \mathtt{com}\rangle \in N.$$

In case $u = \mathtt{com}$ and $\lambda \neq \epsilon$, we choose existing and unique $1 \leq j \leq m$ with $\lambda = \mathtt{-l}_j\lambda'$ for some $\lambda' \in \mathrm{suf}(\mathit{Syn}^*)$, and we take

$$T = \langle\langle \kappa, j\beta_0, 0\rangle, \widehat{\mu}_1, \dots, \widehat{\mu}_m, \mathtt{com}\rangle \in N,$$

where in all cases of $u = \mathtt{com}$, for $1 \leq i \leq m$ we have

$$\widehat{\mu}_i = \begin{cases} \langle\nu_i, \beta_i, g_i\rangle & \text{if } i \neq j \text{ and } \xi_i = \epsilon \\ \langle\xi_i, \gamma_i, h_i\rangle & \text{if } i \neq j \text{ and } \xi_i \neq \epsilon \text{ (i.e. } h_i = 0) \\ \langle\lambda, \gamma_0, 0\rangle & \text{if } i = j \end{cases}$$

[25] According to (n2)–(n4) this restriction is equivalent to the following: if $u = \mathtt{com}$ then $h_i = 0$ for each $1 \leq i \leq m$ with $0 \triangleleft_V^* i$.

Then, for $\mathit{merge}_{U,V} \in F$ as defined below, we finally let

(r1) $T \to \mathit{merge}_{U,V}\big(U,V\big) \in R$ if $u = \mathtt{sim}$,
$T \to \mathit{merge}_{U,V}\big(V,U\big) \in R$ if $u = \mathtt{com}$.

As in the previous subsection we let $x_H,\ x_0,\ x_1,\ldots,x_m,\ y_H,\ y_0,\ y_1,\ldots,y_m$ be pairwise distinct variables. Then we respectively define x and y as the $g_{\max}+1$–tuple $\langle x_0, x_1 \ldots, x_{g_{\max}}\rangle$ and the $h_{\max}+1$–tuple $\langle y_0, y_1, \ldots, y_{h_{\max}}\rangle$. Again we have $g_{\max} = 0$ in case $u = \mathtt{sim}$, and $h_{\max} = h_H = 0$ in case $u = \mathtt{com}$, and again we define $\widehat{\jmath} \in \{0,1\}$ to be such that $\widehat{\jmath} = 0$ iff $j = 0$.

In case $u = \mathtt{sim}$, $\mathit{merge}_{U,V}$ is the function from $\mathit{Phon}^* \times \langle \mathit{Phon}^*\rangle^{h_{\max}+1}$ into the set $\langle \mathit{Phon}^*\rangle^{h_{\max}+\widehat{\jmath}+1}$defined by

$$\langle x,y\rangle \mapsto \begin{cases} \langle x_0y_0, y_1, \ldots, y_{h_{\max}}\rangle & \text{if } \widehat{\jmath} = 0 \\ \langle\ x_0\ \ , y_1, \ldots, y_{h_{\max}}, y_0\rangle & \text{if } \widehat{\jmath} = 1 \end{cases}$$

In case $u = \mathtt{com}$, $\mathit{merge}_{U,V}$ is the function from $\mathit{Phon}^* \times \langle \mathit{Phon}^*\rangle^{g_{\max}+1}$ into the set $\langle \mathit{Phon}^*\rangle^{g_{\max}+\widehat{\jmath}+1}$ defined by

$$\langle y,x\rangle \mapsto \langle y_0x_0, x_1, \ldots, x_{g_{\max}}\rangle$$

Now, for some $1 \le j \le m$, suppose that

$\nu_j \in \mathrm{suf}(\mathtt{-l}_j)$ with $\nu_j = \mathtt{-l}_j\lambda$ for some $\lambda \in \mathit{Licensees}^*$, /cf. (ii) of (smo)/
$l\kappa \in \mathrm{suf}(\mathit{Syn})$ with $l = \mathtt{+L}_j$ and $\kappa \in \mathit{Syn}^*$, /cf. (i) of (smo)/
$\nu_i \in \mathrm{suf}(\mathtt{-l}_i)$ for $1 \le i \le m$ with $i \ne j$ /cf. ($\mathrm{R}_{\mathrm{SMG}}$)/

such that for $1 \le i \le m$ with $i \ne j$ it holds that

$\nu_i = \epsilon$ if $\lambda = \mathtt{-l}_i\lambda'$ with $\lambda' \in \mathit{Syn}^*$. /cf. (ir.2.1)/

