DM + LFG = L_RFG

Ash Asudeh[‡†] · Paul B. Melchin[†] · Dan Siddiqi[†]
[†] School of Linguistics and Language Studies, Carleton University
[‡] Department of Linguistics, University of Rochester

Department of Linguistics
University of Arizona
December 4, 2020[*]

1 Our project

- We are developing a theoretical framework that couples Lexical-Functional Grammar (LFG; Bresnan et al. 2016) with the realizational, morpheme-based approach to word-formation of Distributed Morphology (DM; Halle and Marantz 1993).
- The resulting framework, which we call Lexical-Realizational Functional Grammar (L_RFG), is particularly well-suited to model North American Indigenous languages, which are characterized by *polysynthesis* and *nonconfigurationality*.
- In this talk we will summarize the framework, and demonstrate it with an analysis of Anishinaabemowin/Ojibwe inflection.
- The talk will proceed as follows:
  - ☉ Section 2 outlines the L_RFG framework, comparing and contrasting it to standard DM and standard LFG, and providing details on the exponence function.
  - ○ Section 3 provides a brief introduction to Ojibwe, and a background on relevant aspects of the language's morphosyntax.
  - ○ Section 4 provides a demonstration of our analysis, including the structures of a representative example sentence, as well as presentation and discussion of the templates used and specifications of the Vocabulary Items needed for animate agreement in Ojibwe and for the examples in the handout.
  - ○ Section 5 concludes.
  - ○ Section 6 outlines where we are going next with this project.
  - ○ The two appendices provide structures for additional example sentences, demonstrating most of the Ojibwe agreement morphology under discussion (Appendix A) as well as providing additional discussion of conjunct-order agreement (Appendix B).

[*]This research was supported by SSHRC Insight Development Grant 430-2018-00957 (Siddiqi/Asudeh).

2 The framework

2.1 Motivation

- L_RFG is the offspring of an unlikely marriage between Distributed Morphology as a theory of morphological realization and Lexical-Functional Grammar as a constraint-based theory of syntax and grammatical architecture.
- L_RFG combines the strengths of the two frameworks:
  1. Like LFG, it is a declarative, representational and constraint-based theory that is ideally suited to modelling nonconfigurationality.
  2. Like DM, it provides a realizational, morpheme-based view of word-formation and is good at modelling complex morphological structures including those found in polysynthetic languages, such as many North American Indigenous languages.
- Additionally, because the realizational module, v(ocabulary)-structure, has access to prosodic structure, LrFG has the potential to give non-transderivational (computationally simpler) prosodic explanations for morpheme alignment and surface form phenomena.

2.2 L_RFG and DM

2.2.1 What is DM?

- DM is a framework for morphological theory.
- Like any morphological framework, it assumes an interface with a syntactic module.
- However, unlike other realizational models, DM has always assumed a particular syntactic framework for providing those structures: standard Chomskyan syntax (here called Minimalism for short, even though DM strictly predates Minimalism).
- Given this consistent co-occurrence, it can be hard to easily identify where Minimalism ends and where DM begins.
  - For example, the Y-model is a feature of Minimalism, as is the fact that head-movement is in the PF branch of that Y.
- So what is the essence of DM, without Minimalist or Minimalist-like assumptions about syntax?
  1. **Morpheme-based morphosyntax**
     - In a morpheme-based model of morphology, complex words have internal structure (we set aside here the completely confusing use of "morpheme" in the DM literature and adopt the standard definition of the morphemic hypothesis) and morphology is (often strictly) concatenative.
     - This means that words are not atomic in DM, and paradigms are epiphenomenal.
     - (Stump 2001) calls this property of morphological theories *lexical* (yes, that is also confusing).
     - Parsimony dictates that there should not two equivalent concatenative modules (morphology and syntax). Rather there should be just one (see, for example, Lieber 1992).
     - This property of DM is often called "syntax all the way down" in the DM literature.

2. **Realization**
   - A *realizational* model of morphology (Beard 1995) assumes morphology expresses rather than generates.
   - *Incremental* approaches (using Stump's opposition) to morphology assume that morphology is information-adding.
   - Realizational approaches typically assume completed featural structures as the input to morphological operations.
   - The DM literature uniquely calls this property *late-insertion*, but in reality most contemporary models of morphology are realizational (see Siddiqi and Harley 2016, editors' notes, for discussion).
   - Note that this property, realization, and the previous property, morpheme-based morphosyntax, are not linked properties.
     - For example, A-morphous Morphology (Anderson 1992) and Paradigm Function Grammar (Stump 2001) are word and paradigm realizational models.
     - In Stump's (2001) classification, such models are *Inferential-Realizational*.
     - DM is in contrast *Lexical-Realizational*.
     - Nanosyntax (Starke 2009) is another Lexical-Realizational model of morphology.
3. **Morphology as an interface (no WFRs)**
   - In contemporary DM, the ideal appears to be that morphology is an interface rather than a separate generative component of the grammar.

     > "Morphology has no proprietary categories, but deals only in morphs, understood as pieces of phonological material lexically specified with instructions for their use as exponents of syntactic properties." (Bermúdez-Otero and Luís 2016: 311)
   - This is most relevant in DM's rejection of the *morphome* (Aronoff 1994), the word, and the paradigm.
   - In the DM literature, this property is typically called *non-lexicalist* because the model of syntax that results from an interface with DM necessarily rejects the *Lexicalist Hypothesis/Lexical Integrity Hypothesis* (Chomsky 1970, Lapointe 1980).
   - It is worth a moment to acknowledge that this is a controversial view of DM.
     - In Halle and Marantz (1993), M-structure clearly had proprietary rules.
     - Indeed, a view of DM in the literature today is that the post-spellout, PF branch rules, (such as *Local Dislocation*, *Rebracketing*, *Impoverishment*, *Fusion*, *Fission*, *Enrichment*, and *Readjustment*) constitute a morphological component of the grammar, i.e. a Lexicon.
     - Similarly, a version of Minimalist Syntax that does not assume a DM interface, such as those that appear in most syntax textbooks, does not assume any of these mechanisms.
   - However, it is very important for the present discussion that all these rules serve to address syntax-morphology mismatches.
     - We assume here that such mechanisms are not properties of DM, per se, but rather are properties of the DM-Minimalism interface, which is why they are typically referred to as PF operations.

4. **Three lists**
   - Perhaps the most salient feature of DM that makes it distinct from other models of morphology is that it gave up on the claim that syntactic, semantic, and phonological domains of "word" (or morpheme) align on the same domain (see Marantz 1997 for discussion).
   - Indeed, Distributed Morphology gets its name from this property.
   - Instead it created three distinct lists, one for each domain.
     (a) The *Vocabulary* is the list for phonological properties.
     (b) The *Encyclopedia* is the list for semantic properties.
     (c) The third list is frustratingly without a name agreed upon in the literature.
       – Its members are the formal features that populate syntactic structures (these are the things called "morphemes" by some practitioners of DM).
5. **Elsewhere Principle**
   - DM employs a common type of linguistic rule called the *Elsewhere Principle* (also called the *Paninian Principle*), which is almost standard in morphological theory and dates back to the 70s, when formal morphological theory saw it nascence (see for example Anderson 1969, Kiparsky 1973, and Aronoff 1976).
   - The classic view of the elsewhere condition is seen in Aronoff (1976): a morphological rule with more specific conditions is applied before a general rule.
   - This has long been the main means through which irregular rules block the application of regular rules, for example.
   - In contemporary realizational models, this blocking is achieved through competition, where a more specific form outcompetes a more general form.
   - In the DM literature, this is called the *Subset Principle*.
6. **Underspecification**
   - A fundamental challenge of any morphological theory is syncretism or polysemy.
   - This is the simple phenomenon that morphological forms appear in many different environments, often with slightly different meanings.
   - These are a violation of bi-uniqueness (a unique form maps to a unique meaning; see Harris 2016 for detailed discussion).
   - There are many ways that morphological models can deal with this, such as rules of referral.
   - DM, like many other models, capture this by assuming that a morphological form underdetermines its syntactic and semantic properties.

⇒ In sum, for us, DM without a Minimalism interface is precisely the combination of those six properties, listed here:

1. Morpheme-based morphosyntax (syntax all the way down)
2. Realization (Late Insertion)
3. Morphology as an interface (non-lexicalist)
4. Three lists ("Distributed" Morphology)
5. Elsewhere Principle (Subset Principle)
6. Underspecification

2.2.2 Comparison of L_RFG with standard DM

- DM in L_RFG form is very similar to DM with a Minimalist syntax (DMM), with the key difference that it assumes an interface with LFG as a model of syntax (discussed below).
- How does this make L_RFG different from DMM?

1. L_RFG is a non-derivational, constraint-based model of the grammar.
   - Distributed Morphology is a realizational model of morphology.
   - Conceptually, realizational morphology is akin to harmonic approaches to phonology (such as Optimality Theory; Prince and Smolensky 1993, 2004).
     - The task is to identify the surface representation that best realizes the featural content of a underlying form that has been constrained by certain well-formedness conditions.
     - Indeed, Vocabulary Items themselves, along with the Subset Principle, are the well-formedness conditions that must be satisfied in order to satisfy a legal surface representation.
     - In this way, realizational morphology is inherently non-derivational.
     - Its opposite, incremental morphology, can be derivational.
   - As a model of morphology, aside from the fact that insertion is cyclic in some varieties of DM, there is nothing derivational at all about DM.
   - Setting aside mechanisms such as *Readjustment* which are not discussed here, the six core principles of DM, as described above, describe a model of grammar that assesses the well-formedness of a surface representation (*Vocabulary Insertion*) against the final output of PF-branch operations (at least on a phase by phase basis).
   - Intuitively, a model that assesses the welformedness of representations is better suited to be interfaced to other models that assess the wellformedness of representations.

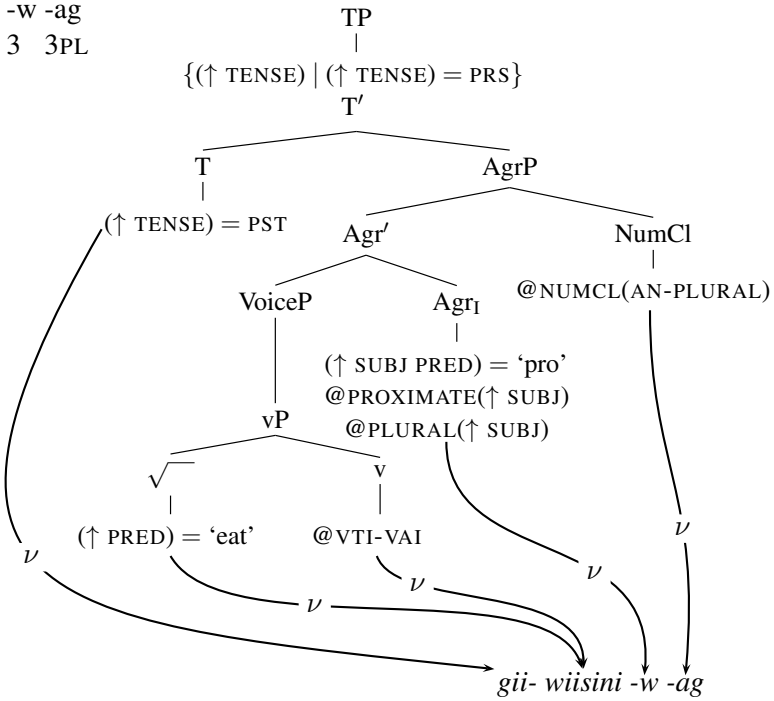
   ⇒ LFG is that. Minimalism is not.
2. L_RFG allows for exponence to be subject to dependencies on several different modules.
   - It is well-known that affixes (and other morphological processes) are not only subject to (morpho)syntactic conditions.
   - Affixation is conditioned by semantics (see, for example, the semantic restrictions *re-* requires of its base) and phonology (see, for example, the phonological restrictions the comparative *-er* and the deadjectivizer *-en* require of their bases).
   - L_RFG is able to capture all three of these types of conditioning on morphological processes precisely because the morphological representation (v(ocabulary)-structure) imposes constraints on the mappings (either directly or indirectly) to not only c-structure, but f-structure, s(emantic)-structure, and p(rosodic)-structure.
   - In contrast, PF in DMM is explicitly blind to LF in the Y model, so meaning directly affecting form (such as the difference between *brothers* and *brethren* or *older* and *elder*) is excluded in DMM.
     - Additionally, surface phonology is ordered after insertion is complete, so output-sensitive morphology (such as the legality of *hasten*, see Halle 1973 for discussion) is difficult or even impossible to obtain absent a DM-OT interface such as proposed by Bye and Svenonius (2012).

2.2.3 L_RFG as a daughter framework of DM

- L_RFG is a variety of DM, despite the different syntax interface, so L_RFG maintains all the key properties of DM.

1. **Morpheme-based morphosyntax**
   - L_RFG directly adopts the *monolistemicity* and *spanning* model of Vocabulary Items developed for DM in Haugen and Siddiqi (2016).
   - Haugen and Siddiqi's model of the vocabulary is neither purely morpheme-based nor word-based, but rather is listeme-based.
   - In L_RFG, the key property for determining what is a Vocabulary Item is not decomposability, as is true in standard DM, but rather listedness.
   - While *Spanning* is not standard in DM, it is definitely part of the DM literature.
   - Spanning is crucial to L_RFG, rather than optional, but otherwise L_RFG's view on morphemes and syntactic structure is virtually the same as in DM.
   - Indeed, L_RFG c-structures are largely the same as syntactic trees found in DM outside of the featural content.

(1) gii- wiisini -w -ag
PST see.VAI 3 3PL
'They ate.'



