Derivational paradigmatic models put to test on some non-canonical phenomena

Nabil Hathout[a] and **Fiammetta Namer**[b]

[a]CLLE, CNRS & Université Toulouse Jean Jaurès
[b]Université de Lorraine & ATILF, CNRS

ISMO 2021. 22-24 September 2021

Goal

Contestants: Four paradigm-based or paradigm-oriented derivational frameworks are compared:

- Construction Morphology (CxM) [Booij, 2010]
- Cumulative Patterns (CP) [Bochner, 1993]
- Paradigmatic Systems (PS) [Bonami & Strnadová, 2019]
- Paradigms vs Discrepancies (ParaDis) [Hathout & Namer, 2016]

Benchmark: The comparisons focus on their ability to:

- account for meaning-form discrepancies
- represent non-canonical derivation (defectiveness, suppletion, doublets)

Data

- We use the following example from [Bonami&Strnadová2019] as dataset for our benchmarking study.
- In each family, members in the same column are in the same semantic relations: (metaphorical) place and agent nouns, relational adjectives, locative verbs.

Data

commerce 'market'	*commerçant* 'shopkeeper'	*commercial* 'commercial'	*commercialiser* 'to (put on) market'
école 'school'	*écolier* 'schoolboy'	*scolaire* 'educational'	*scolariser* 'to send to school'
prison 'prison'	*prisonnier* 'inmate'	*carcéral* 'of prison'	*emprisonner* 'to imprison' *incarcérer* 'to imprison'

Data

- Form-meaning discrepancy is illustrated with the family of *commercialiser*, formally derived from the adjective *commercial*, but semantically defined with respect to the noun *commerce*.

commerce ‘market’	*commerçant* ‘shopkeeper’	*commercial* ‘commercial’	*commercialiser* ‘to (put on) market’
école ‘school’	*écolier* ‘schoolboy’	*scolaire* ‘educational’	*scolariser* ‘to send to school’
prison ‘prison’	*prisonnier* ‘inmate’	*carcéral* ‘of prison’	*emprisonner* ‘to imprison’ *incarcérer* ‘to imprison’

Data

- An example of defectiveness is given by the lack of relational adjective formally derived from *prison*.

commerce 'market'	*commerçant* 'shopkeeper'	*commercial* 'commercial'	*commercialiser* 'to (put on) market'
école 'school'	*écolier* 'schoolboy'	*scolaire* 'educational'	*scolariser* 'to send to school'
prison 'prison'	*prisonnier* 'inmate'	0 *carcéral* 'of prison'	*emprisonner* 'to imprison' *incarcérer* 'to imprison'

Data

- A case of suppletion is given by *scolaire*, *scolariser*, where the latinate stem /skɔl/ is used instead of /ekɔl/.
- Suppletion entails defectiveness: the families of *école* and *scolaire* are defective.

Data

commerce 'market'	*commerçant* 'shopkeeper'	*commercial* 'commercial'	*commercialiser* 'to (put on) market'
école 'school'	*écolier* 'schoolboy'	*scolaire* 'educational'	*scolariser* 'to send to school'
prison 'prison'	*prisonnier* 'inmate'	*carcéral* 'of prison'	*emprisonner* 'to imprison' *incarcérer* 'to imprison'

Data

- {*incarcérer*, *emprisonner*} is an example of verb doublet. *incarcérer* is derived from the suppletive stem /kaʁsɛʁ/ of *prison*.
- In this family, defectiveness, suppletion and doublet.

commerce 'market'	*commerçant* 'shopkeeper'	*commercial* 'commercial'	*commercialiser* 'to (put on) market'
école 'school'	*écolier* 'schoolboy'	*scolaire* 'educational'	*scolariser* 'to send to school'
prison 'prison'	*prisonnier* 'inmate'	*carcéral* 'of prison'	*emprisonner* 'to imprison' *incarcérer* 'to imprison'

Paradigmatic Features and Structures

We also want to test each model wrt to their ability to explicitely represent paradigm-based structures, that is:

- derivational families
- concrete paradigms as superposition of aligned derivational families
- abstract paradigms as the corresponding network of relations between patterns
- paradigmatic levels and generalization of paradigmatic organization

Construction Morphology (CxM), [Booij, 2010]

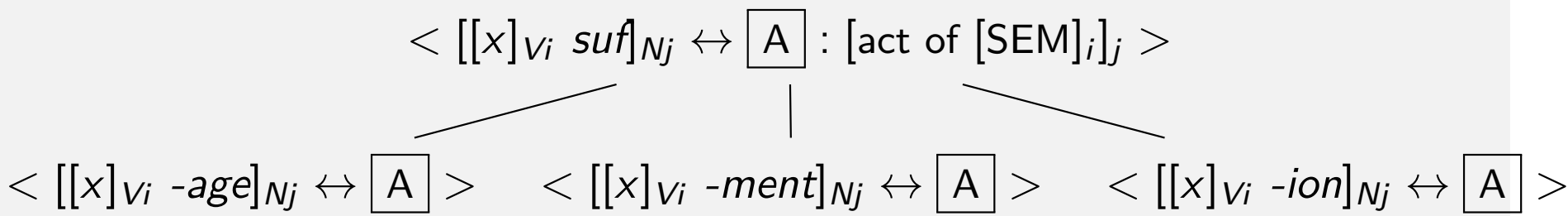
- Constructional schemas associate the formal and semantic properties of (complex) words/word patterns
- They represent both lexemes and word formation processes.
- Coindexations connect the derived lexemes to their base

