Morphological agreement in Minimalist Grammars

To appear in
Formal Grammar: Lecture Notes on Computer Science, Springer

Marina Ermolaeva

University of Chicago, Chicago, USA
`mermolaeva@uchicago.edu`

Abstract. Minimalist Grammars provide a useful tool for modeling natural language syntax by defining grammar fragments in a very precise way. As a formalization of the Minimalist Program, they can accommodate linguistic analyses from the field of generative syntax. However, they have no machinery for encoding agreement; while morphology can be simulated by multiplying lexical items, there is no systematic way to state generalizations and implement actual proposals. This paper extends Minimalist Grammars with morphological features and operations on them. As a proof of concept, I show how Icelandic dative intervention can be encoded in the modified formalism.

Keywords: Minimalist Grammars, Minimalist Syntax, agreement, morphosyntax, Icelandic

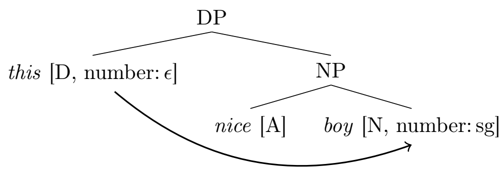
Introduction

Agreement can be defined as the morphological manifestation of dependencies between words. In a basic English sentence like *He walks* the verb agrees in person and number with the subject, and the subject, in turn, receives nominative case from the verb.[1] These dependencies may be nonlocal; for instance, English expletive constructions like *There seems to have arrived a man* exhibit long-distance subject-verb agreement.

Chomsky's Minimalist Program [4][5] treats these phenomena as an effect of a much more general mechanism known as Agree. An explicit theory of feature structures compatible with Chomsky's framework is proposed by Adger in [1]. Lexical items are defined as sets of features, each specified as bearing a value (drawn from some finite set) or being unvalued. Syntax is driven by features: the *probe* of a syntactic operation is an element with an unvalued feature, and the *goal* must bear a matching valued feature. Adger defines feature *valuation* as unification of values (cf. [15]): the unvalued feature on the probe assumes the value of the goal. The three operations are Merge, Move, and Agree. Merge and Move operate on *categorial features* (T, V, N ...) and build new structure. Agree targets *morphosyntactic features* (case, number, person, ...) and forms dependencies between elements of the existing structure.

[1] For the sake of exposition, I assume that case assignment reduces to agreement and that structural case is explicitly assigned by finite verbs. Neither is free of controversy.

Example 1. The phrase *this nice boy* exhibits determiner-noun agreement. The determiner bears an unvalued number feature and dominates the noun, which has a valued feature. This probe-goal configuration allows Agree to apply:



Stabler's Minimalist Grammars (MGs, [16][17]) have been designed as a mathematically rigorous formalization of Minimalist Syntax. The MG formalism is based on operations analogous to Merge and Move. Agree, however, has no counterpart. My goal is to extend MGs in a way that retains the relation to the Minimalist Program, allowing to translate Minimalist proposals involving agreement into the modified formalism.

1 Minimalist Grammars

I begin with the version of Minimalist Grammars defined in [17], with a few tweaks. This formalism employs *chain notation*, reducing syntactically active subtrees of derived trees to tuples of strings.

Definition 2. A minimalist grammar G is a 5-tuple $\langle \Sigma, Syn, Types, Lex, \mathcal{F} \rangle$, where

Σ is a finite set (of pronounced segments),

$Syn = Base$ (nonempty finite set of *categories*)
$\cup \{\texttt{=f} \mid \texttt{f} \in Base\} \cup \{\texttt{=>f} \mid \texttt{f} \in Base\}$ (*selectors*)
$\cup \{\texttt{+f} \mid \texttt{f} \in Base\}$ (*licensors*)
$\cup \{\texttt{-f} \mid \texttt{f} \in Base\} \cup \{\texttt{*f} \mid \texttt{f} \in Base\}$ (*licensees*)
is a set of syntactic features,

$Types = \{::, :\}$, (lexical, derived)

Let the set of *initial chains*[2] $IC = \Sigma^* \times \Sigma^* \times \Sigma^*\ Types\ Syn^*$, and the set of *non-initial chains* $NC = \Sigma^*\ Syn^*$;

$Lex \subset \{\epsilon\} \times \Sigma^* \times \{\epsilon\}\ \{::\}\ Syn^*$, a subset of IC, is a finite set of lexical items (*lexicon*),

$\mathcal{F} = \{merge, move\}$ is a set of structure-building operations:

- *merge* is the union of the following five functions, for $s_s, s_h, s_c, t_s, t_h, t_c \in \Sigma^*$, $\cdot \in \{:, ::\}$, $\texttt{f} \in Base$, $\gamma \in Syn^*$, $\delta \in Syn^+$, $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_l \in NC$ $(0 \le k, l)$,

[2] Angle brackets are used to denote tuples. For any n-tuple or sequence, for $1 \le i \le n$, $T[i]$ denotes the ith component of T. The (finite) product of sets $A_1, A_2, ..., A_n$ $A_1 \times A_2 \times ... \times A_n = \{\langle a_1, a_2, ..., a_n \rangle \mid a_1 \in A_1, a_2 \in A_2, ..., a_n \in A_n\}$. Similarly, their concatenation $A_1 A_2 ... A_n = \{a_1 a_2 ... a_n \mid a_1 \in A_1, a_2 \in A_2, ..., a_n \in A_n\}$.

$$mrg1\colon \quad \frac{\langle \epsilon, s_h, \epsilon\rangle \ :: \ \texttt{=f}\gamma \qquad \langle t_s, t_h, t_c\rangle \ \cdot \ \texttt{f}, \ \beta_1, ..., \beta_l}{\langle \epsilon, s_h, t_s t_h t_c\rangle \ : \ \gamma, \ \beta_1, ..., \beta_l}$$

$$mrg2\colon \quad \frac{\langle s_s, s_h, s_c\rangle \ : \ \texttt{=f}\gamma, \ \alpha_1, ..., \alpha_k \qquad \langle t_s, t_h, t_c\rangle \ \cdot \ \texttt{f}, \ \beta_1, ..., \beta_l}{\langle t_s t_h t_c s_s, s_h, s_c\rangle \ : \ \gamma, \ \alpha_1, ..., \alpha_k, \ \beta_1, ..., \beta_l}$$

$$mrg3\colon \quad \frac{\langle s_s, s_h, s_c\rangle \ \cdot \ \texttt{=f}\gamma, \ \alpha_1, ..., \alpha_k \qquad \langle t_s, t_h, t_c\rangle \ \cdot \ \texttt{f}\delta, \ \beta_1, ..., \beta_l}{\langle s_s, s_h, s_c\rangle \ : \ \gamma, \ \alpha_1, ..., \alpha_k, \ t_s t_h t_c \ : \ \delta, \ \beta_1, ..., \beta_l}$$

$$hmrg1\colon \quad \frac{\langle \epsilon, s_h, \epsilon\rangle \ :: \ \texttt{=>f}\gamma \qquad \langle t_s, t_h, t_c\rangle \ \cdot \ \texttt{f}, \ \beta_1, ..., \beta_l}{\langle \epsilon, t_h s_h, t_s t_c\rangle \ : \ \gamma, \ \beta_1, ..., \beta_l}$$

$$hmrg3\colon \quad \frac{\langle s_s, s_h, s_c\rangle \ \cdot \ \texttt{=>f}\gamma, \ \alpha_1, ..., \alpha_k \qquad \langle t_s, t_h, t_c\rangle \ \cdot \ \texttt{f}\delta, \ \beta_1, ..., \beta_l}{\langle s_s, t_h s_h, s_c\rangle \ : \ \gamma, \ \alpha_1, ..., \alpha_k, \ t_s t_c \ : \ \delta, \ \beta_1, ..., \beta_l}$$

- *move* is the union of the following three functions, for $s_s, s_h, s_c, t \in \Sigma^*$, $\texttt{f} \in Base$, $\texttt{F} \in \{\texttt{-f}, *\texttt{f}\}$, $\gamma, \zeta \in Syn^*$, $\delta \in Syn^+$, and for $\alpha_1, ..., \alpha_k \in NC$ $(0 \leq k)$ satisfying the condition (SMC)[3] there is exactly one $i \in [1, k]$ such that α_i has `-f` or $*$`f` as its first feature,

$$mv1\colon \quad \frac{\langle s_s, s_h, s_c\rangle \ \cdot \ \texttt{+f}\gamma, \ \alpha_1, ..., \alpha_{i-1}, \ t\ \texttt{F}, \ \alpha_{i+1}, ..., \alpha_k}{\langle t s_s, s_h, s_c\rangle \ : \ \gamma, \ \alpha_1, ..., \alpha_{i-1}, \ \alpha_{i+1}, ..., \alpha_k}$$

$$mv2\colon \quad \frac{\langle s_s, s_h, s_c\rangle \ \cdot \ \texttt{+f}\gamma, \ \alpha_1, ..., \alpha_{i-1}, \ t\ \texttt{F}\delta, \ \alpha_{i+1}, ..., \alpha_k}{\langle s_s, s_h, s_c\rangle \ : \ \gamma, \ \alpha_1, ..., \alpha_{i-1}, \ t \ : \ \delta, \ \alpha_{i+1}, ..., \alpha_k}$$

$$mv*\colon \quad \frac{\langle s_s, s_h, s_c\rangle \ \cdot \ \texttt{+f}\gamma, \ \alpha_1, ..., \alpha_{i-1}, \ t\ *\texttt{f}\zeta, \ \alpha_{i+1}, ..., \alpha_k}{\langle s_s, s_h, s_c\rangle \ : \ \gamma, \ \alpha_1, ..., \alpha_{i-1}, \ t \ : \ *\texttt{f}\zeta, \ \alpha_{i+1}, ..., \alpha_k}$$

Definition 3. An *expression* is a member of $Exp = IC\ NC^*$. An expression e is a *complete expression* of category $\texttt{c} \in Base$ iff $e = \langle s_s, s_h, s_t\rangle \ \cdot \ \texttt{c}$, where $\cdot \in \{::, :\}$.