2. **Realization**
   - Exponence in L_RFG works almost identically to Vocabulary Insertion in DM.
   - The crucial difference is that a Vocabulary Item in L_RFG is a more complicated representation than that of DMM as it also contains information relevant to prosodic structure constraints.
   - Exponence in L_RFG is also sensitive to more information than in DMM: it is conditioned also by *meaning constructors* from Glue Semantics (Dalrymple 1999, 2001, Dalrymple et al. 2019, Asudeh 2012) and by f-structures.
   - Finally, exponence in L_RFG is also not a replacement algorithm that discharges features from a derivation.
     - In L_RFG, it is a set of pairwise correspondence functions between representations in v-structure, c-structure, f-structure, and p-structure.
3. **Morphology as an interface**
   - In L_RFG, v-structure is quintessentially non-generative.
   - While DMM has various operations that change the syntax along the PF branch, L_RFG has no such operations.
   - The form of v-structure is entirely determined by the satisfaction of constraints on the mappings with other representations.
   - Morphology is not an output of L_RFG: it is one of many representations described by a given co-description.
   - Additionally, like DM, L_RFG rejects the part of the *Lexical Integrity Hypothesis* that mandates that complex words map to syntactic terminals.
4. **Three lists**
   - L_RFG maintains the tripartite division of wordhood that defines DM.
   - Indeed, L_RFG adds a fourth "special domain" in the sense of Marantz (1997): L_RFG distinguishes between morphological (vocabulary) atomicity and phonological (prosodic) atomicity.
   - In L_RFG, morphological atomicity, phonological atomicity, semantic atomicity, and semantic atomicity do not necessarily align on the same object. Each corresponds to a different representation in the Correspondence Architecture, as described by co-description.
5. **Elsewhere Principle**
   - L_RFG adopts this, though not directly through adopting the Subset Principle of DM.
   - In L_RFG, this falls out of two independently motivated elsewhere constraints, *MostInformative* and *MostSpecific*, where MostInformative is conditioned by meaning and MostSpecific is conditioned by form.
6. **Underspecification**: Yup.

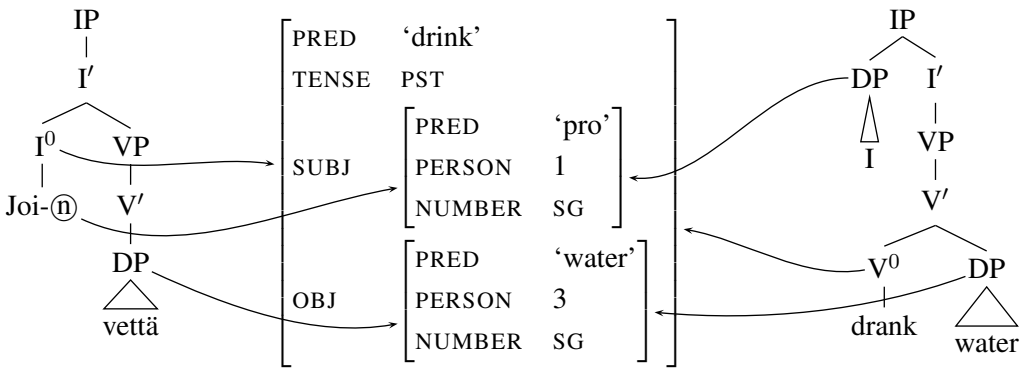
2.3 L_RFG and LFG

2.3.1 What is LFG?

- LFG is a generative theory of grammar that has a number of salient features.

1. The syntactic module is split into sub-modules: c(onstituent)-structure and f(unctional)-structure (Kaplan and Bresnan 1982).
    - A c-structure is a standard phrase structure tree.
    - C-structure represents syntactic information about precedence, dominance, and constituency.
    - An f-structure is a *feature structure*, also known as an *attribute-value matrix*).
    - F-structure represents yet more abstract aspects of syntax, such as grammatical functions, predication, subcategorization, and local and non-local dependencies.
2. The two modules are present simultaneously/in parallel and are related to each other by the the ϕ correspondence function (a.k.a. projection function), which maps c-structure to f-structure.
    - In other words, c-structure elements — nodes in a tree — are mapped to f-structure elements — feature structures — by the ϕ correspondence function.
    - The ϕ function obviously does not map any c-structure node to more than one f-structure node (by the definition of function), but it is many-to-one: multiple c-structure nodes can be mapped to a single f-structure.
3. The ϕ correspondence function thus models feature propagation through a set of equations that model identity between the f-structures of distinct c-structure nodes. It permits information (encoded in features) to propagate between syntactic nodes without the need for movement or feature-checking.
    - For example, a verb and its higher functional projections map to the same f-structure.
        - Thus, the feature requirements of the verb propagate to the clause it heads.
        - This allows head movement to be modelled as categorial ambiguity.
        - Languages that realize finite verbs high in their clause are analyzed as having verbs that are lexically specified as having a functional category and are therefore base-generated higher in the clausal structure.
4. The ϕ-equations — where a set of such equations is typically called an *f-description* (because, at least originally, they described an f-structure) — similarly allow non-local constituents in c-structure to be related to each other through f-structure, which obviates the need for movement.
    - For example, a *wh*-phrase high in the c-structure can be related to the predicate that selects for it lower in the c-structure by stating that the f-structure of the *wh*-phrase is identical to the f-structure of the missing argument.
5. Typologically, c-structure generally models how languages/grammars can *differ*, whereas f-structure models how they are *similar*.
    - For example, consider this simplified case of equivalent f-structures for c-structurally quite distinct English and Finnish expressions of the same proposition (Asudeh and Toivonen 2015: 28, (67)):

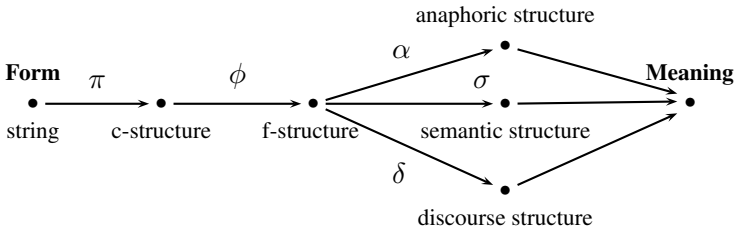
(2)



- Notice that the simplified c-structure above is not annotated with f-descriptions; more on this below.

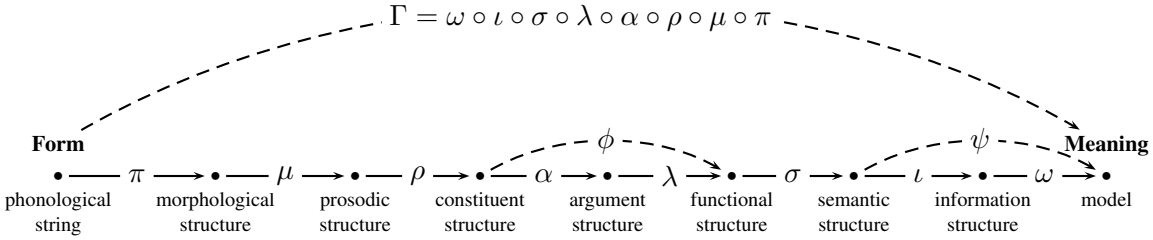
- Kaplan (1987, 1989) subsequently proposed that the correspondence between c-structure and f-structure, $C \xrightarrow{\phi} F$, could be generalized to a full *Correspondence Architecture* (or *Parallel Projection Architecture*), offering the programmatic suggestion in (3). .

(3)



- The actual Correspondence Architecture that emerged over the next couple of decades was the following (Asudeh 2012: 53).

(4)



⇒ LFG now has a modular, Correspondence Architecture:

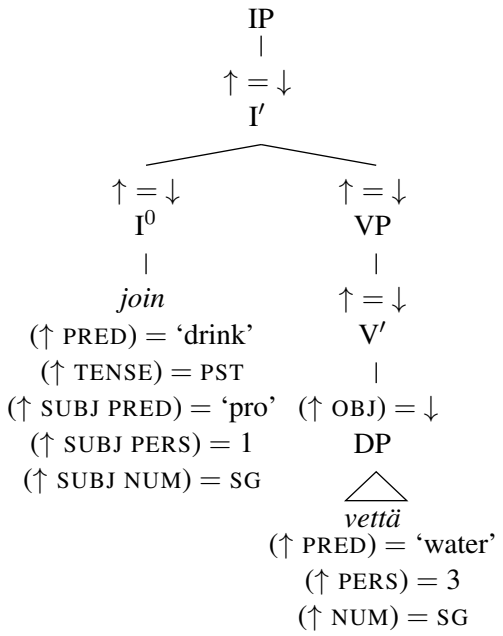
1. Each module models some aspect of the language system as a discrete subsystem, using data structures that are appropriate for modelling that aspect of linguistic information (e.g., non-tangled trees for c-structure, an attribute-value matrix for f-structure).
2. Modules are related to each other by correspondence functions.
3. The mappings between modules easily allow for mismatches between information in the mapped modules (e.g., a pro-drop subject is modelled in f-structure, but there is no corresponding null element in c-structure).
4. The correspondence functions allow a formative to contribute simultaneously to multiple modules.
   - This is called *co-description*.
5. *The modularity of the system allows the modules to be re-arranged and the mappings redefined, while allowing the information inside the module to be modelled as before.*

- Let's return to the Finnish example in (2) above to briefly see how an LFG analysis works.

(5) Join vettä
drink.1SG.PST water
'I drank water.'

  - The c-structure is generated by a set of phrase structure rules that are annotated by f-descriptions; we won't show those here.
  - This results in the following annotated version of the c-structure in (2) above.

(6)



- The information under the terminal nodes, *join* and *vettä* is from the lexical entries, where a lexical entry consists of a triple of a form, a category, and an f-description.
- The f-description is resolved by assigning arbitrary unique f-structure variables to all the arrows as follows:

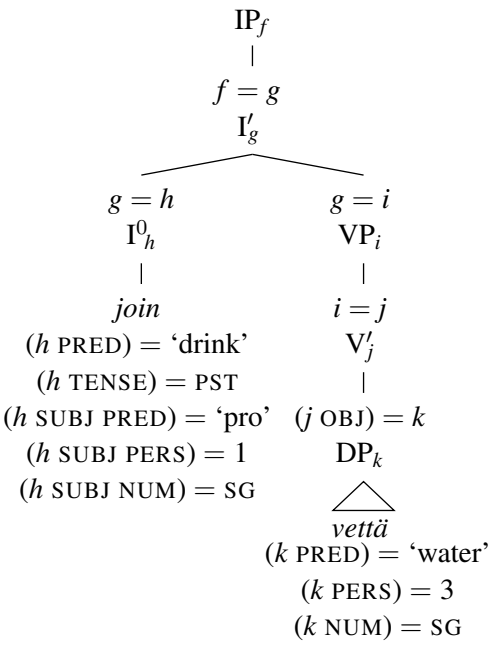
  ↑ := the f-structure of my mother

  ↓ := my f-structure

- This is why the up and down arrows are called *metavariables* in LFG.

- The c-structure in (6) can be assigned variables as follows:

(7)



- These equations are solved and the minimal f-structure that satisfies them — where minimality is a condition on the solution — is the one in (2), repeated here with the labelling from (7).

(8)

$$\begin{array}{l} \\ f \\ g \\ h \\ i \\ j \\ \\ \end{array} \begin{bmatrix} \text{PRED} & \text{'drink'} & \\ \text{TENSE} & \text{PST} & \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{PERSON} & 1 \\ \text{NUMBER} & \text{SG} \end{bmatrix} & \\ \text{OBJ} & k \begin{bmatrix} \text{PRED} & \text{'water'} \\ \text{PERSON} & 3 \\ \text{NUMBER} & \text{SG} \end{bmatrix} & \end{bmatrix}$$

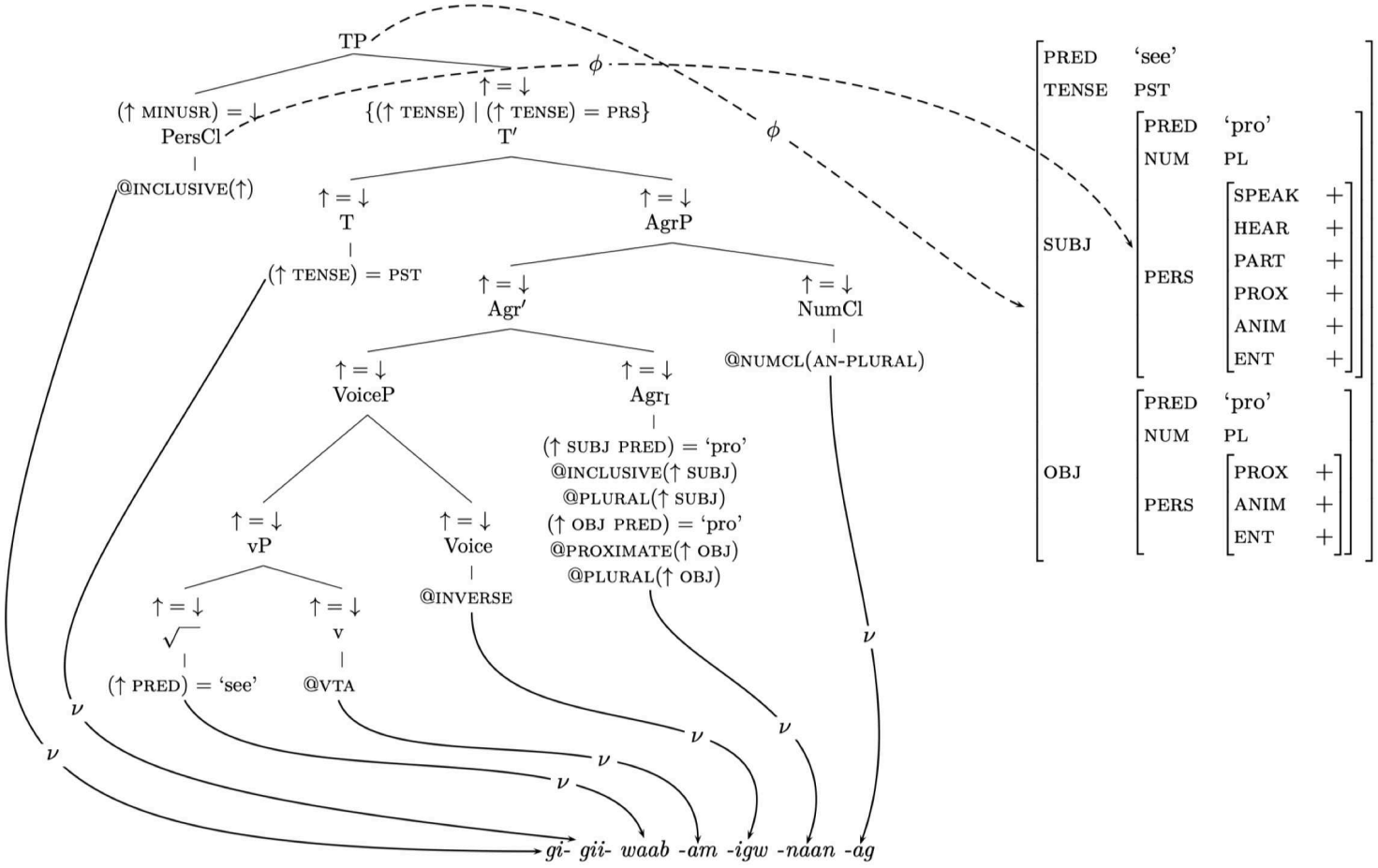
2.4 Comparison of L_RFG with standard LFG

