Slavic meets Semitic: Nominal functional categories as underspecified heads[1]

1 Cross-linguistically parallel homophony?

In *Slavic*:

- the morpheme **K** (inflected for gender, number and case; e.g., Czech *-ek*.M.SG, *-ka*.F.SG, *-ko*.N.SG etc.) is systematically *homophonous* with a variety of functional morphemes:
  - a default diminutive formation[2] that can yield a small degree interpretation, or obtain additional pragmatic readings
  - a nominalizer
  - a conceptually[3] female-denoting morpheme
  - a semantic division/number morpheme (pluria tantum, group formation)

⇒ the same morphological form expresses derivational & inflectional morphology, nominality as a categorical distinction, and nominal features/functional heads throughout the extended nominal domain (GENDER, NUMBER, DEGREE, perhaps PERSON)

A very similar range of nominal functions and interpretations is found in *Semitic*, specifically in Arabic dialects:[4]

⇒ the feminine morpheme (**F**) displays a similar range of functional and semantic interpretations, e.g., in Moroccan and Levantine Arabic (LA), with some modulation:
  - in the division/number domain, **F** also individuates, and
  - because of templatic morphology, **F** does not overtly realize a nominalizing head

How does functional/interpretational variability within the nominal domain map to PF uniformity?

- homophony over a number of functional interpretations within a single language, or even a single family of languages, is not surprising in and of itself

[1]We are indebted to Aya Zarka for the Levantine Arabic data and for her insights and discussions about the language. We thank to the Social Sciences and Humanities Research Council of Canada for funding this research (SSHRC Insight Grant #435-2016-1034, Grammatical vs semantic features: the semantics-morphology mapping, and its consequences for syntax; PI: Kučerová). We further thank to Susana Béjar, Ora Matushansky, Alan Munn, Asya Pereltsvaig, Omer Preminger and the audience at the FASL 2020 for their questions and comments.

[2]Slavic languages display a range of morphologically distinct and lexically specified derivational morphemes expressing a diminutive-like meaning (see, for example, Steriopolo 2009, 2013; Wiltschko & Steriopolo 2008; Krizhman 2019). Here, we are only concerned with the default and fully productive form based on **K** as the only morpheme exhibiting the functional variability in the centre of our research investigation.

[3]We use the term 'conceptual gender' to refer to what the older literature calls 'natural,' 'biological' or 'sex-based' gender. See, e.g., Ackerman 2019 for an argument why the terminological change better reflects the intended denotation.

[4]And partially in Hebrew, although with some important differences we won't discuss here.

- but *parallel systematic* homophony over the same set of functional interpretations, and structural restrictions on their syntactic behaviour and distributional/functional gaps *across language families* requires a structural explanation

We argue against accidental homophony:

- instead, both **K** and **F** are morphological realizations of a feature bundle corresponding to a syncategorematic operator (i*[5]), which operates on features of the projection it modifies/attaches to and treats those features as variables (in the sense of Borer 2005 and Acquaviva 2018)

- where the range of i* functional properties is a *function of its syntactic position*

$\Rightarrow$ different functions and interpretations arise from different attachments sites of i* in the extended nominal domain, instead of a series of semantically specified functional heads (e.g., Fassi Fehri 2016, 2017, 2018a,b), or distinct morphemes (e.g., Borer & Ouwayda 2010), modulo structural economy

- $\Rightarrow$ the underlying syntactic underspecification triggers uniform PF realization despite varied syntactic/semantic behavior, modulo independent differences of the surrounding nominal structures and their morphological realization

2 Facets of functional K and F

2.1 Conceptual gender

- **K** and **F** systematically derive female-denoting nouns from MASC nouns

(1)
a. ředitel 'director.M.SG' $\rightarrow$ ředitel**-ka** 'director-**K**.F.SG, a female director' CZECH
b. dyrektor 'director.M.SG' $\rightarrow$ dyrektor**-ka** 'director-**K**.F.SG, a female director' POLISH
c. kot 'cat.M.SG' $\rightarrow$ kot-ka 'cat-**F:**F.SG, she cat' POLISH
d. far 'mouse.M.SG' $\rightarrow$ far**-a** 'mouse-**F**:F.SG, she mouse' LA
e. daktor 'doctor.M.SG $\rightarrow$ daktor**-a** 'doctor-**F**:F.SG., a female doctor' LA

- only derivations from MASC to FEM are attested[6]

- for Slavic, no derivation of female-denoting nouns from NEUTER nouns

(2)

Gender Change	F	K
MASC $\Rightarrow$ FEM	✓	✓
FEM $\Rightarrow$ MASC	×	×
* $\Rightarrow$ NEUT	n/a	×
NEUT $\Rightarrow$ *	n/a	×

[5]The i* notation is loosely based on the i* heads Wood & Marantz (2015) according to whom i*'s functional interpretation is assigned at the interface based on their syntactic configuration. In the present proposal, the functional interpretation is established within narrow syntax, and i* comes with lexical content (polarity head).

[6]See, e.g., Pesetsky 2013 for a generalization about the markedness profile of conceptual-gender derivations.

