Modelling Exponents

Ash Asudeh — Rochester
Tina Bögel — Konstanz & Frankfurt
Dan Siddiqi — Carleton

July 24 · LFG 2023
lrfg.online

Introduction

- We have been working on L_RFG since 2016 and have been presenting it to this audience since 2020 (Melchin et al. 2020b, Everdell et al. 2021, Asudeh and Siddiqi 2022b).
- In this talk, we turn our attention to the L_RFG theory of morphological *exponence*.
- The theory of exponence is ultimately a theory of *exponents*: an exponent is a morphological representation that serves as the interface between an *exponendum* and a *realization*.
- In L_RFG, an exponent is represented as a *vocabulary structure*, or *v-structure* for short.
- We have had a strong idea of vocabulary structure since the beginning, but we are now putting some meat on the bones.
- The Vocabulary in L_RFG is the mapping from the set of exponenda, the set of lefthand sides of *vocabulary items*, to the set of exponents, the set of righthand sides of vocabulary items, i.e. the set of v-structures.
- A vocabulary item is represented as in (1) below, based on work in progress (Asudeh and Siddiqi forthcoming).[1]
- The tuple in (1) is the representation of an exponendum. It is mapped by ν, the correspondence function from c-structure to v-structure, to its exponent, represented as a vocabulary structure.
  - Note that although the exponence function ν expects a list as its first coordinate, as in $\nu(\langle[\ldots,\alpha,\ldots],\_\_\rangle) = \beta$, we abbreviate this as $\nu(\alpha) = \beta$.

(1) $\langle \underbrace{[C_1,\ldots,C_n]}_{distribution}, \underbrace{F \cup G \cup I}_{function/meaning} \rangle \xrightarrow{v} \big[\quad \big]_{v\text{-}structure}$

  - The first member of a pair is a list of nodes, where each node is represented by its category. Alternatively, one could think of it as a list of pairs of nodes and their labels.
  - The second member of a pair is the union of a set of descriptions of f-structures, F, a set of Glue meaning constructors, G, and a set of descriptions of i-structures, I. Any of these sets may be empty.

[1] Readers with some familiarity of L_RFG may notice that the lefthand side of this vocabulary item is a pair, not a triple, as in Asudeh and Siddiqi (2022a,b). Similarly, there is no longer a 'big Phi', Φ, in the lefthand side of vocabulary items. This is a consequence of certain refinements to the L_RFG architecture that we are not presenting here (Asudeh and Siddiqi forthcoming).

- $\mathrm{L_R FG}$ assumes the morphosyntactic operation of *spanning* (Haugen and Siddiqi 2016, Merchant 2015).
  - *Spans* are represented as lists of c-structure nodes (represented by a list of their associated category labels) in the first coordinate of an exponendum; see (1) above.
  - There are two kinds of spanning in $\mathrm{L_R FG}$:
    - *Vocabulary Spanning*: the case where the category list in the first coordinate of an exponendum has length greater than one; i.e., vocabulary spanning is a matter of listing in the Vocabulary.
    - *Pac-Man Spanning*: the case where some category would be left unexponed and is instead mapped to a neighbouring exponent; i.e., Pac-Man spanning is a matter of the ν-mapping being a total function from the domain of c-structure nodes to the co-domain of v-structures. See §5, page 13, below for an example.

  As a result of the two kinds of spanning, ν must be a many-to-one function.

- In short, a v-structure is modelled as an attribute-value matrix, similarly to f-structure.
- Attributes are symbols, like DEPENDENCE. Values are symbols, strings, v-structures, or sets of symbols.[2]
- On analogy with f-structures and f-descriptions, v-structures are *described* by v-descriptions, a set of defining equations and constraints that picks out the minimal satisfying v-structure, if any, as its model.
- This is the general framework that we will motivate:

(2)
$$\langle\, [\mathrm{C}_1, \ldots, \mathrm{C}_n], \mathrm{F} \cup \mathrm{G} \cup \mathrm{I} \,\rangle \xrightarrow{\nu} \textit{v-s}\begin{bmatrix} \text{PHON(OLOGICAL)REP(RESENTATION)} & \textit{phon. realization \& conditions} \\ \text{P(ROSODIC)FRAME} & \textit{prosodic unit} \\ \text{P(ROSODIC)DOMAIN} & \textit{prosodic unit} \\ \text{DEP(ENDENCE)} & \left\{\text{LT, RT}\right\} \\ \text{CLASS} & \left\{\textit{inflectional classes}\right\} \\ \text{HOST} & \textit{v-s}\begin{bmatrix} \text{IDENT(ITY)} & + \\ \text{PHONREP} & \ldots \\ \text{PFRAME} & \ldots \\ \text{DEP} & \ldots \\ \text{CLASS} & \ldots \end{bmatrix} \end{bmatrix}$$

- This is the vocabulary item (VI) for *-en*, the case we use for exemplification, that we will motivate:[3]