- L_RFG is similar to standard LFG, with changes to the c-structure and its relationship with morphosyntactic elements.
- The terminal nodes of c-structures *are not words*, but instead are f-descriptions (sets of f-structure equations and constraints)
- The c-structure is mapped to a v(ocabulary)-structure, a linearized structure in which vocabulary items (VIs) *expone* (i.e., realize) the features in the terminal nodes, via a correspondence function, ν.
- Formally, v-structure is a list, each member of which is a feature structure defining morphophonological properties relevant to the linear placement and metrical properties of the item.
  - This includes the phonemes/segments, as well as the metrical frame which determines syllable structure, affix/clitic status, and so on.
  - Thus, the v-structure roughly corresponds to the p(honological)-form portion of a lexical entry in the metrical theory of Bögel (2015).[1]
- In this talk, only the strings themselves are relevant, so we make some simplifying assumptions:
  1. We represent the output of the exponence function, ν, simply as a string, not a full VI structure.
  2. We show alignment informally using the standard notational convention of adding a dash to the left or right of the string.
  3. We do not show the $o \circ \rho$-mapping (see Figure 1 below), but instead let the phonological forms stand in for the VI strings (i.e., we conflate the two for simplicity/presentational purposes).
- In sum, vocabulary structure is a morphophonological structure that maps to phonological form via prosodic structure.

[1] We would like to thank Tina Bögel for her insightful comments on this point at the LFG20 conference, and in extensive discussion afterwards. The details of the interaction between v-structure and the phonological string, in particular the effects of the metrical properties of VIs on mismatches in ordering between c-structure and the p-string, are currently being worked out and will be presented in future work in the L_RFG framework.

- Here is an example from Ojibwe (details to follow in Section 4) to demonstrate the basics of an L_RFG analysis.

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



- We complete the v-structure mappings by introducing a new phonological correspondence function, o, which maps from prosodic structure to phonological strings, and treating the ρ mapping as a mapping from vocabulary items to prosodic structures.
- In other words, the output of ρ is the prosodic structure and the output of o is the final result of phonological processes, a set of strings that are based on the prosodic well-formedness conditions of VIs.
- The morphology is responsible for the input to phonology, but phonology does whatever phonology does to create the output, which is not part of morphology per se.
- Given the set of VIs, V, and a set of prosodic structures, P:

(10) $\rho : V \rightarrow P$

- The o correspondence function takes the output of this ρ correspondence function as its input and so maps to the phonological string (o's output) from the prosodic structure that corresponds to the vocabulary item.

- Thus, in this framework, v-structure precedes the phonological string in the Correspondence Architecture (see, e.g., Asudeh 2012: 53), resulting in the revised architecture in Figure 1.
- The output of the grammar, $\langle \Gamma_1, \Gamma_2 \rangle$, for any particular set of input formatives, is a form–meaning pair where the form incorporates prosody (still fed by constituent structure) and the meaning incorporates information structure (still fed by semantic structure).[2]

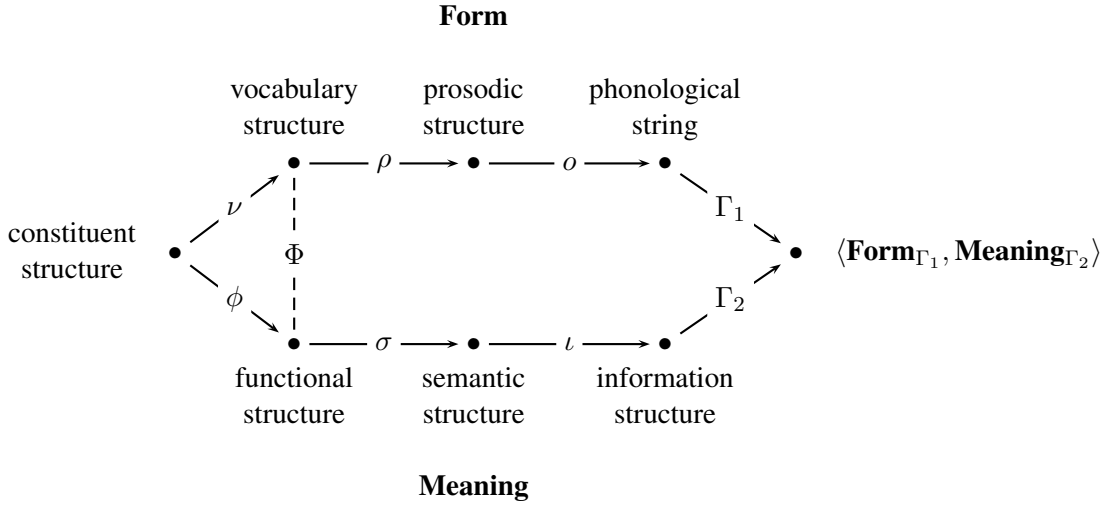


Figure 1: L_RFG's Correspondence Architecture

- We outlined above how, on the form side, information flows through vocabulary structure to prosodic structure to the final phonological string.
- On the meaning side, information flows through f-structure to semantic structure to information structure, which is the final structural aspect of meaning.
  - This is accomplished by adopting Glue Semantics (Dalrymple 1999, 2001, Dalrymple et al. 2019, Asudeh 2012), which allows us to list Glue *meaning constructors* in the f-description.
  - In other words, there is another aspect to our Vocabulary Items below, a third coordinate for meaning, modelled as sets of meaning constructors. We do not show this dimension below, partly for simplicity, and partly because it is work in progress.
- The relationship between terminal nodes and VIs is many-to-one, using the mechanism of *Spanning* (Haugen and Siddiqi 2016, Merchant 2015, Ramchand 2008, Svenonius 2016); i.e. one VI may realize features of multiple terminal nodes.
- The result is similar to the Lexical Sharing model proposed for LFG by Wescoat (2002, 2005, 2007), but maintains the complex internal structures of words as part of syntax.

[2]Note that the *set* of all grammatical form-meaning pairs may have a given form recurring in several pairs, if it is ambiguous, or a given meaning recurring in several pairs, if it is expressible in alternative ways.

2.4.1 L_RFG as a daughter framework of LFG

- The obvious point of contrast between L_RFG and LFG concerns the Lexicalist Hypothesis (Chomsky 1970, Lapointe 1980):

  (11) *Lexicalist Hypothesis*
  No syntactic rule can refer to elements of morphological structure. (Lapointe 1980: 8)

- In LFG this is captured in the *Lexical Integrity Principle*, through formulations like the following:

  (12) *Lexical Integrity*
  Morphologically complete words are leaves of the c-structure tree, and each leaf corresponds to one and only one c-structure node. (Bresnan et al. 2016: 92)

- This statement has two parts:

  1. L_RFG *upholds* the part that states that "each leaf corresponds to one and only one c-structure node".
     - This may contrast with Lexical Sharing (Wescoat 2002, 2005, 2007), in which portmanteau forms like *du* ('of.DEF.MASC.SG') in French appear to correspond to more than one c-structure node. We need to look under the hood carefully, though, to see what the formal definition of Lexical Sharing is rather than simply going by its graphical representation, which may be misleading. We haven't done this work yet.
  2. L_RFG *rejects* the part that states that "morphologically complete words are leaves of the c-structure tree".
     - Clearly, the c-structure leaves/terminals in L_RFG are not "morphologically complete words". The c-structure leaves/terminals are feature bundles that *map* to form, but the form itself is not part of the terminal node.

- However, notice that the notion *morphologically complete word* is left unanalyzed in the definition in (12).

- In fact, it is far from clear that "morphologically complete word" is a coherent notion (see, for example, Anderson 1982).

  - The essential problem is that there are multiple relevant notions of wordhood, and they don't align on a single type of object that we can point to and unambiguously and confidently call a word (Di Sciullo and Williams 1987).[3] In fact, there can be mismatches between the phonological, syntactic, and semantic aspects of words (Marantz 1997).

  1. Portmanteau words are examples of things that are phonologically simple but semantically and syntactically complex.

     (13) Tu bois **du** lait. French
     you drink of.DEF.MASC.SG lait
     'You drink/are drinking milk.'

     (14) **Imma** go. English dialect
     1SG.FUT.PROX go
     'I'm about to go.'

[3] This is a long and broad discussion that we cannot possibly do justice to here.

2. Idiomatic expressions are phonologically and syntactically complex, but not necessarily semantically complex, and never in a way that maps entirely transparently to their phonology and syntax.

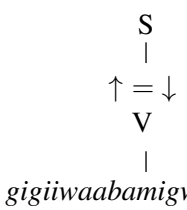
   (15) I read **the shit out of** this book.
   INTENSIFIER
   'I thoroughly read this book.'

3. Units of syntax can be phonologically or semantically dependent on their contexts.

   (16) Je **l'**ai vu. French clitic
   I 3SG.saw
   'I saw it.'

   (17) **The cat**'s been let out of the bag.

- L_RFG thus countenances three criteria for wordhood:
  1. A word as an unanalyzed phonological string (phonological criterion)
  2. A word as a lexicalized string with a non-compositional meaning (semantic criterion)
  3. A word as a syntactic atom (syntactic criterion)
- L_RFG thus assumes that there are three notions of wordhood that sometimes happen to align, but can diverge, i.e., there are mismatches between the three types of wordhood.
- With its focus on mismatches, L_RFG is therefore strongly in the spirit of LFG.
  - L_RFG uses the standard *co-description* mechanism of LFG (for recent exposition, see Dalrymple et al. 2019) to simultaneously state the phonological, syntactic and semantic aspects of formatives.
- Here are some possible points of comfort for an LFGer gazing on L_RFG's familiar yet alien landscape:
  1. L_RFG could be considered to be offering a morphological theory for LFG that had previously been captured by somewhat ad hoc devices like phrase structure rules for word formation; see, e.g., the discussions of Japanese and West Greenlandic in Bresnan et al. (2016). In other words, LFG owes some kind of theory of word structure, which has generally been lacking until recently (see, e.g., Dalrymple 2015, Dalrymple et al. 2019), and L_RFG seeks to pay that debt.
  2. The Vocabulary Items of L_RFG contain much the same information as LFG's lexical entries, but without the commitment that morphophonological form is bundled as part of the lexical entry. It should be easy to specify an algorithm for translating L_RFG's VIs into LFG lexical entries.
  3. Related to the first two points, if one were to want to maintain some version of the Lexicalist Hypothesis, one could view L_RFG as offering a microscopic view of the structure of "words", in particular major categories like verb and noun. For example, the TP node in (9) in some sense *is* the verb, but the L_RFG c-structure shows its internal structure. A standard LFG c-structure for example (9) would instead look like the following (setting the f-description aside). Isn't it normal science to examine things at a smaller and smaller scale?

   (18)
   

2.5 L_RFG's exponence function: ν

- The exponence function ν maps from a pair of arguments to a VI, the exponence of the arguments.
  - The first argument is a list of pre-terminal categories, typically of length 1, which are taken in the linear order they appear in the tree.
  - The second argument is itself a function, Φ, which maps an f-description to the set of f-structures that satisfy the description; i.e. $\Phi(d \in D) = \{f \in F \mid f \models d\}$, where D is the set of valid f-descriptions and F is the set of f-structures.[4]
  - In sum, ν maps **from** a pair whose first argument is a list of c-structure pre-terminal categories and whose second argument is a set of f-structures **to** a structured expression as described above.
- **Conditions on exponence:**
  - Let V be the range of the exponence function ν, the set of VIs (structured expressions); then the following condition on exponence holds.[5]

    (19) Given $\alpha \in A$ and $\beta \in B$, where $A, B \subseteq V$, and a function $[\![\ ]\!]_p$ that returns the conventionalized presuppositions of a given expression,
    If $\bigcup_{a \in A} [\![a]\!]_p = \bigcup_{b \in B} [\![b]\!]_p$
    Then **MostInformative**(α, β)

    – The conventionalized presuppositions of an expression are the set of presuppositions lexically triggered by the expression (Keenan 1971, Beaver 2001, Beaver and Geurts 2014). Presuppositions are propositions. Propositions are sets of possible worlds. Therefore, $[\![\ ]\!]_p$ returns a set of sets of possible worlds.
    – The antecedent of the conditional in (19) therefore collects the conventionalized presuppositions of its arguments in two sets and tests whether the sets are equal.
    – **MostInformative**(α, β) returns whichever of α,β has the most specific f-structure in the set of f-structures returned by Φ applied to the unions of α/β's collected f-descriptions. Formally:

$$\textbf{MostInformative}(\alpha, \beta) = \begin{cases} \alpha \textbf{ if } \exists f \forall g. f \in \pi_2(\nu^{-1}(\alpha)) \wedge g \in \pi_2(\nu^{-1}(\beta)) \wedge g \sqsubset f \\ \beta \textbf{ if } \exists f \forall g. f \in \pi_2(\nu^{-1}(\beta)) \wedge g \in \pi_2(\nu^{-1}(\alpha)) \wedge g \sqsubset f \\ \bot \textbf{ otherwise} \end{cases}$$

    – Thus, the condition in (19) amounts to a combination of the elsewhere condition/subset principle and an economy constraint that enforces spanning when possible.

[4] We thank Ron Kaplan (p.c.) for discussion of this point. Any remaining errors are our own.