2.2 Category change

- Slavic **K** systematically *nominalizes* adjectives, verbs, and possibly prepositions (not enough tokens found)
- **K**-based category change productively generates MASC and FEM nouns, never NEUTER

(3) ***K**-based deadjectival nominals:*

a. sodová (voda) 'soda.ADJ (water)' → sodov-ka 'soda-**K**.F.SG, pop' CZECH
b. mielon-y/-a 'minced.ADJ-.masc/fem' → mielon-ka 'luncheon_meat-**K**.F.SG' POLISH

(4) ***K**-based deverbal nominals:*

a. doplnit 'to complement' → dopln-ěk 'complement-**K**.M.SG, a complement' CZECH
b. podpalić 'to ignite' → podpał-ka 'accelerant–**K**.F.SG' POLISH

(5) ***K**-based deprepositional nominals:*
před (domem) 'in front of (a/the house)' → před-ek 'front–**K**.M.SG, (the) front (of something)' CZECH

In *Semitic*

- because of templatic morphology, no nominalizing **F**

(6)

Category Change	F	K
ADJ ⇒ N_{masc}	??	✓
V ⇒ N_{masc}	??	✓
ADJ ⇒ N_{fem}	??	✓
V ⇒ N_{fem}	??	✓
ADJ ⇒ N_{neut}	n/a	×
V ⇒ N_{neut}	n/a	×

2.3 Noun to Noun Conversion

- **K**-based N-to-N conversion systematically derives MASC nouns from FEM nouns, and vice versa[7]
- no derivations from NEUTER, or forming NEUTER

(7) FEM → MASC:

a. kůra 'tree-bark.F.SG' → kor-ek 'bark-**K**.M.SG, cork (a bottle stopper or the substance)' CZ
b. kora 'tree-bark.F.SG' → kor-ek 'bark-**K**.M.SG' cork (a bottle stopper or the substance)' POLISH

(8) MASC → FEM:

[7]Czech data are based on Dokulil et al, 1986. There is also a handful of nouns that appear to preserve gender, i.e., MASC to MASC, FEM to FEM, and NEUTER to NEUTER. There are no formations from NEUTER or forming NEUTER nouns from other genders. N-to-N conversions without gender changes are discussed in appendix A.1.

a. diplomat.M.SG → diplomat-ka 'diplomat-**K**.F.SG; a briefcase, a female diplomat' CZ
b. dyplomat.M.SG → dyplomat-ka 'diplomat-**K**.F.SG; a briefcase, a female diplomat POLISH
c. stolarz.M.SG 'a carpenter' → stolar-ka 'carpenter-**K**.F.SG; carpentry (not a female carpenter)' POLISH

- if pragmatically plausible, the derivation of grammatically FEM nouns from a MASC base tends to be ambiguous with a conceptual gender formation

(9)

N-to-N Conversion	F	K
$N_{masc} \Rightarrow N_{fem}$	n/a	✓
$N_{masc} \Rightarrow N_{masc}$	n/a	×
$N_{fem} \Rightarrow N_{masc}$	n/a	✓
$N_{fem} \Rightarrow N_{fem}$	n/a	×
$N_{masc} \Rightarrow N_{neut}$	n/a	×
$N_{fem} \Rightarrow N_{neut}$	n/a	×
$N_{neut} \Rightarrow N_{masc}$	n/a	×
$N_{neut} \Rightarrow N_{fem}$	n/a	×
$N_{neut} \Rightarrow N_{neut}$	n/a	×

- N-to-N conversions in Semitic are difficult to characterize because of templatic morphology

In the cases discussed in 2.1–2.3 **K** & **F** affect gender, and behave like a head

2.4 Diminutives, their doubles & friends

- DIM formation by **K** is highly productive from all grammatical genders
- **K**-based formation always preserves the gender value of the base noun[8]

(10) NEUTER → NEUTER:
a. jablko 'apple.N.SG' → jablíč-ko 'apple-**K**.N.SG; a small apple' CZ
b. pudło 'box.N.SG' → pudeł-ko 'box.-**K**.N.SG; a small box' POLISH

(11) FEM → FEM:
a. jáma 'pit.F.SG' → jam-ka 'pit-.**K**.F.SG; a small hole' CZ
b. dziura 'hole.F.SG' → dziur-ka 'hole-**K**.F.SG; a small hole' POLISH

(12) MASC → MASC:
a. słup 'pole.M.SG' → słup-ek 'pole-**K**.M.SG; a small pole' POLISH
b. stůl 'table.M.SG' → stol-ek 'table-.**K**.M.SG; a small table' CZ

- in Semitic, the primary DIM derivation yields a stem-internal alternation (a specific template)

[8]Slavic and Arabic diminutives thus differ from diminutives in German or Dutch that change the gender of the base. The differences go beyond gender: German and Dutch diminutives, unlike their Slavic and Arabic counterparts, individuate mass nouns. Also, to our knowledge, gender changing DIMS cannot double.

- the DIM derivation applies to all genders, and as in Slavic, does not change the gender of the base

(13) FEM → FEM:
  a. daʔera 'circle.F.SG → dowerra 'a small circle.F.SG' LA
  b. bent 'girl.F.SG → bannotta 'a little girl.F.SG' LA

(14) MASC → MASC:
  a. arnab 'rabbit.M.SG' → arnub 'a small rabbit.M.SG' LA
  b. mHammad.M.SG (proper name) → Hammod.M.SG (a familiar 'diminutive' form of the proper name) LA

Double-diminutive formations:

- both Slavic and Semitic exhibit double-diminutive formation
- i.e., an additional diminutive-like morpheme is added to a primary diminutive

In Slavic:

- this derivation involves doubling **K**, with the important clarification that only the outside **K** morphologically displays ϕ-features:

(15) stůl.M.SG 'a table' → stol-ek 'table-**K**.M.SG, a small table' → stol-eč-ek 'table-**K**.M.SG-**K**.M.SG, a very small table' CZ

In Arabic:

- the double-formation combines the stem-internal and a stem-external derivation[9]
- where the stem-external derivation is realized as **F**

(16) a. arnab 'rabbit.M.SG' → arnub 'rabbit.DIM.M.SG' → arnub-i 'rabbit.DIM.M.SG-**F**:SG; a cute small rabbit' LA
  b. mHammad.M.SG (proper name) → Hammod..DIM.M.SG' → Hammod-i.DIM.M.SG-**F**:SG; cute/sweet Hammod' LA
  c. Aya.F.SG (proper name) → Ayooš..DIM.F.SG' → Ayoosh-i.DIM.F.SG-**F**:SG; cute/sweet Aya' LA

- the stem-external morpheme **F** is both semantically and syntactically an *adjunct* (e.g., Wiltschko & Steriopolo 2008 for an argument that some diminutives are structurally adjuncts, while some are functional heads)
- the most telling piece of evidence is that the stem-external **F**, technically FEM.SG., is invisible for agree[10] and that in both systems DIMs can be doubled (or even trippled)

[9] In Hebrew, we also see two specific morphemes, where only the second one if **F** (e.g., DeBelder et al 2019).

[10] When the relevant context is provided, *al-arnub-i* can also mean a female bunny. In this case **F** is a morphological reflex of conceptual gender instead of a double-diminutive. In this case, the nominal triggers feminine agreement:

(i) al-arnub-i nam-et b-Hodn-ii
the-rabbit.DIM.M.SG-**F**:SG sleep.3PST-F.SG in-lap-my
'the she-bunny slept in my lap.'

(17) al-arnub-i nam b-Hodn-ii
the-rabbit.DIM.M.SG-**F**:SG sleep.3M.SG.PST in-lap-my
'the cute bunny slept in my lap.'

Doubling yields:

- additional semantic readings ⇒ higher degree of a small size, and
- additional pragmatic readings ⇒ affectionate (e.g., Jurafsky 1996, Dressler & Barbaresi 1994, Fassi Fehri 2017)

Augmentatives:

- absent in Slavic[11]
- as with diminutives, the stem-internal derivation yields a pure augmentative interpretation (i.e., a large size of the nominal denoted by the base), and the stem-external derivation either adds a higher degree of large size (or importance etc) or adds additional pragmatic readings
- as with diminutives, the stem-external **F** does not change the grammatical gender of the base nominal ⇒ an *adjunct*

(18) raaḥil.M.SG. 'traveler' → raḥḥaal.AUG.M.SG. 'big traveler' → raḥḥaal**-at** raḥḥaal.AUG.M.SG.-**F**:SG 'famous big traveler' MOROCCAN A.; Fassi Fehri 2016: 238, (40)

In the cases discussed in 2.4, **K** and **F** affect degree semantics, and behave like syntactic adjuncts

2.5 Individuation

- **F** productively individuates (e.g., Zabbal 2002, Acquaviva 2008, Ouwayda 2014, Fassi Fehri 2016, 2018); it forms:
  - singulatives, i.e., individuated collective nouns (called batch nouns in Borer & Ouwayda 2010), and
  - individuated events
- individuated nominals become grammatical feminine (i.e., they trigger feminine agreement etc.), i.e., **F** turns a MASC nominal into a FEM
- no MASC-based individuation formation

(19) a. Tabšuur 'chalk' (batch noun) → Tabšuur**-a** 'chalk-**F**:SG, a piece of chalk' LA
b. Saxr 'stone' (batch noun) → Saxr**-a** 'stone-**F**:SG; a piece of stone' LA

[11] Although there are other lexically specified derivational morphemes that yield related meanings. See, e.g., Steriopolo 2009, 2013, Krizhman 2019).

(20) a. raqaṣa raqṣ-an
danced dance-ACC
'he danced some dancing'
b. raqaṣa raqṣ-at-an/ raqṣ-at-ayn
danced dance-**F**:SG-ACC/ dancedance-**F**:SG-DU
'he danced a dance/ two dances'
MOROCCAN A.; Fassi Fehri 2016: 226, (11)

(21) a. raqaṣa raqṣ
danced dance-ACC
'he danced some dancing'
b. raqaṣa raqṣ-a/ raqṣ-at-ayn
danced dance-**F**:SG-ACC/ dancedance-**F**:SG-DU
'he danced a dance/ two dances'
LA

- although certain additional restrictions apply (see, e.g., Borer & Ouwayda 2010),[12] both singulatives and individuated events can be further pluralized

2.6 Group formation

- in Semitic, **F** productively derives group formation (Borer 2005, Ouwyada 2014, Kramer & Winchester 2018)
- from the morpho-syntactic point of view, a MASC noun turns into a FEM singular noun
- unlike singulatives, these group denoting nouns *cannot* be pluralized,[13] suggesting that semantically these are *aggregates*, i.e., maximized units formed from an already individuated content[14]

(22) mtdyyen 'religious.M.SG, a believer' → mtdyn-i 'religious-**F**.SG, a religious group' LA

- in Slavic, group formation by **K** is restricted to numerals, (23-a), quantifiers (Veselovská 2018), and pluralia tantum (Dokulil et al. 1986)
- unlike in Semitic, these group-denoting nominals can be pluralized, with the numerals and quantifiers behaving like regular plurals, and pluralia tantum requiring a counting morphology specific to aggregates (Grimm & Dočekal 2020)

(23) a. dvě děvčata 'two girls' → dvoj-ka děvčat 'two-**K**.F.SG girls.GEN, a group of two girls' CZ

[12]The singulative must be definite or modified by an agreeing numeral (not all Arabic numerals combine with plural nominals; only the numerals that do license pluralized singulatives); adjectival modification is sufficient for pluralized individuated events. Note that the requirement to be further modified is attested with other individuating tools in the languages, for instance, with specific indefinites.