(3)
$$\left\langle\, [\mathrm{v_a}],\ \begin{array}{l}\lambda P \lambda e.\mathbf{cause}(\mathbf{become}(P(e))) : \\ ((\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma) \multimap ((\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma)\end{array} \,\right\rangle \xrightarrow{\nu} v\begin{bmatrix} \text{PHONREP} & /\text{ən}/ \\ \text{PFRAME} & ((\ \ )(\cdot)_\sigma)_{ft} \\ \text{PDOMAIN} & (\ (v\ \text{PFRAME})\ )_\omega \\ \text{DEP} & \text{LT} \\ \text{CLASS} & \text{WEAK} \\ \text{HOST} & \begin{bmatrix} \text{IDENT} & + \\ \text{PHONREP} & /\ldots([\mathbf{son}])[\mathbf{obs}]/ \\ \text{PFRAME} & (\ \rho(v\ \text{HOST})\ )_\sigma \end{bmatrix} \end{bmatrix}$$

[2] Sets can obviously be generalized to contain =any of the other kinds of values.

[3] We adopt the convention of writing the value of a set-valued feature without set-brackets when it is a singleton set; e.g. [CLASS *weak*] instead of [CLASS {*weak*}]. Similarly, in descriptions we will drop the $\in$ feature in paths and write $(v\ \text{DEP}) = \text{LT}$ instead of $(v\ \text{DEP}\ \in) = \text{LT}$ or $\text{LT} \in (v\ \text{DEP})$.

Goals of the talk

1. Develop a model of what is on the righthand side of the exponence function, ν, i.e. a theory of the set of v-structures for some set of exponenda in a Vocabulary.
2. Describe a factorial typology of phonological dependence in formal terms
3. Provide an analysis of the English deadjectivizing verbalizer affix *-en*

1 Phonological features

1.1 Phonological representation

- Conditions on mapping to output phonological form
  - Can be underspecified
    - For example, much of English inflection is probably underspecified for [±voice].
  - Can be a memorized, conditioned list
    - For example, the English indefinite determiners (*a/an*) are listed, phonologically conditioned allomorphs.
    - This is the same approach we would take to French liaison.

1.2 Prosodic frame

- Conditions on mapping to prosody
  - For example, SWEAR-insertion in English is sensitive to foot structure.
  - Similarly, *-um-* infixation (Austronesian) is sensitive to syllable structure.

1.3 Prosodic domain

- This specifies in which prosodic domain the v-structure is integrated into prosody according to some definition of prosodic phrasing at p-structure (Bögel 2015, 2021).
  - For example, using · to represent the p-correspondent of the v-structure in question, English geminates can only appear at [PDOMAIN $(\cdot)_\iota$]
  - Similarly, some Germanic prefixes are metrical, [PDOMAIN $(\cdot)_\omega$], while others are extrametrical, [PDOMAIN $(\cdot),(\ \ )_\omega$]; see §4 #5.
    - We use the comma to represent the unordered concatenation of two intonational units; the following equality therefore holds: [PDOMAIN $(\cdot),(\ \ )_\omega$] = { [PDOMAIN $(\cdot)(\ \ )_\omega$] | [PDOMAIN $(\ \ )_\omega(\cdot)$] }.
    - The actual order of comma cases must be set by the DEPENDENCE feature.

1.4 Dependence

- The direction of the prosodic dependency
- Left, right, or both (infix)[4]
  - {LT} := suffix ("I am dependent to the left")
  - {RT} := prefix ("I am dependent to the right")
  - {LT,RT} := infix ("I am dependent to the left and to the right")
- The presence of this feature entails prosodic/phonological dependence.

2 Morphosyntactic features

2.1 Class

- Inflectional class and other purely morphological selectional properties
  - For example, this is where we would capture verb classes and noun classes, such as Latin declensions and conjugations.

2.2 Host

- The value of the HOST attribute is a v-structure.
  - We assume that the ρ-mapping from v-structure to p-structure is sensitive to the HOST feature.
    - If a v-structure α has a HOST v-structure β, then β's realization in p-structure must be prosodified in the PDOMAIN of α's realization.

      (4) HOST *Mapping*
      For all v-structures v, v':
      $(v \text{ HOST}) = v' \Rightarrow \rho(v') \in \rho(v \text{ PDOMAIN})$
    - We will discuss the ρ-mapping and prosodification a little more in §6 below.
  - Most of the features in HOST are features that we have already encountered: PHONREP, PFRAME, DEP, and CLASS (any of which can be underspecified as usual).
  - The HOST can also be specified for the IDENT(ITY) feature, which is either present with the value $+$ or not present at all. Thus, IDENT is effectively privative.
- Note that the HOST feature does not contain HOST.
- This is captured by the *Principle of Local* HOST *Identification* (LHI) in (13) below.
- The LHI uses the *restriction* operator (Kaplan and Wedekind 1993) to ensure that when a HOST is identified, it brings with it all of its features *except* HOST (if it has one).
- Thus, the LHI ensures that an exponent can include information about its HOST, but not its HOST's HOST.

[4] We assume here that circumfixes can be handled as a prefix/suffix combination, as in finite-state approaches (see, e.g., Beesley and Karttunen 2003). However, Bill Foley (p.c.) has suggested to us that there may be 'true' circumfixes that cannot be handled this way. If so, we could supplement DEP values with values like LEDGE (left edge) and REDGE (right edge).