[5] One difference between our proposal and the Lexical Sharing of Wescoat (2002, 2005, 2007) is the notion, which we'll call *Pac-Man Spanning*, that VIs can span any number of adjacent preterminal nodes, so long as the presuppositions of the exponed expressions are held constant.

3 Ojibwe: Background

3.1 Why look at Ojibwe?

- Ojibwe exhibits many of the features that we hope to be able to model:
  - Nonconfigurationality – word order is very free (i.e., determined by discourse and pragmatic, rather than syntactic, factors)[6]
  - Polysynthesis – complex verb morphology with extensive head-marking
  - A direct-inverse-based agreement system cross-referencing all core arguments
  - Various morphological processes, including verbal reflexives, noun incorporation, applicatives, various kinds of (anti)passives, and more

3.2 Ojibwe primer: Prominence, animacy and obviation

- Ojibwe grammar has many features that are mostly shared with the other Algonquian languages, but some of which are fairly uncommon outside the family:
  - Typical **polysynthetic** morphysyntactic features, including nonconfigurationality, extensive head-marking, and various kinds of incorporation
  - Agreement morphology determined by a **prominence hierarchy**, which involves:
    - A system of grammatical gender based on **animacy**
    - A system of **obviation** distinguishing clause-mate third-person animate arguments
  - A **direct-inverse** system that indicates the relationship between thematic roles and the person hierarchy
  - Two separate inflectional paradigms: **independent order**, found in most matrix clauses, and **conjunct order**, found in subordinate clauses and certain matrix clause contexts
  - Separate (derivational) verb classes based on (i) transitivity and (ii) the animacy of the object (if transitive) or subject (if intransitive)
- Some of these properties warrant some further discussion.
- **Animacy:**
  - Ojibwe grammatical gender is based on animacy (**animate** vs. **inanimate**).
  - All nouns referring to notionally/semantically animate entities are grammatically animate; however, notionally inanimate nouns may be of either gender.
  - Animacy (of the subject or object) determines the verb final suffix (i.e., verb class, *v*) that is used, among other things.
- **Obviation:**
  - Obviation distinguishes third-person animate clausemates: in any clause, one third-person animate argument is **proximate**, and the rest are **obviative**.

[6] When we say that Ojibwe is "nonconfigurational", we do not intend to claim that word order is completely free. We are using the term in the LFG sense (Bresnan et al. 2016), meaning that word order and phrase structure are not used to distinguish grammatical functions like subject and object. Instead, word order is determined by a combination of factors, including obviation and information structure; see Dahlstrom (2017) for extensive discussion and references.

- The choice of which argument is proximate is mainly based on (currently poorly-understood) pragmatic/discourse factors.
- Obviation is marked on nouns and is distinguished in verb agreement.
- Obviative nouns are unspecified for number (except in isolated inflectional contexts), and can be interpreted as singular or plural.

- **The prominence/person hierarchy:**
  - The distribution of agreement affixes, and the choice of direct or inverse morphology, is based on arguments' relative positions in a **prominence/person hierarchy**.
  - This ranks arguments in terms of person, obviation and animacy.
  - The hierarchy is as follows (adapted from Valentine 2001: 268; abbreviations largely follow common Algonquianist practice):[7]

    (20) *Prominence Hierarchy*

2	2nd person
1	1st person
3	3rd person animate proximate
3′	3rd person animate obviative
0	3rd person inanimate

    - It should be noted that, while the ranking of 2 above 1 determines the insertion of the person prefix (at least on the view of Rhodes 1994, Rhodes and Valentine 2015, adopted here; see discussion below), there are other areas of the grammar where 1 appears to be ranked above 2, for instance when determining the insertion of certain agreement morphemes, and others where they appear to be equally ranked (see Section 4.2 and Appendix A).

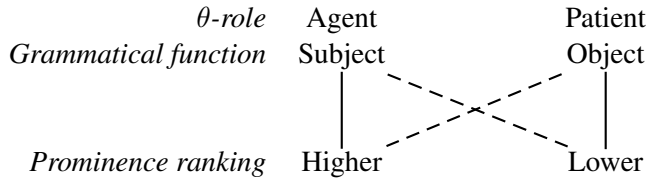
- **Direct/inverse marking:**
  - In transitive clauses, the relationship between the two arguments' relative ranking in the prominence hierarchy and their thematic roles is tracked by the **direct/inverse** morpheme, known as a Theme Sign (analyzed as Voice; e.g., Oxford 2014, 2019):
    - When the agent is the higher-ranked argument and the patient is lower, the verb is marked as **direct**.[8]
    - When the patient is the higher-ranked argument and the agent is lower, the verb is marked as **inverse**.
  - The theoretical status of inversion in Ojibwe is still under debate. One question involves the relationship between inversion and the grammatical functions of subject and object.
  - For some, the agent is always the subject and the patient is always the object (e.g., Valentine 2001, Dahlstrom 2014, Oxford 2019).
    - **Direct:** subject is higher-ranked, object is lower-ranked
    - **Inverse:** subject is lower-ranked, object is higher-ranked

[7] Contra Valentine (2001), we do not include the "unspecified actor" form in the prominence hierarchy; instead, we analyze these forms as instances of a short passive. See Section 4.2.5 for discussion.

[8] Following common practice, we are using the term "agent" to refer to agent-like roles, including causes and many experiencers – i.e., the agent proto-role in the sense of Dowty (1991). Similarly, the term "patient" is used for the proto-role that includes patients, recipients, themes, and so on.

– In the diagram below, the solid lines represent the correspondences in a direct form, and the dashed lines the correspondences in inverse.

(21) **GFs-as-θ-roles analysis**



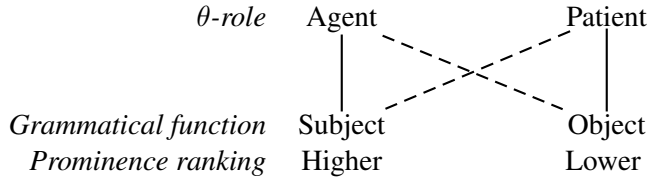
○ For others, the higher-ranked argument is always the subject and the lower-ranked argument is always the object (e.g., Rhodes 1994, 2010).

– **Direct:** subject is agent, object is patient

– **Inverse:** subject is patient, object is agent

– In the diagram below, the solid lines again represent the correspondences in a direct form, and the dashed lines the correspondences in inverse.

(22) **GFs-as-prominence analysis**



○ We adopt the **GFs-as-prominence analysis** (22), where the grammatical functions are defined in terms of the prominence hierarchy.[9]

– This allows us to treat direct/inverse marking as determining the mapping between grammatical functions in the f-structure and thematic argument roles in the semantic structure (Asudeh and Giorgolo 2012, Asudeh et al. 2014).

– It also means that the subject and object have consistent (word-internal) c-structural positions, as with the clausal structure in configurational languages; the alternative would be to have specific positions for the higher and lower arguments, which is more difficult to model.

– See Section 4.2 for a formalization of this analysis.

3.3 Data under consideration

- The data and analysis in this talk is meant to range over and cover the different varieties that linguists consider to be part of the Ojibwe language, including both Nishnaabemwin (such as Odawa) and Anishinaabemowin dialects (such as Southwestern Ojibwe and Algonquin).
- The data are taken mainly from Nichols's (1980) grammar of Southwestern Ojibwe, corroborated with the paradigms in Jones (1977) (Algonquin) and Valentine (2001) (Nishnaabemwin).

[9] While it has been claimed that there is syntactic evidence for the GFs-as-θ-roles analysis (e.g., Dahlstrom 2014, Alsina and Vigo 2017, Oxford 2019), the evidence largely relies on judgements that vary between Algonquian languages, and even between dialects or individual speakers of Ojibwe, as pointed out by Rhodes (1994, 443). It is possible that languages differ as to which is the proper analysis, as is claimed by McGinnis (1999), Alsina and Vigo (2017).

- We include vowels that are omitted in the syncopated (Nishnaabemwin) dialects, and word final /n/, which is often dropped; we are essentially presenting the underlying forms of the morphemes and inflected verbs, though their pronunciation varies widely from one variety to the next.
- Where the inflectional morphemes themselves differ between dialects, we have done our best to present the more conservative forms, consulting the analysis of Proto-Algonquian in Oxford (2014). There are notes on instances of variation in Section 4.

- The current analysis accounts for the full verbal agreement system, including agreement for subjects, primary and secondary objects (SUBJ, OBJ, and OBJ_θ, respectively), both animate and inanimate, in both the independent and conjunct orders.
- We provide the templates (Dalrymple et al. 2004, Asudeh et al. 2013) that are invoked in the analysis, VIs for the set of inflectional morphemes that appear with these verbs, and illustrate by providing c-, f-, and (abbreviated) v-structures for some representative examples.

4 Analysis: Ojibwe inflection

4.1 Example structure

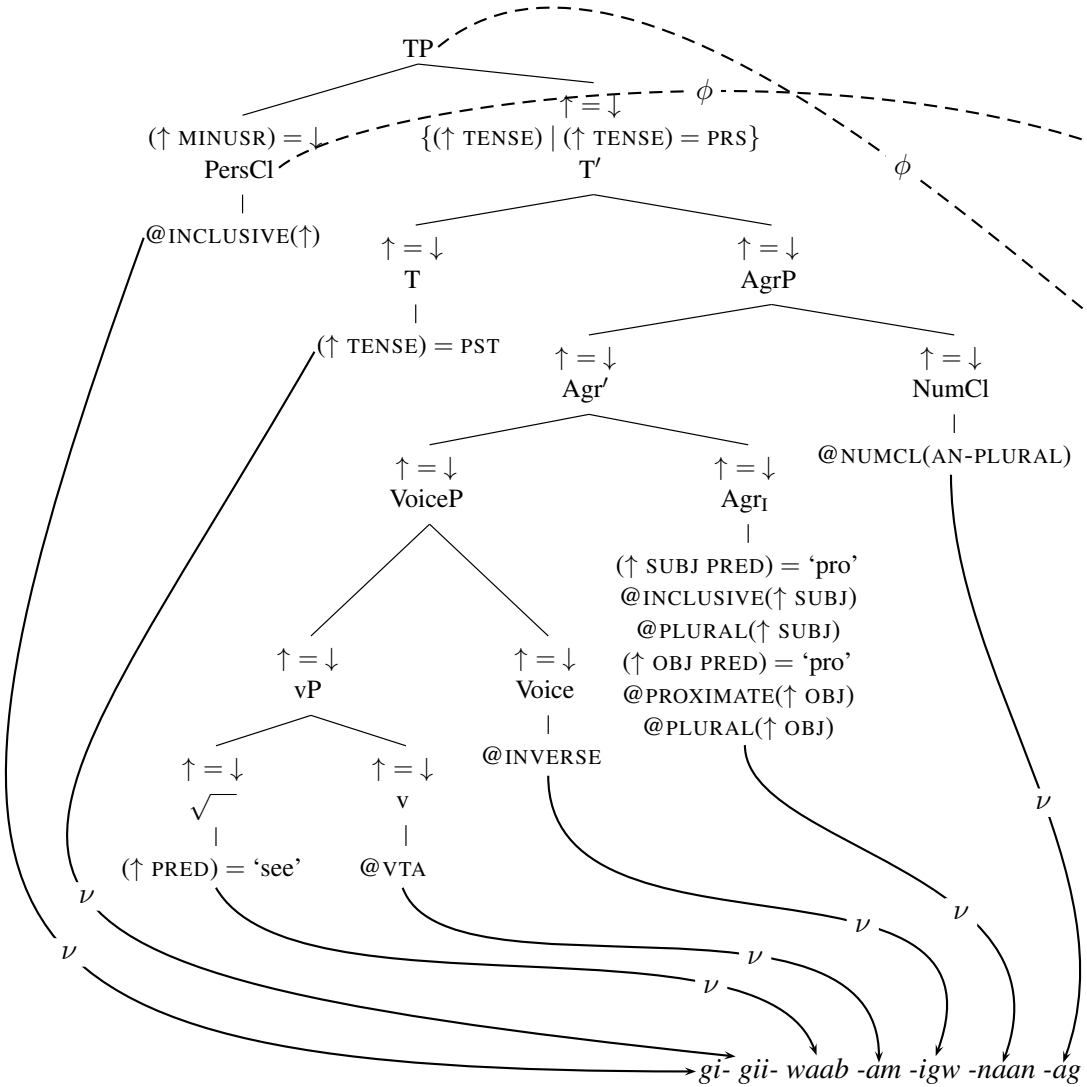
- The following are the c-, f-, and v-structures for a representative example, which was constructed based on the paradigms in Valentine (2001) (more can be found in Appendix A).
- Note that, while we have included templates in the c-structure of the tree, as usual in LFG they are to be interpreted as the full bundle of features abbreviated by the template.
  - In other words, the semantics of template invocation is just substitution (Dalrymple et al. 2004).
  - The c-structure in (25) gives the expanded form of (24), with all templatic information provided by substitution.
  - Thus, while the description for the PersCl node in (24) is written in the c-structure as (23a), it should be read as in (23b):

(23) a. @PLURAL(↑)
@INCLUSIVE(↑)

b. (↑ NUM) = PL
(↑ PERS SPEAK) = +
(↑ PERS HEAR) = +
(↑ PERS PART) = +
(↑ PERS PROX) = +
(↑ PERS ANIM) = +
(↑ PERS ENT) = +