[13]The corresponding form exists but it means, for example, a group of female believers, instead of a plurality of groups of believers.

[14]In Slavic aggregates of this type are formed by NEUTER (Grimm & Dočekal 2019, Arsenijević 2013). We put neuter formations aside because they are orthogonal to the main question of this talk.

b. pár děvčat 'a few girls' → pár-ek děvčat 'couple-**K**.F.SG girls.GEN, a group of two girls' CZ

c. nůž-ky 'knife–**K**.PL, scissors' běž-ky 'run-**K**.PL, cross-country skis', sjezdov-ky 'downhill_ride-**K**.PL, downhill skis' CZ

In the cases discussed in 2.5–2.6 **K** & **F** affect number, and behave like a head

3 The case for i*

The core idea:

- we argue that **K** and **F** are morphological realizations of an underspecified functional head, which we call i* (loosely inspired by the interface-sensitive i* of Wood & Marantz 2015), in the context of an extended nominal projection

⇒ i* is a a *polarity* (POL) operator[15]: a function that takes a specific feature, or group of features of its sister as an argument and returns the negated value of the feature

- since functional head is defined by its features, the output of i* returns the same 'category' as the feature(s) of its sister[16]

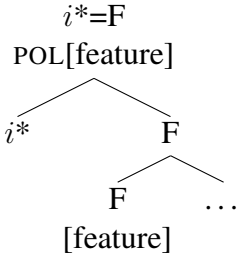
⇒ the functional interpretation of i* is a function of its structural position ⇒ i* takes its core properties from the head whose features it modifies

- ⇒ when i* attaches to a category defining head, then it functions as a category defining head; when it attaches to an individuating head, then it functions as a an individuating head etc.

- since i* is underdefined it can be merged:
  - at any level within the extended nominal domain, as long as the relevant projection contains a feature that is in the domain of the polarity function,
  - to the output of the merge of a head, or a specifier, or as an adjunct

Schematically:

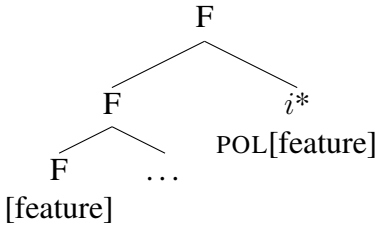
(24) a. When the feature output of i* projects:



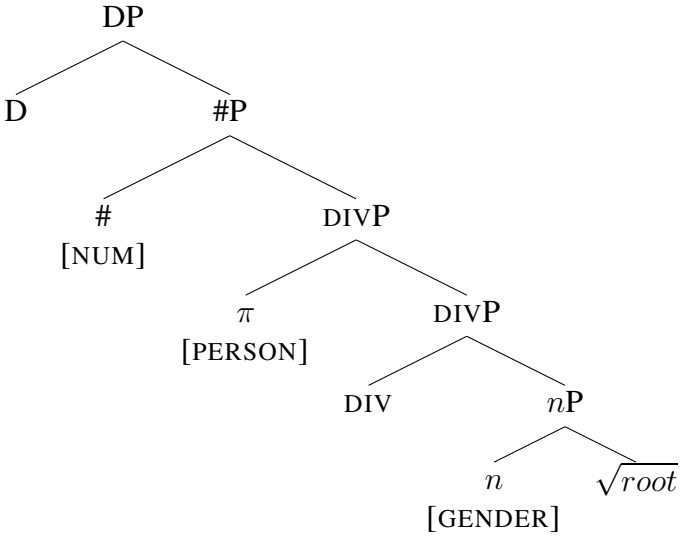
[15]Technically, i* is a syncategorematic operator.

[16]We assume that building is subject to structural economy, i.e., no feature-vacuous structure building is possible. For i* to be licensed, the merge of i* must yield a distinct structure. Since i* is a polarity operator, this economy condition is trivially satisfied as long as the value of the i* feature is distinct from the value of the corresponding feature in the sister projection.

b. When the feature output of i* does not project (adjunct; diminutives only):

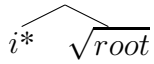


(25) *Default DP structure:*[17]

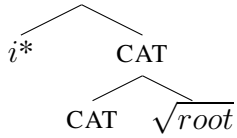


(26) *Attested levels of i* attachment:*

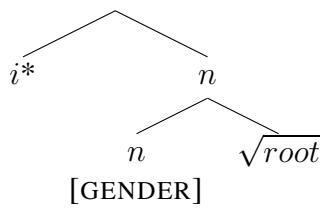
a. LEVEL 0:
N-to-N without gender change (appx. A.1): i*=$\sqrt{root}$



b. LEVEL 1:
category change (nominalizer) (sec. 3.1): i*=CAT



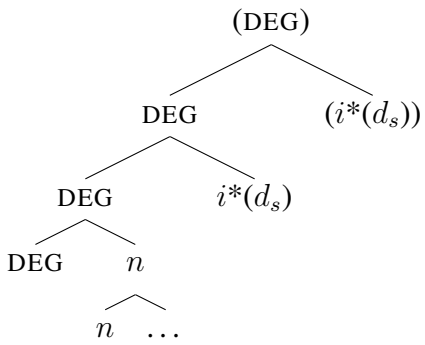
c. LEVEL 2:
N-to-N with gender change (sec. 3.2) i*=n



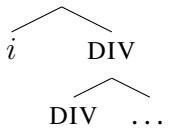
d. LEVEL 3:

[17]For the purposes of this talk the question of whether or not D is a phase head is not directly relevant.

