Comparing Frameworks: A case study of the syntax-morphology interface

WOMP, Princeton University

Terje Lohndal, David Natvig, Michael T. Putnam & Yvonne van Baal

March 22-23, 2024

NTNU

University of Stavanger

PennState

UiT Aurora Centre

Outline

What we hope to achieve in this talk:

- Critically examine an example of a *direct* and an *indirect* approach to the feature-exponency relation
  - Nanosyntax
  - Distributed Morphology
- Case study: Nominal suffixes in varieties of Norwegian
- **Take-home message**: Both approaches are capable of modeling observed alternations, but via rather different architectural assumptions,
  - How to compare and assess their relative strengths and weaknesses?

Modeling realization: An exoskeletal approach

- Emphasis on the way in which syntactic structure determines both the grammatical properties and 'the ultimate fine-grained meanings of lexical items themselves' (Borer 2003: 33).
- A syntax-driven vs. lexicon-driven approach to structure building.
- 'The exoskeletal family'
  - Distributed Morphology (DM)
  - Borer's approach
  - Spanning
  - Nanosyntax
- We will distinguish between a direct mapping and an indirect mapping, using Nanosyntax and Distributed Morphology as examples of such frameworks, respectively (see also Marantz 2023).

An indirect mapping: DM

- An *indirect* mapping system between features and exponents permits a number of operations that can adjust licit syntactic structures post-syntactically prior to the syntax-morphology interface.
- Arregi & Nevins (2012) introduce a number of post-syntactic operations (i.e., *Fission*, *Dissimilation*, *Impoverishment*, *Metathesis*, etc.) that can take place prior to Vocabulary Insertion.

An indirect mapping: Vocabulary Items

- Consider the realization of past tense in English, which we represent with a T(ense)-head endowed with the feature specification of T[+past].
- In order to derive the correct distribution of past tense forms in English, we need to propose additional Vocabulary Items that are contextually specified. A sample of these rules are provided in (1):

(1) Vocabulary Items for T[+past], ordered (Embick 2015: 94)

a. T[+past] $\leftrightarrow$ -t/{$\sqrt{\text{bend}}$, $\sqrt{\text{leave}}$,...}_

b. T[+past] $\leftrightarrow$ -Ø/{$\sqrt{\text{hit}}$, $\sqrt{\text{quit}}$,...}_

c. T[+past] $\leftrightarrow$ -ed

An indirect mapping: Preventing overgeneration

- Two axioms are adopted to prevent overgeneration:
  - Vocabulary Items exist as sets of *lists* for contextually specified exponents that target the same feature (bundles).
  - Listed Vocabulary Items are *ordered* with respect to one another.
- We still need an overarching principle to dictate that the appropriate Vocabulary Item will be selected under the right conditions.
- The strategy opted for here is one that relies on specificity, that is, the insertion of a more specific Vocabulary Item take precedence over those that are less specified (all things being equal).

An indirect mapping: The Subset Principle

- The **Subset Principle** in (2) represents a more precise statement of this specificity condition on Vocabulary Insertion:

  (2) **Subset Principle:** The phonological exponent of a Vocabulary Item is inserted into a position if the item matches all or a subset of the features specified in the terminal morpheme. Insertion does not take place if the Vocabulary Item contains features not present in the morpheme. Where several Vocabulary Items meet the conditions for insertion, the item matching the greatest number of features specified in the term morpheme must be chosen (Halle 1997: 428).

- Thus, the Subset Principle blocks the overapplication of less specific exponency in favor of more specific forms.