2.2.1 Identity

- The IDENT(ITY) feature captures locality conditions on the c-structural and f-structural context of the host.
- If [IDENTITY +] is present in the HOST, then the exponent in question constrains the identity of its host as follows:

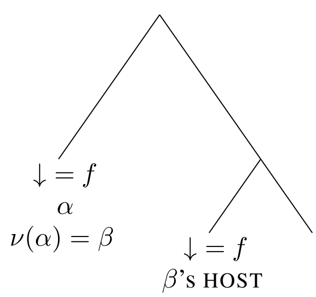
(5) HOST *Identification (Intuition)*
Given β, a v-structure containing the feature [HOST [IDENT +]], and α, a c-structure terminal that β expones (i.e, α is ν-mapped to β), β's HOST is the v-structure that expones the *closest* c-structural terminal to α that maps to the same f-structure as α.

- *Closest* is defined as follows:

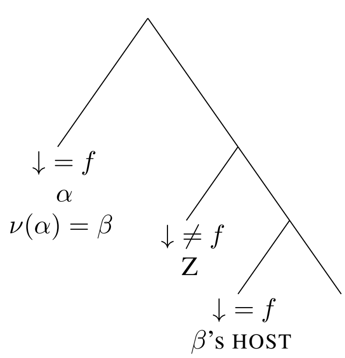
(6) Y is the closest c-structure terminal to X iff
- X c-commands Y; and
- there is no Z such that X c-commands Z and Z c-commands Y.

- The representations in (7–10) sketches two situations in which [IDENT +] is satisfied and two in which it is not. Note that in all cases, α is a c-structure node that corresponds to β, i.e. $\nu(\alpha) = \beta$.

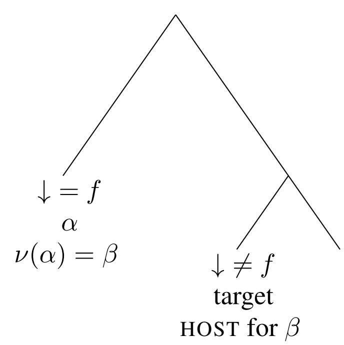
(7) [IDENT +] satisfied:
β's HOST is the closest c-structure node to α that maps to the same f-structure as α.



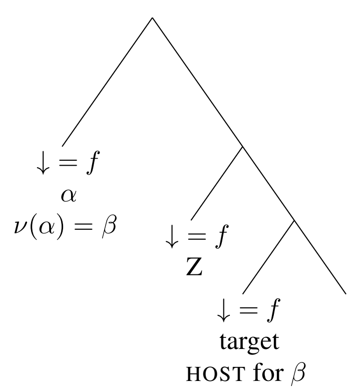
(8) [IDENT +] satisfied:
β's HOST is the closest c-structure node to α that maps to the same f-structure as α.



(9) [IDENT +] not satisfied:
Target HOST for β is the closest terminal to α, but α and β's HOST do not map to same f-structure.



(10) [IDENT +] not satisfied:
Target HOST for β is not the closest c-structure node to α that maps to the same f-structure as α.



- We can use the term *f-domain* for the set of c-structure nodes that map to the same f-structure as some c-structure node α.
  - We define a function to yield a node's f-domain.

  (11) For all c-structure nodes, n, in the set of c-structure nodes N for some c-structure,
  $$\textbf{f-domain}(n) = \{n' \mid n' \in N \wedge \phi(n') = \phi(n)\}$$

- We define a function to calculate the closest c-structure terminal to a node, based on the definition in (6) above.

(12) For all c-structure nodes, n, n', n'', in the set of c-structure terminal nodes T for some c-structure,
$$\textbf{closest}(n, n') \Leftrightarrow \textbf{c-command}(n, n') \wedge \neg[\textbf{c-command}(n, n'') \wedge \textbf{c-command}(n'', n')] \wedge n \neq n'$$

- We can capture the [IDENT +] constraint with the following global constraint on the c-structure/v-structure interface:

(13) *Local* HOST *Identification (LHI)*

For all c-structures, n, n', in the set of c-structure nodes N for some c-structure,
$$(\nu(n')\ \textsc{ident}) = + \Rightarrow \textbf{closest}(n, n')\ \wedge\ n' \in \textbf{f-domain}(n)\ \wedge\ (\nu(n)\ \textsc{host}) = \nu(n')\backslash\textsc{host}$$

- The definition in (13) uses the restriction operator (Kaplan and Wedekind 1993) , $\backslash$, to state that $\nu(n)$'s HOST is the v-structure $\nu(n')$, *except* for any HOST information that $\nu(n')$ may contain.

- Note that this allows us to capture the notion of *bound stems*[5] as in:

(14) habl- Spanish
'talk'

- A bound stem is an exponent that is listed as [IDENT +] on the righthand side of some vocabulary item.

- That is, there are two ways for [IDENT +] to be marked on a v-structure:

1. By being specified as such on the righthand side of a vocabulary item
2. By the v-structure being the HOST for some affix

- As a consequence, *exponence* (the ν-mapping) can be sensitive to [IDENT +] but, *realization* (the ρ-mapping) can be sensitive to [IDENT +] too, and importantly, *even in cases where exponence is not*.

