# $\text{L}_\text{R}$FG for the formally-minded?[*]

Adam Przepiórkowski
ICS Polish Academy of Sciences / University of Warsaw

31st July 2025

## 1 Introduction

$\text{L}_\text{R}$FG is a variant of LFG (Bresnan 1982, Bresnan et al. 2016, Dalrymple et al. 2019, Dalrymple 2023) which does not assume the Lexical Integrity Principle (Bresnan & Mchombo 1995, Asudeh et al. 2013), and instead relies on the "syntax all the way down" approach of Distributed Morphology (DM; Halle & Marantz 1993). It has been in development for some 5–10 years now, depending on whether the starting point is taken to be Asudeh & Siddiqi 2016 or Melchin et al. 2020.

The aim of this unpublished manuscript is to demonstrate that some fundamental mechanisms of $\text{L}_\text{R}$FG are still far from clear and that various definitions are formally incoherent, some to the extent that makes it difficult to grasp the intentions behind the defined concepts. The effect is that learning and adopting $\text{L}_\text{R}$FG on the basis of $\text{L}_\text{R}$FG works published so far is not easy, and may be prohibitive for the more formally-minded linguists. In short, this paper presents a friendly critique of the current state of $\text{L}_\text{R}$FG by its would-be practitioner.

However, I aim at a constructive criticism, where possible, by proposing definitions and formalizations that seem to capture the intentions of $\text{L}_\text{R}$FG developers in a more explicit and precise way. Of course, given the frequent lack of explicitness and precision in the original $\text{L}_\text{R}$FG works and the "moving target" character of $\text{L}_\text{R}$FG, it is very well possible that some of the formalizations suggested below miss the mark.

Since $\text{L}_\text{R}$FG has been dynamically evolving, I concentrate on recent work, published in 2023 or later, especially the proceedings publications Asudeh et al. 2023, Asudeh & Siddiqi 2024, Asudeh et al. 2024, and the abstract for a 2025 workshop Belyaev et al. 2025. At this stage, the paper does not contain an introduction to $\text{L}_\text{R}$FG, i.e., it assumes some familiarity with its basic concepts.

## 2 Formalizations

I start with some formalizations found in recent $\text{L}_\text{R}$FG works, as it is particularly easy to show problems here; more general difficulties, including the lack of formalization of some concepts, are discussed in the following sections.

### 2.1 Host identification

In current $\text{L}_\text{R}$FG, Vocabulary Items (VIs) have the general form given in (1) (Asudeh et al. 2023: 23), exemplified in (2) (Asudeh et al. 2024: 53). $\text{C}_i$s stand for categories of c-structure terminals, F, G, and I stand for functional descriptions, glue constructors (ignored below), and information structure descriptions (absent in the $\text{L}_\text{R}$FG work I've consulted). The exponence function, mapping specifications on the left to v(ocabulary)-structures on the right, is marked as $\nu$.

(1) $\langle \underbrace{[\text{C}_1,\ldots,\text{C}_n]}_{\textit{distribution}}\ ,\ \underbrace{\text{F} \cup \text{G} \cup \text{I}}_{\textit{function/meaning}} \rangle \xrightarrow{\nu} \Big[\quad\Big]_{\textit{v-structure}}$

(2) $\langle\ [\#],\ @\textsc{pl}\ \rangle \xrightarrow{\nu} \begin{bmatrix} \text{PHONREP} & /\text{s}/ \\ \text{DEP} & \text{LT} \\ \text{HOST} & \begin{bmatrix}\text{IDENT} & + \end{bmatrix} \end{bmatrix}$

A given VI for an affix or a clitic may specify, via the HOST value, what kind of host it expects. Typically, such hosts are in some sense adjacent to the item exponed via a given VI; in such cases, the v-structure in the VI is marked as [HOST|IDENT +].

[*]Many thanks to Sebastian Zawada for comments on an earlier version of this unpublished manuscript. I remain solely responsible for the views expressed here and for any remaining problems.

Asudeh et al. 2023: 29–31 define the adjacent host informally in (3)–(4) and formally in (5)–(6) (where f-domain($n$) in (6) is the set of all c-structure nodes that map to the same f-structure as $n$). Unfortunately, while (3) is relatively intuitive, the notion of *closest* in (4) has an important gap, while the formal definitions in (5)–(6) are additionally incoherent.

(3) HOST *Identification (Intuition)*
Given $\beta$, a v-structure containing the feature [HOST [IDENT $+$]], and $\eta$, a c-structure terminal node that $\beta$ expones, $\beta$'s HOST is the v-structure that expones the *closest* c-structural terminal node to $\eta$ that maps to the *same* f-structure as $\eta$.

(4) Y is the closest c-structure node to X iff

- X c-commands Y; and
- there is no Z such that X c-commands Z and Z c-commands Y.

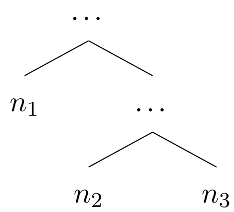
(5) For all c-structure nodes, $n$, $n'$, $n''$, in the set of c-structure terminal nodes $T$ for some c-structure,

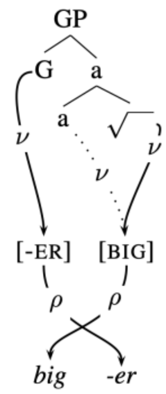
$$\textbf{closest}(n, n') \Leftrightarrow \textbf{c-command}(n, n') \wedge \neg[\textbf{c-command}(n, n'') \ \wedge \textbf{c-command}(n'', n')] \wedge n \neq n'$$

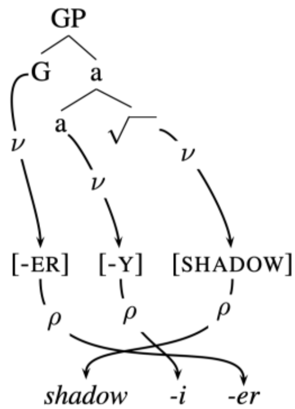
(6) *Local* HOST *Identification (LHI)*
For all c-structure nodes, $n$, $n'$, in the set of c-structure nodes $N$ for some c-structure,

$$(\nu(n') \text{ IDENT}) = + \Rightarrow \textbf{closest}(n, n') \ \wedge \ n' \in \textbf{f-domain}(n) \wedge (\nu(n) \text{ HOST}) = \nu(n')\backslash\text{HOST}$$

The gap in the intended definition of *the closest c-structure* in (4) (and its intended formalization in (5)) can be illustrated by the schematic tree in (7), which has exactly three terminals: $n_1$, $n_2$, and $n_3$.[1] This schematic tree is instantiated by the trees in (8)–(9) from Asudeh & Siddiqi 2024: 74 (shown with further mappings), in which $n_1 = $ G, $n_2 =$ the lower a, and $n_3 = \sqrt{\phantom{x}}$. In brief, the problem is that – according to the definitions in (4)–(5) – $n_1$ in (7) has *no* closest terminal node: it is not $n_2$, as $n_3$ "intervenes" ($n_3$ is c-commanded by $n_1$ and c-commands $n_2$), and it is not $n_3$, as $n_2$ "intervenes" (analogously).