Starred licensees of the form $*$`f` are optionally deleted (by *mv1* or *mv2*) or saved for later (by *mv*$*$). The latter possibility corresponds to intermediate positions of movement. [16] mentions this option of implementing successive cyclic movement; and a version of MGs with starred categorial features is explored in [11]. A formalism with persistent features, optionally erased by syntactic operations, has been shown to be weakly equivalent to standard MGs [18].

MGs offer a limited means of encoding (agglutinative) morphology by assigning separate lexical items to morphemes and constructing morphological words with head movement. Dependencies between words can be enforced by building restrictions into syntactic features.

[3] The SMC (Shortest Move Constraint): is a special case of the requirement that at any given step in the derivation the derived structure contain only finitely many subtrees (chains) which are syntactically active (i.e. have unchecked features).

Example 4. $G = \langle \Sigma_G, Syn_G, Types, Lex_G, \mathcal{F} \rangle$ is an MG. Its lexicon Lex_G contains the following lexical items:

$$
\begin{array}{lll}
\texttt{this.sg.nom} & := \langle \epsilon,\ \textit{this},\ \epsilon \rangle & :: \ \texttt{=n}_{\texttt{3SgN}}\ \texttt{d}\ \texttt{-k}_{\texttt{3SgN}} \\
\texttt{boy.sg.nom} & := \langle \epsilon,\ \textit{boy},\ \epsilon \rangle & :: \ \texttt{n}_{\texttt{3SgN}} \\
\texttt{walk} & := \langle \epsilon,\ \textit{walk},\ \epsilon \rangle & :: \ \texttt{=d}\ \texttt{v} \\
\texttt{prs.3sg} & := \langle \epsilon,\ \textit{-s},\ \epsilon \rangle & :: \ \texttt{=>v}\ \texttt{+k}_{\texttt{3SgN}}\ \texttt{t}
\end{array}
$$

G generates one expression of category `t`, derived as follows:

- *mv1* $\langle \textit{this boy},\ \textit{walk -s},\ \epsilon \rangle : \texttt{t}$
  - *hmrg1* $\langle \epsilon,\ \textit{walk -s},\ \epsilon \rangle : \texttt{+k}_{\texttt{3SgN}}\,\texttt{t},\ \textit{this boy}\,\texttt{-k}_{\texttt{3SgN}}$
    - `prs.3sg` $\langle \epsilon,\ \textit{-s},\ \epsilon \rangle :: \texttt{=>v}\,\texttt{+k}_{\texttt{3SgN}}\,\texttt{t}$
    - *mrg3* $\langle \epsilon,\ \textit{walk},\ \epsilon \rangle : \texttt{v},\ \textit{this boy}\,\texttt{-k}_{\texttt{3SgN}}$
      - `walk` $\langle \epsilon,\ \textit{walk},\ \epsilon \rangle :: \texttt{=d}\,\texttt{v}$
      - *mrg1* $\langle \epsilon,\ \textit{this},\ \textit{boy} \rangle : \texttt{d}\,\texttt{-k}_{\texttt{3SgN}}$
        - `this.sg.nom` $\langle \epsilon,\ \textit{this},\ \epsilon \rangle :: \texttt{=n}_{\texttt{3SgN}}\,\texttt{d}\,\texttt{-k}_{\texttt{3SgN}}$
        - `boy.sg.nom` $\langle \epsilon,\ \textit{boy},\ \epsilon \rangle :: \texttt{n}_{\texttt{3SgN}}$

This toy grammar forces *merge* and *move* to only combine lexical items with compatible morphological features – at the expense of having a separate feature for each combination of morphological properties that may result in a distinct morphological form. All agreement in *this boy walk -s* is local, which makes it easy to refine syntactic features manually. For long-distance agreement dependencies, a better option is to state compatibility restrictions as constraints defined in monadic second-order logic, as shown in [8].

The generative capacity of MGs is sufficient to encode any mildly context-sensitive pattern, so this strategy is adequate for ensuring correct agreement. However, it does not provide a succinct, systematic way of formulating generalizations about morphological dependencies; the relation to the Minimalist Agree operation remains obscure. Furthermore, the mappings from derivation to pronounced form in MGs have been given without special attention to the complications imposed by a detailed model of morphology. In the next section I propose a refinement which allows for straightforward integration with standard models of morphology.

2 Towards agreement

2.1 Bird's-eye view

Bundles and channels MGs treat all features as uninterpretable – in the sense that they all (with the exception of one category feature) must be deleted to form a complete expression. Morphological agreement essentially requires a class of features which are valued in the course of derivation and serve as building blocks of syntactic output. The first step is to redefine lexical items, replacing each sequence of phonological segments with a *bundle* – a set of morphological features. Incidentally, this modification separates syntax from phonology: pronounced segments are no longer present in lexical items and are assumed to be inserted outside syntax.

What about feature valuation? One option is an almost faithful translation of Minimalist Agree [1] into the MG formalism. Agree can be straightforwardly implemented as (covert) movement, allowing lexical items with matching morphological features to exchange information. However, no finite boundary can be imposed on the number of chains with unchecked features in the structure: consider, for instance, a sequence of adjectives modifying a noun, all of which have case requirements yet to be satisfied. This "naive" approach is incompatible with any version of the SMC.

An alternative, explored here, is to use existing syntactic dependencies created by *merge* and *move* to transmit morphological information. Agree is dependent on structure-building operations, which means that agreement is necessarily local. A long-distance morphological dependency between elements X and Y can be represented as a series of local information exchanges across *merge* dependencies involving, step by step, all elements intervening between X and Y.

Expanding on the idea outlined in [12], the flow of morphological information can be controlled by annotating syntactic features with their agreement properties, which can be conveniently thought of in terms of *channels*. For each syntactic feature, one needs to specify whether it accepts information from whatever checks it (*receiving channel*) and which values it transmits to whatever checks it (*emitting channel*).[4] Whenever two syntactic features establish a syntactic dependency, and one of them has a receiving channel, the chain/expression bearing this feature is updated with values specified for the emitting channel of the other feature. Borrowing terminology from linguistic literature, I call this process *downward agreement* if the feature at the receiving end is a selector or licensor, and *upward agreement* if it is a categorial feature or licensee.

Example 5. Recall the lexical item `boy.sg` from Example 4. The phonological exponent *boy* is replaced with a bundle, with ϵ being the *default value*:

$$\texttt{boy.sg} := \left\langle \epsilon,\ \begin{bmatrix} \text{BOY} \\ \texttt{num}{:}\textit{sg} \\ \texttt{per}{:}\textit{3} \\ \texttt{case}{:}\epsilon \end{bmatrix},\ \epsilon \right\rangle :: \texttt{n}_{\leftarrow}^{\begin{bmatrix} \texttt{num}:\textit{sg} \\ \texttt{per}:\textit{3} \end{bmatrix}\rightarrow}$$

The category feature `n` has a receiving channel (denoted by $_{\leftarrow}$), which allows `boy.sg` to receive a case value via upward agreement. The emitting channel on `n` (indicated by $^{\rightarrow}$) transmits number and person to whatever selects `boy.sg`.

[4] A lexical item may transmit different values of the same morphological feature via different channels. Moreover, these values need not be a subset of values in the item's own bundle. Keeping the content of emitting channels unconstrained is useful: for example, a preposition is allowed to transmit lexical case to its complement without morphologically manifesting it itself.

Probes and goals The channel-based agreement system can be refined to bring it more in line with the traditional notion of Agree. One such restriction is mentioned in [5] as a locality condition on goal defined in terms of "closest c-command" and in [1] as the requirement that the features in a probe-goal relation have no other matching feature intervening between them. The channel system has a built-in locality condition: each lexical item interacts directly with the head of the expression it selects/licenses and is not allowed to probe further. For items with multiple selectors/licensors, it is sufficient to require heads to accept (and transmit) agreement information via their *last* receiving channel. In other words, later values overwrite those received earlier. The intuition is simple: if X is selected by Y, the argument of X which is merged last will be the closest goal for Y.