Next, choose elements $g_i \in \{-1\} \cup \{0,\ldots,m\}$ for $1 \le i \le m$ and $\beta_i \in \{1,\ldots,m\}^*$ for $0 \le i \le m$ such that

$U = \langle\langle l\kappa, \beta_0, 0\rangle, \langle \nu_1, \beta_1, g_1\rangle, \ldots, \langle \nu_m, \beta_m, g_m\rangle, \mathtt{com}) \in N$,

and such that, additionally,

$g_J \ne 0$, /cf. (iii) of (smo)/

and, furthermore,

$g_i = g_J$ for each $1 \le i \le m$ with $J \lhd_U^+ i$ or $J \prec_U i$, /cf. (ir.2.3)/

where $J \in \{1,\ldots,m\}$ with $g_J = g_j$ such that

$J \triangleleft_U^* i$ or $J \prec_U i$ for each $1 \leq i \leq m$ with $g_i = g_j$.[26] /cf. (SD4)/

If $\lambda = \epsilon$ we set $k = 0$. If $\lambda \neq \epsilon$ we let $1 \leq k \leq m$ such that $\lambda = \texttt{-1}_k\lambda'$ for some $\lambda' \in$ *Syn**. Furthermore, we choose existing and unique $0 \leq j' \leq m$ such that $\beta_{j'} = \zeta_{j'} J \eta_{j'}$ for some $\zeta_{j'}, \eta_{j'} \in \{1, \ldots, m\}^*$. Hence $j' \triangleleft_U J$, implying that $j' \neq J$.

Then, in case $J = j$ and $k = 0$, we take

$T = \langle\langle\kappa, \beta_J \eta_{j'} \beta, 0\rangle, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, \texttt{com}\rangle \in N$.

In case $J = j$ and $k \neq 0$, we take

$T = \langle\langle\kappa, k \eta_{j'} \beta, 0\rangle, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, \texttt{com}\rangle \in N$.

In case $J \neq j$, we take

$T = \langle\langle\kappa, J \eta_{j'} \beta, 0\rangle, \widehat{\mu}_1, \ldots, \widehat{\mu}_m, \texttt{com}\rangle \in N$,

where in all three cases $\beta = \zeta_{j'}$ if $j' = 0$, and $\beta = \beta_0$ otherwise. Furthermore, for $1 \leq i \leq m$ we have

$$\widehat{\mu}_i = \begin{cases} \langle\lambda, \beta_j, 0\rangle & \text{if } i = k \\ \langle\epsilon, \epsilon, -1\rangle & \text{if } i = j \text{ and } j \neq k \\ \langle\nu_i, \zeta_i, \widetilde{g}_i\rangle & \text{if } i = j' \\ \langle\nu_i, \zeta_i \beta_j \eta_i, 0\rangle & \text{if } i \triangleleft_U j,\ i \neq j' \text{ and } k = 0,^{27} \\ & \text{where } \zeta_i, \eta_i \in \{1, \ldots, m\}^* \text{ with } \beta_i = \zeta_i j \eta_i \\ \langle\nu_i, \zeta_i k \eta_i, 0\rangle & \text{if } i \triangleleft_U j,\ i \neq j' \text{ and } k \neq 0, \\ & \text{where } \zeta_i, \eta_i \in \{1, \ldots, m\}^* \text{ with } \beta_i = \zeta_i j \eta_i \\ \langle\nu_i, \beta_i, 0\rangle & \text{if } J \triangleleft_U^* i \text{ or } J \prec_U i \text{ for } i \neq j \text{ and } i \not\triangleleft_U j\ ^{28} \\ \langle\nu_i, \beta_i, \widetilde{g}_i\rangle & \text{otherwise} \end{cases}$$

where for $1 \leq i \leq m$, $\widetilde{g}_i \in \{-1\} \cup \{0, \ldots, m\}$ is defined by

$$\widetilde{g}_i = \begin{cases} g_i - 1 & \text{if } g_i > g_j = g_J\,^{29} \\ g_i & \text{in all other cases} \end{cases}$$

Now, for the function *move*$_U \in F$ as defined below we let

(r2) $T \rightarrow$ *move*$_U\big(U\big) \in R$.