(7)



(8)



(9)



In the case of the specific trees in (8)–(9), the terminal nodes are G, the lower a, and the root $\sqrt{\phantom{x}}$. In order to determine the realization of G as *-er* (as in *bigger* or *shadowier*), the terminal node closest to G should be determined. But, according to the definitions in (4)–(5), there is no such closest terminal node. If G is X in (4), the lower node a cannot be the closest node Y, as there is Z = $\sqrt{\phantom{x}}$, which is c-commanded by X = G and c-commands Y = a. Similarly, Y cannot be $\sqrt{\phantom{x}}$, as Z = lower a is c-commanded by X = G and c-commands Y = $\sqrt{\phantom{x}}$. Hence, the gap in the definitions.[2]

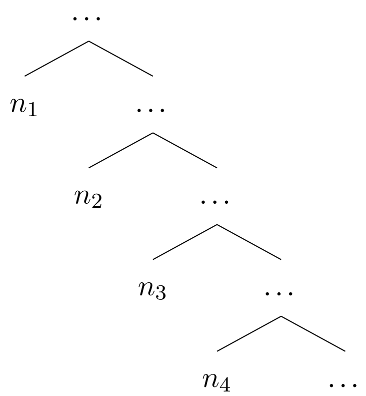
[1] The ellipsis ... signals a single node whose category does not matter here, not a larger tree portion.

[2] The informal definition in (4) is also a little sloppy as it talks about nodes rather than terminal nodes. On this definition, the node closest to X = G is the higher a, and there is no node that would be both closest and terminal.

From the discussion in Asudeh & Siddiqi 2024: 81–82, it is (relatively) clear that the node intended to be closest to G is the lower a, but not $\sqrt{\phantom{x}}$.[3] Hence, one attempt at repairing these definitions would be to include a reference to linear proximity in the tree rather than – or in addition to – c-command, but I will not attempt such a repair here.

The problem with the "formal" definitions (5)–(6) is shared by a number of $\mathrm{L_RFG}$ definitions (as the next section will make clear); it consists of a wrong quantificational structure of the defined notions. In the case of (5), the definition says that, given a c-structure and the set T of all its terminal nodes, a formula must be true for any triple of such terminal nodes $n$, $n'$, $n''$. But for many trees this formula is false, regardless of the intended definition of **closest**. Consider the tree in (10).

(10)



Does **closest**$(n_1, n_3)$ hold? That is, what is the truth value of the right-hand side (RHS) of the formula in (5) when $n = n_1$ and $n' = n_3$? Well, that depends on the choice of $n''$. If $n'' = n_2$, then this RHS is false (because all **c-command** predicates there are true, so the $\neg$[**c-command**$(n, n'')$ $\wedge$ **c-command**$(n'', n')$] conjunct is false). But if we take $n'' = n_4$, then this RHS is true (because **c-command**$(n, n')$ is still true, but now **c-command**$(n'', n')$ is false, so the negated conjunction is true).

Since, for a given left-hand side of the formula, the right-hand side is sometimes true, sometimes false, the whole "definition" in (5), with its 3 universally quantified variables, is simply false – regardless of the intended definition of **closest**.

The problem here is the careless "For all c-structure nodes, $n$, $n'$, $n''$..." in the definition. The right quantificational structure of this definition is:[4]

(11) Let $n$, $n'$, and $n''$ range over the set of terminal nodes in some c-structure. Then **closest** is that relation which makes the following formula true:
$\forall n, n'.$ **closest**$(n, n') \Leftrightarrow$ **c-command**$(n, n') \wedge n \neq n' \wedge$
$\neg\exists n''.$ **c-command**$(n, n'') \wedge$ **c-command**$(n'', n')$

In the notation of Asudeh & Siddiqi 2024, where relations are more explicitly defined using the symbol :=, this can be expressed as follows (also getting rid of the redundant $n \neq n'$ here):

(12) Let $n$, $n'$, and $n''$ range over the set of terminal nodes in some c-structure. Then:
**closest**$(n, n')$ := **c-command**$(n, n') \wedge \neg\exists n''.$ **c-command**$(n, n'') \wedge$ **c-command**$(n'', n')$

Wrong quantificational structure is also one of the problems with the LHI formalization in (6). Because of the initial "For all c-structure nodes, $n$, $n'$...", this formalization says that, for any host $n'$, it is closest to all nodes in the tree, it is in the functional domain of all nodes (i.e., all nodes in the tree map to the same f-structure), and it hosts all nodes (including itself!). This is incoherent. The right – if I understand the intention correctly – formalization is given below:

(13) Let $n$ and $n'$ range over the set of nodes in some c-structure. Then the following holds:
$\forall n.\ (v(n)$ HOST IDENT$) = + \Rightarrow \exists n'.$ **closest**$(n, n') \wedge n' \in$ **f-domain**$(n) \wedge$
$(v(n)$ HOST$) = v(n')\backslash$HOST

[3] For this reason, replacing "c-command" with "asymmetric c-command" in these definitions will not help, as then both the lower a and $\sqrt{\phantom{x}}$ would be closest to G.

[4] This definition inherits from (5) the redundant conjunct $n \neq n'$; it is redundant as it follows from the **c-command**$(n, n')$ conjunct on the commonly assumed definition of "c-command" (see, e.g., `https://en.wikipedia.org/wiki/C-command`).

Note the change in the antecedent of this principle from $(v(n')$ IDENT$)$ in (6) to $(v(n)$ HOST IDENT$)$ in (13). This seems to more directly correspond to the intuition in (3), which starts with a node whose v-structure specifies a host, while (6) starts with the host and looks for the node with v-structure specifying this host. A formula more directly corresponding to (6) is (14).

(14) Let $n$ and $n'$ range over the set of nodes in some c-structure. Then the following holds:
$$\forall n'.\ (v(n')\ \textsc{ident}) = + \Rightarrow \exists n.\ \textbf{closest}(n, n') \wedge n' \in \textbf{f-domain}(n) \wedge (v(n)\ \textsc{host}) = v(n')\backslash\textsc{host}$$

### 2.2 MostInformative

Given a number of VIs competing for the exponence of the same terminal node(s), the ones that are less informative than others lose the competition. Nodes can be more or less informative in various ways. One is the amount of f-structural information they specify. This is regulated by the relation **MostInformative**$_f$, defined in Asudeh & Siddiqi 2024: 78 as in (15), which is an improvement over definitions in a number of earlier publications, which had the structure given in (16).[5]

(15) **MostInformative**$_f$ currently (Asudeh & Siddiqi 2024)

> **MostInformative**$_f(\alpha, \beta)$ returns whichever of $\alpha$,$\beta$ has the most specific f-structure in the set of f-structures returned by $\Phi$ applied to $\alpha/\beta$'s collected f-description.[12]
>
> *Intuition.* Prefer portmanteau forms, whenever possible, on f-structural grounds. Choose the VI that defines an f-structure that contains the greater set of features.
>
> *Formalization.* The proper subsumption relation on f-structures (Bresnan et al. 2016: chap. 5) is used to capture the intuition.
>
> Given two VIs, $\alpha$ and $\beta$,
> $$\textbf{MostInformative}_f(\alpha, \beta) = \begin{cases} \alpha \textbf{ if } \exists f. f \in \Phi(\pi_2(\pi_1(\alpha))) \wedge \forall g. g \in \Phi(\pi_2(\pi_1(\beta))) \rightarrow g \sqsubset f \\ \beta \textbf{ if } \exists f. f \in \Phi(\pi_2(\pi_1(\beta))) \wedge \forall g. g \in \Phi(\pi_2(\pi_1(\alpha))) \rightarrow g \sqsubset f \\ \bot \textbf{ otherwise} \end{cases}$$