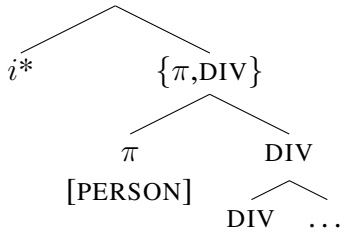
diminutives & doubles (sec. A.2):



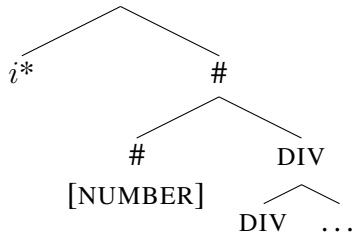
e. LEVEL 4:
individuation & group formation (sec. 3.3 & 3.4): i*=DIV



f. LEVEL 5:
conceptual gender (sec. A.3): i*={π,DIV}



g. LEVEL 6:
collective vs. distributive readings (deflected agreement; 'committee' type) [not discussed here]: i*=#



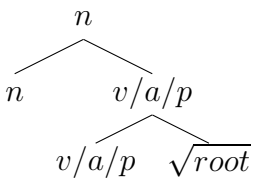
Morphological assumptions:

- for Semitic, we assume that a tri-consonantal template subsumes the structure up to DIV; elements merged above DIV head, including the specifier, are realized outside of the template
- further, we assume that any instantiation of i* associated with a nominal feature (NUMBER, GENDER, PERSON) will be realized as **K** as the morphological nominal default in Slavic, and as **F** in Semitic (unless the features of i* are subsumed within the templatic morphology)

3.1 Category change

- let us demonstrate how i* works on i* as a category changing head, i.e., a nominalizer

(27) Pre-theoretically:



Functional category of the output:

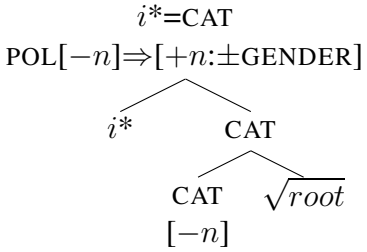
- i* merges to a category head (a, v, possibly p) and turns into a category head by virtue of a feature of the category defining head i* merges with

Feature value of the output:

- for (28) to obey structural economy, i* must output a polarized value of a feature common to the v, a and p categorizing heads

⇒ valued GENDER as the defining feature of n (nominality; e.g., Kramer 2015, Veselovská 2019)

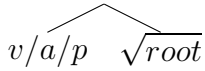
(28)



GENDER feature valuation:

- since either value of [±GENDER] will satisfy the switch to n, we correctly predict that category change yields both MASC and FEM[18]

Root insertion & Encyclopedia:

- roots are associated with indexical information that restricts where in the syntactic structure such a root can be inserted (Acquaviva 2014 and others)
- at spell-out the  structure gets realized by the corresponding root in the context of the relevant categorizing head
- i* gets realized as **K**, concatenated to the root triggering Encyclopedia (lexical semantics) information associated with the root in the context of the i* feature (indexical gender)
- in Semitic, the category changing i* is subsumed under templatic morphology → it triggers vocabulary insertion of a distinct template

[18]Note that i* cannot generate NEUTER because neuter is a complex feature, [−PERSON, −GENDER], see, e.g., Bartošová & Kučerová 2016, 2018, that can only be introduced indexically by the root, or derived higher in the structure when [PERSON is merged. At this level of representation, only [GENDER] can be structurally present.

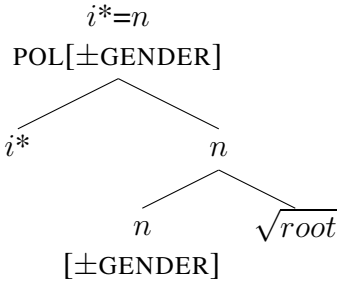
3.2 N-to-N conversion

- a modified version of a category change formation, in which i* merges with n

$\Rightarrow$ instead of manipulating n itself – which would violate economy since the only output is n, i* applies to features of n: a GENDER feature

- i* inherits properties of n
- the polarity content of the i* function operator switches the value of the GENDER feature to its opposite value, i.e, MASC $\Rightarrow$ FEM, or FEM $\Rightarrow$ MASC

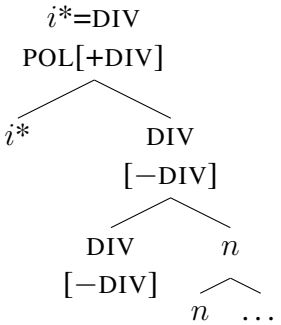
(29)



3.3 Individuation

- nominal roots are by default not individuated (e.g., Borer 2005) $\Rightarrow$ an individuating projection must be merged (DIVP)
- Semitic has a class of genderless unindividuated nominals, so called batch nouns, in which the individuating functional head, DIV, head is set up as [−DIV]
- when i* attaches to a DIV projection, it changes the polarity of the [−DIV] to [+DIV]

(30)



Why gender morphology?

- no specific DIV morphology $\rightarrow$ insertion of the closest nominal morphological realization with a classifier-like property, namely gender (**F**)[19]

Slavic:

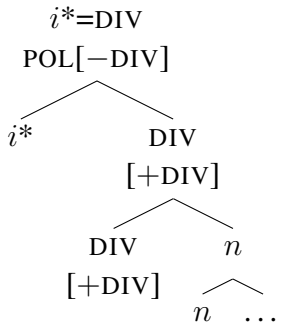
[19]Morphosyntactic evidence suggests that the non-human gender system, i.e., GENDER without PERSON, in Arabic appears privative, i.e., MASC is the absence of gender, valued gender is FEM.

