

# A realisational approach to variable morph ordering

Olivier Bonami<sup>1,2</sup> & Berthold Crysmann<sup>2</sup>

<sup>1</sup>Université Paris-Sorbonne  
Institut Universitaire de France

<sup>2</sup>Laboratoire de Linguistique Formelle (CNRS & U. Paris Diderot)

First AIMM  
Amherst, September 2012

# Introduction

- ▶ Two traditions for dealing with morphotactics
  - ▶ Sequential templates
    - ▶ Standardly used for descriptive purposes, wherever morphotactics is of any complexity
    - ▶ Linear order stated directly
    - ▶ Deviations from a rigid template commonly stated in prose
    - ▶ No agreed upon formal model
  - ▶ Stem-centric morphological composition
    - ▶ Most common approach in generative morphology
    - ▶ Morphological composition starts from the root/stem
    - ▶ Linear order derived from composition structure
    - ▶ Implemented in various formal models, including (Lieber, 1980; Anderson, 1992; Stump, 2001, etc.)
- ▶ Our goals:
  - ▶ Look at a nontrivial set of deviations from simple position class morphology
  - ▶ Provide a formal model of morphotactics where
    - ▶ Linear position is a descriptive primitive
    - ▶ There is no rule ordering of any kind
    - ▶ Morphotactics can be partially separated from exponence
  - ▶ Focus on coverage; limited to inflection.

Aspects of a typology of variable order phenomena

# Aspects of a typology of variable order phenomena

- ▶ Types of conditioning
  - ▶ Phonological (e.g. Paster, 2006; Caballero, 2010; Kim, 2010)
  - ▶ Semantic (e.g. Rice, 2000; Aronoff and Xu, 2010)
  - ▶ Morpho-syntactic (e.g. Stump, 1993)
- ▶ Types of variable placement
  - ▶ Positional disambiguation (parallel position classes; Stump, 1993)
  - ▶ Variable order
    - ▶ Single morph, variation relative to stem (Swahili ambifixal classes; Stump, 1993)
    - ▶ Single morph, variation independent of stem (Fula reversible classes; Stump, 1993)
    - ▶ Multiple morphs, order-reversing (Huave, Fula)
    - ▶ Multiple morphs, order-preserving (Romance pronominal affix clusters)
- ▶ Free permutation among position classes
  - ▶ unconstrained (Chintang prefixes)
  - ▶ partially constrained (Mari declension; French post-stem clusters)

# Positional disambiguation

- ▶ Swahili subject and object agreement are realised in distinct slots (2 and 5)
  - ▶ Inventories for expressing person-number-gender combinations are largely identical across functions and slots  
Only exceptions: 2nd person and 3rd singular M/WA class
  - ▶ Exponents largely underspecified as to grammatical function
  - ▶ Morphotactic position disambiguates to subject or object functions
- ☞ Cf. Stump's (1993) notion of parallel position classes

PER	GEN	SUBJECT		OBJECT	
		SG	PL	SG	PL
1		ni	tu	ni	tu
2		u	m	ku	wa
3	M/WA	a	wa	m	wa
	M/MI	u	i	u	i
	KI/VI	ki	vi	ki	vi
	JI/MA	li	ya	li	ya
	N/N	i	zi	i	zi
	U	u	—	u	—
	U/N	u	zi	u	zi
	KU	ku	—	ku	—

Table: Subject and object prefixes in Swahili

## Variable placement relative to the stem: Swahili

- ▶ Swahili relative agreement markers either precede or follow the stem
  - ▶ suffixation (slot 7) is used with untensed verbs in the affirmative
  - ▶ prefixation (slot 4) is used in all other cases (i.e. either tensed or negative verbs)

1	2	3	4	5	6	7	Gloss
POL	SUBJ	TAM /	REL.	OBJ	STEM	REL.	
MRKR	POL	MRKR	MRKR	MRKR		MRKR	
a	na	<b>ye</b>		soma			'(person) who is reading'
a				soma	<b>ye</b>		'(person) who reads'
a	na	<b>cho</b>	ki	soma			'(book) which he is reading'
a			ki	soma	<b>cho</b>		'(book) which he reads'

- ▶ Schematically:
  - A** **B** **C** **D** **E** STEM
  - A** **B**              **E** STEM **D**

- ▶ Cf. Stump's (1993) notion of **ambifinal position class**

## Variable placement on same side of the stem: Fula

- When they are suffixal, subject and object markers follow tense markers
- Standard (default) order is **subject** before **object**
- Singular objects, however, precede **1st singular subjects** (reversal)

---

mball u <b>don</b> mo	'You (pl.) helped <b>him</b> .'
mball u <b>mi</b> <b>be</b>	'I helped them.'
mball u <b>moo</b> <b>mi</b>	'I helped <b>him</b> .'

---

- Schematically:  
STEM 

A	B	C
---	---	---

  
STEM 

A	C	B
---	---	---
- Cf. Stump's (1993) notion of **reversible position classes**

## Variable placement of sequences: order reversing

### ► Huave verbal affixes (Kim, 2010):

- ▶ Complex morphotactic systems with a wide number of position classes.
- ▶ A number of affixes can occur on either side of the stem.
- ▶ There are exactly two classes of mobile affixes.
- ▶ Where these cooccur on the same side of the stem, ‘mirror effect’:

---

x	i	n	a-jch		‘I will give it.’
1	FT	1.SB	TV-give		
x	i		chut-u	n	‘I will sit.’
1	FT		TV-sit	1.SB	
			chut-u	t u s	‘I sat down.’
			sit	CP ITR 1	

---

► Schematically: A ... B STEM

STEM B ... A

but also:

B STEM ... A

A ... STEM B

## Variable placement of sequences: order preserving

- ▶ Italian pronominal affixes (Monachesi, 1999):
    - ▶ Occur in a fixed order of 6 positions

A	B	C	D	E	F
[ <i>obj</i> ,1sg]: mi	[ <i>loc</i> ]: ci	[ <i>obj</i> , 3, <i>refl</i> ]: si	[ <i>d-obj</i> ,3sg, <i>m</i> ]: lo	[ <i>obj</i> , <i>imp</i> ]: si	[ <i>part</i> ]: ne
...			...		

- ▶ Occur on either side of the stem depending on morphosyntactic context
  - ▶ Whenever two affixes cooccur, they occur in the same order on the same side of the stem.

me lo da te ‘You give it to me.’

da te me lo! ‘Give it to me!’

\* da te lo me!

\* lo me da te!

- Schematically:



## Taking stock

- ▶ Are morphs indexed relative to the left, or relative to the stem?
  - ▶ Where order is fixed, the two options are not distinguishable
  - ▶ Ambifixal classes nicely fit with stem-relative indexing
  - ▶ Reversible position classes (Fula) show that variable placement may not keep distance from the stem constant.
  - ▶ Sequences of mobile affixes constitute decisive evidence: order preservation or reversal?
    - ▶ Mirroring attested (Huave), but both rare and limited in complexity: at most two elements, independently conditioned for each exponents
    - ▶ Order preservation (Italian): more widely attested, more complex, dependent on same condition
    - ☞ Mirroring as the unmarked option severely undermotivated
- ▶ **Our observation:** Italian best described by a system that allows for both types of indexing.



- ▶ Under a double indexing strategy, the stem is the only truly mobile morph in Italian.

## Freely ordered position classes: Chintang

- ▶ Chintang verb prefixes (Bickel et al., 2007)
  - ▶ can be freely permuted
  - ▶ prefixes encode **subject** and **object** agreement, as well as **negation**
  - ▶ Suffixes in Chintang, however, are strictly ordered in position classes

---

u	kha	ma	cop	yokt	e	'They didn't see us.'
u	ma	kha	cop	yokt	e	'They didn't see us.'
kha	u	ma	cop	yokt	e	'They didn't see us.'
kha	ma	u	cop	yokt	e	'They didn't see us.'
ma	u	kha	cop	yokt	e	'They didn't see us.'
ma	kha	u	cop	yokt	e	'They didn't see us.'

---

- ▶ Schematically:  $\left\{ \begin{matrix} A \\ B \\ C \end{matrix} \right\}^* \text{ STEM } [E] [F]$

⇒ Account of free permutation must be integrated with a theory of position classes

# Free order with constraints: Mari declension

- ▶ Nominal declension in Mari (Luutonen, 1997):
  - ▶ The relative position of **case** and **possessor** is variable, and free in some cells

---

	NOPOSS	3.SG.POSS
NOM	muno	munâ-žo
GEN	munâ-n	munâ-žâ-n
ACC	munâ-m	munâ-žâ-m
DAT	munâ-lan	munâ-ž-lan, munâ-lan-že
LAT	mun-eš	mun-eš-əže
ILL	munâ-ško	munâ-škâ-žo
INES	munâ-što	munâ-štâ-žo
MOD	munâ-la	munâ-ž-la, munâ-la-že
COM	munâ-ye	munâ-ž-ye

---

- ▶ Schematically: STEM  $\left\{ \begin{matrix} A \\ B \end{matrix} \right\}^*$ , with pairwise constraints on **A** and **B**.
- ☞ Notice: Mari = Fula + Chintang

# Free order and variable order in the same system

## In informal Parisian French:

- variable order of complement pronominal affixes (5 position classes) depending on mood
- relative order fixed before the stem, almost free after

PRESENT INDICATIVE	IMPERATIVE		
il me les donne	*il les me donne	donne les moi	donne moi les
il m'en donne	*il en me donne	donne m'en/moi-z-en	donnes-en moi
il m'y envoie	*il y m'envoie	envoie m'y/moi-z-y	envoies-y moi
il les leur donne	*il leur les donne	donne les leur	*donne leur les
il les en blâme	*il en les blâme	blâme les en	blâmes-en les
il les y envoie	*il y les envoie	envoie les y	envoies-y les
il leur en parle	*il en leur parle	parle leur(-z-)en	parles-en leur
il leur y parle	*il y leur parle	parle leur(-z-)y	parles-y leur
il y en mange	*il en y mange	?manges-y(-z-)en	?manges-en(-z-)y

- Schematically:  
A   B   C   D   E   STEM   G   H  
STEM   G   H   { A, C, E, , }\*  
D, B }

☞ Notice: French = Swahili + Chintang

A formal framework for variable morphotactics

# Characterizing the framework

- ▶ Similar in spirit to Paradigm Function Morphology
  - ▶ Fully lexicalist: no morphological structure visible to syntax, no syntactic operations or constraints in morphology.
  - ▶ Inferential and realizational
  - ▶ No ordering of morphosyntactic features
  - ▶ Arbitration between rules decided on the basis of Pāṇini's Principle
- ▶ Important differences
  - ▶ Position is a property of exponents, not a property of rules.
  - ▶ Order variation through underspecification
    - ▶ HPSG-style feature logic
    - ▶ Multiple inheritance hierarchies of realization rules
  - ▶ Systematic separation between order and paradigmatic opposition
    - ▶ No rule ordering
    - ▶ No rule blocks
  - ▶ Distinction between realizing a feature and being conditioned by a feature (cf. LFG's distinction between functional and constraining equations).