(16) **MostInformative**$_f$ previously (Asudeh & Siddiqi 2023)

> Given two VIs, $\alpha$ and $\beta$,
> $$\textbf{MostInformative}_f(\alpha, \beta) = \begin{cases} \alpha \textbf{ if } \exists f \forall g. f \in \pi_2(V^i(\alpha)) \wedge g \in \pi_2(V^i(\beta)) \wedge g \sqsubset f \\ \beta \textbf{ if } \exists f \forall g. f \in \pi_2(V^i(\beta)) \wedge g \in \pi_2(V^i(\alpha)) \wedge g \sqsubset f \\ \bot \textbf{ otherwise} \end{cases}$$

In order to decipher the formalization in (15), recall that VIs are mappings from $\langle$distribution, function/meaning$\rangle$ pairs to v-structures:

(1) $$\Big\langle \underbrace{[\mathrm{C}_1,\ldots,\mathrm{C}_n]}_{\textit{distribution}}\ ,\ \underbrace{\mathrm{F} \cup \mathrm{G} \cup \mathrm{I}}_{\textit{function/meaning}} \Big\rangle \xrightarrow{\nu} \Big[\quad\Big]_{\textit{v-structure}}$$

If $\alpha$ is a VI as in (1), then $\pi_1(\alpha)$ is its left-hand side, i.e., the $\langle$distribution, function/meaning$\rangle$ pair. Then, the second element of this pair, which contains functional descriptions, may be referred to as $\pi_2(\pi_1(\alpha))$. Concentrating on such functional descriptions, i.e., on F alone in F $\cup$ G $\cup$ I in (1), $\pi_2(\pi_1(\alpha))$ then refers to functional descriptions in the VI $\alpha$. This functional description may give rise to a number of minimal satisfying f-structures, and the set of these f-structures is denoted as $\Phi(\pi_2(\pi_1(\alpha)))$. So, in effect, the formalization in (15) is saying that the VI $\alpha$ is more informative$_f$ than the VI $\beta$ if one of the f-structures defined in $\alpha$ is properly subsumed by – i.e., is more informative than – all f-structured defined in $\beta$. (The previous version, in (16), had the universal quantifiers in wrong places.)

This definition is almost trivial and transparent compared to the definition of **MostInformative**$_c$ – informativeness with respect to the span of exponed c-structure terminals – given in Asudeh & Siddiqi 2024: 79–80:

[5] The formalization in (15) was proposed (by me) in a discussion with Ash Asudeh and Sebastian Zawada in December 2024.

(17) **MostInformative**$_c(\alpha, \beta)$ takes two sets of vocabulary items $\alpha$,$\beta$ and returns whichever set is smaller.

*Intuition.* Prefer portmanteau forms, whenever possible, on c-structural grounds. Choose the set of VIs that realizes the greater span of c-structure nodes.

*Formalization.* We define functions to aid the presentation, where $c$ is a c-structure, $f$ is an f-structure, and $v$ is a vocabulary item.

> Given a c-structure $c$ and two sets of vocabulary items, $\alpha$ and $\beta$,
> **MostInformative**$_c(\alpha, \beta) =$
> $\alpha = \{x \mid x \text{ is a VI } \wedge \textbf{features}(x) \subseteq \textbf{targets}(c) \wedge \forall y \exists z.[y \in \textbf{categories}(x) \wedge z \in \textbf{labels}(c) \wedge \pi_2(z) = y]\}$
> $\beta = \{x \mid x \text{ is a VI } \wedge \textbf{features}(x) \subseteq \textbf{targets}(c) \wedge \forall y \exists z.[y \in \textbf{categories}(x) \wedge z \in \textbf{labels}(c) \wedge \pi_2(z) = y]\}$
> $\begin{cases} \alpha \textbf{ if } |\alpha| < |\beta| \\ \beta \textbf{ if } |\beta| < |\alpha| \\ \bot \textbf{ otherwise} \end{cases}$

- **features**$(v) := \Phi(\pi_2(\pi_1(v)))$
  *the set of f-structures that VI $v$ defines per the f-description in its left-hand side*[14]
- **categories**$(v) := \pi_1(\pi_1(v))$
  *the category list of VI $v$*
- **targets**$(c) :=$
  $\{f \mid \phi(c) = f \wedge \pi_1(\textbf{labels}(c)) \subseteq \textbf{extendedProj}(f)\}$
  *the set of f-structures that c-structure c defines, such that the nodes in the first-coordinate of the* **labels** *of c are a subset of the* **extendedProj** *of f*
  - **labels**$(c) := \{\langle x, y\rangle \mid x \in \textbf{yield}(c) \wedge y = \lambda(x)\}$
    *a set of pairs where the first member is a node in c-structure c and the second member is the node's label/category*
    - **yield**$(c) := \{n \mid n \text{ is a terminal node in } c\}$
      *the set of terminal nodes in c*
  - **extendedProj**$(f) := \phi^{-1}(f)$
    *the set of c-structure nodes that map to f-structure $f$; the extended projection of $f$ in c-structure*

This is an example of an L$_R$FG definition which, even after considerable time spent trying to untangle it, remains opaque (to me).

The first problem is the general structure of the "formalization" part: it seems to have the structure in (18), which takes two sets and returns the one with the smallest cardinality.

(18) $$\textbf{MostInformative}_c(\alpha, \beta) = \begin{cases} \alpha & \textbf{if} \quad |\alpha| < |\beta| \\ \beta & \textbf{if} \quad |\beta| < |\alpha| \\ \bot & \textbf{otherwise} \end{cases}$$

The definition in (18) is trivial. But then what is the role of the two equations, "$\alpha = \ldots$" and "$\beta = \ldots$" in the middle of this "formalization"? Is this a definition of the domain of **MostInformative**$_c$, i.e., of sets that are comparable by this relation? It is particularly baffling that these two equations have identical right-hand sides, which implies that $\alpha = \beta$; but this clearly is not intended, as then the value of **MostInformative**$_c(\alpha, \beta)$ would always be $\bot$, given that $\alpha = \beta \rightarrow |\alpha| = |\beta|$.

Moreover, the subformula in these equations starting with $\forall y \exists z$ seems to have the same problem as the earlier definition of **MostInformative**$_f$ in (16): as it contains the universal quantifier without implication, it effectively says that all objects $y$ in the considered universe are members of **categories**($x$). But this does not make sense, as the formula also talks about VIs and f-structures, so the universe clearly contains elements which are not categories.

Also the definition of **targets**($c$) is baffling in various respects. What is $\phi(c)$, when $c$ is a c-structure – is this the f-structure corresponding to the root of $c$, or perhaps the set of all f-structures corresponding to all nodes in $c$? In the former case, **targets**($c$) would always be a singleton set, which does not seem to be the intention here. In the latter case, **targets**($c$) would be a singleton set containing the set of such f-structures, which again does not seem what was intended.[6]

Finally, why is $\pi_1$(**labels**($c$)) used in the definition instead of the simpler and apparently co-extensive **yield**($c$)? Is $\pi_1$(**labels**($c$)) meant to be different from **yield**($c$), despite the definitions of these relations?