Another useful restriction is known as the *freezing effect* of feature checking [2]. In essence, it rules out agreement in intermediate positions of successive cyclic movement. This condition is only relevant for starred licensees and can be built into the definition of *mv∗* in a straightforward way.

Feature sharing Long-distance upward agreement (across *merge* dependencies) cannot be reconciled with the requirement that the goal always provide a valued feature to the probe. When this is not the case, the agreement relation between lexical items in an expression has to be recorded so that, when the needed value enters the derivation, both items could be updated simultaneously. Nothing prevents such a relation from spanning multiple chains. The MG formalism distinguishes between the initial chain and non-initial chains but does not record any hierarchical relations. Additional bookkeeping is required to keep track of this information.

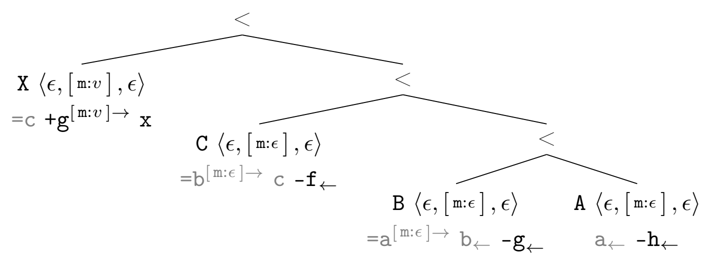
The proposed solution is reminiscent of *feature sharing* [7]. Their version of Agree does not require the goal feature to bear a value: matching features become a shared feature which is valued if either of the coalescing features is valued. I adopt a similar approach by recording for each feature, alongside its value, its *rewritability* – the highest chain that can transmit a value to it. With the SMC in place, every non-initial chain is uniquely identified by the name of its first licensee. Thus, rewritability can be set to `on` (*active*, or sharing value with the initial chain), a licensee name (sharing value with a non-initial chain), or `off` (*inactive*, or not accessible to agreement). I assume that morphological features start out as `off` if valued in the lexicon.

For any chain in an expression, its *subchains* are chains representing its subtrees. Each non-initial chain can be annotated with the sequence of all non-initial chains it is a subchain of (including itself), from the outermost to the most embedded. I will refer to this sequence as *lineage* of the chain. By convention, all lineages end with an `off` value. The set of all lineages for a given grammar is the set of all sequences of elements of *Base* without repetitions, followed by `off`. Lineages are updated throughout the derivation. Whenever a new non-initial chain appears in the expression, the name of its first licensee is prepended to the lineage of the new chain as well as all its subchains. On the other hand, when a chain moves, it ceases to be a subchain of any non-initial chain.[5] Therefore, all chains undergoing movement are stripped of the initial part of their lineage up to and including the relevant licensee.

Long-distance upward agreement succeeds if there is an uninterrupted sequence of channels between the probe and the goal. All that is required is to record, for each morphological feature, where this sequence ends (rewritability) and what path it takes (chain lineage). Thus, the two modifications introduced above are sufficient to keep track of channels connecting chains in an expression.

Example 6. Consider the expression `exp`, shown as a phrase-structure tree:

`exp` :=



The next derivation step will engage the `+g`/`-g` feature pair, transmitting the value v to `B`; v has to percolate to `A`, but not to `C`. This information is lost in the standard chain notation. However, adding rewritabilities and lineages allows to identify chains accessible to agreement – namely, those with `g` in the lineage:

$$\begin{aligned}\texttt{exp}' := {} & \mathit{mrg3}\,(\texttt{X}, \mathit{mrg3}\,(\texttt{C}, \mathit{mrg3}\,(\texttt{B}, \texttt{A}))) = \\ & \langle \epsilon, \left[\,\texttt{m:}v/\texttt{off}\,\right], \epsilon \rangle : \texttt{+g}^{[\,\texttt{m:}v/\texttt{off}\,]\rightarrow}\ \texttt{x}\,(\texttt{off}), \\ & \left[\,\texttt{m:}\epsilon/\texttt{f}\,\right]\ \texttt{-f}_{\leftarrow}\,(\texttt{f off}),\ \left[\,\texttt{m:}\epsilon/\texttt{f}\,\right]\ \texttt{-g}_{\leftarrow}\,(\texttt{f g off}),\ \left[\,\texttt{m:}\epsilon/\texttt{f}\,\right]\ \texttt{-h}_{\leftarrow}\,(\texttt{f g h off})\end{aligned}$$

2.2 Minimalist Grammars with agreement

Definition 7. A minimalist grammar with agreement ($\mathrm{MG_{agr}}$) G is a 5-tuple $\langle Mor, Syn, Types, Lex, \mathcal{F}\rangle$, where

$Mor = \{f : X \to Base_m \times (\{\texttt{on}, \texttt{off}\} \cup Base)\}$ is a set of bundles, where $Base_m$, Val are finite sets such that $\epsilon \in Val$ is the default value, and $Base$ is a nonempty finite set (of syntactic feature names);

$Feat = Syn \times \{\leftarrow, \not\leftarrow\} \times Mor$ is a set of annotated features, where Syn is a set of syntactic features built from $Base$ as specified in Definition 2;

$Types = \{::, :\}$, (lexical, derived)

Let $Lineage = \{s \mid s \in Base^* \;\&\; \text{for } 1 \le i, j \le |s|,\ s_i \ne s_j\}\ \{\texttt{off}\}$. Then the set of initial chains $IC = Mor^* \times Mor^* \times Mor^*\ Types\ Feat^*\ Lineage$, and the set of non-initial chains $NC = Mor^*\ Feat^*\ Lineage$;

[5] If movement is viewed as copying, it is not immediately clear why this should be the case. A system where moving subtrees retain their relation to the original position would be interesting to explore but falls outside the scope of this paper.

$Lex \subset \{\epsilon\} \times Mor^* \times \{\epsilon\}$ $\{::\}$ $Feat^*$ $\{\texttt{off}\}$, a subset of IC, is a finite set of lexical items,

$\mathcal{F} = \{$*merge-agr*, *move-agr*$\}$ is a set of structure-building operations.

Notation 8. Let $M \in Mor$ such that $M = \{\langle\phi_1, \langle v_1, r_1\rangle\rangle, ..., \langle\phi_n, \langle v_n, r_n\rangle\rangle\}$. Then for $i \in [1..n]$, $v_i \in Val$ is the *value* of ϕ_i in M, and $r_i \in \{\texttt{on}, \texttt{off}\} \cup Base$ is its *rewritability*. M can be written as $\begin{bmatrix} \phi_1{:}v_1/r_1 \\ \cdots \\ \phi_n{:}v_n/r_n \end{bmatrix}$ or as $\varnothing$ if $n = 0$.

Notation 9. Let $f = \langle x\texttt{f},\ Y,\ M\rangle \in Feat$ such that $\texttt{f} \in Base$. Then $f_{\mathrm{id}} = \texttt{f}$ is the name of f, $Y \in \{\leftarrow, \not\leftarrow\}$ specifies its receiving channel and $M \in Mor$ its emitting channel. f can be written as $x\texttt{f}_Y^{M\rightarrow}$. Where it does not lead to ambiguity, the receiving channel may be omitted if $Y = \not\leftarrow$, and the emitting channel may be omitted if $M = \varnothing$.

Notation 10. Chain lineages are enclosed in parentheses for better readability.

Notation 11. Let $\texttt{item} = \langle\epsilon, M, \epsilon\rangle\ ::\ \gamma\ (\texttt{off})$, where $M \in Mor$ and $\gamma \in Feat^*$, be a lexical item such that the bundle $M = \begin{bmatrix} \phi_1{:}v_1/r_1 \\ \cdots \\ \phi_n{:}v_n/r_n \end{bmatrix}$. Then $\begin{bmatrix} \textsc{item} \\ \phi_{m+1}{:}v_{m+1}/r_{m+1} \\ \cdots \\ \phi_n{:}v_n/r_n \end{bmatrix}$ can be used as a semi-formal abbreviation for M, where ITEM stands for a subset of lexically valued features $\{\langle\phi_1, \langle v_1, \texttt{off}\rangle\rangle, ..., \langle\phi_m, \langle v_m, \texttt{off}\rangle\rangle\} \subseteq M$.

At the level of bundles, agreement is handled by two functions updating *active* ($\texttt{on}$) features with values provided by the goal. Downward agreement ($agr_\downarrow$) leaves all features in the probe $\texttt{on}$, as the probe is, by definition, part of the initial chain. Upward agreement ($agr_\uparrow$) sets each feature in the probe to $\texttt{on}$ if active in the goal and to a given rewritability value otherwise. An auxiliary function, *act*, reactivates features with rewritability values present in a given sequence, setting them to $\texttt{on}$.