[5] Bound stems are common in languages that require all roots to be inflected, such as Romance languages. Unqualified bound stems are harder to find in languages like English.

Interim summary

(15) $$\langle\, [\mathrm{C}_1, \ldots, \mathrm{C}_n], \mathrm{F} \cup \mathrm{G} \cup \mathrm{I}\,\rangle \;\xrightarrow{\nu}\; v\text{-}s\begin{bmatrix} \textsc{phon(ological)rep(resentation)} & \textit{phon. realization \& conditions} \\ & \text{string} \\ \textsc{p(rosodic)frame} & \textit{prosodic unit} \\ & \text{string wrapped around v-s} \\ \textsc{p(rosodic)domain} & \textit{prosodic unit} \\ & \text{string wrapped around} \\ & \textsc{pframe}\text{ value} \\ \textsc{dep(endence)} & \left\{\textsc{lt}, \textsc{rt}\right\} \\ & \text{set of symbols} \\ \textsc{class} & \left\{\textit{inflectional classes}\right\} \\ & \text{set of symbols} \\ \textsc{host} & v\text{-}s\begin{bmatrix} \textsc{ident(ity)} & + \\ & \text{symbol} \\ \textsc{phonrep} & \ldots \\ \textsc{pframe} & \ldots \\ \textsc{dep} & \ldots \\ \textsc{class} & \ldots \end{bmatrix} \end{bmatrix}$$

3 MostSpecific

- $\mathrm{L_R}$FG posits a constraint on the expression of phonological information, i.e. *morphophonology*, which we have called **MostSpecific**.
- **MostSpecific**(α, β) returns whichever vocabulary item has the most restrictions on its phonological context.
- Let V^o be the co-domain of the exponence function ν in some language L, i.e. the set of exponents/outputs of vocabulary items in L. We write $V^o(\alpha)$ to indicate the co-domain of some particular VI, α — i.e., the output vocabulary structure.
- The proper subsumption relation on feature structures — i.e., v-structures — is used to capture the intuition (below).

  (16) Given two Vocabulary Items, α and β,

  $$\textbf{MostSpecific}(\alpha, \beta) = \begin{cases} \alpha \textbf{ if } V^o(\beta) \sqsubset V^o(\alpha) \\ \beta \textbf{ if } V^o(\alpha) \sqsubset V^o(\beta) \\ \bot \textbf{ otherwise} \end{cases}$$

- The intuition behind **MostSpecific** is to prefer affixes whenever possible. In terms of information encoded in vocabulary items, choose the VI whose output v-structure contains more information, i.e. more features.
- For example, if English comparative *-er*, an affix, and *more*, a free form, are in competition, then **MostSpecific** will select *-er*.
- Similarly, if English verbal inflection *-s* and *does* are in competition, then **MostSpecific** will select *-s*.

4 DEPENDENCY & IDENT: Classifying forms

1. Free form

(17) $\begin{bmatrix} \text{PHONREP} & \ldots \\ \text{PFRAME} & \ldots \end{bmatrix}$

2. Clitic$_{\text{a}}$ (*leaners/simple clitics*)

(18) $\begin{bmatrix} \text{PHONREP} & \ldots \\ \text{PFRAME} & \ldots \\ \text{DEP} & \ldots \end{bmatrix}$

- For example, the English possessive *'s* and auxiliary *'ll* are specified as [DEP LT] because they lean on the preceding element.
- However, they are not fussy about what that element is.[6]

(19) English possessive *'s*

a. The car's fender

b. The car you are in's fender

- We assume on general grounds that *'s* expones the category D.

(20) English "contractions"

a. The person who arrives first'll leave last

b. The person who finds them'll leave last

- We assume on general grounds that *'ll* expones the category T.

3. Clitic$_{\text{b}}$ (phonological clitics)

- There is another kind of clitic whose dependence properties are not determined by v-structure, but rather just by their phonology.
  - For example, in the Frans Plank example, *drink a pint of milk*, the prosodic constituency is (*drinka*) (*pinta*) (*milk*).
  - The phonological dependence of these examples is entirely a product of prosodic structure i) footing together *drink* and the reduced form of the indefinite determiner *a* and ii) footing together *pint* and the reduced form of the preposition *of*.
  - In other words, this kind of prosodic phrasing is captured in p-structure (Bögel 2015, 2021), and simply arises from the fact that the relevant functional words (in this case, *a* and *of*) have /ə/ allomorphs.
- Therefore, the clitic$_{\text{b}}$ variety in fact does not have a DEP feature in v-structure at all, because its surface dependence is no more lexically conditioned than the surface dependence of *drink* or *pint*.
- Thus, the v-structure template for clitic$_{\text{b}}$ is identical to the one for free forms in (17) above.

[6] Note that these always happen to lean on a DP, but the particular element at the right edge varies, and it's this element that is what the clitic's phonological form depends on; for example, it determines voicing assimilation: *the cat's meow* (voiceless) vs. *the car you are in's fender* (voiced).