After spending some time trying to understand the intention of this definition, it is still not clear to me what it is supposed to be saying beyond (18). Such formal sloppiness makes it difficult for formally-minded linguists – the natural addressees of such definitions – to get engaged with L$_{\text{R}}$FG.

I revisit the issue of proper definitions of **MostInformative**$_f$ and **MostInformative**$_c$ in §4, after trying to make explicit the correspondence between c-structure terminals and VIs.

## 3 Node Exponence Principle?

Recent works do not make it at all clear how VIs relate to c-structure, in particular, which VIs may compete for the exponence of which bits of c-structure.

In some of the previous versions of L$_{\text{R}}$FG, at least that in Asudeh & Siddiqi 2023: 886–887, VIs were supposed to define terminal nodes in c-structures, which, incidentally, seems to run counter to the Subset Principle of DM that L$_{\text{R}}$FG claims to follow. This is no longer the case in current versions of L$_{\text{R}}$FG, but functional descriptions in VIs still make use of ↑ and ↓, even though this does not make sense anymore, as these metavariables only make sense when attached to particular c-structure nodes.[7] For example, one of the VIs proposed for Latin in Asudeh et al. 2024 is given in (19) (repeated from (2)), with the macro PL defined in (20).

(19) $\langle$ [#], @PL $\rangle \xrightarrow{\nu} \begin{bmatrix} \text{PHONREP} & \text{/s/} \\ \text{DEP} & \text{LT} \\ \text{HOST} & \begin{bmatrix}\text{IDENT} & + \end{bmatrix} \end{bmatrix}$

(20) PL := ($\uparrow$ PLURAL) = +

This is a minor glitch and I assume that it may be repaired by replacing ↑ ("the f-structure of the immediately dominating c-structure node") with • ("this f-structure"), as in definitions of v-structure in Asudeh et al. 2023: §6.1.

The bigger question is how exactly such descriptions relate to c-structure terminals to which they correspond – in the case of (19), to terminals bearing the category #. The basic intuition, shared with the original DM, is clear: the f-description in a given VI must contain a subset of information provided by the f-structure corresponding to the c-structure node exponed by this VI. That is, in the case of (19), the f-structure of the exponed # node should contain at least the +-valued feature PLURAL, perhaps apart from other features. It seems clear, then, that a vocabulary item $vi$ containing the set of functional descriptions $F$ may only compete for the exponence of a node $n$ such that $\Phi(F) \sqsubseteq \phi(n)$, where $\phi(n)$ is, as usual, the f-structure corresponding to node $n$, $\Phi(F)$ is the minimal f-structure satisfying description $F$, and $\sqsubseteq$ is the (non-defining version of) subsumption relation.[8] Let me try to formalize this as a node exponence principle:[9]

[6]It is incompatible with the statement **features**($x$) $\subseteq$ **targets**($c$), given that **features**($x$) is a set of f-structures rather than a set of sets of f-structures.

[7]It might be said that the use of these metavariables in VIs extends their normal use in the following obvious sense: ↑ in a VI refers to the terminal node that is exponed by this VI. However, one VI may expone a larger number of nodes, in which case it is not clear which of them should be the value of ↑ in this VI.

[8]$\Phi$ is called a *bridging function* in the L$_{\text{R}}$FG literature (e.g., Asudeh & Siddiqi 2023: 885).

[9]In this principle, $vs$ stands for a v-structure description, and $\lambda$ maps c-structure nodes to their categories. Note that these principles only define necessary conditions on exponence, as additional principles also play a role here, especially, the **MostInformative** principles discussed above and defined more precisely in §4 below.

(21) *Node Exponence Principle (NEP)* (version 1 of 7)

A vocabulary item $vi = \langle [C], F\rangle \xrightarrow{v} vs$ may expone a terminal node $n$ only if $\lambda(n) = C$ and $\Phi(F) \sqsubseteq \phi(n)$.

It may seem that this is so intuitive that it does not need to be stated explicitly – and, to the best of my knowledge, it has not been in any $\mathrm{L_R}$FG publications – but it is only the tip of the iceberg of a complete principle of this kind, with things quickly becoming less intuitive.

The immediate complication is that $F$ may correspond to a non-singleton set of minimal f-structures, while (21) assumes that it defines just one such a structure. A permissive extension would require that at least one of the minimal f-structures defined by $F$ subsumes $n$'s f-structure:

(22) *Node Exponence Principle (NEP)* (version 2 of 7)

A vocabulary item $vi = \langle [C], F\rangle \xrightarrow{v} vs$ may expone a terminal node $n$ only if $\lambda(n) = C$ and $\exists f.\, f \in \Phi(F) \wedge f \sqsubseteq \phi(n)$.

(I will deal with the dual problem, of a longer list of categories, at the end of this section.)

Asudeh et al. 2024 introduce another complication, namely, the possibility of enriching functional descriptions in VIs with constraining statements, as in (23), where a nominative node of category K may be exponed only if (the f-structure corresponding to) this node is neuter (i.e., no gender) and singular (i.e., no number).[10]

(23)
$$\left\langle \begin{array}{l} [\mathrm{K}],\ @\textsc{nom} \\ \langle\!\langle \neg(\uparrow \textsc{gender}) \rangle\!\rangle \\ \langle\!\langle \neg(\uparrow \textsc{plural}) \rangle\!\rangle \end{array} \right\rangle \xrightarrow{\nu} \begin{bmatrix} \textsc{phonrep} & \mu \\ \textsc{dep} & \mathrm{LT} \\ \textsc{class} & \mathrm{X}{=}4 \\ \textsc{host} & \begin{bmatrix} \textsc{ident} & + \\ \textsc{class} & \mathrm{X} \end{bmatrix} \end{bmatrix}$$

The version of NEP in (22), with the standard subsumption relation $\sqsubseteq$, is not sufficient here: when a constraining statement such as $\neg(f\ \textsc{gender})$ holds of f-structure $f$, and $f$ subsumes $g$ ($f \sqsubseteq g$), there is no guarantee that $g$ does not contain the GENDER feature.[11]

There are various ways of making NEP handle such constraining equations. One is to replace the standard subsumption with the implicit subsumption employed in LFG and XLE (Crouch et al. 2011) when distributing properties over sets, e.g., in typical LFG analyses of coordination. As discussed in XLE documentation (see the URL in fn. 11), this implicit version of subsumption – call it $\sqsubseteq_c$ (with $c$ standing for *constraining* and *coordination*) – honours constraining equations.

(24) *Node Exponence Principle (NEP)* (version 3 of 7)

A vocabulary item $vi = \langle [C], F\rangle \xrightarrow{v} vs$ may expone a terminal node $n$ only if $\lambda(n) = C$ and $\exists f.\, f \in \Phi(F) \wedge f \sqsubseteq_c \phi(n)$.

Another possibility is to retain the standard subsumption relation $\sqsubseteq$, but split functional descriptions $F$ into defining, $F_d$, and constraining, $F_c$, and add a condition saying that $\phi(n)$ must satisfy all constraining equations in $F_c$, i.e., $\phi(n) \models F_c$:

(25) *Node Exponence Principle (NEP)* (version 4 of 7)

A vocabulary item $vi = \langle [C], F_d, F_c\rangle \xrightarrow{v} vs$ may expone a terminal node $n$ only if

1. $\lambda(n) = C$,
2. $\phi(n) \models F_c$, and
3. $\exists f.\, f \in \Phi(F_d) \wedge f \sqsubseteq \phi(n)$.