Repeated from page 13

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

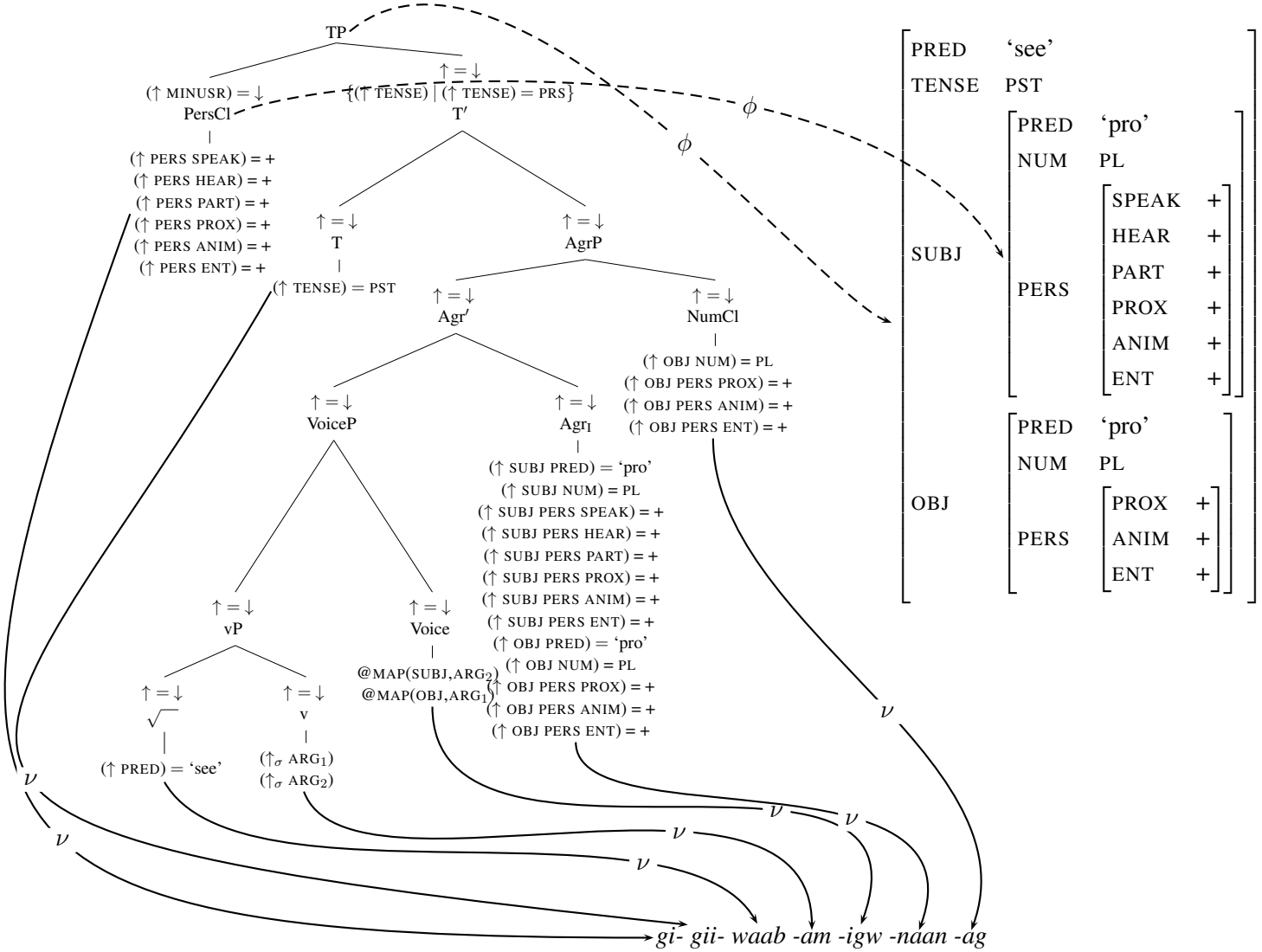


ϕ

ϕ

$$\begin{bmatrix}
\text{PRED} & \text{'see'} & \\
\text{TENSE} & \text{PST} & \\
\text{SUBJ} & \begin{bmatrix}
\text{PRED} & \text{'pro'} & \\
\text{NUM} & \text{PL} & \\
\text{PERS} & \begin{bmatrix}
\text{SPEAK} & + \\
\text{HEAR} & + \\
\text{PART} & + \\
\text{PROX} & + \\
\text{ANIM} & + \\
\text{ENT} & +
\end{bmatrix}
\end{bmatrix} \\
\text{OBJ} & \begin{bmatrix}
\text{PRED} & \text{'pro'} & \\
\text{NUM} & \text{PL} & \\
\text{PERS} & \begin{bmatrix}
\text{PROX} & + \\
\text{ANIM} & + \\
\text{ENT} & +
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}$$

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



4.2 Templates

- As noted above, we make use of the LFG mechanism of *templates* (Dalrymple et al. 2004, Asudeh et al. 2013) to encode bundles of grammatical descriptions that get expressed in the language.
- The templates involved in our analysis can be divided into five groups: those encoding general constraints, those encoding the prominence hierarchy (person/gender), those encoding obviation and number, those encoding verb classes, and those encoding the mapping between grammatical function and argument structure (direction, argument suppression).

4.2.1 Constraints

- Here we provide templates for constraints that determine the distribution of animacy, person, and alignment across grammatical functions and contexts.
- The first two constraints hold in all contexts.[10]
  - The first constraint, which we call the **Transitive Subject Constraint**, ensures that the subject of a clause with an object (either OBJ or OBJ_θ, i.e. PLUSO) must be animate; inanimate subjects are possible only in inanimate clauses (Rhodes 1990, 2010, Valentine 2001):

    (26) *Transitive Subject Constraint*
    @TSC := [(↑ SUBJ) & (↑ PLUSO)] ⇒ [(↑ SUBJ ANIM) = +]

    – This ensures that transitives with an inanimate ARG_1 are inverse, regardless of context (independent or conjunct).[11]
    – It also correctly ensures that verbs with a secondary object (OBJ_θ) must have an animate subject (in Algonquianist terms, correctly predicts that there are AI+O verbs, but no II+O verbs).
  - The second constraint, which we call the **Participant Argument Constraint**, ensures that 1st and 2nd person (i.e., participant) pronominals are possible only as subjects and (direct/primary) objects; secondary objects and obliques must be 3rd person (Rhodes 1990, 2010, Valentine 2001):

    (27) *Participant Argument Constraint*
    @PAC := ¬(↑ PLUSR PERS PART)
- We assume these two constraints are called by the c-structure rule introducing the root node, grouped together in the following template:

(28)
$$\text{@ROOT} := \begin{array}{l}\text{@TSC} \\ \text{@PAC}\end{array}$$

[10] These constraints use the Kibort-Findlay version of Lexical Mapping Theory Bresnan and Kanerva (1989), Bresnan and Moshi (1990), Bresnan and Zaenen (1990), which encodes the [±r] and [±o] features of LMT in explicit disjunctions of grammatical functions. For example, the two [−r] grammatical functions, SUBJ and OBJ, are captured by MINUSR, which is defined as follows: MINUSR = {SUBJ|OBJ}; see Findlay (2016, 2020) for further details and references.

[11] This is already ruled out in independent contexts by (29), but not conjunct contexts with a participant ARG_2.

- The last constraints, the **Prominence Constraints**, capture the different distributions of direct and inverse Voice heads in the independent and conjunct orders:

(29) *Independent Prominence Constraint*
$$@\text{IPC} := [(\uparrow \text{SUBJ}) \;\&\; (\uparrow \text{OBJ})] \Rightarrow \{[(\uparrow \text{SUBJ PERS PART}) = + \;\&\; (\uparrow \text{OBJ PERS PART}) = +] \mid [(\uparrow \text{OBJ PERS}) \sqsubset (\uparrow \text{SUBJ PERS})]\}$$

(30) *Conjunct Prominence Constraint*
$$@\text{CPC} := [(\uparrow \text{SUBJ}) \;\&\; (\uparrow \text{OBJ})] \Rightarrow \{[(\uparrow \{\text{SUBJ}|\text{OBJ}\} \text{ PERS PART}) = + \mid [(\uparrow \text{OBJ PERS}) \sqsubset (\uparrow \text{SUBJ PERS})]\}$$

  - In independent forms, the subject always outranks the object (i.e., the object's PERS features properly subsume those of the subject) unless both the subject and object are participants.
  - In conjunct forms, the subject always outranks the object unless either the subject or object is a participant.
  - I assume that these constraints are specified by the different versions of Agr(P) found in the independent and conjunct orders.
  - The contrast between independent and conjunct order can be captured in templates, defined to a first approximation below in (33).

4.2.2 Prominence templates

- Following Bejar and Rezac (2009) and Oxford (2014), among others, we assume that the person and animacy features are decomposed into a number of privative features.
- Instead of the feature geometries used by the above authors, in our system the implicational relationships between the features are encoded in a set of templates, providing a way to represent the prominence hierarchy without stipulating independent structures beyond those already provided by the LFG framework.

(31) *Prominence hierarchy templates*

Template	*Description*	*Explanation*
INCLUSIVE(f)	(f PERS SPEAK) = + (f PERS HEAR) = + @PARTICIPANT(f)	1st person inclusive
SPEAKER(f)	(f PERS SPEAK) = + @PARTICIPANT(f)	1st person
HEARER(f)	(f PERS HEAR) = + @PARTICIPANT(f)	2nd person
PARTICIPANT(f)	(f PERS PART) = + @PROXIMATE(f)	1 and/or 2
PROXIMATE(f)	(f PERS PROX) = + @ANIMATE(f)	3 and above
ANIMATE(f)	(f PERS ANIM) = + @ENTITY(f)	3′ and above
ENTITY(f)	(f PERS ENTITY) = +	All persons (0 and above)

- **Notes:**
  - Contra Valentine (2001), we exclude unspecified actors from the prominence hierarchy, following the analyses of Rhodes (1990) and Rhodes and Valentine (2015), in which the "unspecified actor" forms are analyzed as a kind of short passive.
    - This is because unspecified actors aren't treated syntactically as a grammatical function; VTA forms with an unspecified actor are inflected as intransitives.
    - The "theme signs" indicating an unspecified actor are treated as passive Voice heads suppressing certain arguments, similar to the reflexive (but with a different kind of suppression – see below).
    - This means that (in independent order forms) there are two homophonous Voice heads *-aa*, one indicating @DIRECT with a 3rd-person animate object, and one indicating an unspecified actor with a 3rd-person animate subject (however, the unspecified actor morpheme is different in conjunct forms).

4.2.3 Number and obviation templates

- We use the following templates to encode singular and plural number, and combinations of number, animacy, and obviation that are encoded in the verbal agreement system.

(32) *Number and obviation templates*

Template	*Description*	*Explanation*
PLURAL(f)	(f NUM) = PL	
SINGULAR(f)	(f NUM) = SG	
INAN-PLURAL(f)	@PLURAL(f) ¬(f PERS ANIM)	Inanimate plurals
AN-PLURAL(f)	@PLURAL(f) @ANIMATE(f) ¬(f PERS PART)	Animate 3rd person plurals
OBVIATIVE(f)	(f OBV) = + @ANIMATE(f) {@SINGULAR(f) | @PLURAL(f)}	Animate obviatives Number is ambiguous

- **Notes:**
  - While the @OBVIATIVE template encodes obviation with animate arguments, which is the canonical form of obviation, it is also possible for obviation to occur with inanimate arguments.
    - Unlike with animates, obviation is not marked on the DP in inanimate obviation; it is marked only in VII verb agreement.
    - Also unlike with animate, obviative inanimate arguments are not treated as number-neutral in either verbal or nominal morphology.
    - We mark these simply as being inanimate (with the @ENTITY template), with the obviation feature (↑ OBV) = +, which is not a PERS feature and which goes unrealized except in VII Agr heads, shown below.
  - The template @AN-PLURAL is used only for 3rd person animate plurals (not for participants), capturing the distribution of the *-ag* morpheme (found both with independent verbs and as a nominal plural marker).

4.2.4 Verb class and order templates

- Traditionally, Algonquianists group verbs into four classes, depending on transitivity and the animacy of one argument: VAI (intransitive, animate subject), VII (intransitive, inanimate subject), VTA (transitive, animate object), and VTI (transitive, inanimate object).
- However, Piggott (1979, 1989) argues that VAI and VTI verb finals (i.e., v heads) should be conflated, and we follow him here, leaving us with three verb class templates.[12]
- The templates for verbal order (independent vs. conjunct) given here are very tentative, subject to revision to capture the subtleties of the distribution of the two orders.

(33) *Verb class and order templates*

Template	*Description*	*Explanation*
VTA	$(\uparrow_\sigma \text{ARG}_1)$ $(\uparrow_\sigma \text{ARG}_2)$	Two semantic arguments
VTI-VAI	$(\uparrow_\sigma \text{ARG}_1)$ @ANIMATE($\uparrow$ SUBJ) $\neg$($\uparrow$ OBJ PERS ANIM)	At least one semantic argument Subject is animate No animate object
VII	$(\uparrow_\sigma \text{ARG}_1)$ $\neg$($\uparrow$ SUBJ PERS ANIM)	At least one semantic argument Subject is inanimate
INDEP-ORDER(f)	@IPC $\neg$(GF f)	Indep. Prominence Constraint Cannot be embedded
CONJ-ORDER(f)	@CPC (GF f)	Conj. Prominence Constraint Must be embedded

- **Notes:**
  - We have removed the specification of $\{(\uparrow \text{OBJ ANIM}) = + \mid \neg(\uparrow \text{OBJ})\}$ from the specification of @VTA, since inverse forms with inanimate objects (i.e., verbs with inanimate ARG_1) use the @VTA verb class.
  - Inverse is ruled out for @VTI-VAI forms by the fact that neither this template nor the inverse Voice head introduces an ARG_2. In contrast, the VTI direct Voice head does introduce ARG_2.
  - It may be that the @ANIMATE($\uparrow$ SUBJ) specification for @VTI-VAI is redundant, in which case it can be removed.
  - While VTI and VAI share a template for verb classification, they are distinguished by the presence of a direct theme sign (Voice head) in VTI contexts.
  - The templates given for verbal order capture the generalization that the independent form is found in (most) root clauses, while the conjunct form is found elsewhere.[13]

4.2.5 Argument structure templates

- The following templates determine the mapping between grammatical functions (in the f-structure) and argument roles (in the s-structure):

[12] Will Oxford (p.c.) has suggested an alternative analysis of these forms, which we aim to pursue.

[13] The actual situation is somewhat more complicated; the conjunct form is also found in main clauses in *wh*-questions, as well as in certain discourse contexts. The templates in (33) should be modified accordingly to account for this; however, distribution of the templates in c-structure rules and VIs should not be affected by this.

