Realizational Morphosemantics in L$_{\text{R}}$FG

Lexical-Realizational Functional Grammar (L$_{\text{R}}$FG; Melchin et al. 2020, Asudeh et al. 2021) is a theoretical framework that couples Lexical-Functional Grammar with the realizational, morpheme-based approach to word-formation of Distributed Morphology (DM; Halle and Marantz 1993). L$_{\text{R}}$FG has been developed to take advantage of LFG's strengths in modelling nonconfigurationality and DM's strengths in modelling complex non-fusional morphology and polysynthesis.

In this talk, we present some initial attempts at an L$_{\text{R}}$FG theory and formalization of *morphosemantics*, i.e. the morphology–semantics interface. We distinguish the domain of morphosemantics from the more general domain of lexical semantics. We take morphosemantics to encompass all and only aspects of meaning that affect the mapping from a semantic representation to a phonological representation. Phenomena that L$_{\text{R}}$FG attributes to the morphology-semantics interface include:

1. Semantically conditioned morphology, e.g. English prefix *re-* (*reread*)
2. Polysemy, e.g. *keep*, *clutch*
3. Lexicalization, e.g. *lousy*, *transmission*
4. Regulars/irregulars, e.g. *brothers/brethren*, *compárable/cómparable*, *productiveness/productivity*

We also compare L$_{\text{R}}$FG's prospects for capturing morphosemantics with those of the realizational morphological approaches for LFG developed by Dalrymple (2015)/Dalrymple, Lowe, and Mycock (2019) and Thomas (2021).

In order to explain morphosemantics in L$_{\text{R}}$FG terms, it is helpful to look at L$_{\text{R}}$FG's version of the (LFG) Correspondence Architecture (Kaplan 1987, 1995), presented in Figure 1. It shows that morphosemantics in L$_{\text{R}}$FG concerns aspects of meaning that condition the mapping to v(ocabulary)-structure, as encoded in Glue Semantics *meaning constructors*.

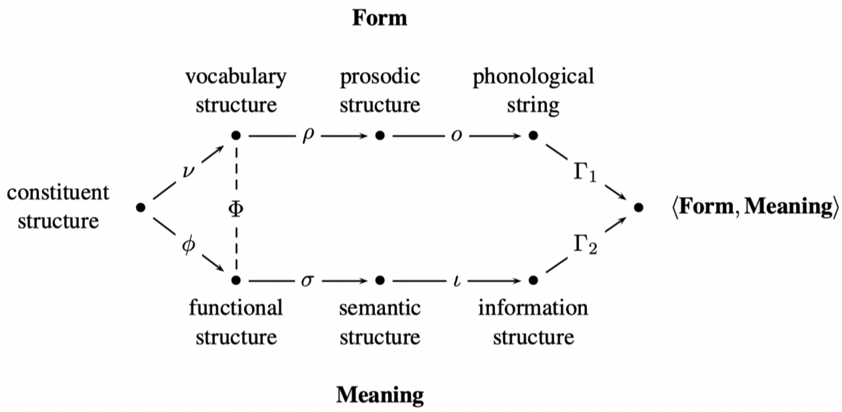

Figure 1: L$_{\text{R}}$FG's Correspondence Architecture

L$_{\text{R}}$FG is syntactically similar to standard LFG, with changes to the c(onstituent)-structure tree and its relationship with morphosyntactic elements. Importantly, the terminal nodes of c-structures in L$_{\text{R}}$FG *are not words*, but instead are LFG *f-descriptions* (and now also Glue Semantics meaning constructors). Vocabulary Items in L$_{\text{R}}$FG work have thus far been defined as ν-mappings from a pair of inputs to a v-structure representation that includes information about form and prosodically-conditioned linearization, although here as before we set all aspects of v-structure but form aside. These input pairs have consisted of a) categories (represented as a list, to allow for *spans* of multi-word expressions and the like) and b) sets of f-structures generated by a given f-description.