[10] Double angle brackets seem to play absolutely no formal role here: "Note that we have used an arbitrary double-angle notation $\langle\!\langle\ \rangle\!\rangle$ to highlight constraining equations (including existential and negative existentials)" (Asudeh et al. 2024: 52). See, however, the discussion of Belyaev et al. 2025 below.

[11] See `https://ling.sprachwiss.uni-konstanz.de/pages/xle/doc/notations.html#N4.2.7`.

But the notion of an f-structure satisfying some descriptions also applies to defining descriptions – from the perspective of satisfaction, the difference between defining and constraining descriptions vanishes. Hence, $F$ does not have to be split into defining and constraining descriptions, and we can instead require that the f-structure $\phi(n)$ satisfies the whole set $F$:

(26) *Node Exponence Principle (NEP)* (version 5 of 7)
A vocabulary item $vi = \langle [C], F \rangle \xrightarrow{v} vs$ may expone a terminal node $n$ only if $\lambda(n) = C$ and $\phi(n) \models F$.

This removes the necessity to refer to particular minimal f-structures defined by $F$; by the nature of the satisfaction relation, if some f-structure $f$ satisfies such an ultimately disjunctive set of functional descriptions, it must satisfy (at least) one of the disjuncts.

Note that in NEP versions 3 (in (24)) and 5 (in (26)), the distinction between defining and constraining statements in $F$ disappears: they are all treated on par (as if they were constraining statements, given that $F$ in VIs cannot contribute to defining f-structures corresponding to c-structure nodes). However, it seems that the difference between the two kinds of descriptions are important for Asudeh et al. (2024) and Belyaev et al. (2025): the defining equations more directly expone a given c-structure node, while the constraining equations play the role of contextual constraints in DM. For example, in (23), the case node K is exponed via a VI in which the particular case – nominative – is described by defining equations encapsulated in the NOM macro (defined in §5 below), and other functional information conditioning the applicability of this VI – about gender and number – is given as constraining statements, additionally marked with double-angle brackets ⟪ ⟫.

This is especially clear in the two contrasting VIs in (27)–(28) from Belyaev et al. 2025, where a VI exponing K contains a call to a macro defining the nominative, @NOM, but an analogous macro defining the inessive in a VI exponing a non-K node, N, is called as ⟪@IN⟫ – despite the declaration in Asudeh et al. 2024 that this notation is only used to highlight constraining equations (see fn. 10).

(27) $\left\langle [\text{K}],\ \begin{array}{c} @\text{NOM} \\ \langle\!\langle(\uparrow \text{DEIXIS})\rangle\!\rangle \\ \langle\!\langle(\uparrow \text{PRED FN}) =_c \text{pro}\rangle\!\rangle \end{array} \right\rangle \xrightarrow{\nu} \textit{-j}$

(28) $\left\langle [\text{N}],\ \begin{array}{c} \langle\!\langle @\text{IN}\rangle\!\rangle \\ \langle\!\langle(\uparrow \text{DEIXIS})\rangle\!\rangle \\ \langle\!\langle(\uparrow \text{PRED FN}) =_c \text{pro}\rangle\!\rangle \end{array} \right\rangle \xrightarrow{\nu} \textit{-m}$

Because of this incompatibility between the use of double-angle brackets in Asudeh et al. 2024 and in Belyaev et al. 2025, and because of the lack of any clear information about the intended formal properties of statements enclosed in such brackets, I assume that the effect of (23) and (27)–(28) would be exactly the same if double angle brackets were removed and the appropriate descriptions treated on par with other functional descriptions in these VIs. For this reason, I do not attempt to formalize this aspect of recent $\text{L}_\text{R}$FG work any further here, and assume that the NEP in (26) suffices here.

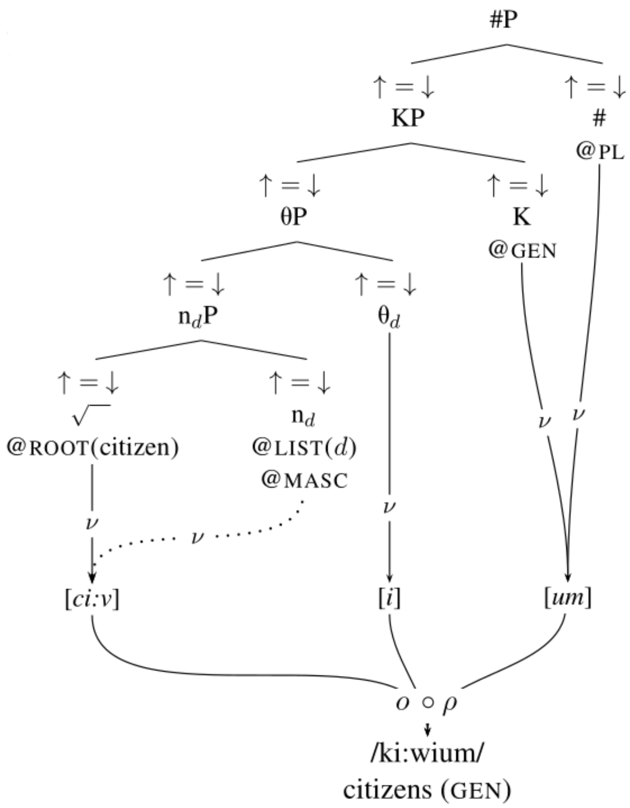
The final complication in the Node Exponence Principle that I would like to discuss in this section is that a single VI may expone a number of adjacent c-structure terminals, by specifying a category list longer than 1, as in the following example from Asudeh et al. 2024: 62, where the VI in (29) expones two nodes in the tree, K and #, as shown in (30), where [*um*] corresponds to terminals bearing categories K and #.[12]

[12]In the f-structure in (30), [GENDER +] is the result of @MASC (i.e., masculine is defined as the presence of GENDER), and [NOMINATIVE +, ACCUSATIVE +, GENITIVE +] is the encoding of the genitive, @GEN, as discussed in §5.

(29) $\langle$ [K,#], @GEN $\rangle \xrightarrow{\nu}$ @PL

$$\begin{bmatrix} \text{PHONREP} & \text{/rum/} \\ \text{DEP} & \text{LT} \\ \text{CLASS} & \text{X=1} \vee \text{X=2} \vee \text{X=5} \\ \text{HOST} & \begin{bmatrix} \text{IDENT} & + \\ \text{CLASS} & \text{X} \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \text{PHONREP} & \text{/um/} \\ \text{DEP} & \text{LT} \\ \text{CLASS} & \text{X=3} \vee \text{X=4} \\ \text{HOST} & \begin{bmatrix} \text{IDENT} & + \\ \text{CLASS} & \text{X} \end{bmatrix} \end{bmatrix}$$

(30)




$$\begin{bmatrix} \text{PRED} & \text{`citizen'} \\ \text{GENDER} & + \\ \text{NOMINATIVE} & + \\ \text{ACCUSATIVE} & + \\ \text{GENITIVE} & + \\ \text{PLURAL} & + \end{bmatrix}$$

This calls for the following extension of NEP:

(31) *Node Exponence Principle (NEP)* (version 6 of 7)

A vocabulary item $vi = \langle [C_1, \ldots, C_k], F \rangle \xrightarrow{v} vs$ may expone terminal nodes $n_1, \ldots, n_k$ only if

1. $n_1, \ldots, n_k$ is a linearly contiguous sequence of nodes in the yield of a c-structure,
2. $\lambda(n_1) = C_1, \ldots, \lambda(n_k) = C_k$,
3. $\phi(n_1) = \ldots = \phi(n_k)$, and
4. $\phi(n_1) \models F$.