# Realization rules and rule interaction

- ▶ Realization rules are triplets of
  - ▶ a description of a **lexeme** (inflection class, stem allomorphs, etc.),
  - ▶ a description of a **morphosyntactic property set**,
  - ▶ a description of a **list of morphs**.

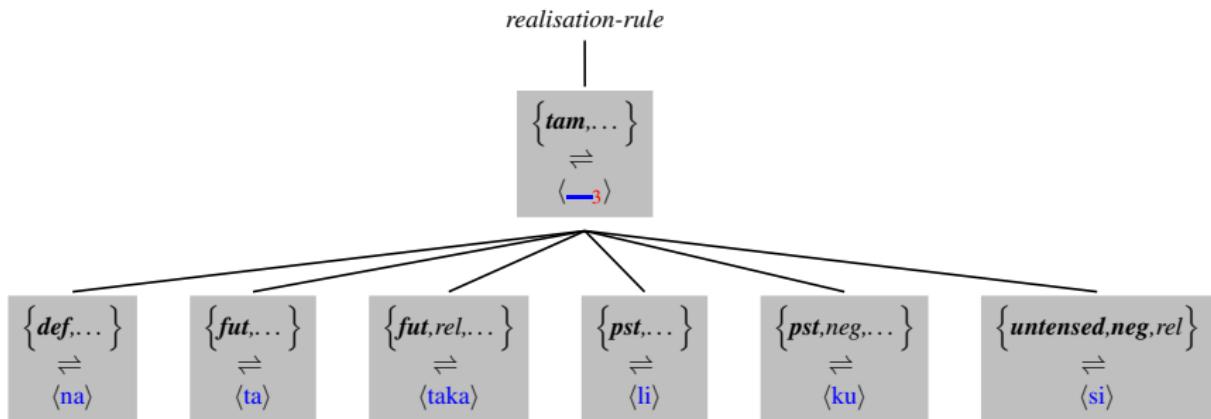
$$\begin{bmatrix} \textit{transitive}, \\ \textit{animate} \end{bmatrix} \bowtie \left\{ \begin{bmatrix} \textit{indicative}, \\ \textit{obj} \\ \textit{1pl}, \\ \textit{MR} & + \end{bmatrix} \right\} \Rightarrow \langle \textit{n} \textit{an}_6 \rangle$$

(Stump, 2001, p. 90 on Potawatomi)

- ▶ The morph list of a word is the combination of the morph lists of all applicable narrowest rules.
- ▶ The phonology of a word is the concatenation of its morphs in the order imposed by their indices.

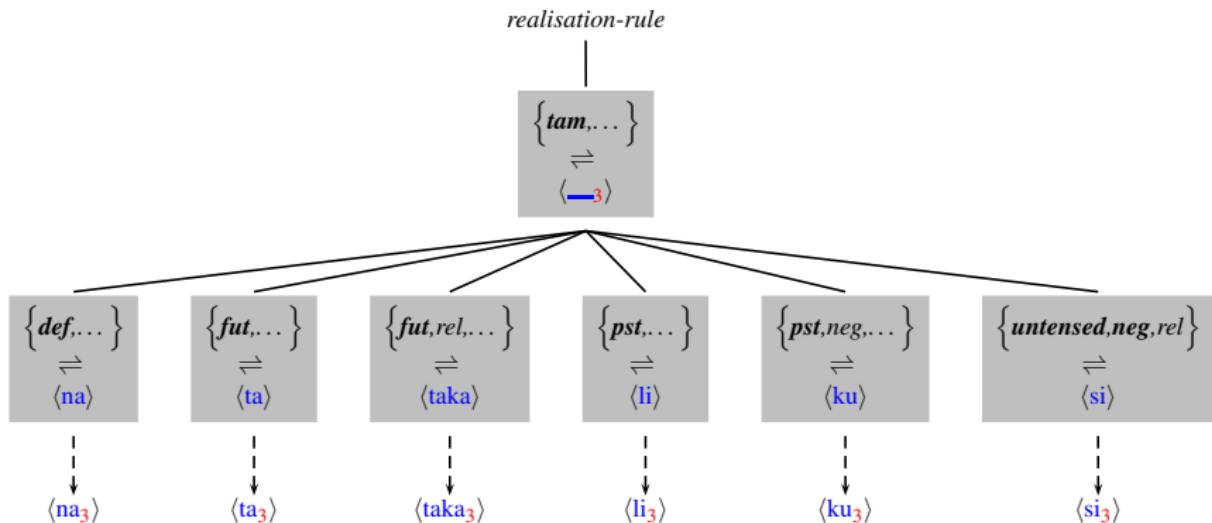
# Underspecification (1): capturing shared rule properties

- Rules are organized in a (monotonic) inheritance hierarchy. This allows for easy encoding of basic morphosyntactic constraints on morphotactics.



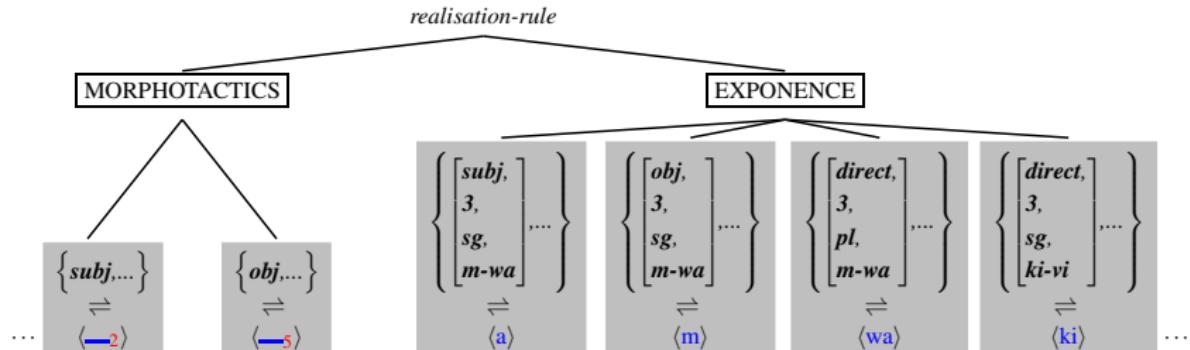
# Underspecification (1): capturing shared rule properties

- Rules are organized in a (monotonic) inheritance hierarchy. This allows for easy encoding of basic morphosyntactic constraints on morphotactics.



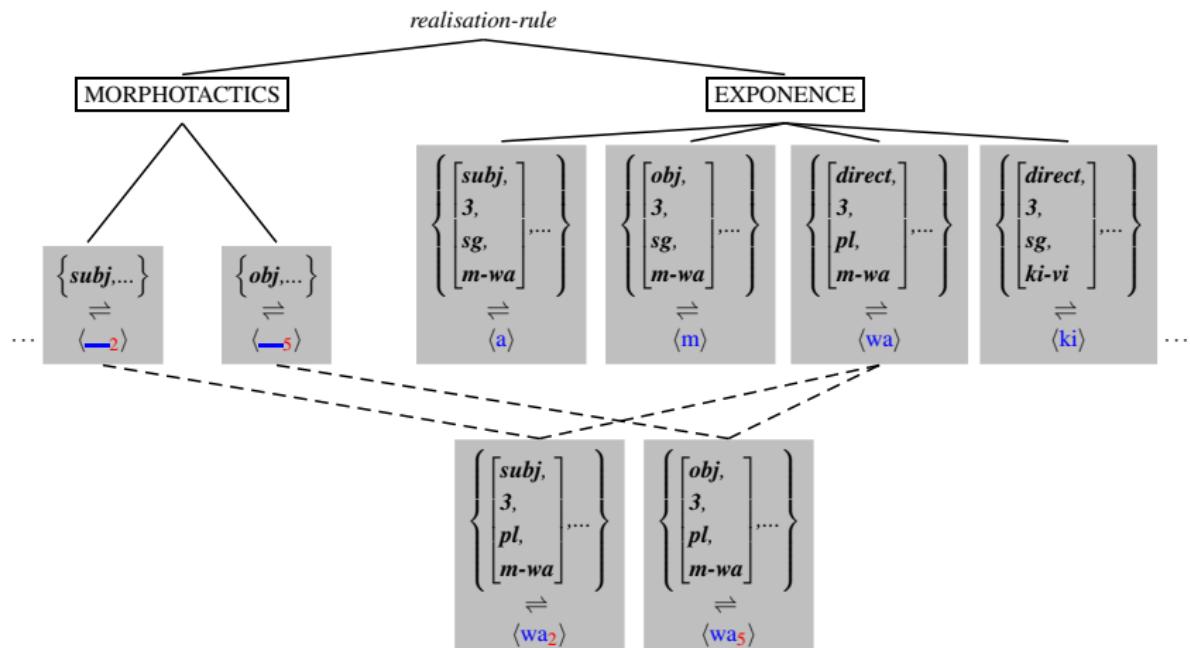
# Underspecification (2): capturing systematic alternation

- ▶ **Online type construction** (Koenig and Jurafsky, 1994; Koenig, 1999) facilitates separation of generalisations on morphotactics from exponente.
- ▶ Ex.: Positional disambiguation (parallel position classes)



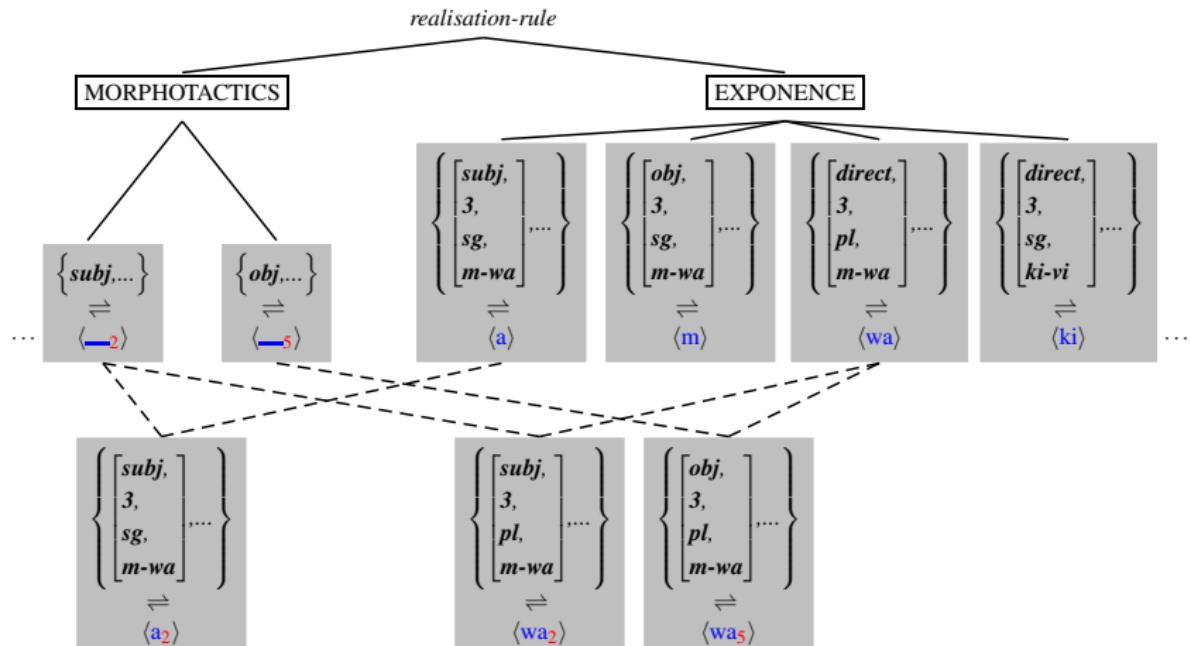
# Underspecification (2): capturing systematic alternation