Consider example (1) from Ojibwe (*Anishinaabemowin*, Algonquian). The analysis of this example in Figure 2 demonstrates the L$_{\text{R}}$FG framework.

(1) gi- gii- waab -am -igw -naan -ag
    2 PST see VTA INV 1PL 3PL
    'They saw us(incl).'

An example of an L$_{\text{R}}$FG Vocabulary Item (VI) for the Tense node in Figure 2 is shown in (2).

(2) $\langle\ [\mathrm{T}],\ \Phi\big\{(\uparrow \textsc{tense}) = \textsc{pst}\big\}\ \rangle \quad \xrightarrow{\nu} \quad \textit{gii}$

The list of categories, shown in red, consists of only a T category. The f-description, shown in purple, is also simple in this case, consisting of the single defining equation $(\uparrow \textsc{tense}) = \textsc{pst}$. The bridging function, Φ, maps this f-description to the set of minimal f-structures that satisfy it (Melchin et al. 2020).

In order to capture *morphosemantics*, we here propose the addition of Glue Semantics meaning constructors to the input of the ν-mapping function, such that the input is a triple rather than a pair. As before, the first two coordinates are a list of categories and a (set of f-structures defined by an) f-description, but we now add a third coordinate that is a set of meaning constructors. These meaning constructors are the Vocabulary Item's contribution to compositional semantics. For example, (3) shows the mapping for the Ojibwe root *waab* and its English counterpart *see*.

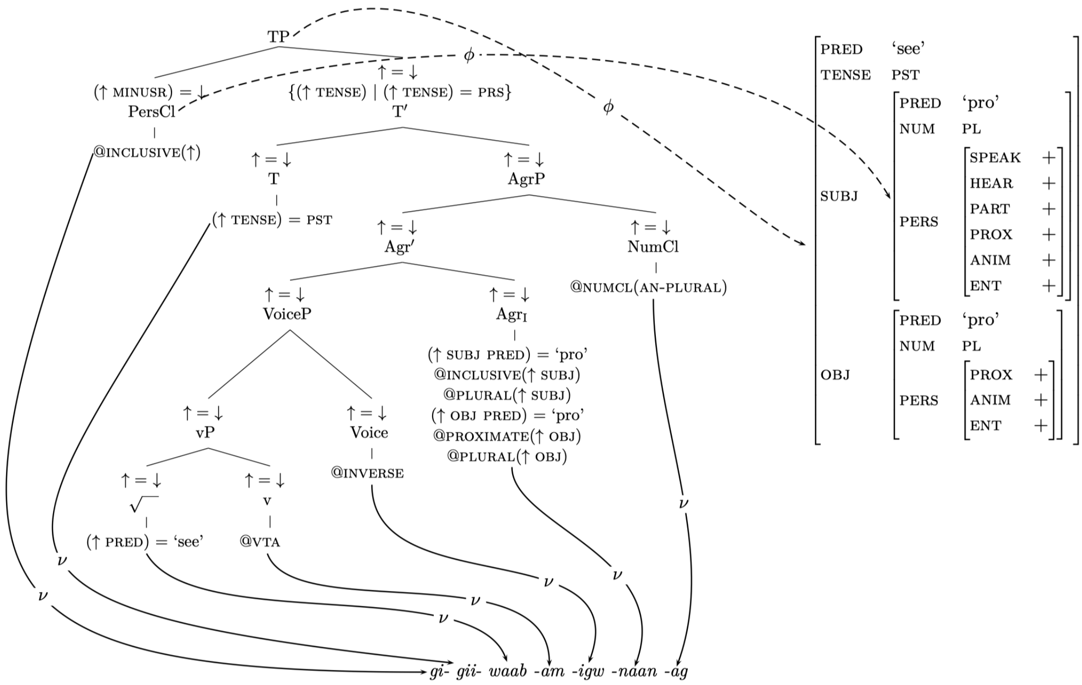

Figure 2: L_RFG c-structure and f-structure for Ojibwe (1)