I make the simplifying assumption here that c-structure terminals exponed via a single VI must map to the same f-structure; see 3. in (31). This makes it possible to retain the simple condition on satisfaction in 4.: $\phi(n_1) \models F$.

This assumption seems to be met in all examples of spanning in $\text{L}_\text{R}$FG literature that I am aware of, but it is not clear that it holds in general. For example, it does not seem to hold in the case of German portmanteau words such as *am* = *an* (preposition) + *dem* (determiner); on standard LFG analyses, when the preposition is semantic, indicating a location, it has a different f-structure than the following nominal projection, including the determiner, so the two c-structure nodes exponed as *am* have different f-structures at least on some – locative – uses. I will not try to extend NEP to such cases here, as it is not clear to me how L$_\text{R}$FG intends to analyse portmanteaus such as German *am*.

Another potential complication is Pac-Man spanning, witnessed in (30), where the VI spanning the $\sqrt{\phantom{x}}$ for 'citizen' also spans the following n$_d$, which does not have its own VI. It is not clear to me whether such freeloaders should be handled by NEP, or by a separate principle regulating Pac-Man spanning: constraints on Pac-Man spanning remain vague in the current L$_\text{R}$FG literature. (I return to Pac-Man spanning at the end of the following section.)

## 4 MostInformative revisited

The intention of the incoherent definition of **MostInformative**$_c$ in (17) was to compare sets of VIs competing for the exponence of the same sequence of terminals – the smallest of these sets is to be preferred. Let me now try to make formally precise what it means for a set of VIs to compete for the exponence of a sequence of terminals:

(32) *Node Exponence Principle (NEP)* (version 7 of 7)

A sequence of $m \geq 1$ vocabulary items
$vi^1 = \langle [C^1_1, \ldots, C^1_{k^1}], F^1 \rangle \xrightarrow{v} vs^1$,
$\ldots$,
$vi^m = \langle [C^m_1, \ldots, C^m_{k^m}], F^m \rangle \xrightarrow{v} vs^m$
may expone a sequence of $k \geq 1$ terminal nodes $n_1, \ldots, n_k$ only if

1. $n_1, \ldots, n_k$ is a linearly contiguous sequence of nodes in the yield of a c-structure,
2. $\lambda(n_1) = C^1_1, \ldots, \lambda(n_k) = C^m_{k^m}$,
3. $\phi(n_1) = \ldots = \phi(n_k)$, and
4. $\phi(n_1) \models F^1 \cup \ldots \cup F^m$.

Conditions 1. and 3. are exactly the same as in (31): they require the sequence of nodes to be contiguous terminals with the same f-structures. Condition 2. in (32) generalizes previous formulations, in requiring that the sequence of categories of nodes $n_1, \ldots, n_k$ is equal to the concatenation of category lists of vocabulary items $vi^1, \ldots, vi^m$; this in particular implies that $k = k^1 + \ldots + k^m$. For example, assuming that category lists are non-empty, 3 terminal nodes with categories $C_1$, $C_2$, and $C_3$ may be exponed by a single VI whose category list is [$C_1$, $C_2$, $C_3$], or by two VIs, one with category list [$C_1$, $C_2$] and the other with [$C_3$] (or one with [$C_1$] and the other with [$C_2$, $C_3$]), or by three VIs, with category lists [$C_1$], [$C_2$], and [$C_3$]. Finally, condition 4. generalizes previous versions of this condition by requiring that the f-structure corresponding to the terminal nodes satisfies the sum of functional descriptions in the VIs.

Given (32), it is easy to properly define both **MostInformative**$_c$ and **MostInformative**$_f$:

(33) Given two sets of VIs, $\alpha$ and $\beta$, competing for the exponence of the same sequence of terminal nodes, prefer **MostInformative**$_c(\alpha, \beta)$, where

$$\textbf{MostInformative}_c(\alpha,\beta) = \begin{cases} \alpha & \textbf{if} \quad |\alpha| < |\beta| \\ \beta & \textbf{if} \quad |\beta| < |\alpha| \\ \bot & \textbf{otherwise} \end{cases}$$

(34) Given two sets of VIs, $\alpha$ and $\beta$, competing for the exponence of the same sequence of terminal nodes, prefer **MostInformative**$_f(\alpha, \beta)$, where

$$\textbf{MostInformative}_f(\alpha,\beta) = \begin{cases} \alpha & \textbf{if} \quad \exists f.\, f \in \mathcal{F}(\alpha) \wedge \forall g.\, g \in \mathcal{F}(\beta) \rightarrow g \sqsubset f \\ \beta & \textbf{if} \quad \exists f.\, f \in \mathcal{F}(\beta) \wedge \forall g.\, g \in \mathcal{F}(\alpha) \rightarrow g \sqsubset f \\ \bot & \textbf{otherwise} \end{cases}$$

with $\mathcal{F}(\gamma)$ referring to the set of minimal f-structures defined jointly by f-descriptions present in the set of VIs $\gamma$:

$$\mathcal{F}(\gamma) = \Phi(\bigcup_{vi \in \gamma} \pi_2(\pi_1(vi)))$$

Again, these definitions should be revised when the exact nature of – and conditions on – Pac-Man spanning are clear. The urgency to define Pac-Man spanning is particularly conspicuous in the case of VIs with the set of f-descriptions empty, as in (35) from Asudeh et al. 2024: 61.

(35) $\langle$ [$\theta_f$], $\varnothing$ $\rangle \xrightarrow{\nu}$ $\begin{bmatrix} \text{PHONREP} & /\text{u}/ \\ \text{DEP} & \text{LT} \\ \text{CLASS} & \text{X}{=}4 \\ \text{HOST} & \begin{bmatrix} \text{IDENT} & + \\ \text{CLASS} & \text{X} \end{bmatrix} \end{bmatrix}$

Without a reasonable constraint on Pac-Man spanning, such VIs risk never being used.

To see why, let us imagine that the NEP in (32) is relaxed so that the concatenation of category lists of the VIs may be a sublist of the categories of the exponed terminals, with the intention that the terminals not corresponding to specific categories in these VIs are exponed Pac-Man-style. The effect of such a relaxation would be that some VIs would never be exponed, as Pac-Man spanning would be preferred. For example, take two contiguous terminals in some c-structure, $n_1$ and $n_2$, such that $n_2$ has the category $\theta_f$, as in the list of categories in (35). Let us call the VI in (35) $vi_2$ and assume that some $vi_1$ competes for the exponence of $n_1$. Then, without an appropriate constraint on Pac-Man spanning, the singleton set $\{vi_1\}$, which would now be allowed to compete for the exponence of the sequence $n_1, n_2$, will win the competition with $\{v_1, v_2\}$, as it is more informative in the sense of **MostInformative**$_c$ ($|\{vi_1\}| < |\{vi_1, vi_2\}|$) and not less informative in the sense of **MostInformative**$_f$ (both sets have the same f-descriptions, given that $vi_2$ has none, so **MostInformative**$_f(\{vi_1\}, \{vi_1, vi_2\}) = \bot$).