- ▶ **Online type construction** (Koenig and Jurafsky, 1994; Koenig, 1999) facilitates separation of generalisations on morphotactics from exponentiation.
- ▶ Ex.: Positional disambiguation (parallel position classes)



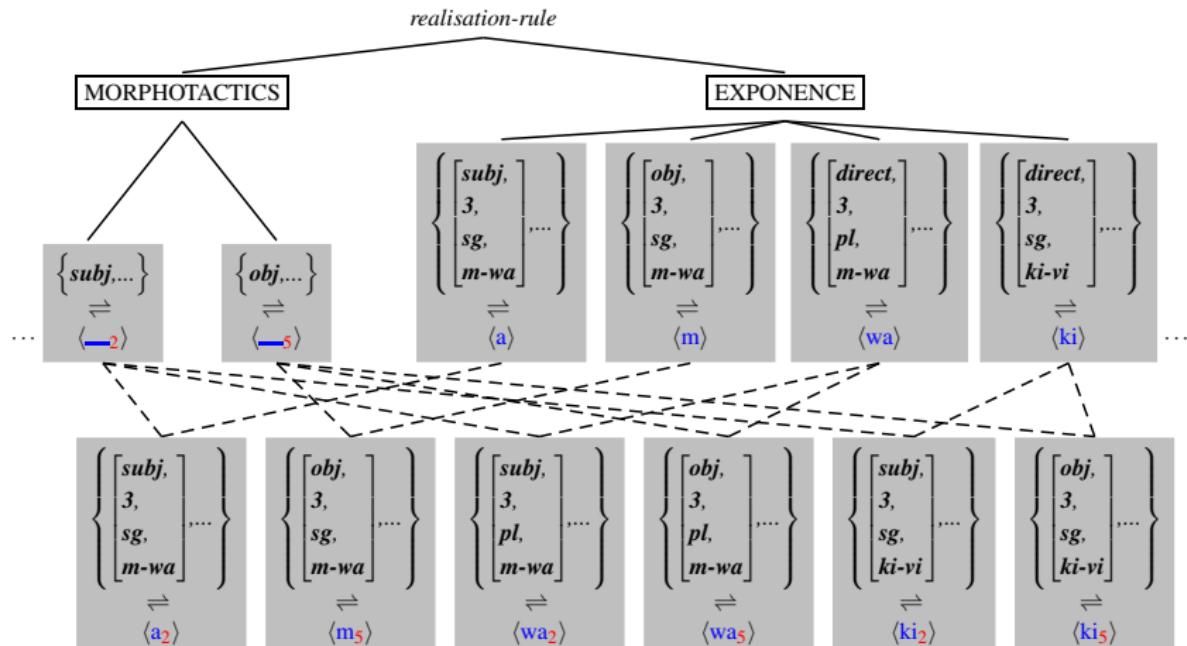
# Underspecification (2): capturing systematic alternation

- ▶ **Online type construction** (Koenig and Jurafsky, 1994; Koenig, 1999) facilitates separation of generalisations on morphotactics from exponentiation.
- ▶ Ex.: Positional disambiguation (parallel position classes)

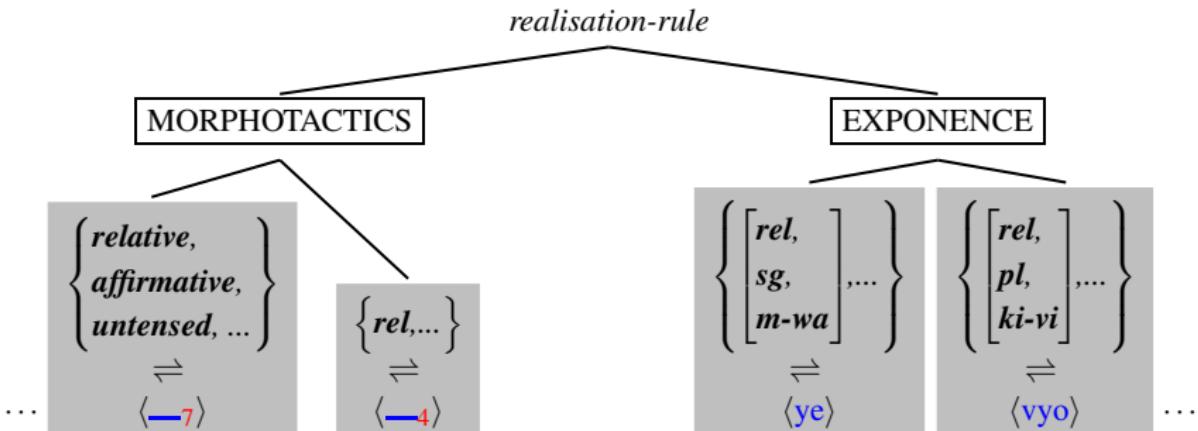


# Underspecification (2): capturing systematic alternation

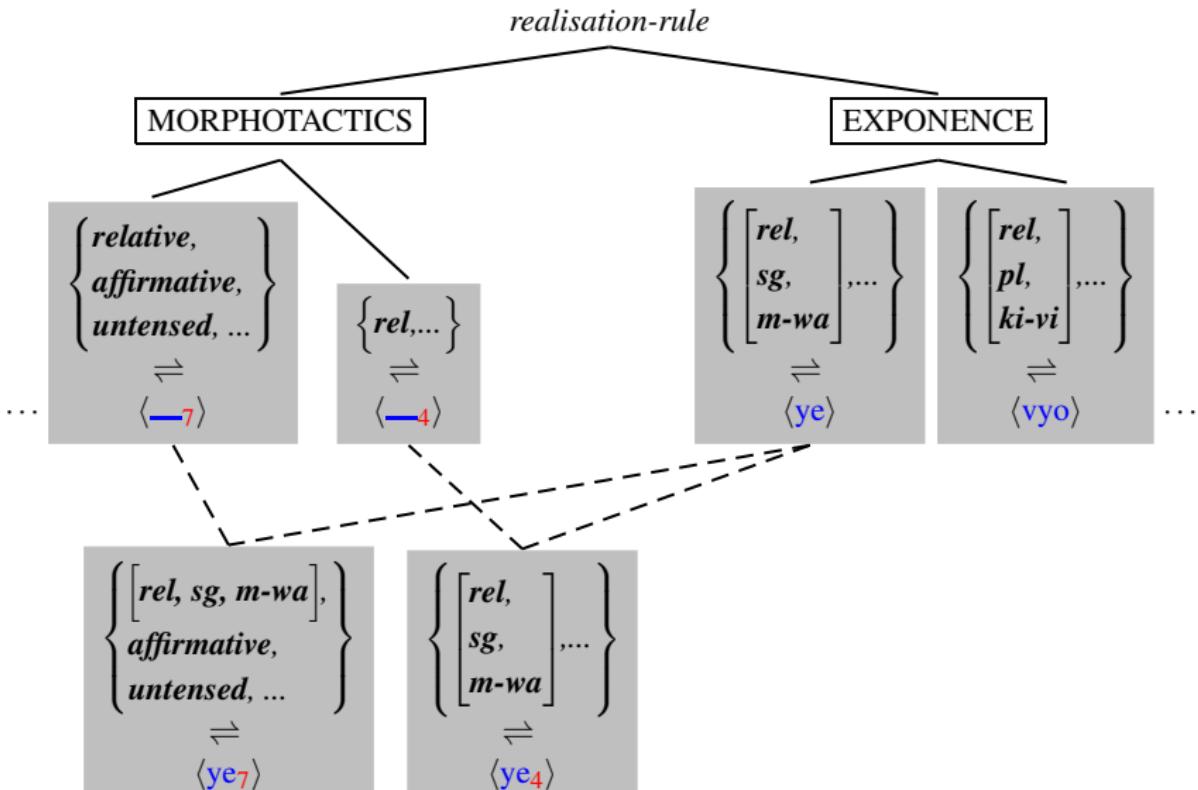
- ▶ **Online type construction** (Koenig and Jurafsky, 1994; Koenig, 1999) facilitates separation of generalisations on morphotactics from exponentiation.
- ▶ Ex.: Positional disambiguation (parallel position classes)