(3) a. $\langle\, [\sqrt{}],\ \Phi\{(\uparrow \text{PRED}) = \text{`see'}\},\ \{\textbf{see} : (\uparrow \text{OBJ})_\sigma \multimap (\uparrow \text{SUBJ})_\sigma \multimap {\uparrow_\sigma}\}\,\rangle \xrightarrow{\nu}$ *waab* Ojibwe

b. $\langle\, [\sqrt{}],\ \Phi\{(\uparrow \text{PRED}) = \text{`see'}\},\ \{\textbf{see} : (\uparrow \text{OBJ})_\sigma \multimap (\uparrow \text{SUBJ})_\sigma \multimap {\uparrow_\sigma}\}\,\rangle \xrightarrow{\nu}$ *see* English

As before, the list of categories is shown in red, the f-description in purple, and now the meaning constructors are shown in blue.

Melchin et al. (2020) specify an L_RFG f-structural condition on exponence (i.e., a condition on the ν-mapping) called **MostInformative**. We here propose a family of such conditions, formalized as functions, expanding the notion of **MostInformative** to c-structure and s(emantic)-structure. Each function takes two VIs as its arguments. In the paper, we provide full formalizations of **MostInformative**$_f$ and **MostInformative**$_c$, but here focus on the morphosemantic constraint, **MostInformative**$_s$.

- **MostInformative**$_f(\alpha, \beta)$ returns whichever of α,β has the most specific f-structure in the set of f-structures returned by Φ applied to (the sets of f-structures defined by) α/β's f-descriptions.
  Intuition: Choose the VI that realizes an f-description that defines an f-structure that contains the greater set of features.
  Formalization: The proper subsumption relation on f-structures is used to capture the intuition.
- **MostInformative**$_c(\alpha, \beta)$ returns whichever of α,β has the longest list of c-structure categories.
  Intuition: Choose the VI that realizes the greater set of categories.
  Formalization: The proper subset relation on lists-as-sets is used to capture the intuition.
- **MostInformative**$_s(\alpha, \beta)$ returns whichever Vocabulary Item has the more specific meaning.
  Intuition: Choose the VI whose denotation is more semantically contentful.
  Formalization: The proper subset relation on set-denoting expressions is used to capture the intuition.

(4) Given two Vocabulary Items, α and β, of the same semantic type T, where T is any type ending in t,

$$\textbf{MostInformative}_s(\alpha, \beta) = \begin{cases} \alpha \textbf{ if } [\![\alpha]\!] \subset [\![\beta]\!] \\ \beta \textbf{ if } [\![\beta]\!] \subset [\![\alpha]\!] \\ \bot \textbf{ otherwise} \end{cases}$$

Note that each version of **MostInformative** can result in a tie, represented by $\bot$. Also note that there are regularities in the mappings/interfaces between structures, so it would be unlikely for all three **MostInformative** constraints to yield $\bot$. We are not currently aware of any empirical case that would merit such an analysis.

Our morphosemantic theory can be demonstrated by the English pair *brothers*/*brethren*, which exemplifies a familiar hyponymy relationship triggered by the coexistence of regular and irregular forms. *Brethren* takes the exceptional plural

-*en*, but unlike the case with *child/children* or *ox/oxen*, *brethren* is not the only plural of *brother* and has a distinct, more specific meaning than the regular plural, *brothers*.

Following Partee and Borschev (2003), we assume that a relational noun like $\sqrt{\text{BROTHER}}$ involves a relation between the nominal entity and some other entity, such as a possessor. The meaning term can be represented as follows:

(5) $\lambda y \lambda x \lambda R.\textbf{male}(x) \wedge R(x, y)$

Notice that, if left unresolved, the relational variable, R, must be filled from context. This is the meaning term for the *obligatory* meaning constructor for $\sqrt{\text{BROTHER}}$.

The relation typically defaults to sibling, so we assume that there is a second, *optional* meaning constructor for $\sqrt{\text{BROTHER}}$ whose meaning term modifies the term above as follows:

(6) $\lambda R.R(\textbf{sibling})$

Thus, the *default* interpretation for singular *brother* is male sibling.