A direct mapping

- Although DM and Nanosyntax share a number of architectural similarties, i.e., they are both late-insertion, realizational approaches to the syntax-morphology interface, upon closer inspection, they also differ significantly.
- Caha (2018) lists three primary differences:
  - The nature of the basic building blocks of linguistic structure
  - The conceptualization of phrasal spellout
  - The existence of a separate module of grammar

A direct mapping: Basic building blocks

- Nanosyntax interprets the syntax as the only component of grammar capable of building complex feature structures.
  - 'Prepackaged' feature bundles (cf. DM) not possible
- Consequences:
  - No pre-syntactic lexicon; lexical items interpreted as hierarchical trees (i.e., L-trees)
  - Movement operations in order to achieve congruence at the syntax-morphology interface, as in (3)

(3) **Spellout algorithm** (Caha 2021: 412; Starke 2018: 245):

a. Merge F and spell out.
b. If (a) fails, move the Spec of the complement and spell out.
c. If (b) fails, move the complement of F and spell out.

A direct mapping: Phrasal Spell-Out

- Singleton 'morphemes' often express multiple features.
- Phrasal spellout governed by the *Superset Principle*:

(4) **Superset Principle:** In case a set of syntactic features does not have an identical match in the lexical repertoire, use a lexical form, which contains a superset of the features present in the syntax (Fábregas & Putnam 2020: 40).

- Adopting this principle opens the door to non-terminal lexical insertion.
- In Nanosyntax, "*phrasal spellout* inserts Vocabulary Items into phrasal nodes" (Caha 2018: 58).

A direct mapping: Separate module

- Nanosyntax calls for the abandonment of postsyntactic operations that change constituency and/or linear order; these are common place in DM.
- Need to ensure that all features are exponed:

(5) **The Exhaustive Lexicalization Principle:** All syntactic features present in the derivation must be matched exhaustively with lexical items (Fábregas 2007).

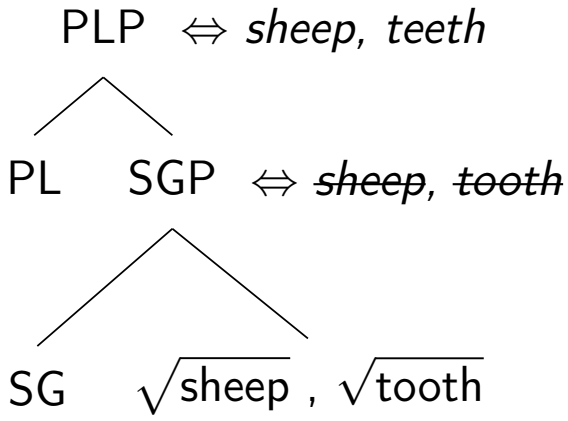
A direct mapping: English plurals

- Nanosyntax reduces allomorphy (in a traditional sense) to the size of L-trees in the mental lexicon (Starke 2004), which pertains to structures for both $\sqrt{\text{roots}}$ and exponents (Blix 2021, Caha 2021, Fischer et al. 2022, Natvig et al. 2023).
- Assumption: The grammatical and semantic distinctions between singular and plural are expressed through the size of trees, such that singular is contained within the plural (Caha 2021).
- We refer to these features as SG and PL for simplicity.

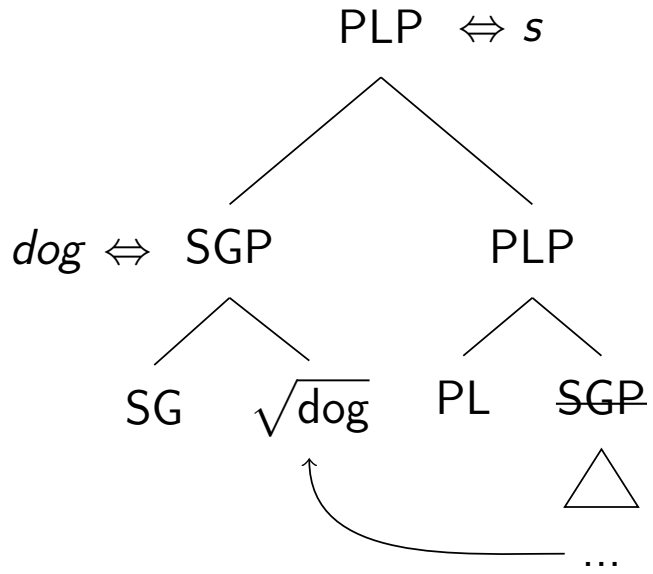
A direct mapping: Cyclic spellout and lexicalization