Once more, we let $g_{\max} = \max\{g_i \,|\, 0 \leq i \leq m\}$, we let x_H, x_0, $x_1, \ldots, x_m$ be pairwise distinct variables, and we let x be the $g_{\max} + 1$–tuple $\langle x_0, \ldots, x_{g_{\max}}\rangle$.

move$_U$ is the function from $\langle$*Phon*$^*\rangle^{g_{\max}+1}$ into $\langle$*Phon*$^*\rangle^{g_{\max}}$ defined by

[26] Because $\mu_j \neq \epsilon$ and $g_j = g_J \neq 0$, it turns out, by (n4), that $g_j > 0$.

[27] Let $0 \leq i \leq m$ with $i \triangleleft_U j$: if $j = J$ then $i = j'$; if $j \neq J$ then $J \triangleleft_U^* i$ or $J \prec_U i$.

[28] Recall that, if $J \triangleleft_U^* i$ or $J \prec_U i$ for some $1 \leq i \leq m$ then $g_i = g_j$ by assumption on U.

[29] Note that $g_i > g_j$ for some $1 \leq i \leq m$ implies that $i \triangleleft_U^+ J$ or $i \prec_U J$.

$$x \mapsto \langle x_{g_J} x_0, x_1, \ldots, x_{g_J - 1}, x_{g_J + 1}, \ldots, x_{g_{\max}} \rangle$$

Terminating rules: For each $\kappa\pi\iota \in$ *Lex* with $\kappa \in$ *Syn**, $\pi \in$ *Phon** and $\iota \in$ *Sem** let

(r3) $T \rightarrow \pi \in R$, where $T = \langle\langle \kappa, \epsilon, 0\rangle, \widehat{\nu}_1, \ldots, \widehat{\nu}_m, \mathtt{sim}\rangle \in N \setminus \{S\}$

with $\widehat{\nu}_i = \langle \epsilon, \epsilon, -1\rangle$ for $1 \leq i \leq m$.

Analogously to the "MG-case," the weak equivalence of G and G_{SMG} (Corollary 4.8) results from two propositions, namely, the following ones:

PROPOSITION 4.6
If $\tau \in$ *RCL*(G_{MG}) then a $T \in N \setminus \{S\}$ and a $p_T \in L_G(T)$ exist such that $\langle T, p_T\rangle$ corresponds to τ according to Definition 4.5.

PROPOSITION 4.7
If $p_T \in L_G(T)$ for some $T \in N \setminus \{S\}$ and $p_T \in O$, then there is a $\tau \in$ *RCL*(G_{MG}) that corresponds to $\langle T, p_T\rangle$ in the sense of Definition 4.5.

The analogy to the respective propositions concerning the given MG G_{MG} is obvious, and again, respective inductions on $k \in \mathbb{N}$ would serve to prove them formally. Instead of presenting corresponding formal proofs of Proposition 4.6 and 4.7 we will briefly discuss the underlying idea of the construction of G, the way in which this idea resembles the one underlying the previous construction as well as the aspects in which it deviates from the former.

The construction of the LCFRS G for the SMG G_{SMG} reflects the definitions of *merge* and *move*$^{\text{s}}$, and the relevance condition (R_{SMG}) as it results from the latter: assume that, on the one hand, there are $\upsilon, \phi \in$ *RCL*(G_{SMG}). On the other hand, suppose $U, V \in N$ and $p_U \in L_G(U)$, $p_V \in L_G(V)$ to be such that $\langle U, p_U\rangle$ and $\langle V, p_V\rangle$ correspond to υ and ϕ, respectively. Then it holds that $\tau =$ *merge*(υ, ϕ) is defined and relevant to G_{SMG} if and only if $U, V \in N$ meet the conditions for occurring on the righthand side of a rewriting rule of the form (r1). In particular, the restrictions imposed on such a U and V prevent the corresponding υ and ϕ from creating a situation as described in (ir.1.1), (ir.1.2) or (ir.1.4). For the corresponding $T \in N$ and $p_T =$ *merge*$_{U,V}(p_U, p_V) \in O$, the pair $\langle T, p_T\rangle$ corresponds to τ.