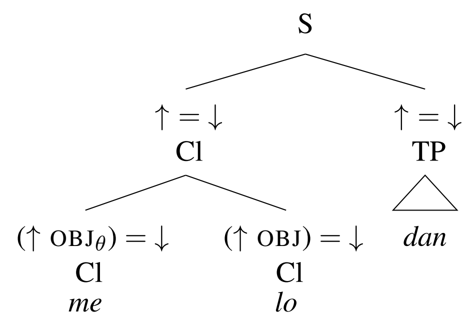
4. Clitic_c (*syntactic clitics*)

- We assume that *syntactic clitics*[7] are associated with a clitic-specific syntactic category, Cl, in the c-structure (Bresnan et al. 2016: 144–145, Arregi and Nevins 2018).
- This differentiates syntactic clitics from simple clitics (clitic_a) above. Indeed, elements of category Cl can be free-standing, affixal, or simple clitics/leaners, depending on their v-structure properties.
- For example, this is how we would treat Romance object clitics, and it is how we have treated certain Ojibwe agreement clitics (Melchin et al. 2020b).[8,9]

(21) me lo d-a-n. Spanish
1.SG 3.SG.MASC give-TV-PL
'They give it to me.'

(22) $\text{Cl} \rightarrow \underset{(\uparrow \text{GF}) = \downarrow}{\text{Cl}} \quad \underset{(\uparrow \text{GF}) = \downarrow}{\text{Cl}}$

(23)



5. Affix

- Affixes arise from the combination of some DEP value and [IDENT +].

(24) $\begin{bmatrix} \text{PHONREP} & \ldots \\ \text{PFRAME} & \ldots \\ \text{DEP} & \ldots \\ \text{HOST} & \begin{bmatrix} \text{IDENT} & + \end{bmatrix} \end{bmatrix}$

- Furthermore, we adopt a prosodic domain account for the different stress and phonotactic restrictions on affixation.
- German prefixes whose domain is [PDOMAIN $(\cdot)_\omega$] are stressed.

  (25) *uralten* ('very old') German prefix; *(úr)(alten)* not *(urálten)*

- German prefixes whose domain is [PDOMAIN $(\cdot)_\iota$] are unstressed (they are extrametrical).

  (26) *gealtert* ('aged') German prefix; *ge(áltert)* not *(gé)(altert)*

[7] We do not make reference to 'special clitics.' We avoid this term simply because it tends to mean somewhat different things in lexicalist LFG circles than it does in DM circles, although definitions overlap. Note that it is not our intent to treat *syntactic clitic* and *special clitic* as equivalent terms. We expect a full theory of special clitics to deploy many of our morphosyntactic categories, including affixes and free forms.

[8] Example (21) is declarative. In the imperative, *den=me=lo*, the clitics appear on the right side of the verb, rather than on the left as in (21), but the clitic constituent retains its order.

[9] The gloss TV stands for "theme vowel."

4.1 Factorial typology over DEP × IDENT

- This yields a factorial typology of major morphological kinds, as shown in Table 1.
- Note that (• FEAT) and ¬(• FEAT) are standard LFG notation for indicating respectively the obligatory presence or absence of feature FEAT in the structure designated by •.

	[• IDENT +]	¬[• IDENT +]
[• DEP]	*affix*	*$clitic_a$* *(leaner/simple clitic)*
¬[• DEP]	*some particles* *some prepositions*	*free form* *$clitic_b$* *(phonological clitic)*

Table 1: A factorial typology of major morphological kinds

5 An example: *-en*

- The English affix *-en*, as in *blacken*, is perfectly productive assuming certain phonological well-formedness conditions:

1. This affix is consistently pronounced as a syllable with a reduced vowel and an alveolar nasal coda.
   ∴ [PHONREP /ən/]

2. The affix is a syllable that is the last in its foot.
   ∴ [PFRAME $(\ldots(\,\cdot\,)_\sigma)_{ft}$]

3. The affix form is subject to local word-level phonotactics.
   ∴ [PDOMAIN $(\,\cdot\,)_\omega$]

4. The affix is dependent to its left; i.e. it is a suffix.
   ∴ [DEP LT]

5. The resulting verb is a weak verb (in the Germanic sense); e.g. it takes *-ed* in the past participle, unlike strong verbs like *take*, which take the past participle affix *-en*. For the purpose of illustration, we identify two classes in English, *weak* and *strong*.[10]
   ∴ [CLASS WEAK]

7. The affix 'lowers' to the head of the complement of the affix.
   ∴ [HOST IDENT +]

8. The affixed form must meet phonological and prosodic conditions on the host.
   - The output form of the base must be no longer than one syllable and end in an obstruent, optionally preceded by a sonorant (per Halle 1973).[11]
     - For example, *soften* is legal despite a seemingly illegal base, because the final /t/ in the base is not present in the output [sɑfən].
     - We know this is a phonological constraint on the host and not a general phonological rule in English, because unaffixed forms with similar phonology are legal (e.g., **dryen* but *lion*, **dimmen* but *women*).