(34) *Templates for argument mapping*

Template	*Description*	*Explanation*
DIRECT	@MAP(SUBJ,ARG_1) @MAP(OBJ,ARG_2)	Subject $\mapsto$ agent Object $\mapsto$ patient
INVERSE	@MAP(SUBJ,ARG_2) @MAP(OBJ,ARG_1)	Subject $\mapsto$ patient Object $\mapsto$ agent
REFLEXIVE	@SUPPRESS(ARG_2,BIND(ARG_1))	Intransitive, binding relationship
SHORT-PASSIVE	@SUPPRESS(ARG_1,CLOSE-OFF)	Intransitive, agent existentially bound

- **Notes:**
  - We adopt certain templates from the account of lexical mapping in Findlay (2016, 2020):
    - @MAP(x,y) indicates that grammatical function x maps to argument role y.
    - @SUPPRESS(x,y) indicates that argument role x receives no mapping; y is a template indicating how role x is interpreted.
    - @BIND(z) indicates that argument role z is reflexively bound (by role x indicated in the @SUPPRESS template); Findlay (2020) encodes this reflexive binding in a meaning constructor associated with the @BIND template.
    - @CLOSE-OFF indicates that argument role x (indicated in the @SUPPRESS template) is existentially bound.
  - Thus, @REFLEXIVE indicates that the ARG_2 role is not associated with a grammatical function, but is coreferential with ARG_1.
  - @SHORT-PASSIVE, using the template definition from Findlay (2020), indicates that the ARG_1 role is not associated with a grammatical function, but is existentially bound.
    - The construction we are analyzing as a short passive is referred to in the Algonquianist tradition as the "unspecified actor" form (Valentine 2001); however, we follow Rhodes and Valentine (2015) in analyzing it as a passive.[14]
    - Unlike many languages, Ojibwe has only the short passive (where the agent role is unexpressed); there is no long passive in the language (where the subject is "demoted" to an oblique; Rhodes and Valentine 2015).

[14] Specifically, it corresponds to the "passive I" construction of Rhodes and Valentine (2015); while we do not provide an analysis of the "passive II" construction here, it seems amenable to the same kind of analysis.

4.3 Vocabulary Items

- Here we list the VIs involved in Ojibwe agreement inflection.

4.3.1 Voice heads

- With the exception of the reflexive morpheme (which is traditionally called a verb final), these are traditionally referred to as "theme signs".
- The main voice heads involved in the agreement system are given below:

(35) *Voice heads*

a. Direct Voice heads

$\langle$ [Voice], $\Phi\left\{ \begin{array}{l} @\text{DIRECT} \\ @\text{ADDRESSEE}(\uparrow \text{OBJ}) \end{array} \right\} \rangle \xrightarrow{\nu}$ *-in*

$\langle$ [Voice], $\Phi\left\{ \begin{array}{l} @\text{DIRECT} \\ @\text{PARTICIPANT}(\uparrow \text{OBJ}) \end{array} \right\} \rangle \xrightarrow{\nu}$ *-i*

$\langle$ [Voice], $\Phi\left\{ \begin{array}{l} @\text{DIRECT} \\ \neg(\uparrow \text{OBJ PERS ANIM}) \\ (\uparrow_\sigma \text{ARG}_2) \end{array} \right\} \rangle \xrightarrow{\nu}$ *-am, -oo, -i*

b. Passive Voice heads

$\langle$ [Voice], $\Phi\left\{ \begin{array}{l} @\text{SHORT-PASSIVE} \\ @\text{PARTICIPANT}(\uparrow \text{SUBJ}) \end{array} \right\} \rangle \xrightarrow{\nu}$ *-igoo*

$\langle$ [Voice], $\Phi\left\{ \begin{array}{l} @\text{SHORT-PASSIVE} \\ @\text{ANIMATE}(\uparrow \text{SUBJ}) \\ @\text{CONJ-ORDER}(\uparrow) \end{array} \right\} \rangle \xrightarrow{\nu}$ *-in*

c. Other Voice heads

$\langle$ [Voice], $\Phi\{ @\text{ANIMATE}((\uparrow_\sigma \text{ARG}_2)_{\sigma^{-1}}) \} \rangle \xrightarrow{\nu}$ *-aa*

$\langle$ [Voice], $\Phi\{ @\text{INVERSE} \} \rangle \xrightarrow{\nu}$ *-igw*

$\langle$ [Voice], $\Phi\{ @\text{REFLEXIVE} \} \rangle \xrightarrow{\nu}$ *-idizo*

- **Notes:**
  - The direction markers are unchanged from before, except the requirement of the inverse marker for an animate object has been removed, since these appear in inverse contexts with an inanimate object; furthermore, the VTI theme sign (realized as *-am, -oo,* or *-i*) has been added.
  - As mentioned in Section 4.2.4, the v heads with the @VTI-VAI template do not introduce ARG_2; in this context, it is introduced by the VTI direct Voice head, which is realized in various contexts as *-am, -oo,* or *-i* (morphological allomorphy conditioned by the choice of v head present).
  - The form *-aa* is underspecified, showing up as a direct form when the object is 3rd-person animate, and a passive form when the subject is 3rd-person animate (though in the conjunct order, the *-in*

suffix in (35c) plays this role). These two roles have in common that the grammatical function that maps to ARG_2 is animate (object in direct voice contexts, subject in the passive).

4.3.2 Agr heads

- This is the category traditionally referred to as "central agreement suffixes".
- They are divided into two sets: one found in independent-order contexts (38), and one found in conjunct-order contexts (39).
  - We analyze these as two separate syntactic categories, called by c-structure rules to head the AgrP projection.
  - Specifically, they are called by an Agr′ rule defined as follows:

(36) $\text{Agr}' \rightarrow \text{VoiceP} \quad \{ \underset{@\text{INDEP-ORDER}(\uparrow)}{\text{Agr}_\text{I}} \mid \underset{@\text{CONJ-ORDER}(\uparrow)}{\text{Agr}_\text{C}} \}$

- Many of the independent Agr forms have separate allomorphs that arise when (a) there is a PLUSO element present, but (b) there is no animate OBJ present.
  - In other words, it surfaces in transitives with an inanimate object, or in non-ditransitive contexts with a OBJ_θ .
  - This phenomenon is known as *n-registration* (Rhodes 1990), since the relevant morphemes contain /n/ and it registers a certain argument structure configuration.
  - The distribution of the n-registration Agr VIs can be characterized by the following template encoding the relevant constraints:

(37) $@\text{NREG} := (\uparrow \text{PLUSO})$
$\neg(\uparrow \text{OBJ PERS ANIM})$

  - For clarity, the VIs for Agr_I are divided into two groups: those without @NREG in (38a) and those with @NREG in (38b).

(38) *Independent Agr forms*

a. Non-@NREG forms

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}(\uparrow \textsc{minusr}) = \%\textsc{gf}\\ @\textsc{speaker}(\%\textsc{gf})\\ @\textsc{plural}(\%\textsc{gf})\\ \{(\uparrow \textsc{obj pers part}) \mid \neg(\uparrow \textsc{obj})\}\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-min}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}(\uparrow \textsc{minusr}) = \%\textsc{gf}\\ @\textsc{participant}(\%\textsc{gf})\\ @\textsc{plural}(\%\textsc{gf})\\ \{(\uparrow \textsc{obj pers part}) \mid \neg(\uparrow \textsc{obj})\}\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-m}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}\neg(\uparrow \textsc{subj pers part})\\ \neg(\uparrow \textsc{pluso})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-w}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}@\textsc{speaker}(\uparrow \textsc{subj})\\ @\textsc{plural}(\uparrow \textsc{subj})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-naa}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}@\textsc{proximate}(\uparrow \textsc{subj})\\ @\textsc{plural}(\uparrow \textsc{subj})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-waa}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}(\uparrow \textsc{subj obv}) = +\\ \{\neg(\uparrow \textsc{subj pers anim}) \mid\\ (\uparrow \textsc{obj})\\ \neg(\uparrow \textsc{obj pers prox})\}\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-ini}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\{\ @\textsc{short-passive}\ \}\,\rangle \quad \xrightarrow{\nu} \quad \textit{-m}$$

b. @NREG forms

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}@\textsc{proximate}(\uparrow \textsc{subj})\\ @\textsc{plural}(\uparrow \textsc{subj})\\ @\textsc{nreg}\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-naawaa}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\left\{\begin{array}{l}@\textsc{animate}(\uparrow \textsc{obj}_\theta)\\ @\textsc{singular}(\uparrow \textsc{obj}_\theta)\\ @\textsc{nreg}\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-nan}$$

$$\langle\, [\mathrm{Agr_I}],\quad \Phi\{\ @\textsc{nreg}\ \}\,\rangle \quad \xrightarrow{\nu} \quad \textit{-n}$$

- **Notes:**
  - The distribution of the morpheme *-min* differs across dialects. In certain dialects, including some of the Nishnaabemwin dialects characterized by Valentine (2001) and the Southwestern Ojibwe dialect recorded in Nichols (1980), it is found in any form that does not have an animate OBJ; see Goddard (2007) for a diachronic analysis of these morphemes.
  - The suffix *-w* signals a 3rd-person animate argument in intransitive forms (formerly analyzed as part of the *-wag* and *-wan* suffixes); in many cases it is absent due to phonological rules that delete glides word-finally and in certain consonant-adjacent contexts.
  - The morpheme *-naawaa* is an allomorph of *-waa* that shows up only in @NREG contexts; the other

transitive plural suffix *-naa* does not have a separate @NREG form.

- The morphemes *-nan* and *-n* are @NREG morphemes that do not have separate non-@NREG counterparts: *-nan* appears when there is an animate singular OBJ_θ, and *-n* is an elsewhere form.
- *-ini* indexes obviation in the subject when either the subject is inanimate, or the subject is animate and the object is also obviative.
- The second suffix *-m* is the unspecified-actor form found with intransitive predicates.

- The conjunct Agr forms can be sorted into four groups: those that realize just the subject (39a); those that realize just one MINUSR function, but unspecified for which (39b);[15] those that realize both arguments (39c); and the passive form, which realizes neither argument (39d).[16]

(39) *Conjunct Agr forms*

a. Marking just subject

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\left\{\begin{array}{l}@\text{SPEAKER}(\uparrow \text{SUBJ})\\ @\text{SINGULAR}(\uparrow \text{SUBJ})\end{array}\right\}\rangle \quad \xrightarrow{\nu} \quad \textit{-aan}$$

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\left\{\begin{array}{l}@\text{PARTICIPANT}(\uparrow \text{SUBJ})\\ @\text{SINGULAR}(\uparrow \text{SUBJ})\end{array}\right\}\rangle \quad \xrightarrow{\nu} \quad \textit{-an}$$

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\left\{\begin{array}{l}@\text{PROXIMATE}(\uparrow \text{SUBJ})\\ @\text{SINGULAR}(\uparrow \text{SUBJ})\end{array}\right\}\rangle \quad \xrightarrow{\nu} \quad \textit{-d}$$

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\left\{\begin{array}{l}@\text{PROXIMATE}(\uparrow \text{SUBJ})\\ @\text{PLURAL}(\uparrow \text{SUBJ})\end{array}\right\}\rangle \quad \xrightarrow{\nu} \quad \textit{-waad}$$

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\{\ @\text{OBVIATIVE}(\uparrow \text{SUBJ})\ \}\rangle \quad \xrightarrow{\nu} \quad \textit{-nid}$$

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\{\ @\text{ENTITY}(\uparrow \text{SUBJ})\ \}\rangle \quad \xrightarrow{\nu} \quad \textit{-g}$$

$$\langle\, [\text{Agr}_\text{C}],\quad \Phi\left\{\begin{array}{l}@\text{OBVIATIVE}(\uparrow \text{SUBJ})\\ (\uparrow \text{SUBJ OBV}) = +\end{array}\right\}\rangle \quad \xrightarrow{\nu} \quad \textit{-nig}$$

[15] Note that, while they realize features of only one argument, the forms in (39a) and (39b) can appear in transitive forms, as long as there is no compatible form in (39c) realizing features of both.

[16] Here we omit an additional morpheme that appears in most instances where there is a 3rd-person plural argument, *-waa*; see Appendix B for discussion of this morpheme.

b. Marking one MINUSR

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}(\uparrow \text{MINUSR}) = \%\text{GF}\\ @\text{SPEAKER}(\%\text{GF})\\ @\text{PLURAL}(\%\text{GF})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-aang}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}(\uparrow \text{MINUSR}) = \%\text{GF}\\ @\text{INCLUSIVE}(\%\text{GF})\\ @\text{PLURAL}(\%\text{GF})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-ang}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}(\uparrow \text{MINUSR}) = \%\text{GF}\\ @\text{PARTICIPANT}(\%\text{GF})\\ @\text{PLURAL}(\%\text{GF})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-eg}$

c. Marking both MINUSR

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}@\text{SPEAKER}(\uparrow \text{SUBJ})\\ @\text{SINGULAR}(\uparrow \text{SUBJ})\\ @\text{ANIMATE}(\uparrow \text{OBJ})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-ag}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}@\text{PARTICIPANT}(\uparrow \text{SUBJ})\\ @\text{SINGULAR}(\uparrow \text{SUBJ})\\ @\text{ANIMATE}(\uparrow \text{OBJ})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-ad}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}@\text{SPEAKER}(\uparrow \text{SUBJ})\\ @\text{PLURAL}(\uparrow \text{SUBJ})\\ @\text{ANIMATE}(\uparrow \text{OBJ})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-angid}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}@\text{ANIMATE}(\uparrow \text{SUBJ})\\ @\text{SPEAKER}(\uparrow \text{OBJ})\\ @\text{PLURAL}(\uparrow \text{OBJ})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-amind}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}@\text{ANIMATE}(\uparrow \text{SUBJ})\\ @\text{PARTICIPANT}(\uparrow \text{OBJ})\\ @\text{SINGULAR}(\uparrow \text{OBJ})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-g}$

$\langle\, [\text{Agr}_C],\quad \Phi\left\{\begin{array}{l}@\text{SPEAKER}(\uparrow \text{SUBJ})\\ @\text{SINGULAR}(\uparrow \text{SUBJ})\\ @\text{PARTICIPANT}(\uparrow \text{OBJ})\\ @\text{PLURAL}(\uparrow \text{OBJ})\end{array}\right\} \,\rangle \quad \xrightarrow{\nu} \quad \textit{-agog}$

d. Unspecified actor

$\langle\, [\text{Agr}_C],\quad \Phi\{\ @\text{SHORT-PASSIVE}\ \}\,\rangle \quad \xrightarrow{\nu} \quad \textit{-ng}$

- **Notes:**
  - The form *-amind*, realizing forms with a 3rd-person animate subject and a 1st-person plural object, is found only in more conservative dialects, including Nipissing Algonquin (Oxford 2014, 2019). In more innovative dialects, including Nishnaabemwin (Valentine 2001) and Southwestern Ojibwe (Nichols 1980), the form *-angid* is found in this context. In these dialects, the VI for *-angid* would

include an f-description using local names for both MINUSR functions, such that one of the core arguments is 1st-person plural, and the other is 3rd-person animate.