Spellout occurs at each cycle in the derivation (phrasal spellout) via superset mapping conditions. We get the following S-trees:

(6)



(7)



A direct mapping: Underlying L-trees

- The following structures indicate the underlying L-trees for each form that lexicalize generated structures (S-trees)

(8)
a. [PL[SG[$\sqrt{\text{sheep}}$]]] $\Leftrightarrow$ *sheep*
b. [SG[$\sqrt{\text{tooth}}$]] $\Leftrightarrow$ *tooth*
c. [PL[SG[$\sqrt{\text{tooth}}$]]] $\Leftrightarrow$ *teeth*
d. [SG[$\sqrt{\text{dog}}$]] $\Leftrightarrow$ *dog*
e. [PL] $\Leftrightarrow$ *s*

Nominal suffixes in Norwegian

- Norway has a vast number of dialects, and these dialects often differ in terms of their nominal suffixes. Here we concentrate on the standard written Nynorsk variety.
- There is substantial dialect variation in Norwegian. Some dialects do not distinguish between the three genders in the plural, i.e., they have many syncretic forms.

Nominal suffixes in Norwegian: Segmenting morpemes

- In Table 1, each suffix encodes multiple features.

	Singular		Plural	
Gender	**Indefinite**	**Definite**	**Indefinite**	**Definite**
Masculine	bil 'car'	bil-en	bil-ar	bil-ane
Feminine	dør 'door'	dør-a	dør-er	dør-ene
Neuter	hus 'house'	hus-et	hus-∅	hus-a

Table 1: Nominal suffixes in Nynorsk

- In Table 2, each morpheme encodes a unique feature.

	Singular		Plural	
Gender	**Indefinite**	**Definite**	**Indefinite**	**Definite**
Masculine	bil 'car'	bil-en	bil-a-r	bil-a-ne
Feminine	dør 'door'	dør-a	dør-e-r	dør-e-ne
Neuter	hus 'house'	hus-et	hus-∅	hus-a

Table 2: An alternative segmentation of morphemes in Nynorsk

Nominal suffixes in Norwegian: The structure of DPs

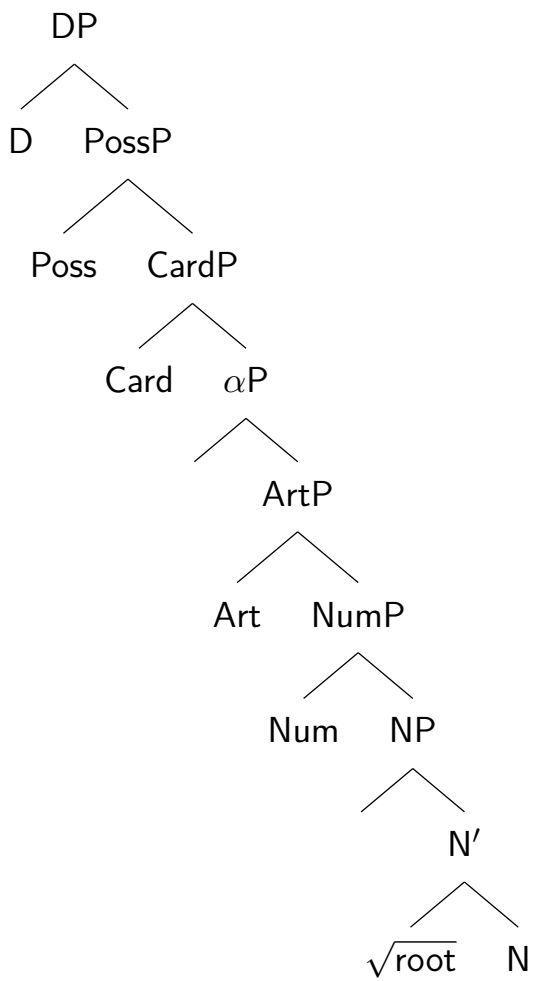
(9) a. determiner/possessive/demonstrative > numeral > adjective > noun + suffix > possessive

b. dei fire stor-e bil-ane mine.
DEF.PL four big-DEF car-DEF.PL my.PL
'my four big cars.'