   ∴ [HOST PHONREP /. . . ([**son**])[**obs**]/]
   ∴ [HOST PFRAME $(\ldots)_\sigma$]

9. The affix is a deadjectivizing verbalize. As is common in Distributed Morphology, we assume multiple subvarieties of categories, such as subvarieties of little v (for example, this is how we would capture theme vowel selection in Spanish). The fact that *-en* is deadjectivizing is a consequence of c-structural head adjunction of little a to the particular little v that *-en* is the exponent of. The use of adjunction allows the selectional history to be transmitted through the c-structure:[12]

(27) $\mathrm{v_a} \;\rightarrow\; \underset{\uparrow=\downarrow}{\mathrm{v_a}} \quad \underset{\uparrow=\downarrow}{\mathrm{a}}$

(28) $$\left\langle\, [\mathrm{v_a}],\ \lambda P \lambda e.\mathbf{cause}(\mathbf{become}(P(e))) : ((\uparrow_\sigma \textsc{event}) \multimap \uparrow_\sigma) \multimap ((\uparrow_\sigma \textsc{event}) \multimap \uparrow_\sigma) \,\right\rangle \xrightarrow{\nu} v\begin{bmatrix} \textsc{phonrep} & /\text{ən}/ \\ \textsc{pframe} & ((\quad)(\,\cdot\,)_\sigma)_{ft} \\ \textsc{pdomain} & (\,(v\ \textsc{pframe})\,)_\omega \\ \textsc{dep} & \textsc{lt} \\ \textsc{class} & \textsc{weak} \\ \textsc{host} & \begin{bmatrix} \textsc{ident} & + \\ \textsc{phonrep} & /\ldots([\mathbf{son}])[\mathbf{obs}]/ \\ \textsc{pframe} & (\,\rho(v\ \textsc{host})\,)_\sigma \end{bmatrix} \end{bmatrix}$$

[10] This is meant to be illustrative of the feature CLASS. Contemporary English probably does not have active CLASS features; rather, forms with the past participle affix *-en* are simply irregular.

[11] We are presenting an unadulterated version of Halle's (1973) theory, but we are aware of complications, such as the well-formedness of *crispen*, which we set aside here.

[12] This phrase-structural approach replaces the feature TYPE in the previous brief presentation of v-structure in Asudeh and Siddiqi (2022a,b). This allows us to capture an attested transitive property of this kind of selection that TYPE failed to capture (Oleg Belyaev, p.c.; Belyaev 2023).

- Using • to represent "this v-structure" and · to represent "the p-structure correspondent of this v-structure," i.e. $\rho(\bullet)$, the equivalent description is shown in (29).

(29)
$(\bullet\ \text{PHONREP}) = /\text{ən}/$
$(\bullet\ \text{PFRAME}) = ((\ \ )(\ \cdot\ )_\sigma)_{ft}$
$(\bullet\ \text{PDOMAIN}) = (\ \rho(\bullet\ \text{PFRAME})\ )_\omega$
$(\bullet\ \text{DEPENDENCE}) = \text{LT}$
$(\bullet\ \text{CLASS}) = \text{WEAK}$
$(\bullet\ \text{HOST IDENT}) = +$
$(\bullet\ \text{HOST PHONREP}) =_c /\ldots([\textbf{son}])[\textbf{obs}]/$
$(\bullet\ \text{HOST PFRAME}) =_c (\ \rho(\bullet\ \text{HOST})\ )_\sigma$

- We can capture the general capacity to specify HOST content through this template:[13]

(30) HOST(X,PR,PF,D,C) :=
$\text{X} = + \Rightarrow (\bullet\ \text{HOST IDENTITY}) = +$
$\text{PR} \neq \textbf{Id} \Rightarrow (\bullet\ \text{HOST PHONREP}) =_c \text{PR}$
$\text{PF} \neq \textbf{Id} \Rightarrow (\bullet\ \text{HOST PFRAME}) =_c \text{PF}$
$\text{D} \neq \textbf{Id} \Rightarrow (\bullet\ \text{HOST DEP}) =_c \text{D}$
$\text{C} \neq \textbf{Id} \Rightarrow (\bullet\ \text{HOST CLASS}) =_c \text{C}$

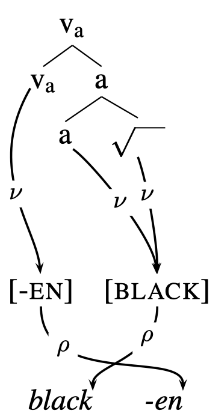
- With (30) in hand, we can rewrite (29) as:

(31)
$(\bullet\ \text{PHONREP}) = /\text{ən}/$
$(\bullet\ \text{PFRAME}) = ((\ \ )(\ \cdot\ )_\sigma)_{ft}$
$(\bullet\ \text{PDOMAIN}) = (\ \rho(\bullet\ \text{PFRAME})\ )_\omega$
$(\bullet\ \text{DEPENDENCE}) = \text{LT}$
$(\bullet\ \text{CLASS}) = \text{WEAK}$
@HOST($+$, $/\ldots([\textbf{son}])[\textbf{obs}]/$, $(\ \ )_\sigma$, __, __)

- Any underspecified argument to a template is understood as an instance of the appropriate **Id** identity element (see footnote 13).

- Note that the re-ordering of the affix and host happens at p(rosodic)-structure, via the ρ correspondence function.