- When the VI *-g*, marking a 3rd-person animate subject and a 2nd-person plural object, is present, the 2nd-person object direct Voice head *-in* is realized by an allomorph, *-ih*, which appears only in this context; the sequence of consonants /hg/ coalesces to [k] (Oxford 2019).
- Unlike independent-order and nominal inflection, inanimate arguments are unspecified for number in conjunct-order agreement.
- The phenomenon of n-registration is absent in conjunct-order Agr forms, which make reference only to SUBJ and OBJ.

4.3.3 Agreement clitics

- Ojibwe has two sets of agreement clitics that appear *only in independent-order contexts*: a set of proclitics that index the person of (usually) the subject, and a set of enclitics that index number and obviation of third-person arguments (usually the object) in certain contexts.
- The person proclitics (category PersCl) are introduced in Spec-TP in a node annotated $(\uparrow \text{MINUSR}) = \downarrow$; it indexes the person of either SUBJ or OBJ, whichever is higher on the relevant prominence hierarchy (here using the feature HEAR rather than SPEAK for the highest point in the hierarchy, meaning 2nd person outranks 1st person).

(40) *Person proclitics*

$\langle$ [PersCl], $\Phi\{$ @HEARER($\uparrow$) $\}\rangle \xrightarrow{\nu}$ *gi-*

$\langle$ [PersCl], $\Phi\{$ @PARTICIPANT($\uparrow$) $\}\rangle \xrightarrow{\nu}$ *ni-*

$\langle$ [PersCl], $\Phi\left\{\begin{array}{l}\text{@ANIMATE}(\uparrow)\\((\text{SUBJ}\uparrow)\ \text{PLUSO})\end{array}\right\}\rangle \xrightarrow{\nu}$ *o-*

- **Notes:**
  - The label is changed from before.
  - The specification of *o-* was changed from @PROXIMATE($\uparrow$) to @ANIMATE($\uparrow$), reflecting that this form occasionally appears when both subject and object are obviative.
  - Note that the 3rd-person proclitic *o-* does not appear in intransitive forms (forms with neither OBJ nor OBJ_θ); there the Agr_I suffix *-w* appears instead.
- The number enclitics appear on a node in the specifier of AgrP,[17] which is annotated $\uparrow = \downarrow$; the @NUMCL template indicates which grammatical function's features are being specified, as defined in (42).

(41) *Number/obviation enclitics*

$\langle$ [NumCl], $\Phi\{$ @NUMCL(AN-PLURAL) $\}\rangle \xrightarrow{\nu}$ *-ag*

$\langle$ [NumCl], $\Phi\{$ @NUMCL(INAN-PLURAL) $\}\rangle \xrightarrow{\nu}$ *-an*

$\langle$ [NumCl], $\Phi\{$ @NUMCL(OBVIATIVE) $\}\rangle \xrightarrow{\nu}$ *-an*

[17] In a fuller exposition of Ojibwe verbal inflection, which includes negation and modality, this will be revised so that these enclitics appear in spec-ModP, as they follow the modal suffixes. However, since we are omitting modal suffixes in this analysis, we will leave them here for now.

- **Notes:**
  - These morphemes mark number/obviation of OBJ if there is an OBJ present; of OBJ_θ if there is an OBJ_θ but no OBJ; and of SUBJ if there is neither PLUSO function present.
  - This is encoded in the @NUMCL template, defined as follows:

(42) @NUMCL(template) := {[(↑ OBJ) & @template(↑OBJ)] |
[¬(↑ OBJ) & @template(↑OBJ_θ)] |
[¬(↑ PLUSO) & @template(↑SUBJ)]}

  - Alternatively, the context could be specified in the VIs themselves.
  - In all cases, they only index features of third-person arguments.
  - Homophonous morphemes are used to mark animate plural, inanimate plural, and (animate) obviation in nouns as well, though it's unclear if we will be able to use the same VIs for this context.

4.3.4 Other VIs used

- In addition to the agreement morphemes listed above, the following morphemes appear in the examples in Section 4.1 and Appendix A.

(43) ⟨ [T], Φ{ (↑ TENSE) = PST } ⟩ $\xrightarrow{\nu}$ *gii-*

⟨ [$\sqrt{}$], Φ{ (↑ PRED) = 'see' } ⟩ $\xrightarrow{\nu}$ *waab*

⟨ [v], Φ{ @VTA } ⟩ $\xrightarrow{\nu}$ *-am*

⟨ [$\sqrt{}$, v], Φ{ (↑ PRED) = 'eat' ; @VTI-VAI } ⟩ $\xrightarrow{\nu}$ *wiisini*

  - For most verbs in Ojibwe, the verb root and the v morpheme indicating the verb class are separate morphemes, as with *waab* and *-am*.
  - However, the verb meaning 'eat' has suppletive forms for the three compatible verb classes (i.e., depending on transitivity and animacy of the object): *amw* 'eat.VTA', *miij* 'eat.VTI', *wiisini* 'eat.VAI'.
  - This is analyzed as the verb exponing a span including both $\sqrt{}$ and v.
  - We see the intransitive form *wiisini* in (45) below.

5 Conclusion

1. We have outlined a new grammatical framework, Lexical-Realizational Functional Grammar, for a fully constraint-based realizational morphosyntax.
   - L_RFG is the offspring of Distributed Morphology, a theory of morphological realization, and Lexical-Functional Grammar, a constraint-based theory of syntax and grammatical architecture.
2. We compared L_RFG to its parent frameworks, after having identified, in our view, the key conceptual foundations of these frameworks.
3. We offered a detailed analysis of Ojibwe agreement in L_RFG.

- L_RFG combines the strengths of its two parent frameworks:
  1. Like LFG, it is a declarative, representational and constraint-based theory that is ideally suited to modelling nonconfigurationality.
  2. Like DM, it provides a realizational, morpheme-based view of word-formation and is good at modelling complex morphological structures including those found in polysynthetic languages, such as many North American Indigenous languages.

6 Future research

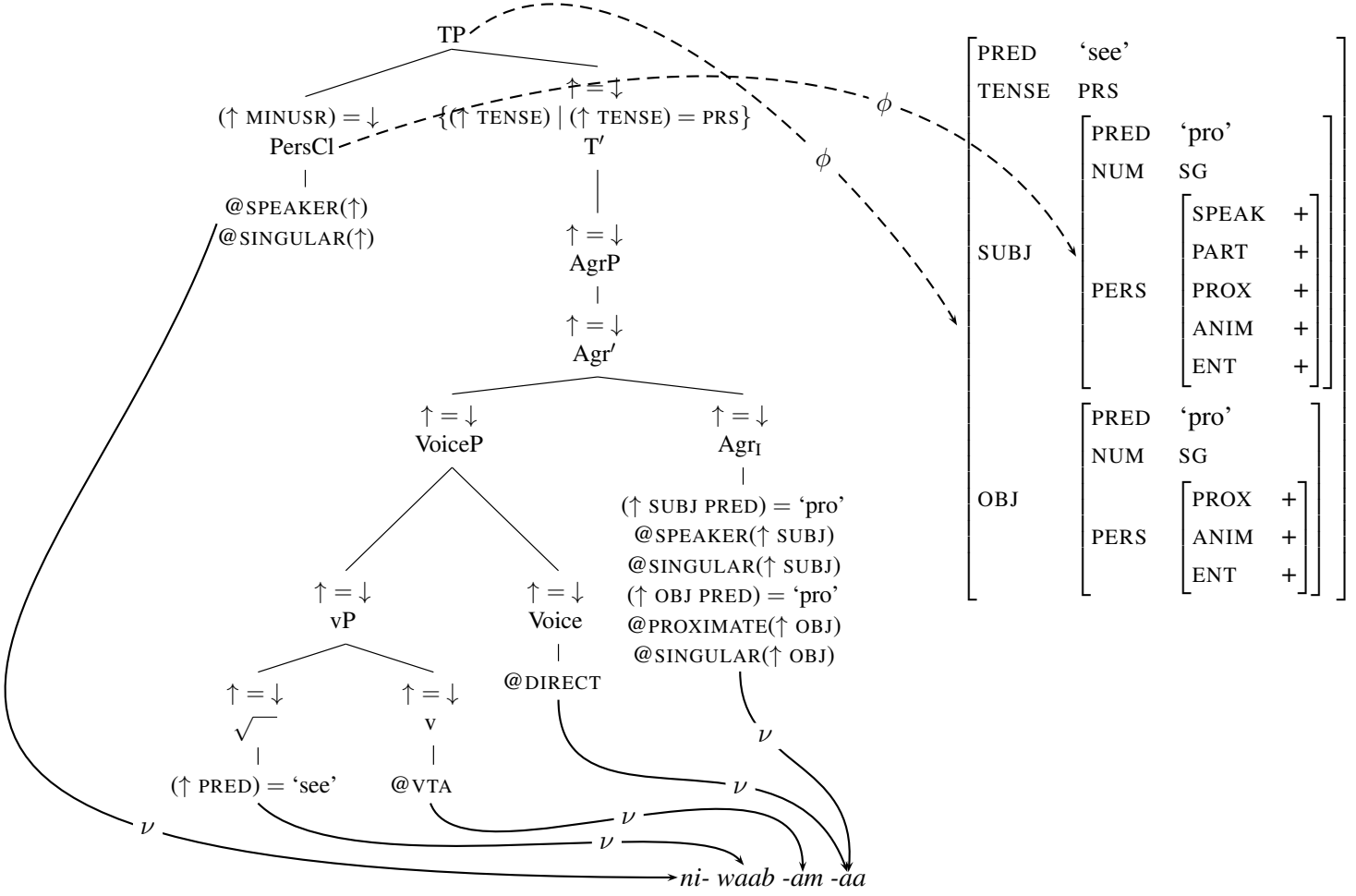
- We are currently in discussions with various researchers about phenomena that are otherwise difficult to account for, but which seem amenable to an L_RFG-style analysis, including:
  - Oleg Belyaev's work on case inflection in Ossetic (Iranian)
  - Bronwyn Bjorkman's work on clitic-induced doubling in Ingush (Nakh-Dagestanian) and Breton (Celtic)
  - Michael Everdell's work on the argument-adjunct distinction and the interpretation of floating quantifiers in O'dam/Tepehuan (Uto-Aztecan)
- We also aim to develop an L_RFG analysis of Archi, as per the challenge set out in Bond et al. (2016).
- We are also starting a collaboration with Tina Bögel on L_RFG's prosodic structure and the mapping to this structure from vocabulary structure.
- Another line of future research involves developing a more complete theory of portmanteaux, to facilitate analyses of fusional languages such as those found in Europe (e.g., English and French).

Appendices

A More examples

- Here we show more representative examples, demonstrating most of the templates and VIs introduced in Section 4

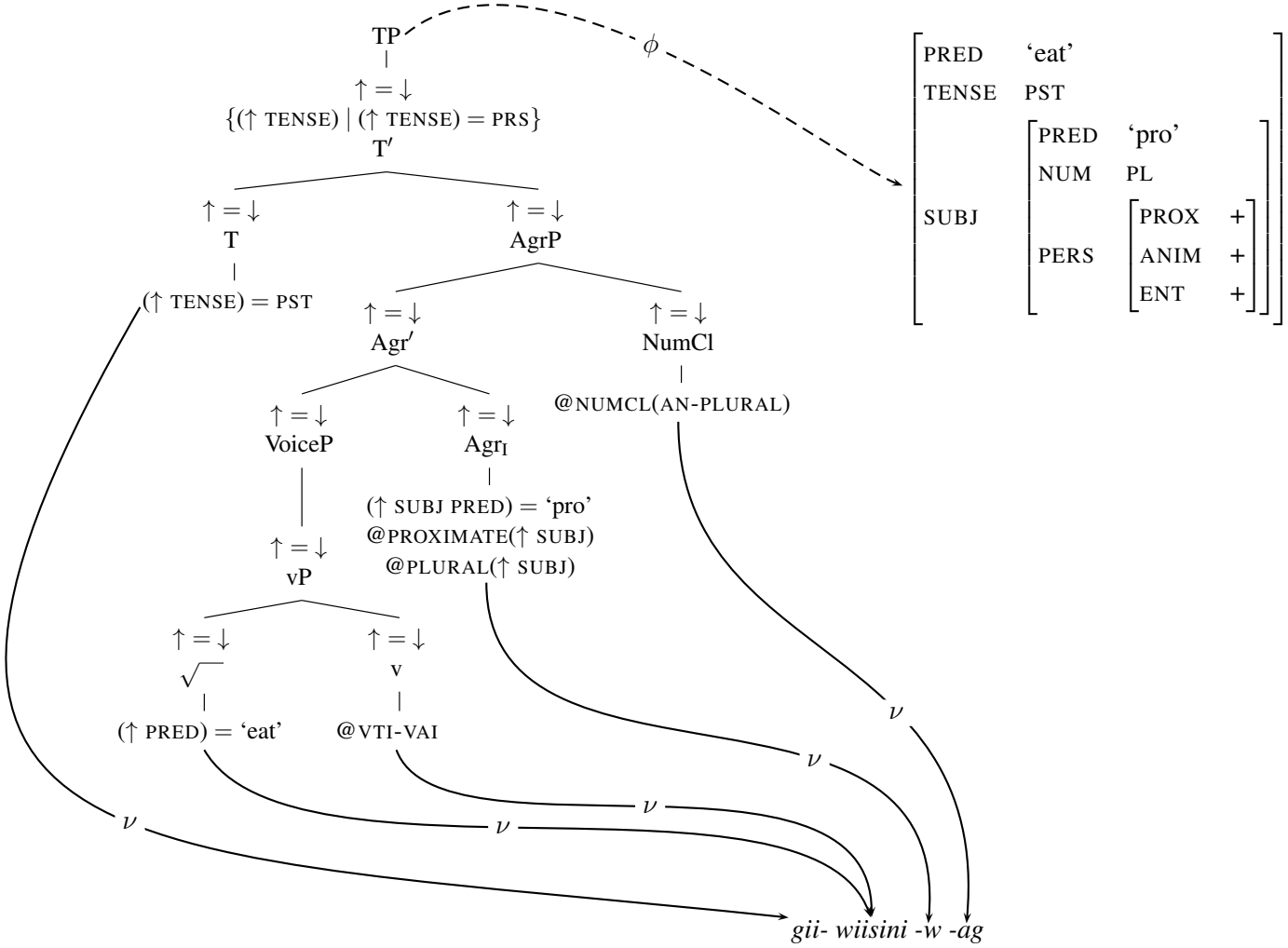
(44) ni- waab -am -aa
1 see VTA DIR
'I see him/her.'[18]



[18]This example includes the phenomenon referred to as Pac-Man Spanning (see footnote 5), in which a VI (here, *-aa*) spans an adjacent preterminal node (here, Agr) for which no other VI is available. As shown in 4.3.2, there are no VIs specified for the category Agr_I that are compatible with singular number – all of the Agr_I VIs are specified as (↑ NUM) = PL. This means that, while the VI *-aa* is specified only for the category Voice and not for Agr_I, there is no v-structure that is more informative than one in which it also realizes the adjacent Agr head. In this way, (pre)terminal nodes that are necessary in the c-structure but for which there is no VI available can still serve as an input to the exponence function without the need for (stipulated) empty categories.