- Based on Julien's (2005) extensive work, and van Baal et al.'s (in press) small adjustments, the structure of Norwegian nominal phrases is argued to be as shown on the next slide.

Nominal suffixes in Norwegian: The structure of DPs

(10)



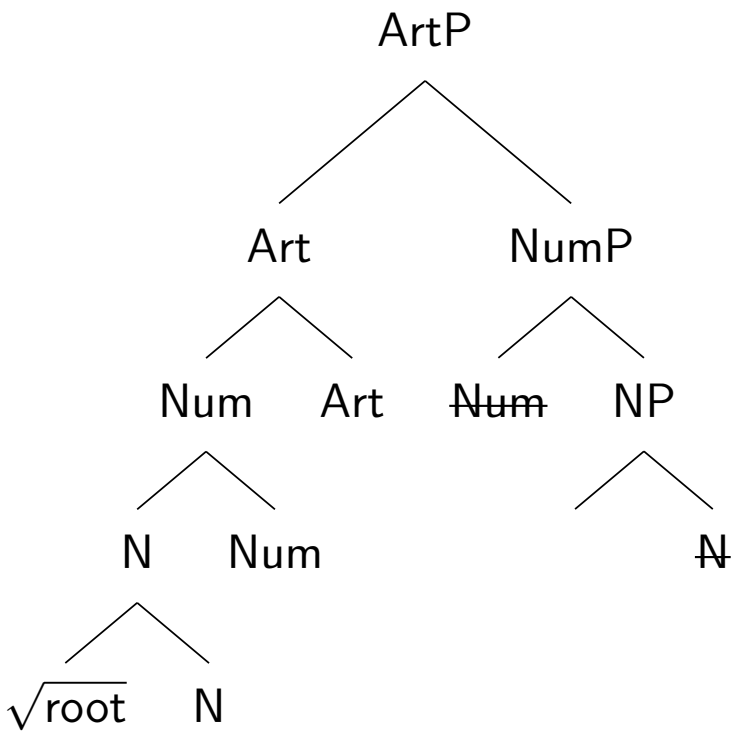
Late-insertion analyses of the Norwegian data

- We now consider two possible analyses of the data. The first is an *indirect* mapping between features and exponents whereas the other is a *direct* mapping analysis.
- A key distinction between the two approaches involves the relationship between generated (morpho)syntactic structures and the representational content of the exponents to which these structures are mapped.

An indirect mapping analysis - syntactic structure

- For an analysis of the nominal suffixes, the relevant part of the tree is from ArtP and downwards. In Julien's (2005) analysis, there is a series of head movements in all noun phrases.

(11)



An indirect mapping analysis - mechanisms and features

- After syntax has created the structure above, there will be a process of Vocabulary Insertion. However, certain post-syntactic operations can take place first.
- **Fusion**: combines features into a feature bundle.
- For Norwegian, the gender, number, and definiteness features that are each present in their own location in the syntax, are bundled together into a single feature bundle. These feature bundles will then be matched onto Vocabulary Items through the Subset Principle, unless other post-syntactic operations alter the feature bundles.
- An indirect mapping approach can work with binary features (e.g., [+DEF] and [-DEF]) or with privative features. Julien (2005: 18-19) works with privative features, and assumes that the absence of a [DEF] feature leads to an indefinite interpretation.