- *general individuation feature* (in parallel to general number; e.g., Corbett 2000, Wiltschko 2008)[20]
- general individuation feature is compatible both with [+DIV] and [−DIV] ⇒ not in the domain of application of i*[21]

3.4 Group formation

- group formation interpretation in Semitic results from the same structure like the individuation structure but in this case a valued PERSON feature forces an individuated structure
- i* then returns [−DIV]
- this derived structure cannot be pluralized because plural requires an individuated, i.e., [+DIV] structure → the merge of another layer of i* is blocked by structural economy because the output of the iterated merge would be equal to the merge before the first i* was merged (structural economy violation)
- as with individuation, since there is no designated morphological realization of the DIV feature, morphology realizes the i* by its closest relative, i.e., gender

(31)



Slavic:

- group formation blocked by the general individuation feature
- what abouts group formations based on quantifiers, numerals and pluralia tantum?

⇒ group formations arise via other means (see appendix A.4)

4 Conclusions and open questions

- we presented an empirical study that provides evidence that there is a class of structural building operations that are in an important structural sense underspecified

[20] A move motivated by work on semantic properties of number in Slavic nominals (e.g., Grimm & Dočekal 2020, building on Krifka 1995.

[21] But see appendix C for a discussion of individuation in mass nouns.

- and that a variety of seemingly varied functional projections can be unified under the underspecification approach
- there are of course many open questions, such as why these particular syntactic interpretations and why they wildly corroborate, and whether we can find i*-like behavior in other projections as well
- the so-called reflexive in Slavic might be a good candidate for i* in vP

A Other cases of i* (not discussed for reasons of time)

A.1 i* at the root level: N-to-N conversions without a gender change

- there is a class of N-to-N conversions that does not change gender
- not frequent but all 4 genders are attested

(32) FEM → FEM

a. kniha 'book.F.SG' → kníž-ka 'book-**K**.F.SG, a book (less formal)' CZECH
b. košile 'shirt.F.SG' → košil-ka 'shirt-**K**.F.SG' an undershirt CZECH

(33) NEUT → NEUT

a. rameno 'shoulder.N.SG → ramín-ko.**K**.N.SG 'hanger' CZECH
b. biuro 'office.N.SG' → biur-ko 'desk-**K**.N.SG' POLISH

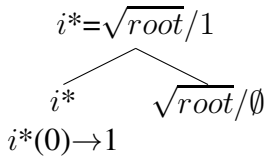
(34) MASC INAN → MASC INAN
cukier 'sugar.M.SG/MASS' → cukier-ek 'sugar-**K**.M.SG, a piece of candy' POLISH

(35) MASC ANIM → MASC ANIM
chlap.M.SG 'guy' → chláp-ek 'guy-**K**.M.SG, dude'

- i* directly merges with the root
- since roots do not have syntactic features, i* cannot function on syntactic features – i* can only operate on indexical information associated with encyclopedic entry properties of the root[22]
- when i* attaches to the root, technically it attaches to the first merge position/place holder for the late inserted root, i.e., 0
- the output of i*(0) = 1, triggering the insertion of level 1 encyclopedia entry, instead of the default meaning (i.e., **K** derivations are always semantically based on the lexical semantics of the 0 level root)

[22]We assume that if the same roots associate with more than one lexical semantic meaning, the meanings are attached hierarchically in an array.

(36)

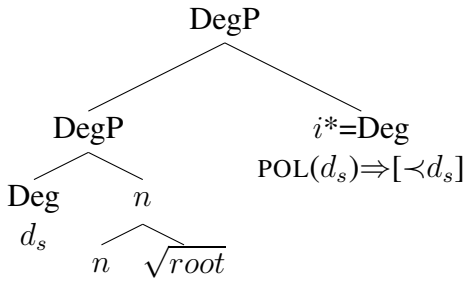


- here, i* manipulates a variable because the root position is a placeholder for late insertion

A.2 Diminutives

- each noun phrase can in principle contain a degree phrase (Morzycki 2009)
- we argue, following much work on Slavic diminutives (e.g., Wiltschko & Steriopolo 2007), that DIM is an adjunct to a DegP

(37)



- technically, i* changes the default POS heading DEG to NEG

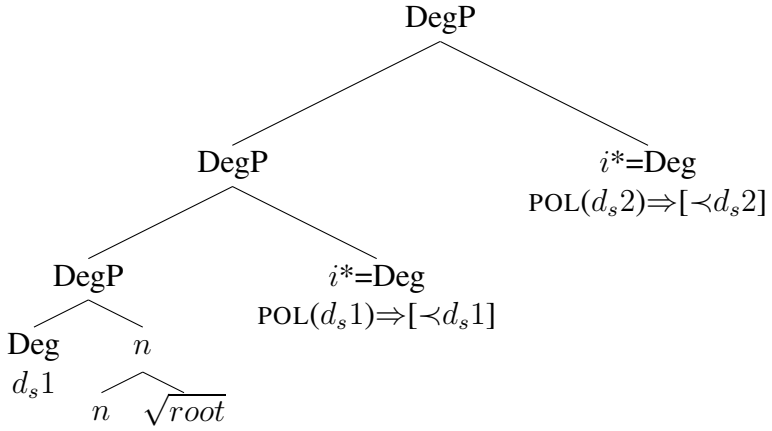
(38) a. $[\![\text{POS}]\!]$ = λ g$_{<e,d>}$ λx.standard (g)$\preceq$g(x)
b. $[\![\text{NEG}]\!]$ = λ g$_{<e,d>}$ λx.standard (g)$\succeq$g(x)