For $\upsilon \in$ *RCL*(G_{SMG}), $U \in N$ and $p_U \in L_G(U)$ such that $\langle U, p_U\rangle$ corresponds to υ similar considerations arise w.r.t. the move–operator: $\tau =$ *move*$^{\text{s}}(\upsilon)$ is defined and relevant to G_{SMG} if and only if U meets the conditions for occurring on the righthand side of a rule of the form (r2). Here, the respective restrictions guarantee in particular that the corresponding υ does not provide a case of (ir.2.1) or (ir.2.3). For the corresponding $T \in N$ and $p_T =$ *move*$_U(p_U) \in O$, the pair $\langle T, p_T\rangle$ corresponds to τ.

Note that, as far as the set functions in G are concerned, the construction of G contrasts with that in the MG–case in a particular detail which we want to emphasize here. For this reason suppose that, by an application of *merge* or *move*$^{\text{s}}$, some $\phi \in$ *Exp*(*Feat*) becomes a specifier of some $\upsilon \in$ *CL*(G_{SMG}). In virtual terms of the

LCFRS G, we then do not distinguish between the case that ϕ has an (unchecked) licensee and the case that it has not. This is different to the "MG–case" and due to the specific character of *move*$^{\mathrm{s}}$ as it deviates from *move*: suppose that ϕ has some licensee `-x`. In the SMG–case, we do not have to foresee the possibility within the LCFRS G that at some later derivation step ϕ or a remnant of ϕ will move "on its own" in order to check the corresponding instance of `-x`. If υ serves to derive another expression $\upsilon' \in CL(G_{\mathrm{SMG}})$ such that this instance of `-x` enters into a configuration which allows checking it off, then ϕ or its corresponding remnant will by no means belong to $Comp^{+}(\upsilon')$, since ϕ is a specifier of υ.

Note further that "precedence" in terms of $\prec_U$ as it is generally coded in a corresponding nonterminal U from $N \setminus \{S\}$ by means of (n2) is used only in order to define rules of the form (r2), and only in the SMG–case, i.e., to have "precedence" in terms of $\prec_U$ available is only relevant for simulating the structure building operator *move*$^{\mathrm{s}}$.

COROLLARY 4.8
$\pi \in L(G)$ iff $\pi \in L(G_{SMG})$ for each $\pi \in Phon^{*}$.

The proof of this corollary is very similar to that of Corollary 4.4. We leave to the reader the explicit adaptations that are necessary.

5 Final Remarks

We have shown that each MG as well as each SMG as defined in [18] is weakly equivalent to an LCFRS of a particular kind, namely, an LCFRS$_{1,2}$ in the sense of Definition 2.4, i.e. an LCFRS $G = \langle N, O, F, R, S\rangle$ with $rank(G) = 2$ such that, if $A \rightarrow f(B, C) \in R$ then $d_G(B) = 1$. This result is of special interest, since, conversely, it can be shown that each such LCFRS is weakly equivalent to both a corresponding MG as well as a corresponding SMG ([10]). Consequently, the MG–type and the SMG–type as defined in [18] are shown to determine the same class of derivable string languages, thereby confirming a conjecture explicitly stated in [18].

MGs provide a restricted type of UMGs.[30] The latter are known to have the same weak generative power as LCFRSs ([11, 13], [5]). It is known that each LCFRS is weakly equivalent to some LCFRS of rank 2 (see e.g. [15]). Nevertheless it seems to be an open problem, whether our result implies that MGs also provide a proper restriction of UMGs in terms of derivable string languages. Our conjecture, indeed, is that this is the case given the highly restricted nature of the interleavings allowed by an LCFRS$_{1,2}$.