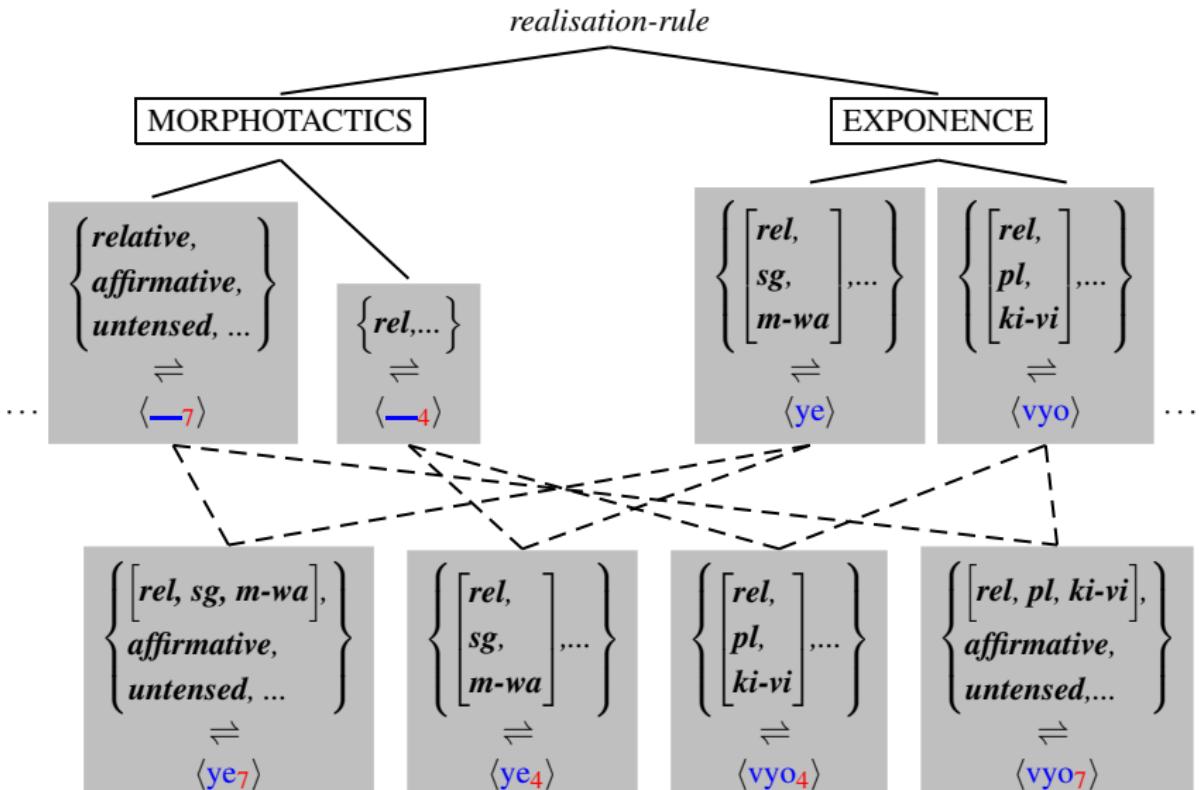
# Swahili ambifocal relative affixes



# Swahili ambifixal relative affixes

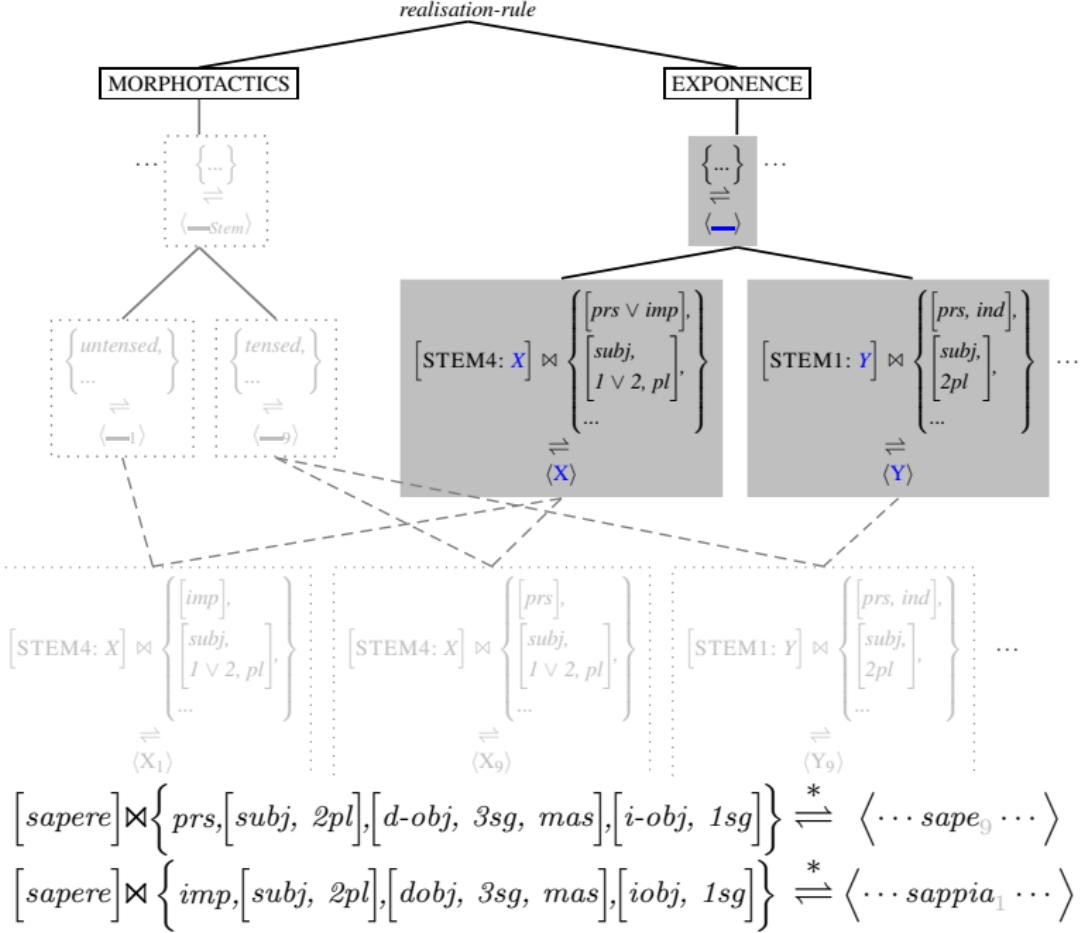


# Swahili ambifixal relative affixes



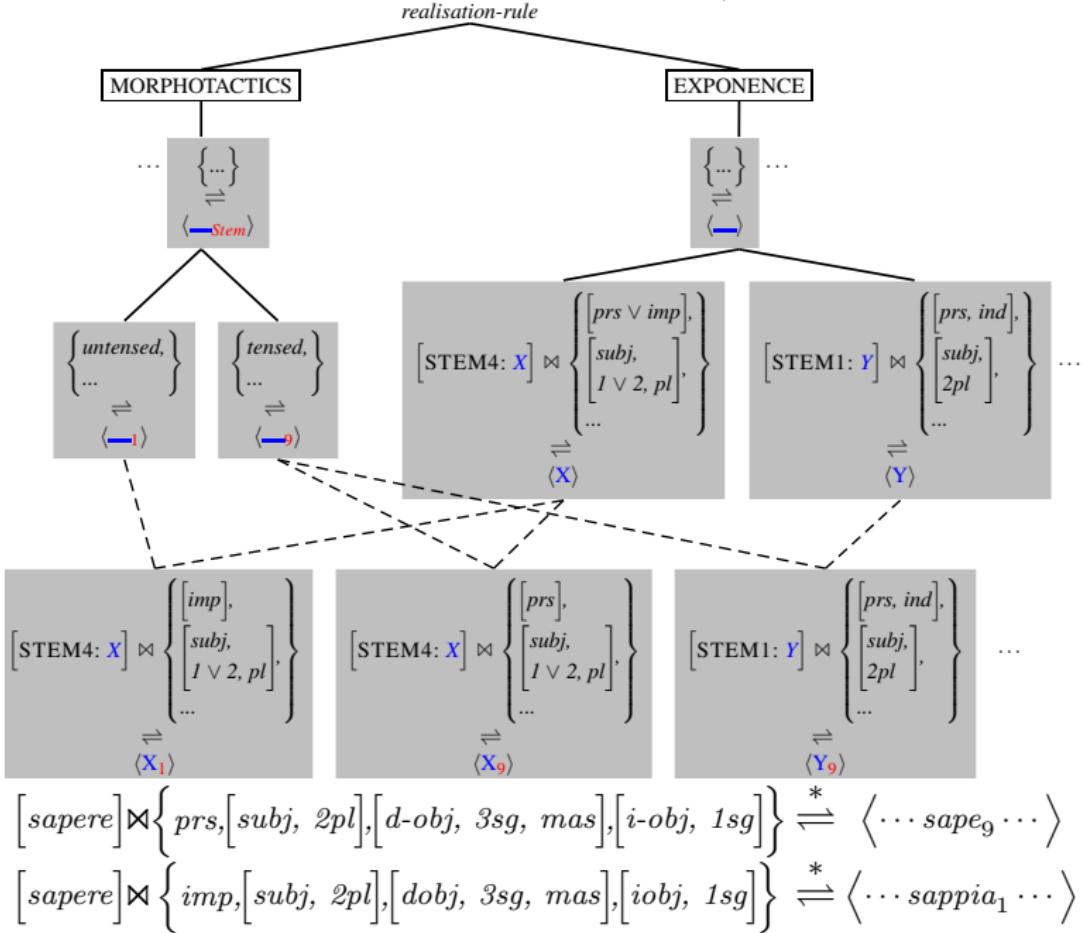
# Italian I: stem selection

(Montermini and Boyé, 2012)

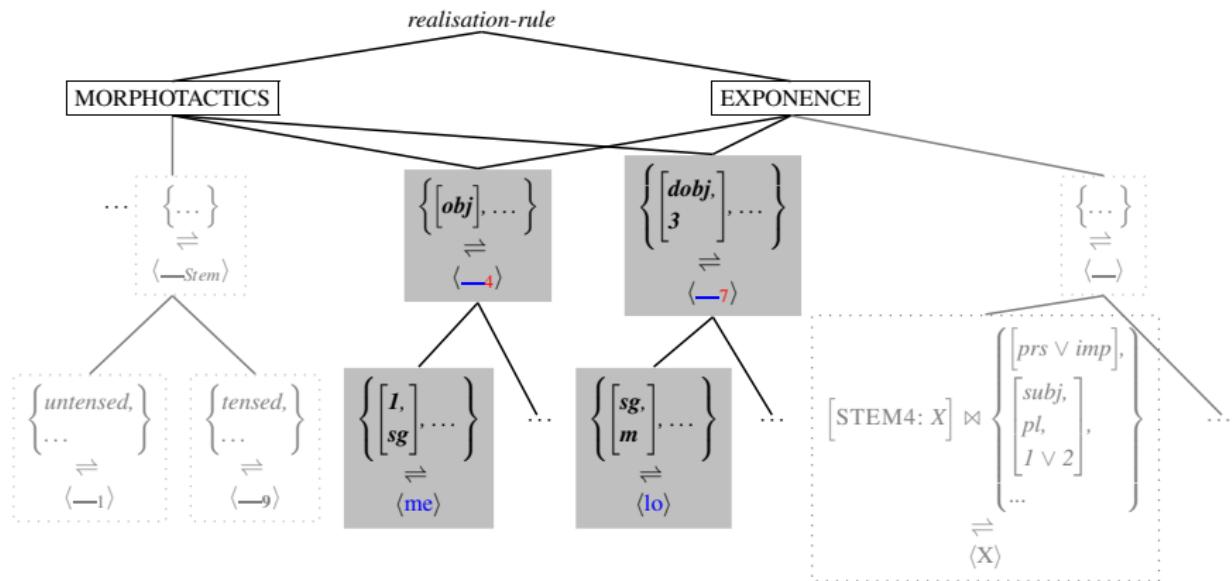


# Italian I: stem selection

(Montermini and Boyé, 2012)