Definition 12. $agr_\downarrow : (Mor \times Mor) \rightarrow Mor$ is a function such that for $P, M \in Mor$, for $\phi \in Base_m$ the result of downward agreement of P with M is as follows:

$$agr_\downarrow(M, P, \phi) \equiv P^{\downarrow M}(\phi) = \begin{cases} \langle M(\phi)[1], \texttt{on}\rangle & \text{if } M(\phi) \text{ is defined} \\ & \text{and } P(\phi)[2] = \texttt{on}; \\ P(\phi) & \text{otherwise.} \end{cases}$$

Definition 13. $agr_\uparrow : ((\{\texttt{off}\} \cup Base) \times Mor \times Mor) \rightarrow Mor$ is a function such that for $P, M \in Mor$, $\texttt{re} \in \{\texttt{off}\} \cup Base$, $\phi \in Base_m$ the result of upward agreement of P with M (setting to $\texttt{re}$ the rewritability of any feature in P whose value is not to be shared with the feature in M) is as follows:

$$agr_\uparrow(\texttt{re}, M, P, \phi) \equiv P_{\texttt{re}}^{\uparrow M}(\phi) = \begin{cases} \langle M(\phi)[1], \texttt{on}\rangle & \text{if } P(\phi)[2] = \texttt{on} \\ & \text{and } M(\phi)[2] = \texttt{on}; \\ \langle M(\phi)[1], \texttt{re}\rangle & \text{if } P(\phi)[2] = \texttt{on} \\ & \text{and } M(\phi)[2] = \texttt{off}; \\ \langle P(\phi)[1], \texttt{re}\rangle & \text{if } M(\phi) \text{ is undefined;} \\ P(\phi) & \text{otherwise.} \end{cases}$$

Definition 14. *act* : $(Base^{+} \times Mor) \rightarrow Mor$ is a function such that for $P \in Mor$, $L \in Base^{+}$, $\phi \in Base_m$:

$$act(L, P, \phi) \equiv act_L(P, \phi) = \begin{cases} \langle P(\phi)[1], \mathtt{on} \rangle & \text{if } P(\phi)[2] \in L; \\ P(\phi) & \text{otherwise.} \end{cases}$$

The definitions are extended to apply to objects other than bundles:

Notation 15. For $M \in Mor$, $\mathtt{re} \in \{\mathtt{off}\} \cup Base$, $L \in Base^{+}$, $fun \in \{agr_{\downarrow}(M), agr_{\uparrow}(\mathtt{re}, M), act(L)\}$:

for $f \in Feat,\ fun(f) = \langle f[1], f[2], fun(f[3]) \rangle$;
for $(x_1, ..., x_n) \in Mor^{*},\ fun(x_1, ..., x_n) = fun(x_1), ..., fun(x_n)$;
for $(x_1, ..., x_n) \in Feat^{*},\ fun(x_1, ..., x_n) = fun(x_1), ..., fun(x_n)$;
for $c = s\, \gamma\, (A)\ \in\ NC$ such that $s \in Mor^{*}$ and $\gamma \in Feat^{*}$,
$fun(c) = fun(s)\, fun(\gamma)\, (A)$.

Finally, *merge* and *move* are redefined to accommodate agreement. The new rules manipulate lineages as well as bundles. Note that $P^{\downarrow\varnothing} = P$, but $P^{\uparrow\varnothing}_{\mathtt{re}} \neq P$: upward agreement with an empty bundle sets the rewritability of all features in the goal to $\mathtt{re}$. This special case corresponds to lack of agreement in the absence of a receiving channel and in intermediate positions of movement.

Definition 16. *merge-agr* is the union of the following five functions, for s_s, s_h, s_c, t_s, t_h, t_c, $t_1, ..., t_l \in Mor^{*}$, $\cdot \in \{:, ::\}$, $\mathtt{f}, \mathtt{x}_1, ..., \mathtt{x}_l \in Base$, $\gamma, \zeta \in Feat^{*}$, $\delta_1, ..., \delta_l \in Feat^{+}$, $M, N \in Mor$, $L_1, ..., L_l \in Lineage$, $X, Y \in \{\leftarrow, \not\leftarrow\}$, $\alpha_1, ..., \alpha_k$, $t_1\ \delta_1\ \mathtt{x}_1 L_1, ..., t_l\ \delta_l\ \mathtt{x}_l L_l\ \in NC$ $(0 \leq k, l)$,

mrg1-agr:

$$\frac{\langle \epsilon, s_h, \epsilon \rangle\ ::\ \mathtt{=f}_X^{M\rightarrow}\gamma\,(\mathtt{off}) \qquad \langle t_s, t_h, t_c \rangle\ \cdot\ \mathtt{f}_Y^{N\rightarrow}\,(\mathtt{off}),\ t_1\,\delta_1\,(\mathtt{x}_1 L_1), ..., t_l\,\delta_l\,(\mathtt{x}_l L_l)}{\langle \epsilon, s_h{}^{\downarrow\hat{N}}, (t_s t_h t_c)^{\uparrow\hat{M}}_{\mathtt{off}} \rangle\ :\ \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ (t_1\,\delta_1\,(\mathtt{x}_1 L_1))^{\uparrow\hat{M}}_{\mathtt{x}_1}, ..., (t_l\,\delta_l\,(\mathtt{x}_l L_l))^{\uparrow\hat{M}}_{\mathtt{x}_l}}$$

mrg2-agr:

$$\frac{\langle s_s, s_h, s_c \rangle\ :\ \mathtt{=f}_X^{M\rightarrow}\gamma\,(\mathtt{off}),\ \alpha_1, ..., \alpha_k \qquad \langle t_s, t_h, t_c \rangle\ \cdot\ \mathtt{f}_Y^{N\rightarrow}\,[\mathtt{off}],\ t_1\,\delta_1\,(\mathtt{x}_1 L_1), ..., t_l\,\delta_l\,(\mathtt{x}_l L_l)}{\langle (t_s t_h t_c)^{\uparrow\hat{M}}_{\mathtt{off}} s_s{}^{\downarrow\hat{N}}, s_h{}^{\downarrow\hat{N}}, s_c{}^{\downarrow\hat{N}} \rangle\ :\ \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ \alpha_1{}^{\downarrow\hat{N}}, ..., \alpha_k{}^{\downarrow\hat{N}},\ (t_1\,\delta_1\,(\mathtt{x}_1 L_1))^{\uparrow\hat{M}}_{\mathtt{x}_1}, ..., (t_l\,\delta_l\,(\mathtt{x}_l L_l))^{\uparrow\hat{M}}_{\mathtt{x}_l}}$$

mrg3-agr:

$$\frac{\langle s_s, s_h, s_c \rangle\ \cdot\ \mathtt{=f}_X^{M\rightarrow}\gamma\,(\mathtt{off}),\ \alpha_1, ..., \alpha_k \qquad \langle t_s, t_h, t_c \rangle\ \cdot\ \mathtt{f}_Y^{N\rightarrow} g\zeta\,(\mathtt{off}),\ t_1\,\delta_1\,(\mathtt{x}_1 L_1), ..., t_l\,\delta_l\,(\mathtt{x}_l L_l)}{\langle s_s{}^{\downarrow\hat{N}}, s_h{}^{\downarrow\hat{N}}, s_c{}^{\downarrow\hat{N}} \rangle\ :\ \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ \alpha_1{}^{\downarrow\hat{N}}, ..., \alpha_k{}^{\downarrow\hat{N}},\ (t_s t_h t_c\, g\zeta\,(g_{\mathrm{id}}\mathtt{off}))^{\uparrow\hat{M}}_{g_{\mathrm{id}}},\ (t_1\,\delta_1\,(g_{\mathrm{id}}\mathtt{x}_1 L_1))^{\uparrow\hat{M}}_{g_{\mathrm{id}}}, ..., (t_l\,\delta_l\,(g_{\mathrm{id}}\mathtt{x}_l L_l))^{\uparrow\hat{M}}_{g_{\mathrm{id}}}}$$

hmrg1-agr:

$$\frac{\langle \epsilon, s_h, \epsilon \rangle\ ::\ \mathtt{=>f}_X^{M\rightarrow}\gamma\,(\mathtt{off}) \qquad \langle t_s, t_h, t_c \rangle\ \cdot\ \mathtt{f}_Y^{N\rightarrow}\,(\mathtt{off}),\ t_1\,\delta_1\,(\mathtt{x}_1 L_1), ..., t_l\,\delta_l\,(\mathtt{x}_l L_l)}{\langle \epsilon, t_{h}{}^{\uparrow\hat{M}}_{\mathtt{off}} s_h{}^{\downarrow\hat{N}}, (t_s t_c)^{\uparrow\hat{M}}_{\mathtt{off}} \rangle\ :\ \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ (t_1\,\delta_1\,(\mathtt{x}_1 L_1))^{\uparrow\hat{M}}_{\mathtt{x}_1}, ..., (t_l\,\delta_l\,(\mathtt{x}_l L_l))^{\uparrow\hat{M}}_{\mathtt{x}_l}}$$