Full version of example on page 6

(45) gii- wiisini -w -ag
PST see.VAI 3 3PL
'They ate.'



B On conjunct order agreement

- In the account of conjunct-order Agr heads given in Section 4.3.2, we omitted an additional morpheme that follows the Agr_C heads in most instances when there is a 3rd-person animate plural argument, namely the suffix *-waa*.
- We are not yet sure how to analyze this morpheme, but our current tentative analysis involves splitting conjunct Agr agreement into two separate agreement heads:
  - The first, which we call Agr_C, is realized in all conjunct-order forms, and indicates agreement with one or both of the core MINUSR arguments (subject and object; conjunct-order forms never show @NREG agreement or indicate OBJ_θ in any way); this includes the forms given in (39).
  - The second, which we call $\hat{\text{A}}\text{gr}_{\text{C}2}$, has only one realization, *-waa*, and indicates agreement with a 3rd-person animate plural core argument.
  - As indicated by the "roof" notation ($\hat{\text{X}}$), $\hat{\text{A}}\text{gr}_{\text{C}2}$ is a non-projecting head (Toivonen 2001, 2003); we analyze it as forming a complex head with the Agr_C head (see Oxford 2018 for an analogous analysis in a Minimalist DM framework, using fission), captured by the following c-structure rule:

    (46) $\quad \text{Agr}_\text{C} \quad \rightarrow \quad \text{Agr}_\text{C} \quad (\hat{\text{A}}\text{gr}_{\text{C}2})$

  - In addition to *-waa*, there is a spanned, portmanteau VI *-waad* which spans the Agr_C and $\hat{\text{A}}\text{gr}_{\text{C}2}$ heads in (many) instances where the subject is 3rd-person animate plural form (currently listed in (39a) as realizing only Agr_C).
- However, this analysis has a few problems:
  - The rule in (46) in its current form is recursive, and therefore overgenerates, predicting it should be possible to have any number of *-waa* morphemes present indexing the same argument (this is obviously not the case). We are not currently sure how to fix this.
  - There is a homophonous *-waa* morpheme in the independent-order Agr category (Agr_I); ideally this state of affairs should be avoided by positing a single, underspecified *-waa* morpheme. However, we are not sure if the two can be unified. (Furthermore, it may be that the two differ in their allomorphic properties, in which they should in fact be considered separate morphemes. We are still looking into this.)

References

Alsina, Alex, and Eugenio M. Vigo. 2017. Fixing LFG to account for direct-inverse agreement: The case of Plains Cree. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG'17 Conference*, 24–44. Stanford, CA: CSLI Publications.

Anderson, Stephen R. 1969. West Scandanavian Vowel Systems and the Ordering of Phonological Rules. Ph.D. thesis, MIT.

—. 1982. Where's Morphology. *Linguistic Inquiry* 13(4): 571–612.

—. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

—. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.

Asudeh, Ash. 2012. *The Logic of Pronominal Resumption*. Oxford: Oxford University Press.

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

Asudeh, Ash, and Gianluca Giorgolo. 2012. Flexible Composition for Optional and Derived Arguments. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG12 Conference*, 64–84. Stanford, CA: CSLI Publications.

Asudeh, Ash, Gianluca Giorgolo, and Ida Toivonen. 2014. Meaning and Valency. In *Proceedings of the LFG14 Conference*, 68–88. Stanford, CA: CSLI Publications.

Asudeh, Ash, and Ida Toivonen. 2015. Lexical-Functional Grammar. In Bernd Heine and Heiko Narrog, eds., *The Oxford Handbook of Linguistic Analysis*, 373–406. Oxford: Oxford University Press, 2nd edn.

Beard, Robert. 1995. *Lexeme Morpheme Base Morphology*. Albany, NY: SUNY Press.

Beaver, David I. 2001. *Presupposition and Assertion in Dynamic Semantics*. Stanford, CA: CSLI Publications.

Beaver, David I., and Bart Geurts. 2014. Presupposition. In Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University, Winter 2014 edn.

Bejar, Susana, and Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40: 35–73.

Bermúdez-Otero, Ricardo, and Ana R. Luís. 2016. A View of the Morphome Debate. In *The Morphome Debate*, 309–340. Oxford: Oxford University Press.

Bögel, Tina. 2015. The Syntax-Prosody Interface in Lexical Functional Grammar. Ph.D. thesis, University of Konstanz.

Bond, Oliver, Greville G. Corbett, Marina Chumakina, and Dunstan Brown, eds. 2016. *Archi: Complexities of Agreement in Cross-Theoretical Perspective*. Oxford: Oxford University Press.

Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler. 2016. *Lexical-Functional Syntax*. Chichester, UK: Wiley-Blackwell.

Bresnan, Joan, and Jonni M. Kanerva. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20(1): 1–50.

Bresnan, Joan, and Lioba Moshi. 1990. Object Asymmetries in Comparative Bantu Syntax. *Linguistic Inquiry* 21(2): 147–185.

Bresnan, Joan, and Annie Zaenen. 1990. Deep Unaccusativity in LFG. In Katarzyna Dziwirek, Patrick Farrell, and Errapel Mejias-Bikandi, eds., *Grammatical Relations: A Cross-theoretical Perspective*, 45–57. Stanford, CA: CSLI.

Bye, Patrik, and Peter Svenonius. 2012. Non-concatenative Morphology as Epiphenomenon. In Jochen Trommer, ed., *The Morphology and Phonology of Exponence*, 427–495. Oxford: Oxford University Press.

Chomsky, Noam. 1970. Remarks on Nominalization. In Roderick A. Jacobs and Peter S. Rosenbaum, eds., *Readings in English Transformational Grammar*, 184–221. Waltham, MA: Ginn and Company.

Dahlstrom, Amy. 2014. *Plains Cree Morphosyntax*. Abdington-on-Thames: Routledge.

—. 2017. Seeking consensus on the fundamentals of Algonquian word order. In Monica Macaulay, ed., *Papers of the Forty-Fifth Algonquian Conference*, 1–13. Albany, NY: SUNY Press.

Dalrymple, Mary, ed. 1999. *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. Cambridge, MA: MIT Press.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

—. 2015. Morphology in the LFG Architecture. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG15 Conference*, 65–83. Stanford, CA: CSLI Publications.

Dalrymple, Mary, Ronald M. Kaplan, and Tracy Holloway King. 2004. Linguistic Generalizations over Descriptions. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG04 Conference*, 199–208. Stanford, CA: CSLI Publications.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, eds. 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Dalrymple, Mary, John J. Lowe, and Louise Mycock. 2019. *The Oxford Reference Guide to Lexical Functional Grammar*. Oxford: Oxford University Press.

Di Sciullo, Anna Maria, and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3): 547–619.

Findlay, Jamie Y. 2016. Mapping theory without argument structure. *Journal of Language Modelling* 4(2): 293–338.

—. 2020. Lexical Mapping Theory and the anatomy of a lexical entry. Paper presented at the LFG20 virtual conference.

Goddard, Ives. 2007. Reconstruction and History of the Independent Indicative. In H. C. Wolfart, ed., *Papers of the Thirty-Eighth Algonquian Conference*, 207–271. Winnipeg, MB: Papers of the Algonquian Conference.

Halle, Morris. 1973. Prolegomena to a Theory of Word-Formation. *inguistic Inquiry* 4(1): 3–16.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Samuel Jay Keyser and Ken Hale, eds., *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.

Harris, Alice. 2016. *Multiple Exponence*. Oxford: Oxford University Press.

Haugen, Jason D., and Daniel Siddiqi. 2016. Towards a restricted realizational theory: Multimorphemic monolistemicity, portmanteaux, and post-linearization spanning. In *Morphological Metatheory*, 343–386. Amsterdam: John Benjamins.

Jones, David. 1977. *A basic Algonquin grammar for teachers of the language at Maniwaki, Quebec*. Maniwaki: River Desert Band Council.

Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson, and Paul Bennett, eds., *Linguistic Theory and Computer Applications*, 149–181. London: Academic Press. Reprinted in Dalrymple et al. (1995: 339–367).

—. 1989. The Formal Architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen, eds., *Proceedings of ROCLING II*, 3–18. Reprinted in Dalrymple et al. (1995: 7–27).

Kaplan, Ronald M., and Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In Joan Bresnan, ed., *The Mental Representation of Grammatical Information*, 29–130. Cambridge, MA: MIT Press.

Keenan, Edward L. 1971. Two Kinds of Presuppositions in Natural Language. In Charles J. Fillmore and D. Terence Langendoen, eds., *Studies in Linguistics Semantics*, 45–52. Holt, Rinehart & Winston.

Kiparsky, Paul. 1973. Elsewhere in Phonology. In Stephen Anderson and Paul Kiparsky, eds., *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston.

Lapointe, Steven. 1980. A Theory of Grammatical Agreement. Ph.D. thesis, University of Massachusetts, Amherst.

Lieber, Rochelle. 1992. *Deconstructing Morphology*. Chicago: University of Chicago Press.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2): 201–225.

McGinnis, Martha. 1999. Is there syntactic inversion in Ojibwa? In Leora Bar-el, Rose-Marie Déchaine, and Charlotte Reinholtz, eds., *Papers from the Workshop on Structure & Constituency in North American Languages*, 101–118. MIT Occasional Papers in Linguistics 17.

Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46(2): 273–303.

Nichols, John David. 1980. Ojibwe Morphology. Ph.D. thesis, Harvard University, Cambridge, MA.

Oxford, Will. 2014. Microparameters of Agreement: A diachronic perspective on Algonquian verb inflection. Ph.D. thesis, University of Toronto, Toronto.

—. 2018. Fission in Algonquian and the status of morphological templates. To appear in Proceedings of WSCLA 23, ed. by Daniel Reisinger, UBC Working Papers in Linguistics.

—. 2019. Inverse marking and Multiple Agree in Algonquin. *Natural Language and Linguistic Theory* 37(3): 955–996.

Piggott, Glyne L. 1979. Verb classification in Ojibwa. *Montreal Working Papers in Linguistics* 12: 149–186.

—. 1989. Argument structure and the morphology of the Ojibwa verb. In Donna B. Gerdts and Karin

Michelson, eds., *Theoretical Perspectives on Native American Languages*, 176–208. Albany, NY: SUNY Press.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Tech. rep., RuCCS, Rutgers University, New Brunswick, NJ. Technical Report #2.

—. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

Ramchand, Gillian Catriona. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.

Rhodes, Richard A. 1990. Ojibwa secondary objects. In Katarzyna Dziwirek, Patrick Farrell, and Errapel Mejías-Bikandi, eds., *Grammatical Relations: A Cross Theoretical Perspective*, 401–414. Palo Alto, CA: CSLI Stanford.

—. 1994. Agency, inversion, and thematic alignment in Ojibwe. In Susanne Gahl, Andy Dolbey, and Christopher Johnson, eds., *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society*, 431–446.

—. 2010. Relative root complement: a unique grammatical relation in Algonquian syntax. In Jan Wohlgemuth and Michael Cysouw, eds., *Rara & Rarissima: Collecting and Interpreting Unusual Characteristics of Human Languages*, 305–324. Berlin: Mouton de Gruyter.

Rhodes, Richard A., and J. Randolph Valentine. 2015. Transitivity in Ojibwe. In Andrej Malchukov and Bernard Comrie, eds., *Valency Classes in the World's Languages : Case Studies from New Guinea, Australia, and the Americas, and Theoretical Outlook*, 1205–1264. Berlin: Mouton de Gruyter.

Siddiqi, Daniel, and Heidi Harley, eds. 2016. *Morphological Metatheory*. Amsterdam: John Benjamins.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36.

Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Oxford: Oxford University Press.

Svenonius, Peter. 2016. Emergent extended projections. Paper presented at MIT on November 4, 2016.

Toivonen, Ida. 2001. The Phrase Structure of Non-Projecting Words. Ph.D. thesis, Stanford University, Stanford, CA.

—. 2003. *Non-Projecting Words: A Case Study of Swedish Particles*. Dordrecht: Kluwer Academic.

Valentine, J. Randolph. 2001. *Nishnaabemwin Reference Grammar*. Toronto: University of Toronto Press.

Wescoat, Michael T. 2002. On Lexical Sharing. Ph.D. thesis, Stanford University.

—. 2005. English nonsyllabic auxiliary contractions: An analysis in LFG with lexical sharing. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG05 Conference*, 469–486. Stanford, CA: CSLI.

—. 2007. Preposition-determiner contractions: An analysis in Optimality-Theoretic Lexical-Functional Grammar with lexical sharing. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG07 Conference*, 439–459. Stanford, CA: CSLI.