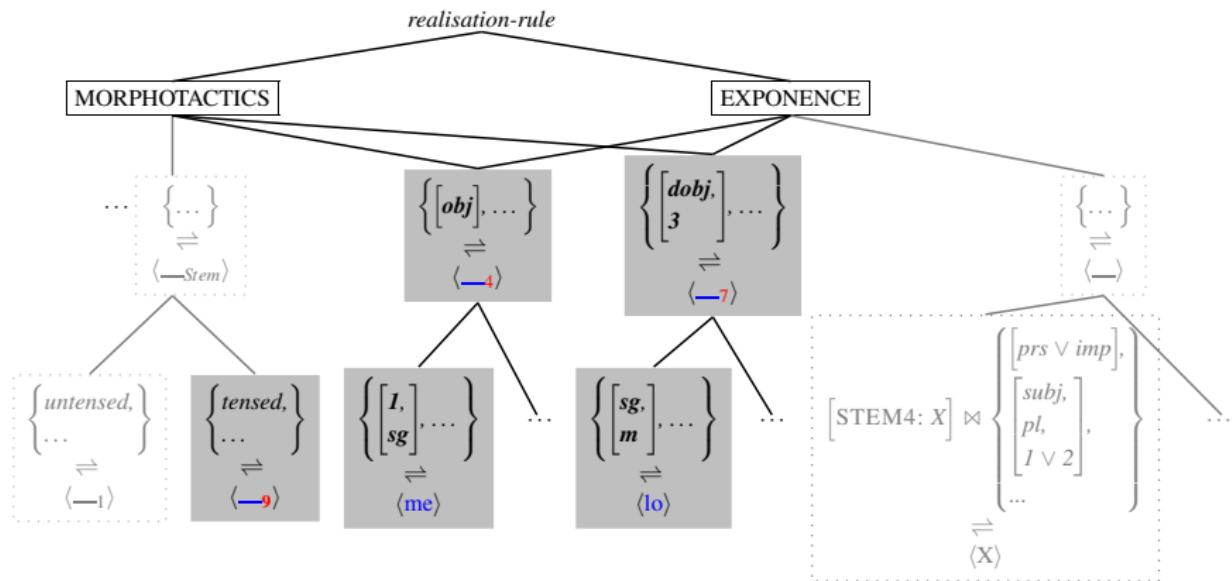


# Italian II: clitic placement



$[sapere] \bowtie \left\{ prs, [subj, 2pl], [d-obj, 3sg, mas], [i-obj, 1sg] \right\} \stackrel{*}{\Rightarrow} \langle \dots me_4 \dots lo_7 \dots sape_9 \dots \rangle$   
 $[sapere] \bowtie \left\{ imp, [subj, 2pl], [dobj, 3sg, mas], [iobj, 1sg] \right\} \stackrel{*}{\Rightarrow} \langle \dots sappia_1 \dots me_4 \dots lo_7 \dots \rangle$

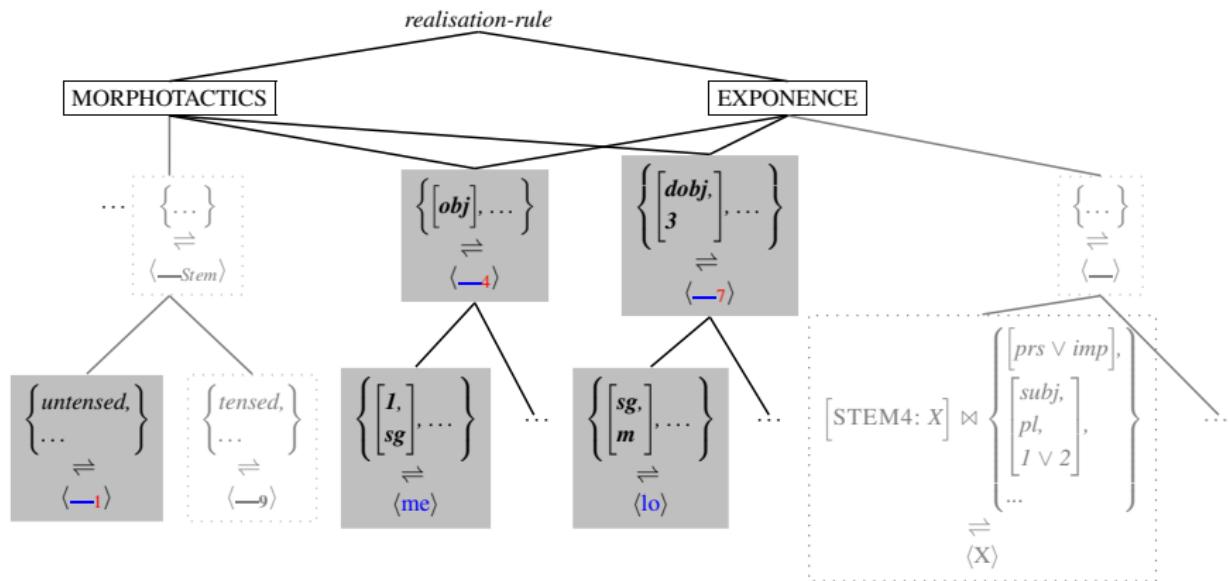
# Italian II: clitic placement



$[sapere] \bowtie \{ prs, [subj, 2pl], [d-obj, 3sg, mas], [i-obj, 1sg] \} \stackrel{*}{\Rightarrow} \langle \dots me_4 \dots lo_7 \dots sape_9 \dots \rangle$

$[sapere] \bowtie \{ imp, [subj, 2pl], [dobj, 3sg, mas], [iobj, 1sg] \} \stackrel{*}{\Rightarrow} \langle \dots sappia_1 \dots me_4 \dots lo_7 \dots \rangle$

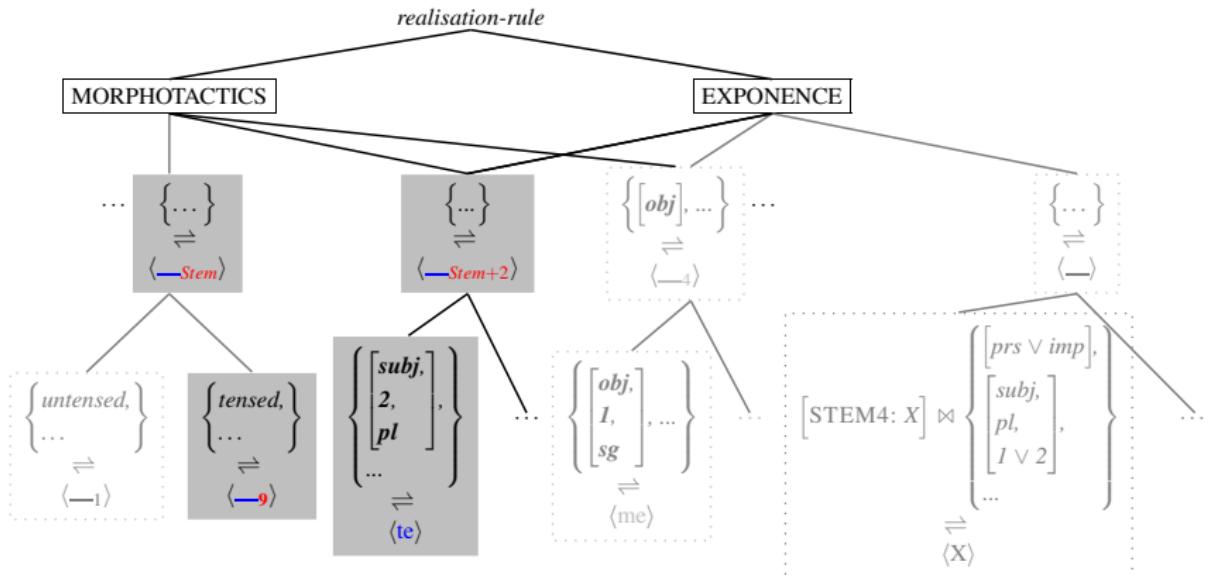
# Italian II: clitic placement



$$[\textit{sapere}] \bowtie \left\{ \textit{prs}, [\textit{subj}, \textit{2pl}], [\textit{d-obj}, \textit{3sg}, \textit{mas}], [\textit{i-obj}, \textit{1sg}] \right\} \stackrel{*}{\Rightarrow} \langle \dots \textit{me}_4 \dots \textit{lo}_7 \dots \textit{sape}_9 \dots \rangle$$

$$[\textit{sapere}] \bowtie \left\{ \textit{imp}, [\textit{subj}, \textit{2pl}], [\textit{dobj}, \textit{3sg}, \textit{mas}], [\textit{iobj}, \textit{1sg}] \right\} \stackrel{*}{\Rightarrow} \langle \dots \textit{sappia}_1 \dots \textit{me}_4 \dots \textit{lo}_7 \dots \rangle$$

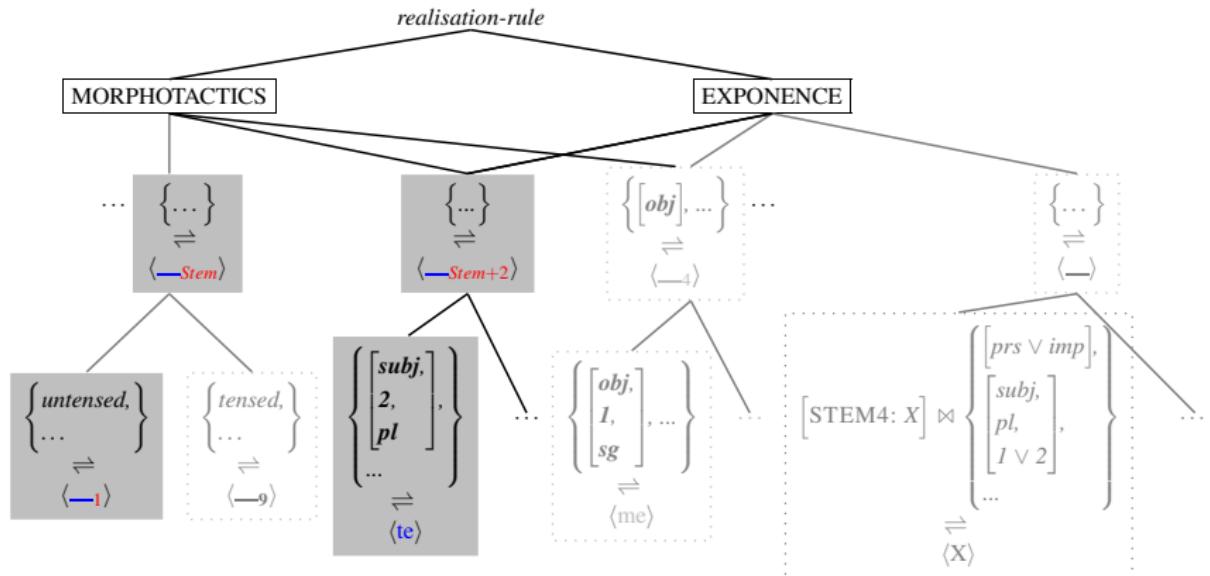
# Italian III: stems and tense/agreement affixes



$[sapere] \bowtie \{ prs, [subj, 2pl], [d-obj, 3sg, mas], [i-obj, 1sg] \} \stackrel{*}{\Rightarrow} \langle me_4, lo_7, sape_9, te_{11} \rangle$

$[sapere] \bowtie \{ imp, [subj, 2pl], [dobj, 3sg, mas], [iobj, 1sg] \} \stackrel{*}{\Rightarrow} \langle sappia_1, te_3, me_4, lo_7 \rangle$