- The L_RFG c-structure with additional ρ-mapping indicated is:

(32)


[13] Note that we take the element **Id** to be whatever the appropriate *identity* element is for the argument in question. That is, an underspecified argument to a template returns whatever element is appropriate to combine with the value type in question to yield no change to the value. In the case of v-structure values, **Id** is the empty v-structure, since this can be thought of as unifying with any v-structure α to yield α. In the case of string values, such as the values of PHONREP and PFRAME, **Id** is the empty string, since this concatenates with any string α to yield α. In the case of set values, such as the values of DEP and CLASS, **Id** is the empty set, since this unions with any set A to return A.

- The less marked alternative is a zero-marked form.
- L_{R}FG does not employ zero affixes.
- Zero-marking in L_{R}FG is a result of the fact that *Pac-Man Spanning* is always competing with overt exponence.
- Pac-Man Spanning is the result of the three **MostInformative** constraints (Asudeh and Siddiqi 2022a,b, forthcoming) preferring portmanteaus, whenever the HOST requirements of *-en* are not satisfied.

(33)

Pac-Man spanning	*-en* affixation
to orange	to redden
to yellow	to blacken
* to red	* to orangen
* to black	* to yellowen

6 Mapping to Prosody

- The essence of our morphological analysis of *blacken* is captured by (28) and (32) above.
- However, now it is time to say more about the ρ-mapping, which we base on, e.g., Bögel (2015, 2021).
- That is, given (28), how should the actual output of the ρ-mapping, the p-structure, be represented? Similarly, how should the p-structure of the HOST, *black*, be represented?
- First, recall our principle (4), repeated here:

(4) HOST *Mapping*

For all v-structures v, v':
$(v \text{ HOST}) = v' \;\Rightarrow \rho(v') \in \rho(v \text{ PDOMAIN})$

- Given (4), the mapping must be as follows, where the p-structure is represented as a *p-diagram* (see Bögel 2015, 2021), as in Figure 1.
  - Note that we assume that the SEGMENTS in the p-diagram are mapped to the output *p(honological)-string* by the o-mapping from p-structure to the p-string in our Correspondence Architecture (Asudeh and Siddiqi 2022b, forthcoming).
  - This is why the mapping arrow from PHONREP in each v-structure is annotated $o \circ \rho$.

7 Conclusion

- We have shown how, in L_{R}FG an *exponent* is a vocabulary structure that is ν-mapped from an *exponendum*.
- The exponendum is the lefthand side and the exponent is the righthand side of a *vocabulary item*, a listed mapping in the Vocabulary.
- An exponent is in turn ρ-mapped to a p(rosodic)-structure, its *realization*:

(34) $\textit{exponendum} \xrightarrow{\nu} \textit{exponent} \xrightarrow{\rho} \textit{realization}$