An indirect mapping analysis - Vocabulary Items

- For the eight different suffixes found in Nynorsk, the following Vocabulary Items (VIs) can be proposed:
  - Art[PL, DEF, M] $\leftrightarrow$ -ane
  - Art[PL, DEF, F] $\leftrightarrow$ -ene
  - Art[PL, M] $\leftrightarrow$ -ar
  - Art[PL, F] $\leftrightarrow$ -er
  - Art[DEF, M] $\leftrightarrow$ -en
  - Art[DEF, N] $\leftrightarrow$ -et
  - Art[DEF] $\leftrightarrow$ -a
  - Elsewhere $\leftrightarrow$ -Ø
- The majority of these VIs form a combination of the full feature bundle created by syntax (e.g., [PL, DEF, M]), or a subset thereof (e.g., [PL, M]) and are therefore the only VI that can be inserted.
- There are two VIs that will be inserted for several feature combinations: *-a* and -Ø.

An indirect mapping analysis - the suffix *-a*

- The suffix *-a* is found on definite singular F nouns (e.g. *bok-a* 'the book') as well as definite plural N nouns (e.g. *hus-a* 'the houses').
  - Both contain a [DEF] feature, and we propose that the VI matches [DEF] with *-a*. Since the other VIs are more specific (and for example include a plural or gender feature), these will be inserted in the other contexts.
- For definite plural neuter nouns to match with this VI, another post-syntactic operation is necessary. If syntax creates the feature bundle [PL, DEF, N], the most specific VI would lead to insertion of the suffix *-et*, contrary to fact.
  - We propose an Impoverishment operation, which deletes the [N] feature in the context of a [PL] feature. The result of this Impoverishment is that the feature bundle becomes [PL, DEF] and then the only VI that can be inserted under the Subset Principle is *-a*.

An indirect mapping analysis - the zero suffix

- The other suffix that occurs in multiple contexts is the zero suffix.
- Given the Elsewhere Vocabulary Item, a zero suffix will be inserted in all cases where no more specific VI can be inserted.
- This applies to singular, indefinite nouns, because all other VIs contain either a [DEF] or a [PL] feature not present in the syntax of singular indefinite phrases.
- The zero suffix is also inserted for indefinite, plural, neuter nouns: although there are VIs that include a [PL] feature, these also include a [M] or [F] feature and can hence not be inserted.
- In summary, the VIs and Impoverishment rule described here account for the Nynorsk data.

A direct mapping analysis - cartography

- Nanosyntactic syntactic structures for Norwegian nominals discussed above; 'build out' both NumP and ArtP.
  - Singular: [Num_1]
  - Plural: [Num_2[Num_1]]
  - Indefinite: [Art_1]
  - Definite: [Art_2[Art_1]]
- We draw on the same logic for contrasting gender distinctions (cf. Caha 2021):
  - Neuter: [Ref]
  - Masculine: [Class[Ref]]
  - Feminine: [F[Class[Ref]]]

A direct mapping analysis

$\sqrt{\text{root}}$	Ref	Num_1	Num_2	Art_1	Art_2
hus				∅	
hus				e	
hus				∅	
hus				a	

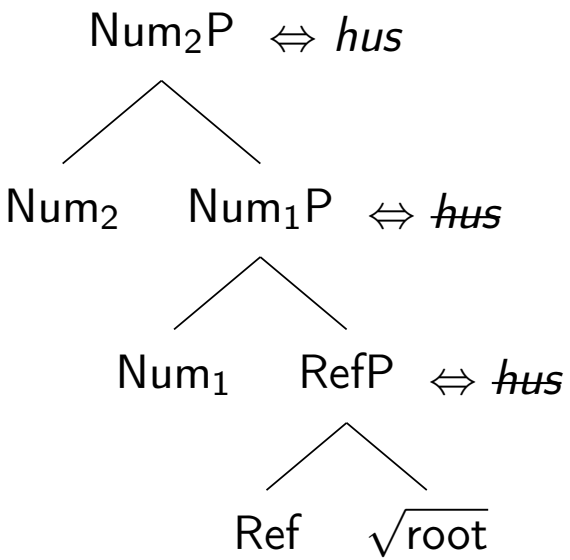
$\sqrt{\text{root}}$	Ref	Class	Num_1	Num_2	Art_1	Art_2
bil					∅	
bil					en	
bil		a			r	
bil		a			ne	