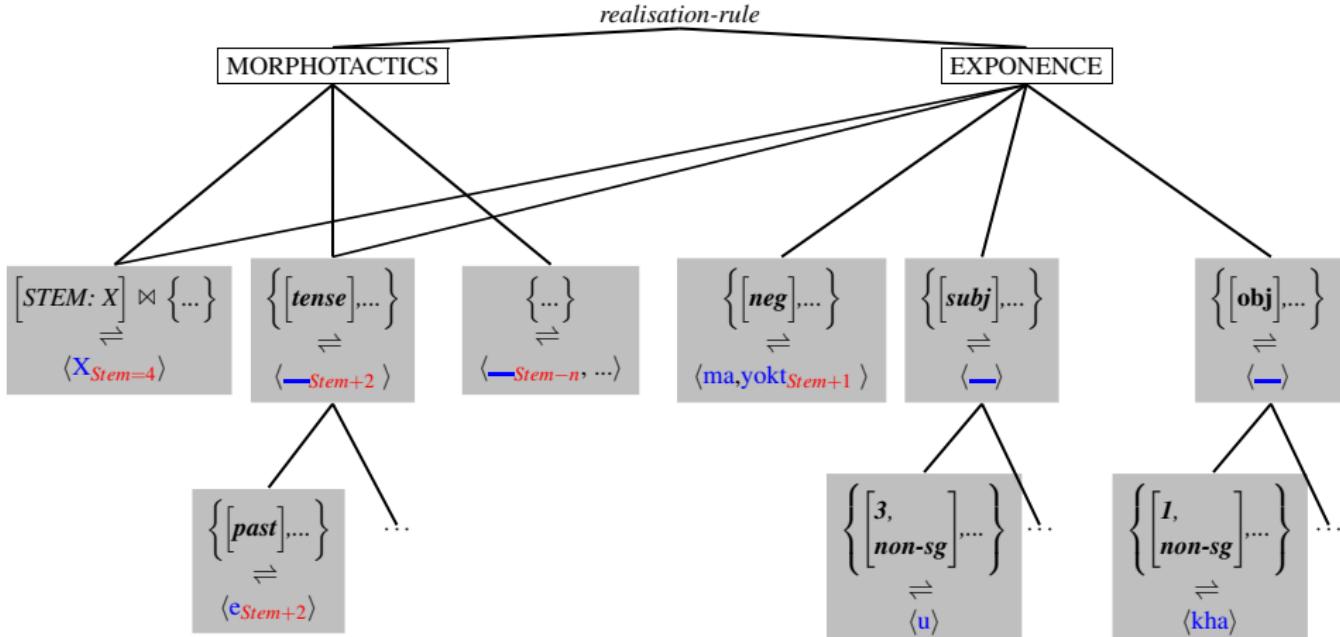
# Italian III: stems and tense/agreement affixes



$$[\textit{sapere}] \bowtie \left\{ \textit{prs}, [\textit{subj}, \textit{2pl}], [\textit{d-obj}, \textit{3sg}, \textit{mas}], [\textit{i-obj}, \textit{1sg}] \right\} \stackrel{*}{\Rightarrow} \langle \textit{me}_4, \textit{lo}_7, \textit{sape}_9, \textit{te}_{11} \rangle$$

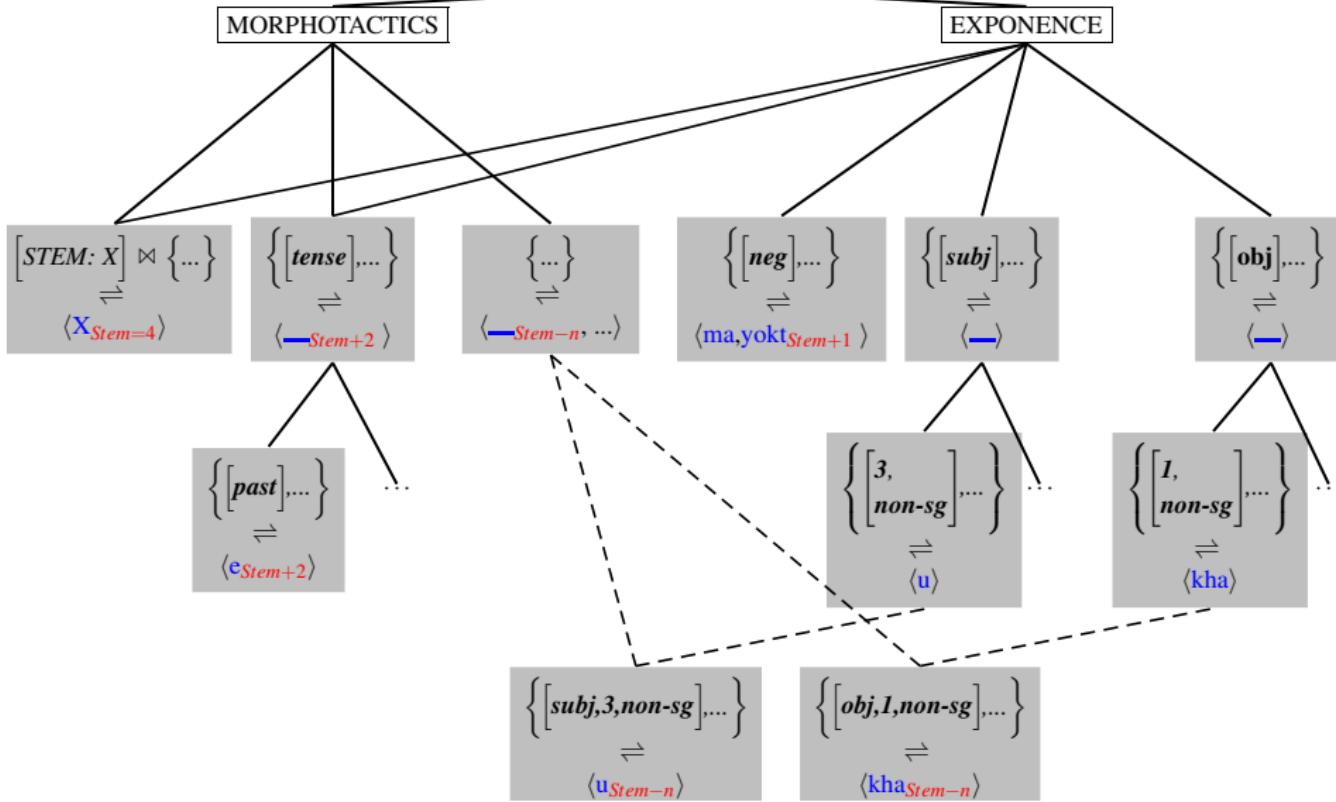
$$[\textit{sapere}] \bowtie \left\{ \textit{imp}, [\textit{subj}, \textit{2pl}], [\textit{dobj}, \textit{3sg}, \textit{mas}], [\textit{iobj}, \textit{1sg}] \right\} \stackrel{*}{\Rightarrow} \langle \textit{sappia}_1, \textit{te}_3, \textit{me}_4, \textit{lo}_7 \rangle$$

# Chintang



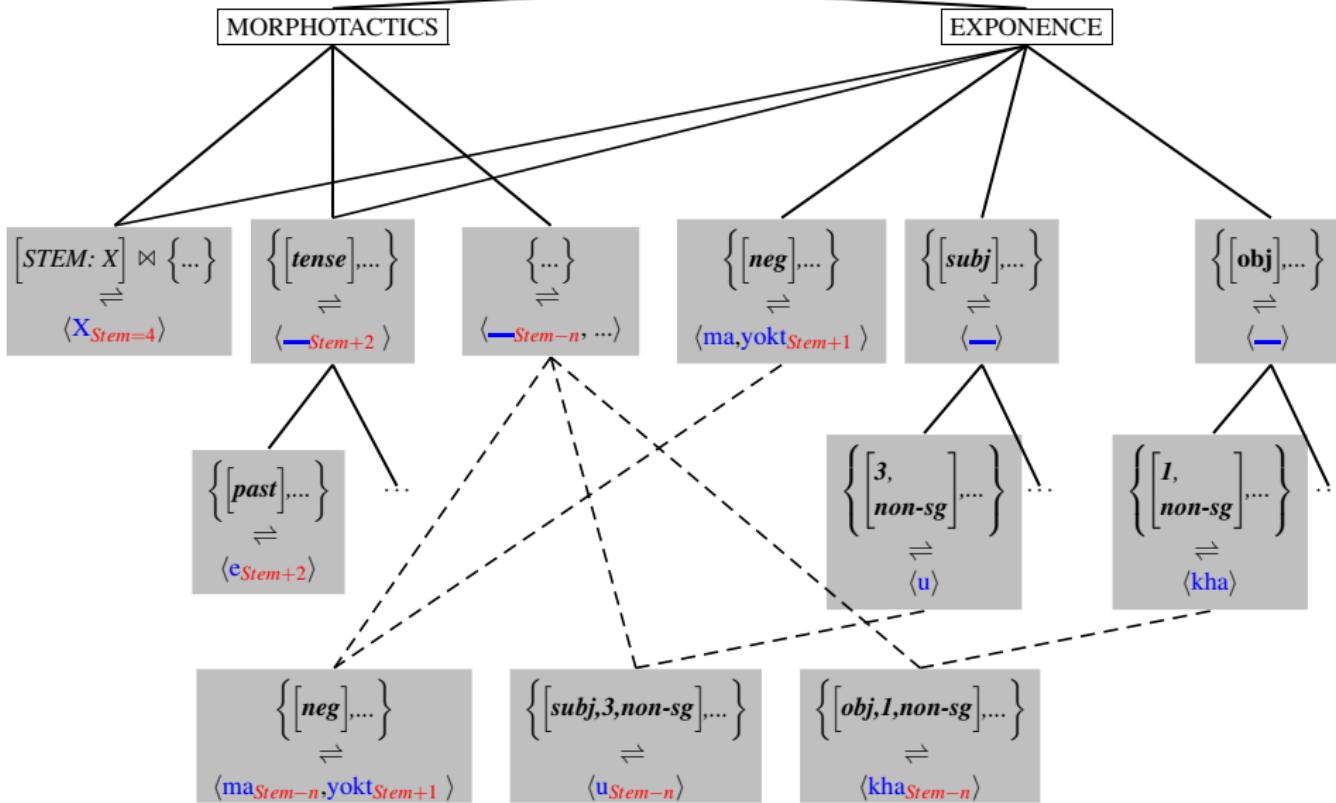
# Chintang

*realisation-rule*



# Chintang

*realisation-rule*



# Conclusions

- ▶ Empirical claims:
  - ▶ Languages use a variety of morph indexing strategies
    - ▶ Absolute positions
    - ▶ Position relative to stem
    - ▶ (partially) free positioning
    - ▶ Second position: position relative to first realized morph
  - ▶ These strategies commonly interact within a single system
- ▶ Main theoretical claim: all these strategies can be modelled as varieties of positional underspecification, using inheritance hierarchies.
- ▶ Properties of the proposed framework:
  - ▶ lexicalist, inferential, realizational, Pāṇinian
  - ▶ because it is constraint-based, can easily scale up to phonological and semantic conditions on order, where these are found
  - ▶ morph-based, but actually postulates less structure in morphology than existing amorphous frameworks
    - ▶ no recursive rule application
    - ▶ no extrinsic indexing of rules
    - ▶ morph boundaries exist but are invisible both to realization rules and to postlexical phonology

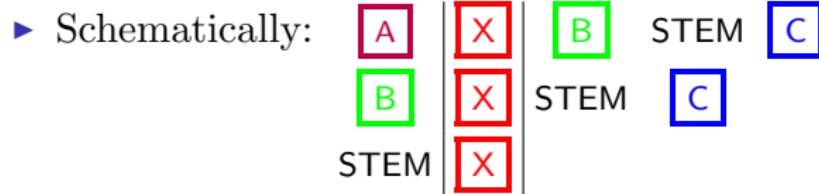
## Second position affixes: Sorani Kurdish

- ▶ Sorani Kurdish endoclitic person markers (Samvelian, 2007):
  - ▶ Realize either subject agreement or object pronouns depending on tense and transitivity
  - ▶ Linearize inside the verb whenever it is the first word of the VP
  - ▶ Normally linearize in second position

---

1	2	3	4	
		nard=jân	im	'they sent me'
na=jân		nard	im	'they did not send me'
	da=jân	nard	im	'they were sending me'
na=jân	da	nard	im	'they were not sending me'

---



- ▶ Illustrates a third type of indexing: relative to the first realized element

## Sorani Kurdish: two complications

- The endoclitics are really alternating between second position and an ordinary suffixal position class, depending on morphosyntactic conditions.