hmrg3-agr:

$$\frac{\langle s_s, s_h, s_c \rangle\ \cdot\ \mathtt{=>f}_X^{M\rightarrow}\gamma\,(\mathtt{off}),\ \alpha_1, ..., \alpha_k \qquad \langle t_s, t_h, t_c \rangle\ \cdot\ \mathtt{f}_Y^{N\rightarrow} g\zeta\,(\mathtt{off}),\ t_1\,\delta_1\,(\mathtt{x}_1 L_1), ..., t_l\,\delta_l\,(\mathtt{x}_l L_l)}{\langle s_s{}^{\downarrow\hat{N}}, t_{h}{}^{\uparrow\hat{M}}_{\mathtt{off}} s_h{}^{\downarrow\hat{N}}, s_c{}^{\downarrow\hat{N}} \rangle\ :\ \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ \alpha_1{}^{\downarrow\hat{N}}, ..., \alpha_k{}^{\downarrow\hat{N}},\ (t_s t_c\, g\zeta\,(g_{\mathrm{id}}\mathtt{off}))^{\uparrow\hat{M}}_{g_{\mathrm{id}}},\ (t_1\,\delta_1\,(g_{\mathrm{id}}\mathtt{x}_1 L_1))^{\uparrow\hat{M}}_{g_{\mathrm{id}}}, ..., (t_l\,\delta_l\,(g_{\mathrm{id}}\mathtt{x}_l L_l))^{\uparrow\hat{M}}_{g_{\mathrm{id}}}}$$

$$\text{where} \quad \hat{M} = \begin{cases} M & \text{if } Y = \leftarrow \\ \varnothing & \text{if } Y = \not\leftarrow \end{cases} \quad \text{and} \quad \hat{N} = \begin{cases} N & \text{if } X = \leftarrow \\ \varnothing & \text{if } X = \not\leftarrow \end{cases}$$

Movement rules have to address the additional complication of classifying chains. Each non-initial chain undergoes upward agreement (if it has the moving chain in its lineage) or downward agreement (as a subchain of the initial chain).

Definition 17. *move-agr* is the union of the following three functions, for s_s, s_h, $s_c, t_1, ..., t_k \in Mor^*$, $\mathtt{f} \in Base$, $\mathtt{F} \in \{\mathtt{-f}, *\mathtt{f}\}$, $\gamma, \zeta \in Feat^*$, $\delta_1, ..., \delta_k \in Feat^+$, $M, N \in Mor$, $A \in Base^*$, $B, L_1, ..., L_l \in Lineage$, $X, Y \in \{\leftarrow, \not\leftarrow\}$, and for $\alpha_1, ..., \alpha_k \in NC$ $(0 \leq k)$ such that for $j \in [1,k]$ $\alpha_j = t_j\ \delta_j\ L_j$, satisfying (SMC): there is exactly one $i \in [1,k]$ such that α_i has $\mathtt{-f}$ or $*\mathtt{f}$ as its first syntactic feature

mv1-agr:

$$\frac{\langle s_s, s_h, s_c \rangle \cdot \mathtt{+f}_X^{M\rightarrow}\gamma\,(\mathtt{off}),\ \alpha_1, ..., \alpha_{i-1},\ t_i\,\mathtt{F}_Y^{N\rightarrow}\,(A\mathtt{f}\ \mathtt{off}),\ \alpha_{i+1}, ..., \alpha_k}{\langle (act_{A\mathtt{f}}(t_i))^{\uparrow\hat{M}}_{\mathtt{off}} s_s{}^{\downarrow\hat{N}}, s_h{}^{\downarrow\hat{N}}, s_c{}^{\downarrow\hat{N}} \rangle \,:\, \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ \alpha_1{}', ..., \alpha_{i-1}{}',\ \alpha_{i+1}{}', ..., \alpha_k{}'}$$

where for $j \in [1,k],\ j \neq i$

$$\alpha_j{}' = \begin{cases} (act_{L_j{}'\mathtt{f}}(t_j\ \delta_j\ (\mathtt{x}L_j{}'')))^{\uparrow\hat{M}}_{\mathtt{x}} & \text{if } L_j = L_j{}'\mathtt{fx}L_j{}'' \text{ such that} \\ & L_j{}' \in Base^*, \mathtt{x} \in Base, L_j{}'' \in Lineage \\ (t_j\ \delta_j\ (L_j))^{\downarrow\hat{N}} & \text{otherwise} \end{cases}$$

mv2-agr:

$$\frac{\langle s_s, s_h, s_c \rangle \cdot \mathtt{+f}_X^{M\rightarrow}\gamma\,(\mathtt{off}),\ \alpha_1, ..., \alpha_{i-1},\ t_i\,\mathtt{F}_{\leftarrow}^{N\rightarrow}g\zeta\,(A\mathtt{f}B),\ \alpha_{i+1}, ..., \alpha_k}{\langle s_s{}^{\downarrow\hat{N}}, s_h{}^{\downarrow\hat{N}}, s_c{}^{\downarrow\hat{N}} \rangle \,:\, \gamma^{\downarrow\hat{N}}\,(\mathtt{off}),\ \alpha_1{}', ..., \alpha_{i-1}{}',\ (act_{A\mathtt{f}}(t_i\,g\zeta\,(g_{\mathrm{id}}B)))^{\uparrow\hat{M}}_{g_{\mathrm{id}}},\ \alpha_{i+1}{}', ..., \alpha_k{}'}$$

where for $j \in [1,k],\ j \neq i$

$$\alpha_j{}' = \begin{cases} (act_{L_j{}'\mathtt{f}}(t_j\ \delta_j\ (g_{\mathrm{id}}L_j{}'')))^{\uparrow\hat{M}}_{g_{\mathrm{id}}} & \text{if } L_j = L_j{}'\mathtt{f}L_j{}'' \text{ such that} \\ & L_j{}' \in Base^*, L_j{}'' \in Lineage \\ (t_j\ \delta_j\ (L_j))^{\downarrow\hat{N}} & \text{otherwise} \end{cases}$$

mv∗-agr:

$$\frac{\langle s_s, s_h, s_c \rangle \cdot \mathtt{+f}_X^{M\rightarrow}\gamma\,(\mathtt{off}),\ \alpha_1, ..., \alpha_{i-1},\ t_i\,*\mathtt{f}_Y^{N\rightarrow}\zeta\,(A\mathtt{f}B),\ \alpha_{i+1}, ..., \alpha_k}{\langle s_s, s_h, s_c \rangle \,:\, \gamma\,(\mathtt{off}),\ \alpha_1{}', ..., \alpha_{i-1}{}',\ (act_A(t_i\,*\mathtt{f}_Y^{N\rightarrow}\zeta\,(\mathtt{f}B)))^{\uparrow\varnothing}_{\mathtt{f}},\ \alpha_{i+1}{}', ..., \alpha_k{}'}$$

where for $j \in [1,k],\ j \neq i$

$$\alpha_j{}' = \begin{cases} (act_{L_j{}'}(t_j\ \delta_j\ (\mathtt{f}L_j{}'')))^{\uparrow\hat{M}}_{\mathtt{f}} & \text{if } L_j = L_j{}'\mathtt{f}L_j{}'' \text{ such that} \\ & L_j{}' \in Base^*, L_j{}'' \in Lineage \\ (t_j\ \delta_j\ (L_j))^{\downarrow\hat{N}} & \text{otherwise} \end{cases}$$

$$\text{where} \quad \hat{M} = \begin{cases} M & \text{if } Y = \leftarrow \\ \varnothing & \text{if } Y = \not\leftarrow \end{cases} \quad \text{and} \quad \hat{N} = \begin{cases} N & \text{if } X = \leftarrow \\ \varnothing & \text{if } X = \not\leftarrow \end{cases}$$

Example 18. $G' = \langle \Sigma_{G'}, Syn_{G'}, Types, Lex_{G'}, \mathcal{F} \rangle$ is a $\mathrm{MG}_{\mathrm{agr}}$. Its lexicon $Lex_{G'}$ contains the following lexical items:

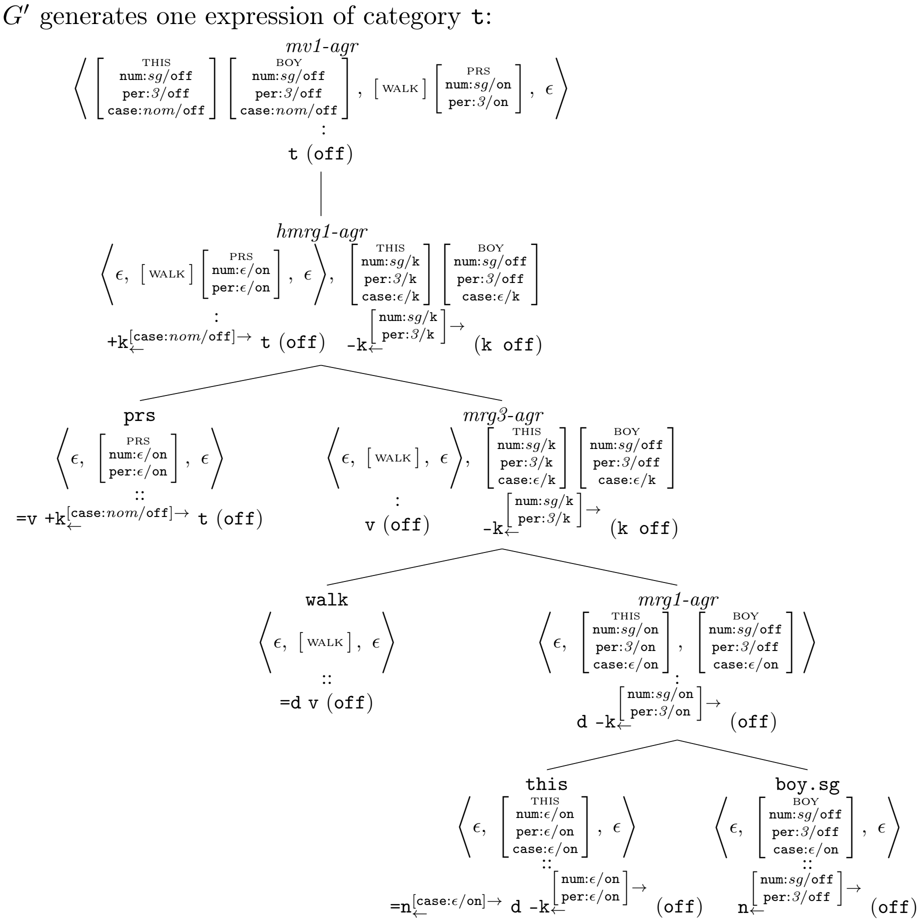
$$\texttt{this} := \left\langle \epsilon, \begin{bmatrix} \textsc{this} \\ \texttt{num}{:}\epsilon/\texttt{on} \\ \texttt{per}{:}\epsilon/\texttt{on} \\ \texttt{case}{:}\epsilon/\texttt{on} \end{bmatrix}, \epsilon \right\rangle :: \texttt{=n}^{[\texttt{case}:\epsilon/\texttt{on}]\rightarrow}_{\leftarrow}\ \texttt{d}\ \texttt{-k}^{\begin{bmatrix}\texttt{num}:\epsilon/\texttt{on}\\ \texttt{per}:\epsilon/\texttt{on}\end{bmatrix}\rightarrow}_{\leftarrow}\ (\texttt{off})$$

$$\texttt{boy.sg} := \left\langle \epsilon, \begin{bmatrix} \textsc{boy} \\ \texttt{num}{:}sg/\texttt{off} \\ \texttt{per}{:}3/\texttt{off} \\ \texttt{case}{:}\epsilon/\texttt{on} \end{bmatrix}, \epsilon \right\rangle :: \texttt{n}^{\begin{bmatrix}\texttt{num}:sg/\texttt{off}\\ \texttt{per}:3/\texttt{off}\end{bmatrix}\rightarrow}_{\leftarrow}\ (\texttt{off})$$

$$\texttt{walk} := \left\langle \epsilon, \begin{bmatrix} \textsc{walk} \end{bmatrix}, \epsilon \right\rangle :: \texttt{=d v } (\texttt{off})$$

$$\texttt{prs} := \left\langle \epsilon, \begin{bmatrix} \textsc{prs} \\ \texttt{num}{:}\epsilon/\texttt{on} \\ \texttt{per}{:}\epsilon/\texttt{on} \end{bmatrix}, \epsilon \right\rangle :: \texttt{=v +k}^{[\texttt{case}:nom/\texttt{off}]\rightarrow}_{\leftarrow}\ \texttt{t}\ (\texttt{off})$$

G' generates one expression of category $\texttt{t}$:



3 Case study: dative intervention in Icelandic

3.1 Data

A certain class of Icelandic constructions exhibits an interesting agreement pattern. The verb agrees in number with its nominative object inside a small clause (SC). However, this agreement seems optional: as an alternative, the verb may appear in the default 3.SG form. Furthermore, agreement can be disrupted by a dative experiencer intervening between the verb and the nominative object, in which case only the default verb form is possible. Only some experiencers cause this effect (1), while others are transparent for agreement (2):

(1) a. Það **finnst** fáum börnum [$_{SC}$ tölvurnar ljótar].
EXPL **find.SG** few children.DAT computers.DEF.NOM ugly.NOM

b. *Það **finnast** fáum börnum tölvurnar ljótar.
EXPL **find.PL** few children.DAT computers.DEF.NOM ugly.NOM

'Few children find the computers ugly.' [13, p.54–55]

(2) a. Það **finnst** mörgum stúdentum tölvurnar ljótar.
EXPL **find.SG** many students.DAT computers.DEF.NOM ugly.NOM

b. Það **finnast** mörgum stúdentum tölvurnar ljótar.
EXPL **find.PL** many students.DAT computers.DEF.NOM ugly.NOM

'Many students find the computers ugly.' [10, p.1000]

The intervention effect can occur even if the dative undergoes wh-movement and no longer linearly intervenes between the verb and the nominative object:

(3) a. Hvaða stúdent **finnst** tölvurnar ljótar?
which student.DAT **find.SG** computers.DEF.NOM ugly.NOM

b. Hvaða stúdent ??**finnast** tölvurnar ljótar?
which student.DAT **find.PL** computers.DEF.NOM ugly.NOM

'Which student finds the computers ugly?' [10, p.1001]

A generalization dealing with these examples is proposed in [13]: dative experiencers are transparent for agreement just in case they can undergo Object Shift – a movement to the specifier of *v*. The ability of a DP to shift is an independent property of the quantifier. Furthermore, [13] argues that agreement with the nominative is, in fact, deterministic: obligatory iff the experiencer *has* shifted, impossible otherwise. Object Shift is string-vacuous in examples like (2a). However, VP-level adverbs are expected to precede an in-situ dative and follow a shifted dative. The former configuration is compatible only with default agreement (4), while the latter only with normal agreement (5):

(4) a. Það **finnst** fljótt mörgum köttum mýsnar góðar.
EXPL **find.SG** quickly many cats.DEF.DAT mice.DEF.NOM tasty

b. Það ??/***finnast** fljótt mörgum köttum mýsnar góðar.
EXPL **find.PL** quickly many cats.DEF.DAT mice.DEF.NOM tasty

'Many cats quickly find the mice tasty.' [13, p.63]

(5) a. Það ??/***finnst** mörgum köttum fljótt mýsnar góðar.
EXPL **find.SG** many cats.DEF.DAT quickly mice.DEF.NOM tasty

b. Það **finnast** mörgum köttum fljótt mýsnar góðar.
EXPL **find.PL** many cats.DEF.DAT quickly mice.DEF.NOM tasty

'Many cats quickly find the mice tasty.' [13, p.63]

This reasoning is extended to fronted dative experiencers, including wh-arguments (3), which are also required to undergo Object Shift for agreement to succeed. Importantly, Object Shift does not alter the relation between the T(ense) head, which morphologically manifests agreement, and the nominative object. This observation can be reconciled with the traditional notion of Agree by assuming that the primary locus of agreement is lower than T – namely, v. Object Shift removes the dative from the *probing domain* of v, allowing v to probe the nominative object. T inherits the relevant features from v (via Agree with other heads intervening between T and v). This approach to Agree reduces an instance of long-distance agreement to a series of local dependencies – not unlike the $\mathrm{MG}_{\mathrm{agr}}$ formalism.