$\sqrt{\text{root}}$	Ref	Class	F	Num_1	Num_2	Art_1	Art_2
dør						∅	
dør						a	
dør				e		r	
dør				e		ne	

A direct mapping analysis

- Consider $\sqrt{\text{hus}}$, a neuter noun, which lexicalizes the feature [Ref] and only [Ref]. In addition, $\sqrt{\text{hus}}$ spells out both [Num_1] for singular and [Num_2] for plural. In short, the form *hus* corresponds to the lexicalization of the entire S-tree up to and including [Num_2], at every stage of the derivation.

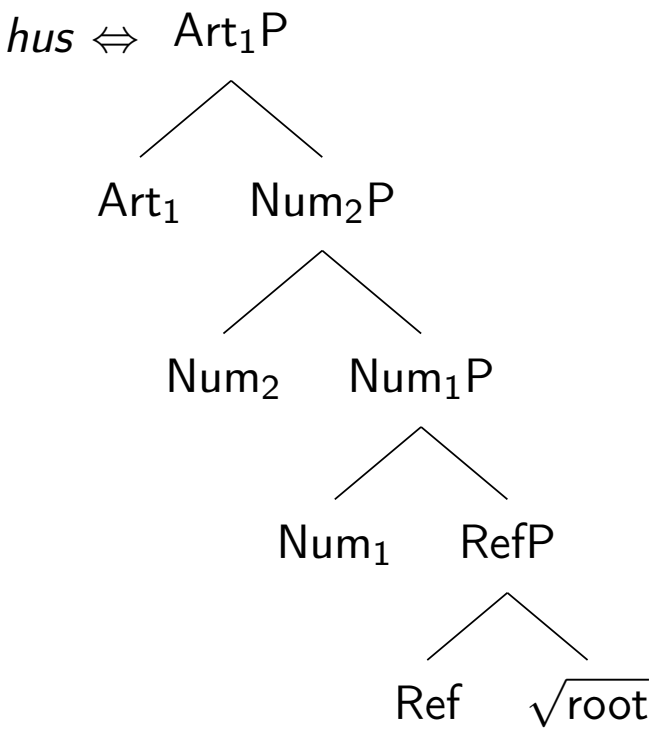
(12)



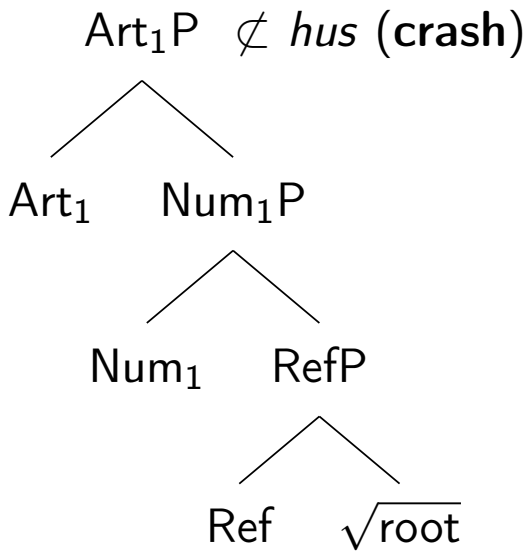
A direct mapping analysis

Neuter (with ArtPs)

(13)



(14)



A direct mapping analysis: masc. and fem.