The conjecture is likewise supported by a look at what may be seen as the crucial difference between MGs, respectively SMGs, and UMGs. Let us conclude by examine this difference in some more detail: in contrast to the general possibilities provided by the UMG–type, an MG neither employs any kind of *head movement* nor *covert phrasal*

[30] Recall our convention from the introductory part that we refer to an MG of the type as originally defined in [17] as an *unrestricted MG (UMG)*, while we use the term *minimalist grammar* and its short cut *MG* only in order to refer to an MG of the revised type introduced in [18].

movement, and an additional condition is imposed on the move–operator as to which maximal projection may move *overtly* into the highest specifier position, namely, condition (iii) of (mo), the definition of *move*. However, as far as it concerns the derivable string languages, we do not have to take care of the precise definitions of *head movement* and *covert phrasal movement* as given w.r.t. UMGs, since it has been shown ([11, 13], [5]) that, in terms of weak equivalence, UMGs can be defined as MGs in the sense of Definition 3.3 dropping condition (iii) of (mo). That is to say, within an UMG, the (overtly) moving maximal projection ϕ which has licensee `-x` may be located anywhere within the expression υ which has licensor `+x`. This equivalence result can be strengthened, since [13] even shows that in terms of weak generative capacity nothing gets lost, when additionally the label of each lexical item of an UMG is demanded to be from $\mathit{Select}_\epsilon(\mathit{Select} \cup \mathit{Licensors})_\epsilon \mathit{Base}\,\mathit{Licensees}^*\mathit{Phon}^*\mathit{Sem}^*$, analogously to the SMG–definition.[31] That is to say, if we restrict the UMG–type to the possibility of applying only overt phrasal movement and creating only single specifiers by means of corresponding labels of the lexical items, the specified class of derivable languages is not restricted at all.[32] The same is true of the MG–type. Of course this type is *a priori* restricted to overt phrasal movement by definition, but as shown in [10], for each MG G there is a weakly equivalent MG G' such that the label of each lexical item of G' is from $\mathit{Select}_\epsilon(\mathit{Select} \cup \mathit{Licensors})_\epsilon \mathit{Base}\,\mathit{Licensees}^*\mathit{Phon}^*\mathit{Sem}^*$. The proof in [10] is done by showing that each $\mathrm{LCFRS}_{1,2}$ can be converted into a weakly equivalent MG of this kind. Crucially, the resulting MG G' also fulfills property (*).

(*) Whenever, for some $\upsilon \in \mathit{CL}(G')$ and `-x` $\in$ *Licensees*, there is some maximal projection $\phi \in \mathit{MaxProj}(\upsilon)$ that has licensee `-x` then $\phi \in \mathit{Comp}^+(\upsilon)$.

This property in connection with the specific labeling of the lexical items guarantees that G' is likewise interpretable as an SMG without changing the closure of G', the set $\mathit{CL}(G')$. More precisely, an $\upsilon \in \mathit{CL}(G')$ fulfills condition (iii) of (mo) iff it fulfills condition (iii) of (smo), and w.r.t. (iii) of (smo), the corresponding $\phi \in \mathit{MaxProj}(\upsilon)$ that has licensee `-x` is identical to the existing $\chi \in \mathit{Comp}^+(\upsilon)$ with $\phi \in \mathit{Spec}^*(\chi)$; to put it differently, extraction is always strictly out of the "rightmost" path. But exactly for this reason, property (*) also seems to reveal the crucial difference to UMGs, since the analogous "best case scenario" for UMGs seems to be the provable fact that for each UMG there is a weakly equivalent UMG G' fulfilling (**).

(**) For each $\upsilon \in \mathit{CL}(G')$, there is some $\alpha \in \mathit{Spec}(\beta)$ for some $\beta \in \mathit{Comp}^+(\upsilon)$ such that, whenever for some `-x` $\in$ *Licensees*, there is some maximal projection $\phi \in \mathit{MaxProj}(\upsilon)$ that has licensee `-x` then $\phi \in \mathit{Comp}^+(\upsilon)$ or $\phi \in \mathit{Comp}^+(\alpha)$.

---

[31] In [13] this is shown by converting an arbitrary LCFRS of rank 2 into a weakly equivalent UMG of this particular kind. In [8] a corresponding proof is given by converting an arbitrary UMG which only allows overt phrasal movement to take place into such an UMG.

[32] This also proves that the chain based account to minimalist grammars in terms of connected forests as presented in [19] provides a "restricted," but weakly equivalent type of UMGs.

References

[1] Michael Brody. Mirror theory. Manuscript, University College London, London, 1997.

[2] Noam Chomsky. *The Minimalist Program*. MIT Press, Cambridge, MA, 1995.