- i* changes the point of reference to be below the standard minimal value
- gender is transparent since i* is an adjunct and adjuncts do not project

Doubling

- double even triple diminutives are possible since i* can apply recursively to reset the scale to the minimal value of its input
- a double DIM formation obeys structural economy only if it yields additional interpretations (Sichel & Wiltschko 2018) ⇒ double DIM yields a new degree scale

(39)



- when the second instantiation of i* of Deg merges, we again get recursion, i.e., the structure enforces spell-out
- which gives the right morphology for Semitic ⇒ the primary DIM is stem internal, the second iteration is a suffix

Open question:

- why does i* surface as **F** in Semitic?

Pragmatic readings

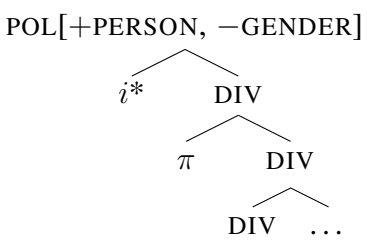
- both single and double DIM not only yield a new degree (based on the bounded nominal projection) but can also be mapped to a pragmatic reading
- pragmatic readings can constitute affection, or derogation (see, e.g., Fontin 2011, Fassi Fehri 2016, 2018)
- these are not a direct product of feature interaction of i* but rather a mapping of its morphosyntactic effects to the interfaces

A.3 Conceptual Gender

- conceptual gender being introduced by a higher functional head (e.g., Pesetsky 2013, Kramer 2015), in connection with with a PERSON feature (e.g., Heim 2008, Sudo 2012, Kučerová 2018)
- a human-denoting masculine[23] ⇒ a complex feature: [+PERSON, −GENDER]
- this instantiation of i* attaches to a structure that already contains PERSON feature
- PERSON (π) is merged in the spec of DIV (den Dikken 2019) ⇒ i* attaches DIVP after the specifier is merged
- since PERSON is associated with D, i* primarily manipulates the only manipulates n features and , i* effectively manipulates the GENDER part of the feature

⇒ POL[+PERSON, −GENDER] ⇒ [+PERSON, +GENDER], i.e., MASC noun turns into a FEM-denoting noun

[23] In the language systems investigated here, morpho-syntactically unmarked masculine becomes marked if it denotes a human, as a result of a diachronic change introducing animacy/humanness into the nominal system. While human MASC gender exhibits distinct morphological behaviour (masculine sound plurals only for humans in Arabic etc.), we are not aware of any morpho-syntactic differences between human-denoting and non-human denoting feminine nouns, i.e., while there is [+PERSON, −GENDER] in narrow syntax, there is no [+PERSON, +GENDER]. The female-denoting interpretation arises only via presuppositions associated with the PERSON feature.

(40)


- since i* is above DIV, in Semitic, i* triggers a morphological insertion of a suffix (**F**)

[−PERSON, −GENDER]?

- it is not immediately clear why the system does not output **K** or **F** as [−PERSON, −GENDER] as well, partially because we do not fully understand how the polarity function operates on complex features
- in Slavic, [−PERSON, −GENDER] is attested at the DIV level: it corresponds to NEUTERs forming semantic aggregates but in this case the input is an inanimate count nominal and the output is an aggregate, i.e., it does not look like the type of feature manipulation we associate with i*

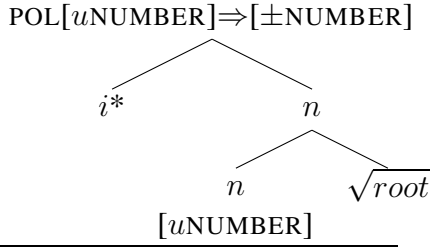
A.4 Group formation with quantifiers and alike?

Numerals and quantifiers:

- nominalized structures, with i* introducing [+n: GENDER], and feeding into additional functional projection
- both numerals and quantifiers imply a semi-lattice structure ⇒ group-formation like semantic interpretation

Pluralia tantum

- N-to-N conversion where i* manipulates unvalued number on n: POL[uNUMBER] ⇒ [vNUMBER]
- when i* returns [+NUMBER], the merge of i* triggers late insertion of roots specified as [+PL], i.e., pluralia tantum[24]
- wrt to the higher functional projections pluria tantum behave like aggregates and can only be pluralized by using an aggregate specific counting morphology

(41)


[24]When i* returns [−NUMBER], the merge of i* is expected to trigger late insertion of roots specified as [−PL]. These forms don't seem be to be attested but they might be indistinguishable from regular count nominals or excluded by lexicon economy.

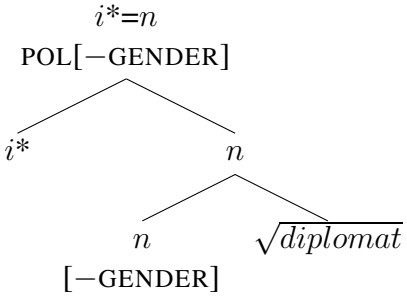
B Structural homophony?