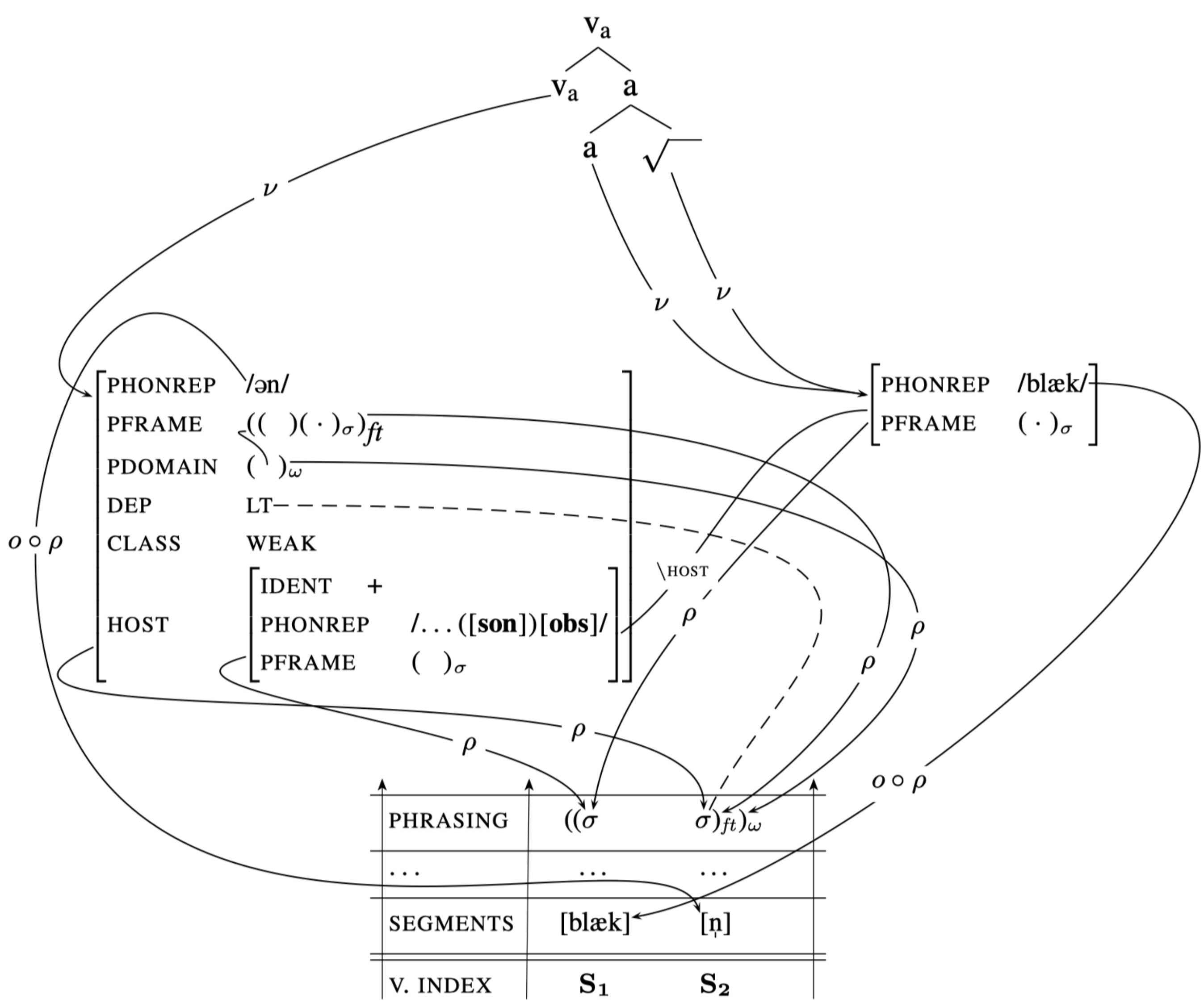


Figure 1: C-structure, v-structures, p-structure, and mappings for *blacken*

References[14]

Arregi, Karlos, and Andrew Nevins. 2018. Beware Occam's Syntactic Razor: Morphotactic Analysis and Spanish Mesoclisis. *Linguistic Inquiry* 49(4): 625–683.

Asudeh, Ash, Tina Bögel, and Daniel Siddiqi. forthcoming. The morphology-phonology interface.

Asudeh, Ash, Paul B. Melchin, and Dan Siddiqi. 2021. Constraints All the Way Down: DM in a Representational Model of Grammar. Ms., Carleton University and University of Rochester. To appear in the *Proceedings of the 37th West Coast Conference on Formal Linguistics*.

Asudeh, Ash, and Daniel Siddiqi. 2022a. Morphology in LFG. To appear in Dalrymple (2022).

—. 2022b. Realizational Morphosemantics in $\mathrm{L_R}$FG.

—. forthcoming. *Constraint-based Morphosyntax*. Cambridge: Cambridge University Press. With Oleg Belyaev, Bronwyn Bjorkman, Tina Bögel, Michael Everdell, Paul B. Melchin, and Will Oxford.

Beesley, Kenneth R., and Lauri Karttunen. 2003. *Finite State Morphology*. Stanford, CA: CSLI Publications.

Belyaev, Oleg. 2023. Ossetic verb morphology in $\mathrm{L_R}$FG. Presented in the LFG 2023 workshop, *Revisiting Lexical Integrity*, July 24, University of Rochester.

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

—. 2021. Function Words at the Interface: A Two-Tier Approach. *Languages* 6(4): 197.

Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler. 2016. *Lexical-Functional Syntax*. Malden, MA: Wiley-Blackwell, 2nd edn.

Dalrymple, Mary, ed. 2022. *The Handbook of Lexical Functional Grammar*. Language Science Press. Forthcoming.

Everdell, Michael, and Paul B. Melchin. 2021. Control the sentence, subordinate the pronoun: on the status of controlled versus non-controlled complement clauses in O'dam. In *Proceedings of the 39th West Coast Conference on Formal Linguistics*. Cascadilla Press. Forthcoming.

Everdell, Michael, Paul B. Melchin, Ash Asudeh, and Daniel Siddiqi. 2021. Beyond c-structure and f-structure: On the argument-adjunct distinction in O'dam. In Miriam Butt, Jamie Y. Findlay, and Ida Toivonen, eds., *Proceedings of the LFG21 Conference*, 125–145. Stanford, CA: CSLI Publications.

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

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

Kaplan, Ronald M., and Jürgen Wedekind. 1993. Restriction and Correspondence-Based Translation. In *Proceedings of the 6th Meeting of the EACL*. European Chapter of the Association of Computational Linguistics, University of Utrecht.

[14]Most of these are $\mathrm{L_R}$FG work and are available at lrfg.online.

Melchin, Paul B., Ash Asudeh, and Dan Siddiqi. 2020a. Ojibwe Agreement in a Representational, Morpheme-based Framework. In Angelica Hernández and M. Emma Butterworth, eds., *Proceedings of the 2020 Canadian Linguistic Association*. Canadian Linguistic Association.

—. 2020b. Ojibwe Agreement in Lexical-Realizational Functional Grammar. In Miriam Butt and Ida Toivonen, eds., *Proceedings of the LFG20 Conference*, 268–288. Stanford, CA: CSLI Publications.

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

Siddiqi, Daniel. 2021. On the Taxonomy of Root Suppletion. In *Proceedings of the 39th West Coast Conference on Formal Linguistics*. Cascadilla Press. Forthcoming.