## 5 Morphosyntactic features in f-structures

In earlier versions of $\text{L}_\text{R}$FG, at least in Asudeh & Siddiqi 2023: 886–887 (other early papers are usually vague here), information specified in VIs could enter the f-structure. If so, current $\text{L}_\text{R}$FG seems to be more in the spirit of DM, with VIs specifying how certain nodes are to be realized (exponed), rather than adding information to such nodes. That is, such VIs implement the Subset Principle of DM, as opposed to earlier $\text{L}_\text{R}$FG, which seemed more in the spirit of the Superset Principle of Nanosyntax (e.g., Caha 2009) (or perhaps in the spirit of analyses combining the two approaches, e.g., Zompì 2023).

However, on this more faithfully DM approach, it is not yet clear how VIs interact with grammar proper – an issue that Asudeh et al. 2024 attempts to start to deal with. Here I only discuss one problem, namely, ensuring that f-structures created via c-structure rules contain all the information needed for exponence.

Recall (19) from Asudeh et al. 2024, repeated below for convenience, together with the definition of the PL macro in (20).

(19) $\langle$ [#], @PL $\rangle \xrightarrow{\nu}$ $\begin{bmatrix} \text{PHONREP} & /\text{s}/ \\ \text{DEP} & \text{LT} \\ \text{HOST} & \begin{bmatrix} \text{IDENT} & + \end{bmatrix} \end{bmatrix}$

(20) PL $:=$ ($\uparrow$ PLURAL) $= +$

Given the realizational character of current $\text{L}_\text{R}$FG, (19) may compete for exponing a node only if this node is specified as [PLURAL $+$] in the grammar.

This creates a complication, as c-structure rules usually do not specify the grammatical number of a nominal constituent. For example, there is no need to specify in the grammar whether the nominal object of a preposition is singular or plural; in standard LFG, this information comes from the lexical entry of the nominal head. But wait, there are no lexical entries like this in $\text{L}_\text{R}$FG anymore, they are replaced by VIs. Does this mean that VIs should pass information to c- and f-structures after all, as apparently in Asudeh & Siddiqi 2023?

The solution proposed in Asudeh et al. 2024 is to define what they call "bang macros", to be included in c-structure rules. The intuition is that such a bang macro is a disjunction of all possible values of a morphosyntactic feature, for example, for Latin case:

(36) CASE! $:=$ $\{$@NOM$|$@VOC$|$@ACC$|$@GEN$|$@DAT$|$@ABL$\}$

Here, @NOM, @ACC, etc., call macros defining the f-structure representations of particular cases (to be defined below). Definitions of NUM! and GEND! are similar, but rely on the singular number being defined as the lack of plural and on the neuter gender being defined as the lack of masculine or feminine, so the disjunctions are shorter than might be expected (actually no disjunction, in the case of number):

(37) NUM! $:=$ @PL

(38) GEND! $:=$ $\{$@MASC$|$@FEM$\}$

Given such macros, adding an obligatory call @CASE! to a c-structure rule defines all possible case values, adding an optional call @GEND! defines all three gender values (with lack of gender interpreted as neuter), etc. This is illustrated with the following rules from Asudeh et al. 2024: 57–58:[13,14]

(39) $\text{KP} \quad \rightarrow \quad \begin{array}{c} \uparrow = \downarrow \\ \theta\text{P} \end{array} \quad \begin{array}{c} \uparrow = \downarrow \\ \text{K} \\ @\textsc{case}! \end{array}$

(40) $\text{n}_x\text{P} \quad \xrightarrow{9} \quad \begin{array}{c} \uparrow = \downarrow \\ \sqrt{\ } \\ @\textsc{root}(\_\_) \end{array} \quad \begin{array}{c} \uparrow = \downarrow \\ \text{n}_{x \in \{a,b,c,d,e,f,g,v,w\}} \\ @\textsc{list}(x) \\ (\ @\textsc{gend}!\ ) \end{array}$

(41) $\#\text{P} \quad \rightarrow \quad \begin{array}{c} \uparrow = \downarrow \\ \text{KP} \end{array} \quad \left( \begin{array}{c} \uparrow = \downarrow \\ \# \\ @\textsc{num}! \end{array} \right)$

By virtue of such expanded c-structure rules, each plural nominal constituent will have a head of category # mapping to an f-structure containing the [PLURAL +] feature (as well as CASE and other features). Hence, by the *Node Exponence Principle*, a VI containing the f-description (↑ PLURAL) = + (as in (19)) may enter competition for exponing this # node. So far, so good. However, this approach is problematic when such morphosyntactic features are also defined elsewhere in the grammar, as is normally the case with CASE.

In order to deal with case syncretisms, $\text{L}_\text{R}$FG defines case hierarchies via a cascade of macros, as in (36), which encodes Caha's (2009) case hierarchy for Slavic.

(42) a. NOM := (↑ NOM) = +
b. ACC := @NOM
(↑ ACC) = +
c. GEN := @ACC
(↑ GEN) = +
d. LOC := @GEN
(↑ LOC) = +
e. DAT := @LOC
(↑ DAT) = +
f. INS := @DAT
(↑ INS) = +

The effect is that each case contains features corresponding to all cases lower in the hierarchy, for example:

(43) nominative $\equiv \begin{bmatrix} \textsc{nom} & + \end{bmatrix}$

(44) genitive $\equiv \begin{bmatrix} \textsc{nom} & + \\ \textsc{acc} & + \\ \textsc{gen} & + \end{bmatrix}$

Now, the problem is that particular cases may be assigned in the grammar regardless of calls to bang macros in rules such as (39). For example, a grammatical rule may assign the nominative to the subject, as in the simplistic (45), or a given root may require that its argument be in the dative.[15]

(45) $\text{IP} \quad \rightarrow \quad \begin{array}{c} \text{KP} \\ \downarrow = (\uparrow \textsc{subj}) \\ @\textsc{nom} \end{array} \quad \begin{array}{c} \text{I}' \\ \uparrow = \downarrow \end{array}$

[13] A confusing convention introduced in Asudeh et al. 2024 is that "annotations that are about the relationship between c-structure and f-structure, i.e. the $\phi$-mapping, are written above the category, whereas annotations for exponence, i.e. c-structure exponenda, are written below the category." It is confusing because it is not clear from this vague description whether annotations above the category have different formal properties than the ones below the category. Below I assume that they do not, i.e., that all these annotations behave as in standard LFG.

[14] (40) is a metarule, defining 9 different rules (for different values of $x$). In (41) the macro call @NUM! is not optional, as the whole constituent is optional, and the assumption is that this constituent is present iff the nominal is plural.

[15] It is not fully clear how such grammatical properties of roots may be expressed in current $\text{L}_\text{R}$FG, but perhaps the macros LIST and ROOT of Asudeh et al. 2024 could be extended to this purpose.

Unfortunately, the rule in (45) is incompatible with the rule in (39). This is because, a KP that is assigned the nominative in (45) may be further assigned an arbitrary case in (39), given that nominative is contained in every case according to (42).