$\sqrt{\text{root}}$	Ref	Num_1	Num_2	Art_1	Art_2
hus				∅	
hus				e	
hus				∅	
hus				a	

$\sqrt{\text{root}}$	Ref	Class	Num_1	Num_2	Art_1	Art_2
bil					∅	
bil					en	
bil		a			r	
bil		a			ne	

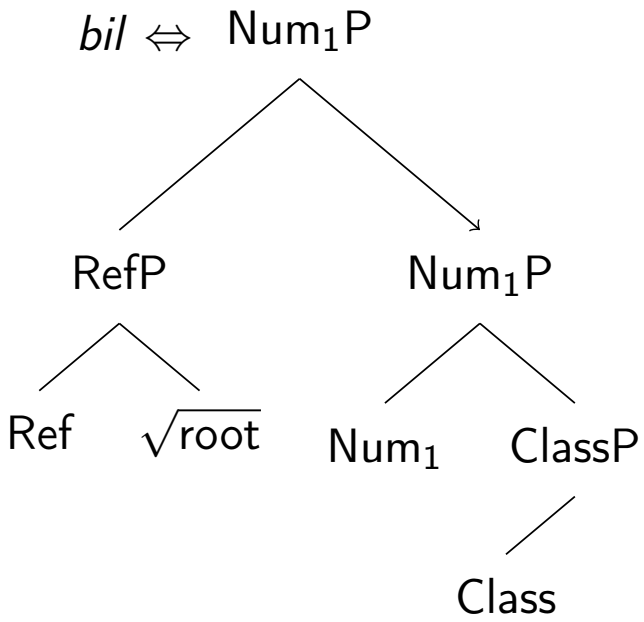
$\sqrt{\text{root}}$	Ref	Class	F	Num_1	Num_2	Art_1	Art_2
dør						∅	
dør						a	
dør				e		r	
dør				e		ne	

A direct mapping analysis

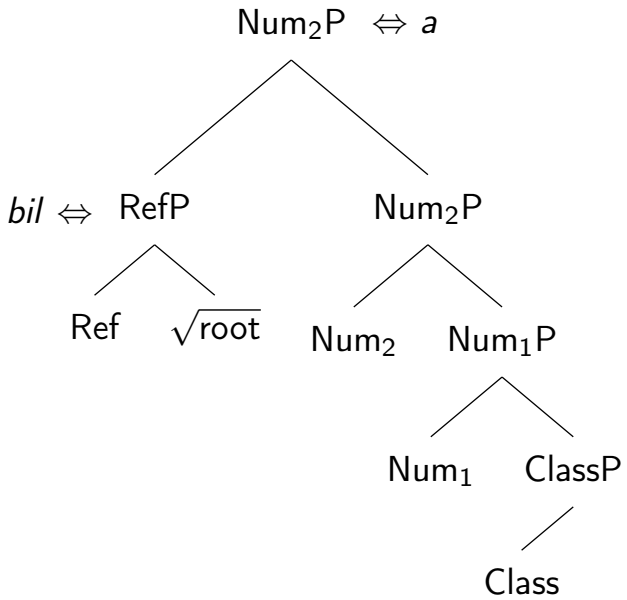
- Regardless of whether or not this turns out to be the correct analysis, one – or both – $\sqrt{\text{roots}}$ must 'shrink' in their lexicalization capacity form singular to plural. Two competing proposals:
  - Backtracking (Caha 2021): A set of operations where previous cycles are undone and licit spell outs are replaced with movement, step by step, until a lexicalizable treelet is created.
  - Partial overwrite (Blix 2021): L-trees for $\sqrt{\text{roots}}$ are stored with branching tree structures, such that $\sqrt{\text{roots}}$ are able to lexicalize S-trees following movement operations.