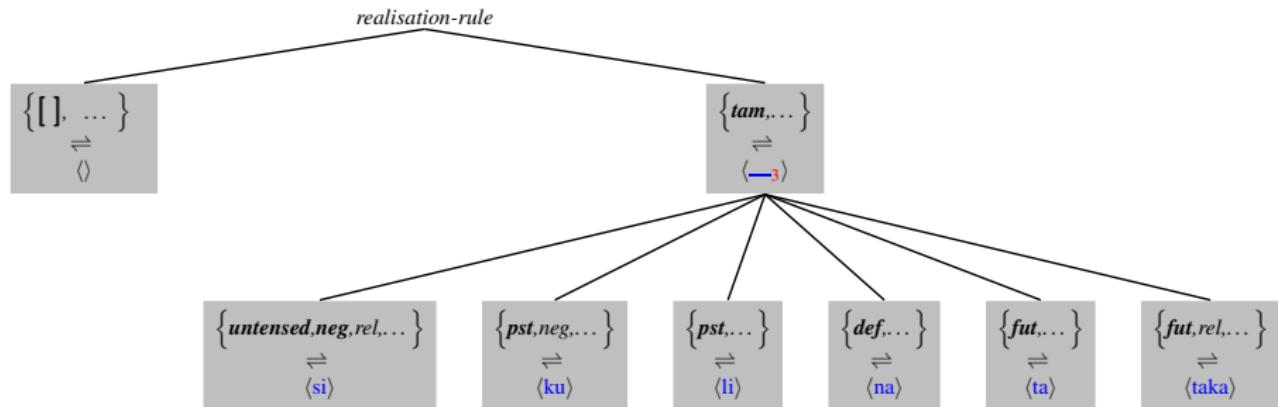
1	2	3	4	
		nard=jân	im	'they sent me'
na=jân	nard	im		'they did not send me'
	nard	-im=î	'he/she sent me'	
na	nard	-im=î	'he/she did not send me'	

- There are two recently morphologized perfect periphrases that apparently disrupt the second position effect (Walther, to appear)

1	2	2b	3	4	
		nard	u=jân	im	'they have sent me'
na=jân	nard	u		im	'they haven't sent me'
	nard	bu=jân	im		'they have sent me'
na=jân	nard	bu	im		'they haven't sent me'

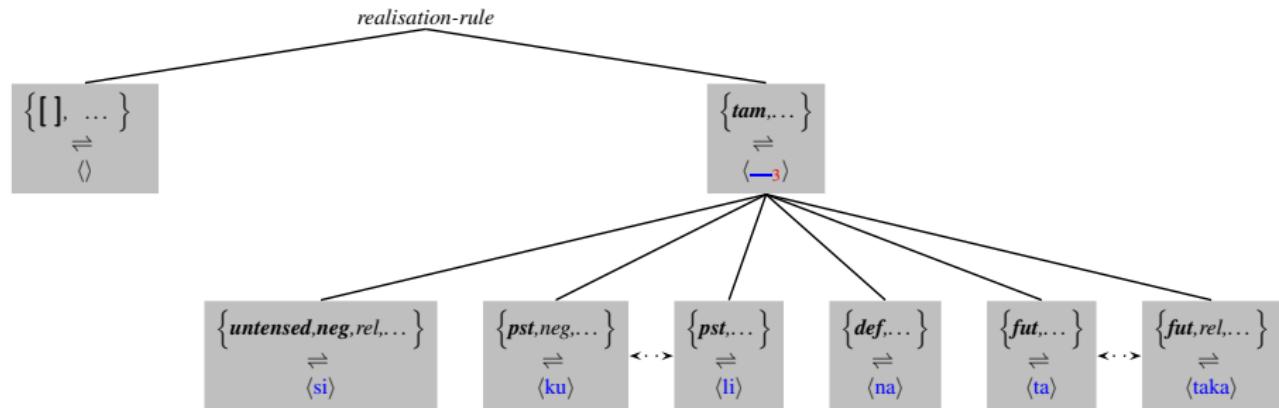
# Pāṇinian competition

Can be integrated in a monotonic HPSG grammar through a closure operation on realization rules (Crysman and Bonami, to appear)



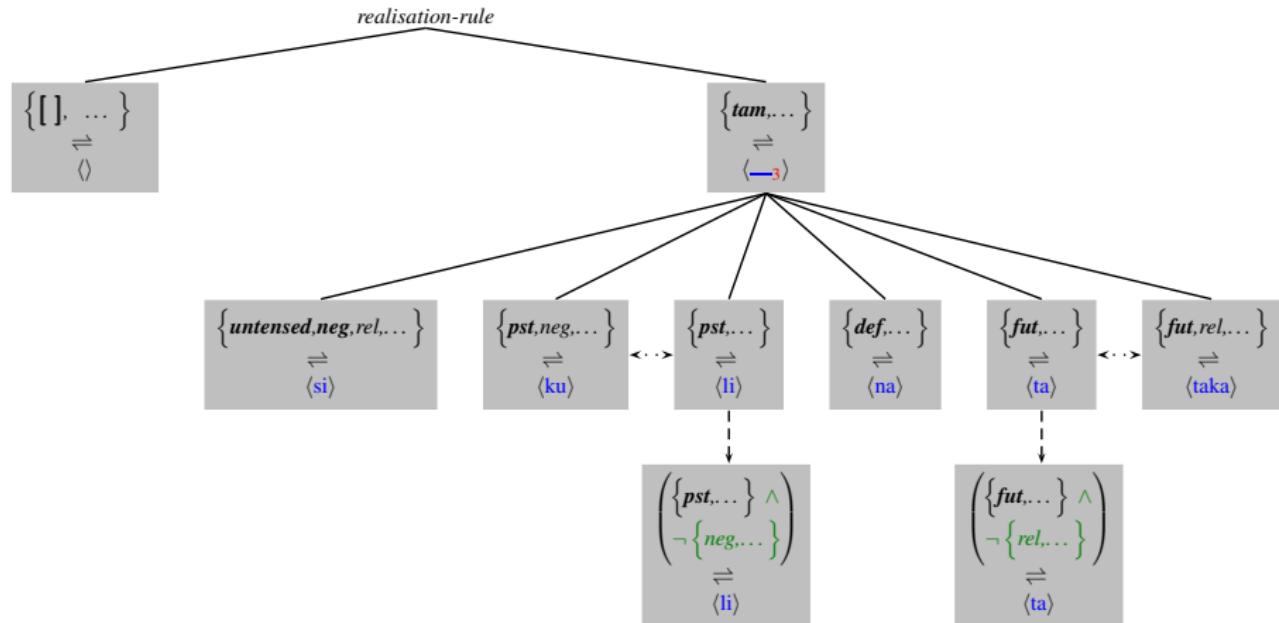
# Pāṇinian competition

Can be integrated in a monotonic HPSG grammar through a closure operation on realization rules (Crysman and Bonami, to appear)



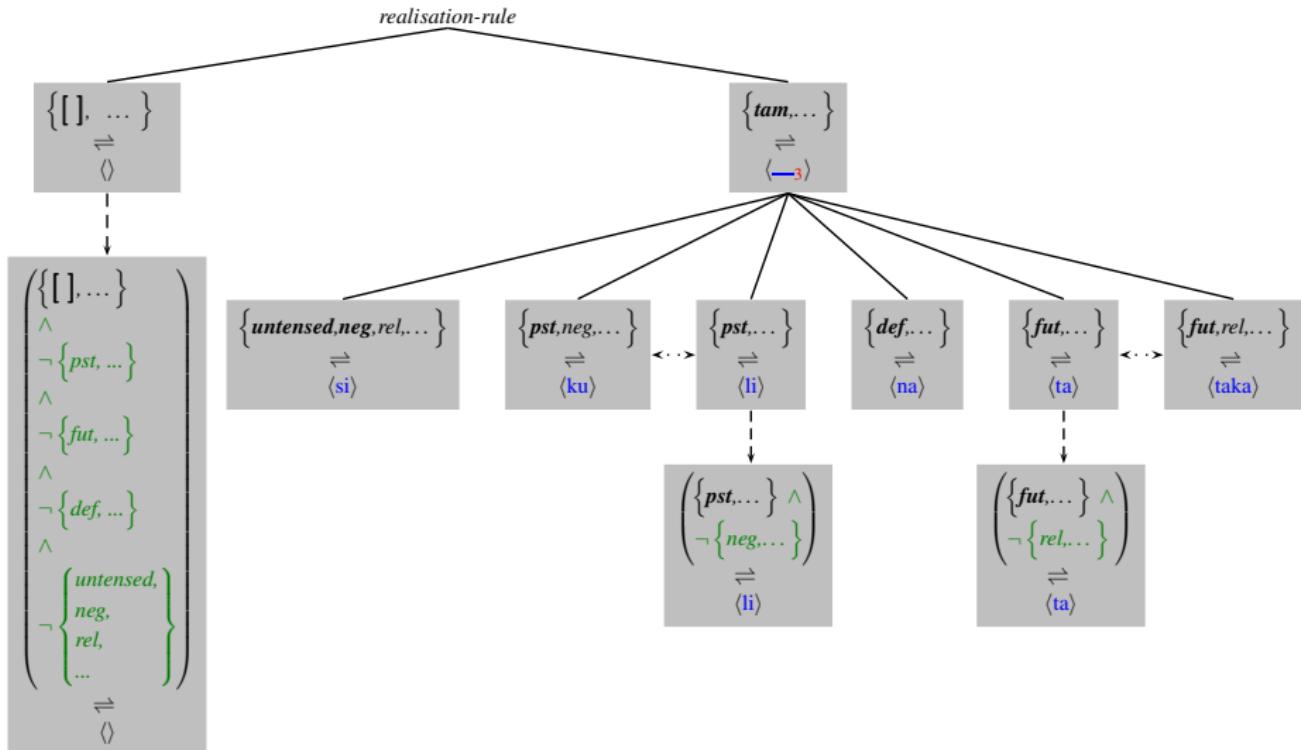
# Pāṇinian competition

Can be integrated in a monotonic HPSG grammar through a closure operation on realization rules (Crysman and Bonami, to appear)