However, as the term in (6) is optional, R in (5) can instead be instantiated contextually/pragmatically, for example as **close.friend** (where culturally appropriate, which is evidence of its pragmatic nature). Indeed, *brother* can also be the singular of *brethren*, with the relevant meaning, as in the favoured reading (outside of other context) of a monk saying about another monk at the same monastery:

(7) My brother spoke out of turn.

We assume that the regular plural morpheme -*s* just expresses a plural meaning as a modifier, $\lambda P.{}^*P$, following Link (1983). Thus, the *default* interpretation of plural *brothers* is male siblings.

In contrast, *brethren* obligatorily expresses the following relational meaning constructor in addition to the general meaning in (5) and the meaning of the plural:

(8) $\lambda R.R(\textbf{member.of.same.order})$

So *brethren* denotes the members of an all-male order. (For speakers for whom the group must be a religious order, the meaning can be suitably further restricted.)

Given the default interpretation of male siblings for $\sqrt{\text{BROTHER}}+\text{PL}$, there are no grounds for $\textbf{MostInformative}_s$ to choose either one of *brothers/brethren* over the other (i.e., it returns $\perp$, which means the constraint is not decisive). This is because the set of members of an all-male order is not a subset of the set of male siblings or vice versa. However, if the relation R in $\sqrt{\text{BROTHER}}+\text{PL}$ is left unspecified, then the set of members of an all-male group *is* a subset of the set of males that bear some relation to something, so $\textbf{MostInformative}_s$ would choose *brethren* over underspecified *brothers*, which would be the result of applying the regular plural meaning to the meaning in (5) and not choosing the option in (6).

The upshot, then, is that in a context where *brethren* can be used, *brothers* can only be used a) with the same meaning as *brethren*, but due to a contextually specified R — in that case $\textbf{MostInformative}_s$ would again return $\perp$, since the constraint is based on *proper* subsets and the two sets are equal in this case; or b) *brothers* must have some contextually available meaning that is not a proper superset of the meaning of *brethren*.

We make a correct prediction about morphosemantics here. *Brothers* can be used with the same meaning as *brethren* when the meaning is contextually available, as when a monk might equivalently say (9) or (10).

(9) My brethren will make sure you are comfortable.

(10) My brothers will make sure you are comfortable.

However, the latter utterance could instead have other contextual meanings. Thus, if the monk wished to communicate *specifically* that the members of the order will ensure the addressee's comfort, *brethren* would be a better choice than the more ambiguous *brothers*, because *brethren* has a more specific meaning.

References

Asudeh, Ash, Paul B. Melchin, and Daniel Siddiqi. 2021. Constraints all the way down: DM in a representational model of grammar. In *Proceedings of WCCFL 39*.

Dalrymple, Mary. 2015. Morphology in the LFG architecture. In *Proceedings of the LFG15 Conference*.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, eds. 1995. *Formal issues in LFG*.

Dalrymple, Mary, John J. Lowe, and Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20*.

Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In *Linguistic theory and computer applications*.

—. 1989. The formal architecture of Lexical-Functional Grammar. *Journal of Information Science and Engineering* 5.Revised version published in Dalrymple et al. (1995: 7–27).

—. 1995. The formal architecture of Lexical-Functional Grammar. In Dalrymple et al. (1995: 7–27). Earlier version published as Kaplan (1989).

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use, and interpretation of language*.

Melchin, Paul B., Ash Asudeh, and Dan Siddiqi. 2020. Ojibwe agreement in Lexical-Realizational Functional Grammar. In *Proceedings of the LFG20 conference*.

Partee, Barbara H., and Vladimir Borschev. 2003. Genitives, relational nouns, and argument-modifier ambiguity. In *Modifying adjuncts*, vol. 4 of *Interface Explorations*.

Thomas, Amanda. 2021. Divergence and mismatches in Portuguese morphology and syntax. D.Phil. thesis, University of Oxford.