[3] Philippe de Groote, Glyn Morrill, and Christian Retoré, editors. *Logical Aspects of Computational Linguistics (LACL '01)*, Lecture Notes in Artificial Intelligence Vol. 2099. Springer, Berlin, Heidelberg, 2001.

[4] Henk Harkema. A recognizer for minimalist grammars. In *Proceedings of the Sixth International Workshop on Parsing Technologies (IWPT 2000)*, Trento, pages 111–122, 2000.

[5] Henk Harkema. A characterization of minimalist languages. In de Groote et al. [3], pages 193–211.

[6] Henk Harkema. *Parsing Minimalist Languages*. PhD thesis, University of California, Los Angeles, CA, 2001.

[7] Gregory M. Kobele. Formalizing mirror theory. Manuscript, University of California, Los Angeles, CA, 2001.

[8] Gregory M. Kobele and Jason Kandybowicz. A normal form theorem for minimalist grammars. Manuscript, University of California, Los Angeles, CA, 2001.

[9] Hilda Koopman and Anna Szabolcsi. *Verbal Complexes*. MIT Press, Cambridge, MA, 2000.

[10] Jens Michaelis. Observations on strict derivational minimalism. *Electronic Notes in Theoretical Computer Science*, 53. Proceedings of the joint conference of the sixth conference on Formal Grammar and the seventh meeting of the Association for Mathematics of Language (FGMOL '01), Helsinki, 2001.

[11] Jens Michaelis. Derivational minimalism is mildly context–sensitive. In M. Moortgat, editor, *Logical Aspects of Computational Linguistics (LACL '98)*, Lecture Notes in Artificial Intelligence Vol. 2014, pages 179–198. Springer, Berlin, Heidelberg, 2001. Also available at `http://www.ling.uni-potsdam.de/~michael/papers.html`.

[12] Jens Michaelis. *On Formal Properties of Minimalist Grammars*. PhD thesis, Potsdam University, Potsdam, 2001.

[13] Jens Michaelis. Transforming linear context–free rewriting systems into minimalist grammars. In de Groote et al. [3], pages 228–244. Also available at `http://www.ling.uni-potsdam.de/~michael/papers.html`.

[14] Carl J. Pollard. *Generalized Phrase Structure Grammars, Head Grammars, and Natural Language*. PhD thesis, Stanford University, Stanford, CA, 1984.

[15] Owen Rambow and Giorgio Satta. Independent parallelism in finite copying parallel rewriting systems. *Theoretical Computer Science*, 223:87–120, 1999.

[16] Hiroyuki Seki, Takashi Matsumura, Mamoru Fujii, and Tadao Kasami. On multiple context–free grammars. *Theoretical Computer Science*, 88:191–229, 1991.

[17] Edward P. Stabler. Derivational minimalism. In C. Retoré, editor, *Logical Aspects of Computational Linguistics (LACL '96)*, Lecture Notes in Artificial Intelligence Vol. 1328, pages 68–95. Springer, Berlin, Heidelberg, 1997.

[18] Edward P. Stabler. Remnant movement and complexity. In G. Bouma, G.–J. M. Kruijff, E. Hinrichs, and R. T. Oehrle, editors, *Constraints and Resources in Natural Language Syntax and Semantics*, pages 299–326. CSLI Publications, Stanford, CA, 1999.

[19] Edward P. Stabler. Minimalist grammars and recognition. In C. Rohrer, A. Roßdeutscher, and H. Kamp, editors, *Linguistic Form and its Computation*, pages 327–352. CSLI Publications, Stanford, CA, 2001.

[20] Edward P. Stabler and Edward L. Keenan. Structural similarity. In *Algebraic Methods in Language Processing.* Proceedings of the 16th Twente Workshop on Language Technology (TWLT 16) joint with the 2nd AMAST Workshop on Language Processing, Iowa City, IA, pages 251–265, 2000.

[21] K. Vijay–Shanker, David J. Weir, and Aravind K. Joshi. Characterizing structural descriptions produced by various grammatical formalisms. In *25th Annual Meeting of the Association for Computational Linguistics (ACL '87)*, Stanford, CA, pages 104–111. ACL, 1987.

[22] David J. Weir. *Characterizing Mildly Context–Sensitive Grammar Formalisms*. PhD thesis, University of Pennsylvania, Philadelphia, PA, 1988.