An immediate solution is to retain bang macros in rules such as (39), and only use constraining descriptions in other places in the grammar. For example, a series of constraining macros could be defined as in (46) and used in the grammar as in (47).

(46) a. $\text{NOM?} := \neg(\downarrow \text{ACC})$

b. $\text{ACC?} := (\downarrow \text{ACC}) \wedge \neg(\downarrow \text{GEN})$

c. $\text{GEN?} := (\downarrow \text{GEN}) \wedge \neg(\downarrow \text{LOC})$

d. $\text{LOC?} := (\downarrow \text{LOC}) \wedge \neg(\downarrow \text{DAT})$

e. $\text{DAT?} := (\downarrow \text{DAT}) \wedge \neg(\downarrow \text{INS})$

f. $\text{INS?} := (\downarrow \text{INS})$

(47) $\text{IP} \rightarrow \underset{\substack{\downarrow = (\uparrow \text{SUBJ}) \\ @\text{NOM?}}}{\text{KP}} \quad \underset{\uparrow=\downarrow}{\text{I}'}$

This, however, greatly complicates some relatively natural analyses of case-related phenomena. For example, in Przepiórkowski 2025 I take advantage of the $\text{L}_\text{R}$FG encoding of case in (42) to analyse the infamous mixed agreement / case assignment (MACA) pattern of Slavic numerals, illustrated with Russian in (48).

(48)

| | ‘five chairs’ | |
|---|---|---|
| position | ‘five’ | ‘chairs’ |
| NOM | *pjat’*.NOM | *stul’ev*.GEN |
| ACC | *pjat’*.ACC | *stul’ev*.GEN |
| GEN | *pjati*.GEN | *stul’ev*.GEN |
| LOC | *pjati*.LOC | *stuljax*.LOC |
| DAT | *pjati*.DAT | *stuljam*.DAT |
| INS | *pjat’ju*.INS | *stuljami*.INS |

The gist of the analysis is that, in such numeral–noun constructions, the noun is always assigned the genitive case, i.e., it is always at least [NOM +, ACC +, GEN +], but its case value is also subsumed by that of the numeral. If the numeral is, say, accusative, i.e., [NOM +, ACC +], this subsumption has no effect, as the noun is already specified for these two attributes. But when the numeral is, say, dative, the defining effect of the subsumption is that the noun gets the additional [LOC +, DAT +] features. This analysis would not be possible if all actual – non-bang – grammatical case assignments were via constraining statements, as seems to be required by the use of the bang macros in rules such as (39) above.

In summary, since most of the $\text{L}_\text{R}$FG literature so far has concentrated on exponence and word-level phenomena, interaction with the actual grammar – syntactic rules above words and syntactic properties of lexical items – is neglected, which does not facilitate the adoption of $\text{L}_\text{R}$FG by a working syntactician.

## 6 In place of conclusion

These notes stem from my failed attempt to adopt $\text{L}_\text{R}$FG for my work.[16] When trying to understand the framework, I struggled with the “moving target” character of $\text{L}_\text{R}$FG: different sets of assumptions in different papers, without any clear information about what has changed and why. Another major problem was the formal sloppiness, with a surprisingly large number of incoherent definitions. A related problem was the lack of definitions of some of the most fundamental mechanisms of $\text{L}_\text{R}$FG: while some effort has been devoted to attempts at defining notions such as “most informative” and “most specific”, which help decide which of the VIs entering a competition for exponence are to be preferred, there is – at least in the literature I am aware of – no attempt at answering the more fundamental question of which VIs may enter such a competition at all. This might seem relatively intuitive in the case of some VIs with one category and a simple functional description, but it quickly becomes unclear when more complex sets of functional descriptions are considered (including disjunctive and constraining descriptions), when larger sets of VIs compete against each other for the exponence of a sequence of terminal nodes, and when Pac-Man spanning is also allowed.

[16] The original title of this paper was “Frustrations of a would-be $\text{L}_\text{R}$FG practitioner”.

I hope that this unpublished manuscript, which attempts to clarify some of these issues and formally define some of these concepts, will be of some use both for $L_R$FG developers and for formally-minded linguists struggling to make sense of $L_R$FG.

## References


Asudeh, Ash, Bronwyn Bjorkman, Frances Dowle, Neil Myler, Daniel Siddiqi & Lisa Sullivan. 2024. Fusional morphology, metasyncretism, and secondary exponence: A morphemic, realizational approach to Latin declension. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG '24 conference*, 45–68. Konstanz: Publikon. https://lfg-proceedings.org/lfg/index.php/main/article/view/49/55.

Asudeh, Ash, Tina Bögel & Daniel Siddiqi. 2023. Modelling exponents. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG '23 conference*, 23–44. Konstanz: Publikon. https://lfg-proceedings.org/lfg/index.php/main/article/view/38/26.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with Lexical Integrity. *Journal of Language Modelling* 1(1). 1–54.

Asudeh, Ash & Daniel Siddiqi. 2016. Realizational-Lexical Morphology for LFG. Presented at the Analyzing Morphological Systems Conference, Lyon, France, 20 June, 2016.

Asudeh, Ash & Daniel Siddiqi. 2023. Morphology in LFG. In Dalrymple (2023) 855–901. https://langsci-press.org/catalog/book/312.

Asudeh, Ash & Daniel Siddiqi. 2024. The story of *er*. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG '24 conference*, 69–89. Konstanz: Publikon. https://lfg-proceedings.org/lfg/index.php/main/article/view/47/53.

Belyaev, Oleg, Danil Alekseev, Ash Asudeh, Bronwyn Bjorkman, Frances Dowle, Nadeem Siddiqi & Lisa Sullivan. 2025. Ossetic nominal inflection: Between morphology and syntax. Abstract for the "Arguments and Challenges for Lexicalism" workshop at LFG '25. https://drive.google.com/file/d/1L41Xgcr7fdBGUjcScZaqnmehiS72CM6o/view.

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 3–86. Cambridge, MA: MIT Press.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax* Blackwell Textbooks in Linguistics. Wiley-Blackwell 2nd edn.

Bresnan, Joan & Sam Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13. 181–254.

Caha, Pavel. 2009. *The nanosyntax of case*: University of Tromsø Ph.D. dissertation.

Crouch, Dick, Mary Dalrymple, Ron Kaplan, Tracy King, John Maxwell & Paula Newman. 2011. XLE documentation. https://ling.sprachwiss.uni-konstanz.de/pages/xle/doc/xle_toc.html.

Dalrymple, Mary (ed.). 2023. *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press. https://langsci-press.org/catalog/book/312.

Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press. doi:https://doi.org/10.1093/oso/9780198733300.001.0001.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, vol. 24 Current Studies in Linguistics, 111–176. Cambridge, MA: MIT Press.

Melchin, Paul, Ash Asudeh & Daniel Siddiqi. 2020. Ojibwe agreement in Lexical-Realizational Functional Grammar. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*, 268–288. Stanford, CA: CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2020/lfg2020-mas.pdf.

Przepiórkowski, Adam. 2025. Mixed agreement vs. case assignment patterns. Unpublished manuscript, ICS PAS / University of Warsaw.

Zompì, Stanislao. 2023. **ABA in multidimensional paradigms: A Max/Dep-based account*. Cambridge, MA: Massachusetts Institute of Technology Ph.D. dissertation.