3.2 Grammar Fragment

Let each determiner phrase transmit its number value via its category feature channel and the default value via its licensee channel. Then the difference between *shiftable* and *non-shiftable* DPs can be reduced to the distinction between a starred licensee (∗`k`) and a plain licensee (`-k`).[6]

$$\texttt{many}\sim \quad := \left\langle \epsilon, \left[\genfrac{}{}{0pt}{}{\text{MANY}\sim}{\texttt{num:}\mathit{pl}/\texttt{off}}\right], \epsilon \right\rangle \; :: \; \texttt{d}^{[\texttt{num:}\mathit{pl}/\texttt{off}]\rightarrow} \; *\texttt{k}^{[\texttt{num:}\epsilon/\texttt{off}]\rightarrow} \quad (\texttt{off})$$

$$\texttt{few}\sim \quad := \left\langle \epsilon, \left[\genfrac{}{}{0pt}{}{\text{FEW}\sim}{\texttt{num:}\mathit{pl}/\texttt{off}}\right], \epsilon \right\rangle \; :: \; \texttt{d}^{[\texttt{num:}\mathit{pl}/\texttt{off}]\rightarrow} \; \texttt{-k}^{[\texttt{num:}\epsilon/\texttt{off}]\rightarrow} \quad (\texttt{off})$$

The verb `find` selects a small clause (containing an object DP) and an experiencer DP, receiving a number value from the former. There is no agreement with the dative experiencer, so the relevant selector `=d` has no receiving channel.

$$\texttt{find} \quad := \left\langle \epsilon, \left[\,\text{FIND}\,\right], \epsilon \right\rangle \; :: \; \texttt{=sc}_{\leftarrow} \; \texttt{=d} \; \texttt{V}^{[\texttt{num:}\epsilon/\texttt{on}]\rightarrow} \quad (\texttt{off})$$

$$\texttt{SC} \quad := \left\langle \epsilon, \left[\genfrac{}{}{0pt}{}{\text{SC}}{\texttt{num:}\mathit{pl}/\texttt{off}}\right], \epsilon \right\rangle \; :: \; \texttt{sc}^{[\texttt{num:}\mathit{pl}/\texttt{off}]\rightarrow} \quad (\texttt{off})$$

The crucial point in the derivation is `AgrO` which has two receiving channels. `v` and $\texttt{v}_{\texttt{shift}}$ pass information along; additionally, $\texttt{v}_{\texttt{shift}}$ provides the landing site of Object Shift. Thus, `T` eventually receives the last value transmitted to `AgrO`.

$$\texttt{AgrO} \quad := \left\langle \epsilon, \left[\,\text{AgrO}\,\right], \epsilon \right\rangle \; :: \; \texttt{=>V}_{\leftarrow} \; \texttt{+k}_{\leftarrow} \; \texttt{agrO}^{[\texttt{num:}\epsilon/\texttt{on}]\rightarrow} \quad (\texttt{off})$$

[6] For space reasons, DPs and small clauses are treated as atomic units. They can be decomposed to model internal agreement in detail. In particular, it is possible to embed DPs within another functional projection, connecting intervention and case (cf. [14], i.a.); this would allow the verb to assign dative to its argument and manipulate the agreement properties of its outer layer at the same time.

$$\mathtt{v} \;:=\; \left\langle \epsilon, [\,\text{v}\,], \epsilon \right\rangle \;::\; \mathtt{=>agrO}_{\leftarrow}\ \mathtt{v}^{[\mathtt{num}:\epsilon/\mathtt{on}]\rightarrow} \quad (\mathtt{off})$$

$$\mathtt{v}_{\mathtt{shift}} \;:=\; \left\langle \epsilon, [\,\text{V}_{\text{SHIFT}}\,], \epsilon \right\rangle \;::\; \mathtt{=>agrO}_{\leftarrow}\ \mathtt{+k}\ \mathtt{v}^{[\mathtt{num}:\epsilon/\mathtt{on}]\rightarrow} \quad (\mathtt{off})$$

$$\mathtt{T} \;:=\; \left\langle \epsilon, \left[{}^{\text{T}}_{\mathtt{num}:\epsilon/\mathtt{on}}\right], \epsilon \right\rangle \;::\; \mathtt{=>v}_{\leftarrow}\ \mathtt{t} \quad (\mathtt{off})$$

Abstracting away from person and case, the eight lexical items defined above suffice[7] to model the number contrast between (1) and (2). Any experiencer can check its licensee in the specifier of `AgrO`. In this case, `AgrO` receives the default number value, giving rise to (1a) and (2a).

Example 19. The derivation of (1a) proceeds as follows:

$$\mathtt{i1} := \textit{mrg1-agr}\,(\mathtt{find}, \mathtt{sc}) =$$
$$\left\langle \epsilon, [\,\text{FIND}\,], \left[{}^{\text{SC}}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right] \right\rangle : \mathtt{=d}\,\mathtt{V}^{[\mathtt{num}:\mathit{pl}/\mathtt{on}]\rightarrow}\,(\mathtt{off})$$

$$\mathtt{i2} := \textit{mrg3-agr}\,(\mathtt{i1}, \mathtt{few}{\sim}) =$$
$$\left\langle \epsilon, [\,\text{FIND}\,], \left[{}^{\text{SC}}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right] \right\rangle : \mathtt{V}^{[\mathtt{num}:\mathit{pl}/\mathtt{on}]\rightarrow}\,(\mathtt{off}),$$
$$\left[{}^{\text{FEW}\sim}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right]\ \mathtt{-k}^{[\mathtt{num}:\epsilon/\mathtt{off}]\rightarrow}\,(\mathtt{off})$$

$$\mathtt{i3} := \textit{hmrg1-agr}\,(\mathtt{AgrO}, \mathtt{i2}) =$$
$$\left\langle \epsilon, [\,\text{FIND}\,][\,\text{AgrO}\,], \left[{}^{\text{SC}}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right] \right\rangle :: \mathtt{+k}_{\leftarrow}\ \mathtt{agrO}^{[\mathtt{num}:\mathit{pl}/\mathtt{on}]\rightarrow}\,(\mathtt{off}),$$
$$\left[{}^{\text{FEW}\sim}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right]\ \mathtt{-k}^{[\mathtt{num}:\epsilon/\mathtt{off}]\rightarrow}\,(\mathtt{k\ off})$$

$$\mathtt{i4} := \textit{mv1-agr}\,(\mathtt{i3}) =$$
$$\left\langle \left[{}^{\text{FEW}\sim}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right], [\,\text{FIND}\,][\,\text{AgrO}\,], \left[{}^{\text{SC}}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right] \right\rangle : \mathtt{agrO}^{[\mathtt{num}:\epsilon/\mathtt{on}]\rightarrow}\,(\mathtt{off})$$

$$\mathtt{i5} := \textit{hmrg1-agr}\,(\mathtt{v}, \mathtt{i4}) =$$
$$\left\langle \epsilon, [\,\text{FIND}\,][\,\text{AgrO}\,][\,\text{v}\,], \left[{}^{\text{FEW}\sim}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right]\left[{}^{\text{SC}}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right] \right\rangle : \mathtt{v}^{[\mathtt{num}:\epsilon/\mathtt{on}]\rightarrow}\,(\mathtt{off})$$

$$\mathtt{i6} := \textit{hmrg1-agr}\,(\mathtt{T}, \mathtt{i5}) =$$
$$\left\langle \epsilon, [\,\text{FIND}\,][\,\text{AgrO}\,][\,\text{v}\,]\left[{}^{\text{T}}_{\mathtt{num}:\epsilon/\mathtt{on}}\right], \left[{}^{\text{FEW}\sim}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right]\left[{}^{\text{SC}}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right] \right\rangle : \mathtt{t}\,(\mathtt{off})$$

[7] These items support partial derivations up to `T`, where morphological dependencies are resolved. Assuming that expletives are merged above TP and move to the specifier of CP [3], the following addition enables full CP derivations of (1)–(3):

$$\mathtt{which}{\sim} \;:=\; \left\langle \epsilon, \left[{}^{\text{WHICH}\sim}_{\mathtt{num}:\mathit{pl}/\mathtt{off}}\right], \epsilon \right\rangle \;::\; \mathtt{d}^{[\mathtt{num}:\mathit{pl}/\mathtt{off}]\rightarrow}\ \mathtt{-k}^{[\mathtt{num}:\epsilon/\mathtt{off}]\rightarrow}\ \mathtt{-wh} \quad (\mathtt{off})$$

$$\mathtt{T}_{\mathtt{expl}} \;:=\; \left\langle \epsilon, [\,\text{T}_{\text{EXPL}}\,], \epsilon \right\rangle \;::\; \mathtt{=t}\ \mathtt{=expl}\ \mathtt{t} \quad (\mathtt{off})$$

$$\mathtt{Expl} \;:=\; \left\langle \epsilon, [\,\text{EXPL}\,], \epsilon \right\rangle \;::\; \mathtt{expl}\ \mathtt{-wh} \quad (\mathtt{off})$$

$$\mathtt{C} \;:=\; \left\langle \epsilon, [\,\text{C}\,], \epsilon \right\rangle \;::\; \mathtt{=t}\ \mathtt{+wh}\ \mathtt{c} \quad (\mathtt{off})$$

The second option is only available for shiftable experiencers. The $*\texttt{k}$ feature can "survive" movement to the specifier of `AgrO` to be checked later, in the specifier of $\texttt{v}_{\texttt{shift}}$ (whose `+k` has no receiving channel): this movement corresponds to Object Shift. In this case, `AgrO` will never receive the default agreement value and will instead transmit whatever value came from the small clause, resulting in the verb agreeing with its nominative argument.