A direct mapping analysis - masculine

(15)

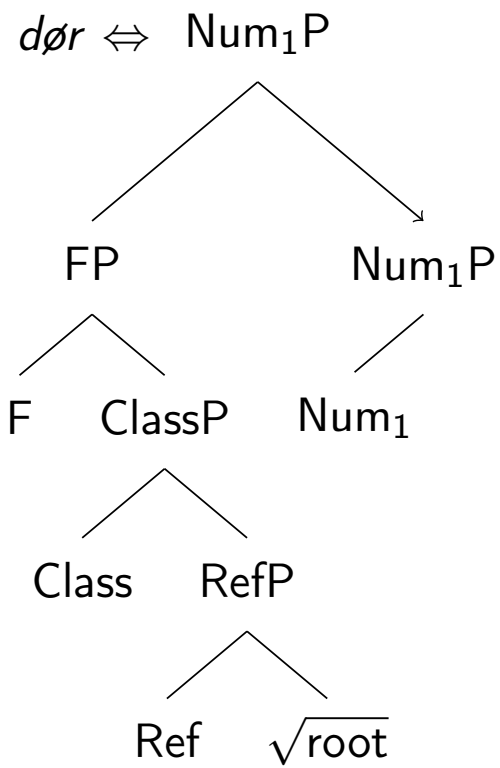


(16)

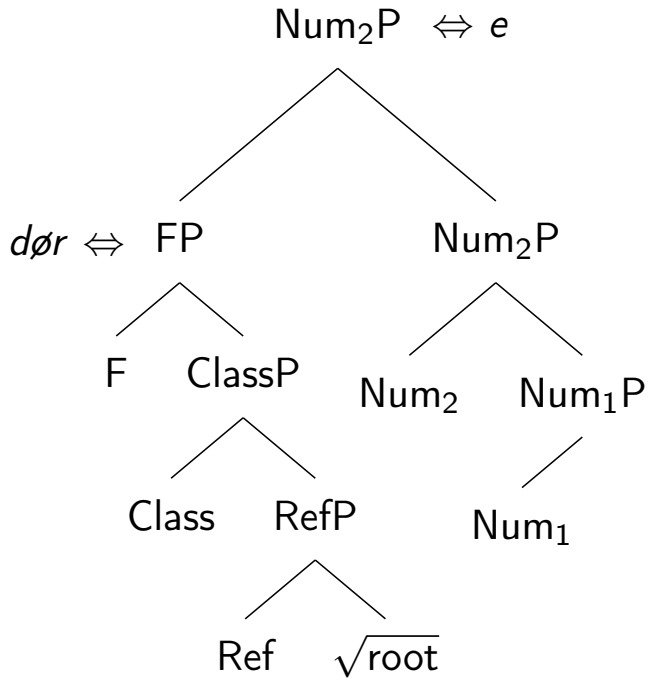


A direct mapping analysis - feminine

(17)



(18)



Framework comparison

- Both approaches are capable of handling the data we have considered; however, they do so through very different mechanisms and they also enforce certain theoretical assumptions.
- DM readily admits zero heads such as null categorizers and null morphemes. Such heads are usually not allowed in Nanosyntax, and various scholars, e.g., Borer (2014), have argued against them.
- Furthermore, Nanosyntax does not have a notion of a 'morphological module', which means that all aspects of morphology need to be encoded in the syntax. Again, this is different in DM where it is possible to insert morphemes after the syntactic derivation is finished. For instance, this enables an analysis whereby declension class markers are inserted post-syntactically (e.g., Kramer 2015, Lohndal & Westergaard 2021).

Conclusions

- On a direct mapping analysis, a particular exponent can in principle correspond to a range of features.
- On an indirect mapping analysis, a particular feature corresponds to one exponent (modulo various adjustments that may affect this relationship).
- The two approaches also make fundamentally different assumptions about the content of the mental lexicon.
    - Whereas an indirect approach like Distributed Morphology assumed a lexicon that consists of three different lists, a direct approach like Nanosyntax eschews any notion of a pre-syntactic lexicon. Instead, lexical items are hierarchical tree structures.

Thank you!

Questions and comments very welcome.