- the proposal predicts a high degree of structural ambiguity
- different structures ⇒ different syntactic behaviour

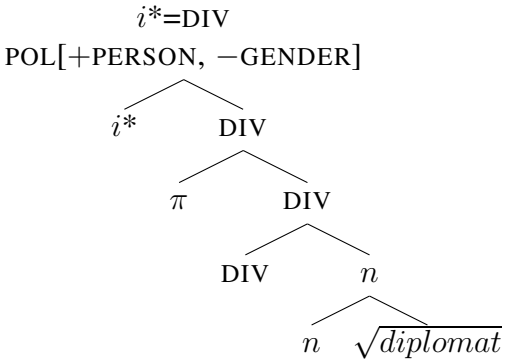
Case I:

- for example, nominal with **K** being ambiguous between a female-denoting **K** and N-to-N converted **K** should correspond to two distinct structures:

(42) a. diplomat-ka 'diplomat-**K**.F.SG; a briefcase'



b. diplomat-ka 'diplomat-**K**.F.SG; a female diplomat'

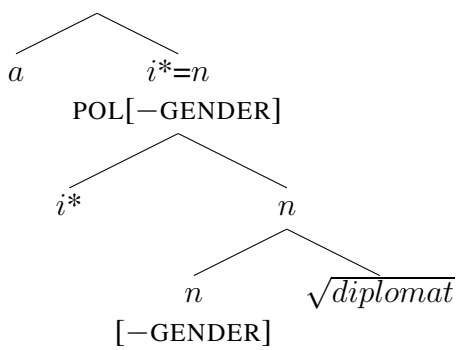


Predictions:

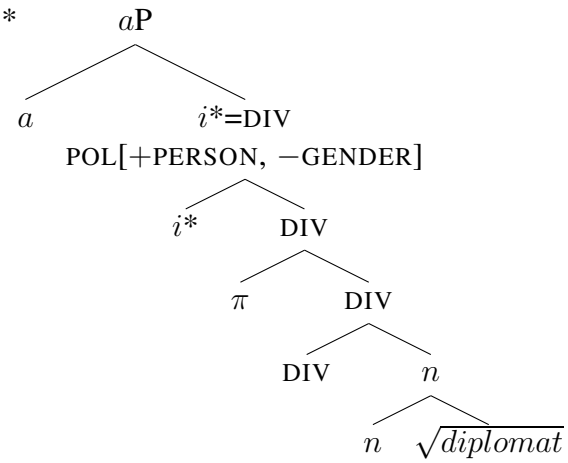
- since in this case i* attaches within functional domain proper, unlike in the case of category change and N-to-N conversion, no derivational morphology can apply to a nominal with a female-denoting **K**

(43) diplomat-k-ový 'diplomat-**K**-ADJ' → only related to a briefcase, not to a female diplomat

(44) a. ✓



b. *



Case II:

- double-diminutive **F** ambiguous with female-denoting **F**
- only the latter one triggers agreement because only the latter one is a head that projects its gender feature:

(45) a. al-arnub-i nam-et b-Hodn-ii
the-rabbit.DIM.M.SG-**F**:SG sleep.3PST-F.SG in-lap-my
'the she-bunny slept in my lap.'
b. al-arnub-i nam b-Hodn-ii
the-rabbit.DIM.M.SG-**F**:SG sleep.3M.SG.PST in-lap-my
'the cute bunny slept in my lap.'

Case III:

- group-forming **F** ambiguous with female-denoting **F**
- only the latter can be pluralized

(46) a. mtdyyen 'religious.M.SG, a believer' → mtdyn-i 'religious-**F**.SG, a religious group'
b. *mtdyn-at 'religious-**F**.PL, a plurality of religious groups'

(47) a. mtdyyen 'religious.M.SG, a believer' → mtdyn-i 'religious-**F**.SG, a female believer'
b. *mtdyn-at 'religious-**F**.PL, a plurality of female believers'

C Individuation of mass nouns?

- **F** is also supposed to individuate mass nouns but in this case there always is an additional change in the templatic morphology and a shift in lexical meaning, indicating additional structure building and/or additional derivational morphology (similar to specific derivational morphology in Slavic)

(48) sokkar 'sugar.MASS' → sokkareyy-i 'sugar-**F**:SG; a sugar bowl'

- in fact, the template we see here is a template used for names of instruments, even from count nouns:

(49) foren 'oven.M.SG → forneyy-i 'oven-**F**:F.SG; a portable electric oven' LA

- this might not be a pure formation by **F**
- the empirical situation in Slavic is somewhat more complex
- note that for a mass noun to be individuated, the mass noun must be containerized; once we get into a denotation of containers it is not straightforward to distinguish between containarization as a result of individuation (a default container?) and a noun-to-noun conversion with a shift in the lexical semantics of of the nominal base
- if we apply stringent criteria on the lexical semantics of the container, it seems that **K** never individuates
- instead, both DIM derived by **K** and its base are systematically ambiguous between mass and count interpretations

(50) a. Na stole bylo mnoho cukru
on table was much sugar.SG/MASS
'There was much sugar on the table.'
b. Na stole bylo mnoho cukrů
on table was many sugar.PL
'There were many kinds of sugar/ pieces of sugar (cubes, packets of sugar) on the table.'

(51) a. Na stole bylo mnoho cukř-íku
on table was much sugar.SG/MASS-**K**
'There was much cute/sweet/delicious sugar on the table.'
b. Na stole bylo mnoho cukř-íků
on table was many sugar.**K**.PL
'There were many kinds of adorable sugar/ small pieces of sugar (cubes, packets of sugar) on the table.'

Note:

- the pattern in Slavic and Semitic is consistent with the generalization that DIM individuates mass nouns only if it also changes gender of its base, i.e., it is not an adjunct (e.g., Dutch and German, see, e.g., Borer (2005), p. 92, ft. 6)