Example 20. The derivation of (2b) proceeds as follows:

$$
\begin{aligned}
\texttt{a1} &:= \textit{mrg1-agr}\,(\texttt{find}, \texttt{sc}) \;= \\
&\quad \left\langle \epsilon, \left[\,\textsc{find}\,\right], \left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{=d}\,\texttt{V}^{[\texttt{num}:\,pl/\texttt{on}]\rightarrow}\,(\texttt{off}) \\
\texttt{a2} &:= \textit{mrg3-agr}\,(\texttt{a1}, \texttt{many}{\sim}) \;= \\
&\quad \left\langle \epsilon, \left[\,\textsc{find}\,\right], \left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{V}^{[\texttt{num}:\,pl/\texttt{on}]\rightarrow}\,(\texttt{off}), \\
&\quad \left[\substack{\textsc{many}\sim\\ \texttt{num}:\,pl/\texttt{off}}\right] \;*\texttt{k}^{[\texttt{num}:\,\epsilon/\texttt{off}]\rightarrow}\,(\texttt{k off}) \\
\texttt{a3} &:= \textit{hmrg1-agr}\,(\texttt{AgrO}, \texttt{a2}) \;= \\
&\quad \left\langle \epsilon, \left[\,\textsc{find}\,\right]\left[\,\textsc{AgrO}\,\right], \left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{+k}_{\leftarrow}\,\texttt{agrO}^{[\texttt{num}:\,pl/\texttt{on}]\rightarrow}\,(\texttt{off}), \\
&\quad \left[\substack{\textsc{many}\sim\\ \texttt{num}:\,pl/\texttt{off}}\right] \;*\texttt{k}^{[\texttt{num}:\,\epsilon/\texttt{off}]\rightarrow}\,(\texttt{k off}) \\
\texttt{a4} &:= \textit{mv}{*}\textit{-agr}\,(\texttt{a3}) \;= \\
&\quad \left\langle \epsilon, \left[\,\textsc{find}\,\right]\left[\,\textsc{AgrO}\,\right], \left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{agrO}^{[\texttt{num}:\,pl/\texttt{on}]\rightarrow}\,(\texttt{off}), \\
&\quad \left[\substack{\textsc{many}\sim\\ \texttt{num}:\,pl/\texttt{off}}\right] \;*\texttt{k}^{[\texttt{num}:\,\epsilon/\texttt{off}]\rightarrow}\,(\texttt{k off}) \\
\texttt{a5} &:= \textit{hmrg1-agr}\,(\texttt{v}_{\texttt{shift}}, \texttt{a4}) \;= \\
&\quad \left\langle \epsilon, \left[\,\textsc{find}\,\right]\left[\,\textsc{AgrO}\,\right]\left[\,\textsc{v}_{\textsc{shift}}\,\right], \left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{+k}\,\texttt{v}^{[\texttt{num}:\,pl/\texttt{on}]\rightarrow}\,(\texttt{off}), \\
&\quad \left[\substack{\textsc{many}\sim\\ \texttt{num}:\,pl/\texttt{off}}\right] \;*\texttt{k}^{[\texttt{num}:\,\epsilon/\texttt{off}]\rightarrow}\,(\texttt{k off}) \\
\texttt{a6} &:= \textit{mv1-agr}\,(\texttt{a5}) \;= \\
&\quad \left\langle \left[\substack{\textsc{many}\sim\\ \texttt{num}:\,pl/\texttt{off}}\right], \left[\,\textsc{find}\,\right]\left[\,\textsc{AgrO}\,\right]\left[\,\textsc{v}_{\textsc{shift}}\,\right], \left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{v}^{[\texttt{num}:\,pl/\texttt{on}]\rightarrow}\,(\texttt{off}) \\
\texttt{a7} &:= \textit{hmrg1-agr}\,(\texttt{T}, \texttt{a6}) \;= \\
&\quad \left\langle \epsilon, \left[\,\textsc{find}\,\right]\left[\,\textsc{AgrO}\,\right]\left[\,\textsc{v}_{\textsc{shift}}\,\right]\left[\substack{\text{T}\\ \texttt{num}:\,pl/\texttt{on}}\right], \left[\substack{\textsc{many}\sim\\ \texttt{num}:\,pl/\texttt{off}}\right]\left[\substack{\text{SC}\\ \texttt{num}:\,pl/\texttt{off}}\right] \right\rangle \;:\; \texttt{t}\,(\texttt{off})
\end{aligned}
$$

4 Discussion

In this paper, I have developed a modification of Minimalist Grammars that accommodates morphological agreement, redefining syntactic operations over bundles of morphological features. The Agree operation of Chomsky's Minimalist Program is reduced to local transmission of information over syntactic dependencies. In order to demonstrate the practicality of the new formalism, I have used it to express a precise, formalized analysis of Icelandic dative intervention inspired by the proposal in [13]. The key element of the generalization

– namely, the link between Object Shift and agreement – has been translated into a grammar fragment, which can be expanded further to incorporate more insights from Minimalist syntax (see e.g. fn. 6, 7).

MGs with agreement output sequences of bundles and can interface with any sufficiently explicit theory of morphology. For instance, they are compatible with formalizations of Distributed Morphology [9] that take a sequence of feature structures as syntactic input: [19] and, more recently, [6].

While proving the equivalence of MGs with agreement and unmodified MGs is outside the scope of this paper, a few observations on the matter are in order. Agreement is transmitted across dependencies established by structure-building operations. This limits the number of goals in any given expression to the number of chains, which, in turn, is guaranteed to be finite by the SMC. The problem of long-distance upward agreement, which is the only remaining source of "nonlocality" in the grammar, is addressed by adopting an approach similar to Feature Sharing: informally speaking, each item that requires a feature value via upward agreement pushes the responsibility for obtaining it to whatever item immediately selects or licenses it. Thus, at any given point in the derivation, there are finitely many *different* morphological feature values which can be updated or transmitted to other items. This makes it possible to convert an $\mathrm{MG}_{\mathrm{agr}}$ into an equivalent MG, reformulating lexical items over bundles in terms of unanalyzable elements (corresponding to bundles of valued features) and recasting agreement transformations as compatibility constraints.

Acknowledgments My thanks to Karlos Arregi and Greg Kobele for their advice regarding both the linguistic and the formal aspects of this paper, and to the anonymous reviewers for their helpful comments and suggestions.

References

1. Adger, D.: A minimalist theory of feature structure. Features: Perspectives on a key notion in linguistics pp. 185–218 (2010)
2. Bošković, Ž.: On successive cyclic movement and the freezing effect of feature checking. In: Hartmann, J., Hegedüs, V., van Riemsdijk, H. (eds.) Sounds of silence: Empty elements in syntax and phonology. Elsevier North Holland, Amsterdam (in press)
3. Bowers, J.: Transitivity. Linguistic Inquiry 33(2), 183–224 (2002)
4. Chomsky, N.: The Minimalist Program. MIT Press, Cambridge, MA (1995)
5. Chomsky, N.: Minimalist Inquiries: the framework. In: Martin, R., Michaels, D., Uriagereka, J. (eds.) Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik, pp. 89–156. MIT Press, Cambridge, MA (2000)
6. Ermolaeva, M., Edmiston, D.: Distributed morphology over strings. Poster presented at PLC 41, Philadelphia (2017)
7. Frampton, J., Gutmann, S.: Agreement is feature sharing (2000), unpublished manuscript
8. Graf, T.: Local and Transderivational Constraints in Syntax and Semantics. Ph.D. thesis, UCLA (2013)

9. Halle, M., Marantz, A.: Distributed morphology and the pieces of inflection. In: Hale, K., Keyser, S. (eds.) The view from building 20, pp. 111–176. The MIT Press (1993)
10. Holmberg, A., Hróarsdóttir, T.: Agreement and movement in icelandic raising constructions. Lingua 114(5), 651–673 (2004)
11. Kobele, G.M.: Generating Copies: An investigation into structural identity in language and grammar. Ph.D. thesis, UCLA (2006)
12. Kobele, G.M.: A derivational approach to phrasal spellout (2012), slides presented at BCGL 7
13. Kučerová, I.: Long-distance agreement in icelandic: locality restored. The Journal of Comparative Germanic Linguistics 19(1), 49–74 (2016)
14. Rezac, M.: Phi-agree and theta-related case. In: Harbour, D., Adger, D., Bejar, S. (eds.) Phi theory. Phi Features across Modules and Interfaces, pp. 83–130. The MIT Press (2008)
15. Shieber, S.M.: An introduction to unification-based approaches to grammar. CSLI lecture notes (1986)
16. Stabler, E.P.: Derivational minimalism. In: Retoré, C. (ed.) Logical Aspects of Computational Linguistics: First International Conference, LACL '96 Nancy, France, September 23–25, 1996 Selected Papers, pp. 68–95. Springer Berlin Heidelberg, Berlin, Heidelberg (1997)
17. Stabler, E.P.: Recognizing head movement. In: Proceedings of the 4th International Conference on Logical Aspects of Computational Linguistics. pp. 245–260. LACL '01, Springer-Verlag, London, UK, UK (2001)
18. Stabler, E.P.: Computational perspectives on minimalism. Oxford handbook of linguistic minimalism pp. 617–643 (2011)
19. Trommer, J.: Morphology consuming syntax' resources: generation and parsing in a minimalist version of distributed morphology. In: Proceedings of the ESSLI Workshop on Resource Logics and Minimalist Grammars (1999)