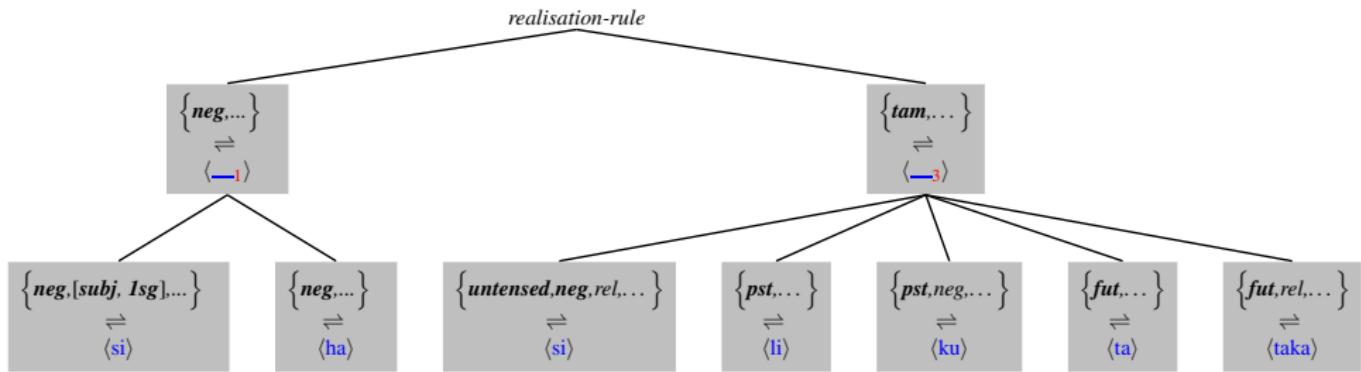
# Pāṇinian competition

Can be integrated in a monotonic HPSG grammar through a closure operation on realization rules (Crysman and Bonami, to appear)



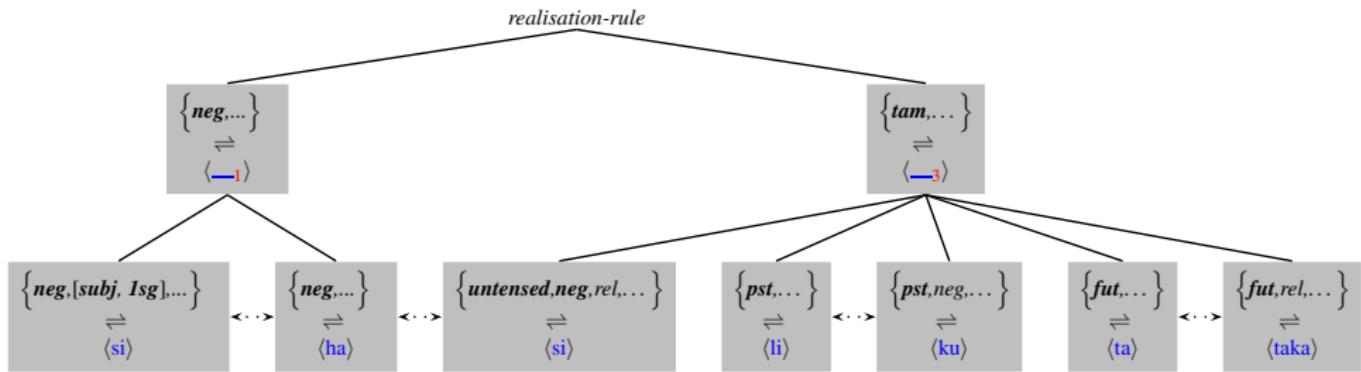
## Pāṇinian competition (II)

The determination of competitors relies on a distinction between features **expressed** by a rule (***bold italics***) and features merely **conditioning** the application of a rule.



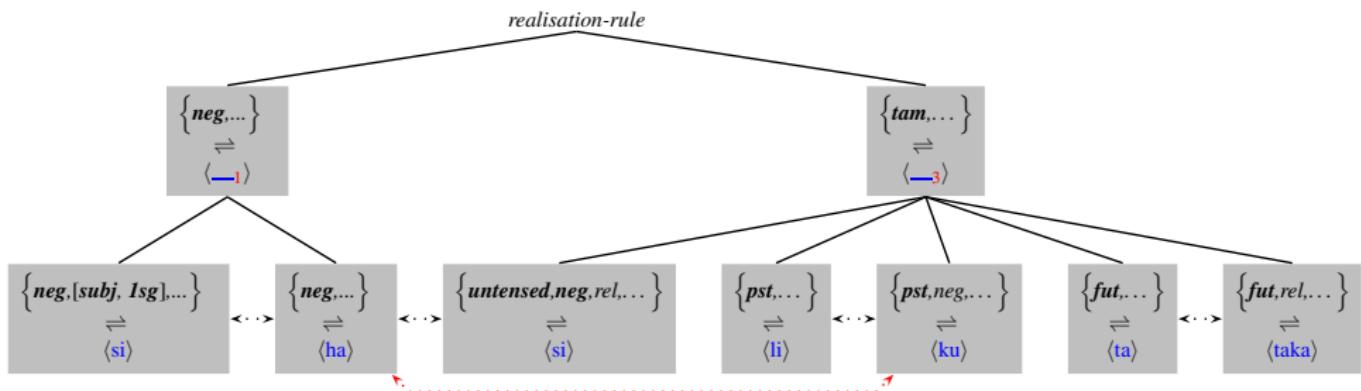
# Pāṇinian competition (II)

The determination of competitors relies on a distinction between features **expressed** by a rule (***bold italics***) and features merely **conditioning** the application of a rule.



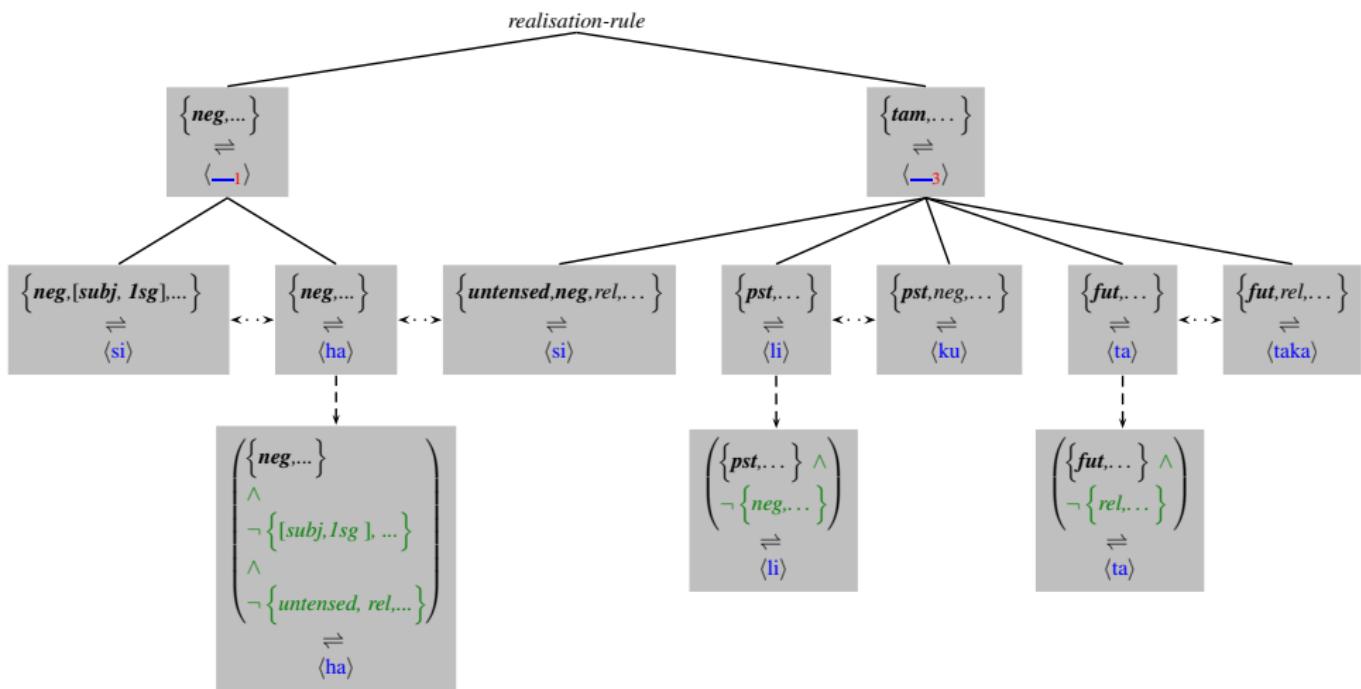
## Pāṇinian competition (II)

The determination of competitors relies on a distinction between features **expressed** by a rule (***bold italics***) and features merely **conditioning** the application of a rule.



# Pāṇinian competition (II)

The determination of competitors relies on a distinction between features **expressed** by a rule (***bold italics***) and features merely **conditioning** the application of a rule.



- Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Aronoff, M. and Xu, Z. (2010). ‘A realization optimality-theoretic approach to affix order’. *Morphology*, 20:388–411.
- Bickel, B., Banjade, G., Gaenzle, M., Lieven, E., Paudya, N. P., Rai, I. P., Manoj, R., Rai, N. K., and Stoll, S. (2007). ‘Free prefix ordering in Chintang’. *Language*, 83:43–73.
- Caballero, G. (2010). ‘Scope, phonology and morphology in an agglutinating language: Choguita Rarámuri (Tarahumara) variable suffix ordering’. *Morphology*, 20:165–204.
- Crysman, B. and Bonami, O. (to appear). ‘Establishing order in type-based realisational morphology’. In S. Müller (ed.), *Proceedings of HPSG 2012*. Stanford: CSLI publications.
- Kim, J. (2010). ‘Phonological and morphological conditions on affix order in Huave’. *Morphology*, 20:133–163.
- Koenig, J.-P. (1999). *Lexical relations*. Stanford: CSLI Publications.
- Koenig, J.-P. and Jurafsky, D. (1994). ‘Type underspecification and on-line type construction’. In *Proceedings of WCCFL XIII*. 270–285.
- Lieber, R. (1980). *On the Organization of the Lexicon*. Ph.D. thesis, MIT. Published in 1990 by Garland Publishing, New York.
- Luutonen, J. (1997). *The variation of morpheme order in Mari declension*. Helsinki: Suomalais-Ugrilainen Seura.
- Monachesi, P. (1999). *A lexical approach to Italian cliticization*. Stanford: CSLI Publications.
- Montermini, F. and Boyé, G. (2012). ‘Stem relations and inflection class assignment in Italian’. *Word Structure*.
- Paster, M. (2006). *Phonological conditions on affixation*. Ph.D. thesis, University of California Berkeley.
- Rice, K. (2000). *Morpheme Order and Semantic Scope. Word Formation in the Athabaskan Verb*. Cambridge: Cambridge University Press.
- Samvelian, P. (2007). ‘What Sorani Kurdish absolute prepositions tell us about cliticization’. In *Texas Linguistics Society IX*. 265–285.
- Stump, G. T. (1993). ‘Position classes and morphological theory’. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1992*. Dordrecht: Kluwer, 129–180.
- (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Walther, G. (to appear). “fitting into morphological structure : accounting for Sorani Kurdish endoclitics”. In *Proceedings of MMM8*.