laveur $\leftrightarrow$ **laver**	**WF process**
$<$ [[*lav*]$_{Vi}$ eur]$_{Nj}$ $\leftrightarrow$ [he who [wash]$_i$]$_j$ $>$	$<$ [[*x*]$_{V\,i}$ eur]$_{N\,j}$ $\leftrightarrow$ [he who [SEM]$_i$]$_j$ $>$
$<$ [*lav*]$_{Vi}$ $\leftrightarrow$ [wash]$_i$ $>$	$<$ [*x*]$_{Vi}$ $\leftrightarrow$ [SEM]$_i$ $>$

Construction Morphology (CxM), [Booij, 2010]

- Constructions are part of a hierarchical lexicon
- Inheritance relations can be used to describe affix rivalry

Inheritance Relations



Construction Morphology (CxM), [Booij, 2010]

- Second order schemas ($\approx$) are used to describe indirect relations

laveur 'washer' / _lavage_ 'washing'

$< [[x]_{Vi}\text{ -eur}]_{Nj} \leftrightarrow [\text{he who } [\text{SEM}]_i]_j > \approx < [[x]_{Vi}\ age]_{Nk} \leftrightarrow [\text{act of } [\text{SEM}]_i]_k >$

- We can represent derivational families and create paradigmatic representations by generalizing second orders schemas

Paradigmatic Representation: {laver, laveur, lavage}

$$< [x]_{Vi} \leftrightarrow [\text{SEM}]_i > \approx < [[x]_{Vi}\text{ -eur}]_{Nj} \leftrightarrow [\text{he who } [\text{SEM}]_i]_j > \approx$$
$$< [[x]_{Vi}\ age]_{Nk} \leftrightarrow [\text{act of } [\text{SEM}]_i]_k >$$

Cumulative Patterns (CP), [Bochner,1993]

- Derivational families are Cumulative Sets (CSs), where words are connected in undirected relations,

Two CSs: *laver* 'to wash', *saler* 'to salt'

{*laver*, *lavage*, *laveur*} {*saler*, *salage*, *saleur*}

- Relations in CSs are instances of relations between patterns
- Semantic and formal sharing in CSs is ensured by variable sharing in the corresponding abstract relation.

Cumulative Patterns (CP), [Bochner,1993]

Relation between abstract patterns

$$\begin{bmatrix} /X/ \\ V \\ `Z' \end{bmatrix} \leftrightarrow \begin{bmatrix} /X\text{ -age}/ \\ N \\ `\text{act of } Z' \end{bmatrix}, \begin{bmatrix} /X/ \\ V \\ `Z' \end{bmatrix} \leftrightarrow \begin{bmatrix} /X\text{ -eur}/ \\ N \\ `\text{he who } Z' \end{bmatrix},$$

$$\begin{bmatrix} /X\text{ -age}/ \\ N \\ `\text{act of } Z' \end{bmatrix} \leftrightarrow \begin{bmatrix} /X\text{ -eur}/ \\ N \\ `\text{he who } Z' \end{bmatrix}$$

Cumulative Patterns (CP), [Bochner,1993]

- Relations btw abstract patterns form a complete graph. Therefore they can be aggregated in an abstract set: Cumulative Pattern (CP).

Cumulative pattern

$$\left\{ \begin{bmatrix} /X/ \\ V \\ \text{‘Z’} \end{bmatrix}, \begin{bmatrix} /X\text{ -age}/ \\ N \\ \text{‘act of Z’} \end{bmatrix}, \begin{bmatrix} /X\text{ -eur}/ \\ N \\ \text{‘he who Z’} \end{bmatrix} \right\}$$

- Relations between patterns are evaluated by an *evaluation metric* in terms of productivity.
- Hierarchy relations between CPs are possible, but their relevance is limited because of the loss of predictability
- Affix rivalry is not represented because it causes a loss of predictability

Paradigmatic Systems (PS), [Bonami&Strnadová,2019]

- Morphological families are complete graphs of morphologically related words
- Morphological paradigms are superposition of morphological families whose elements are connected by the same relations of content-based contrasts.
- Families are all the same size
- There is a systematic content-based contrast between all pairs of cells.
  - contrast of forms is secondary

Verb	Agent_N	Action_N
laver	laveur	lavage
former	formateur	formation
gonfler	gonfleur	gonflement

ParaDis: Families

/lav/	/lavaʒ/	/lavœʁ/

V	N

‘to wash’	‘act of washing’/‘act performed by a washer’	‘he who washes’/‘he who performs washing’

♢ Family is the basic unit
♢ Families are generalized to all the levels relevant to morphology
♢ 4 independent dimensions: formal, categorial, semantic, morphological

LAVER	LAVAGE	LAVEUR

ParaDis: Paradigms

F1	F2	F3
/f_1/	/f_1-aʒ/	/f_1-œʁ/
/lav/ (F1,1)	/lavaʒ/ (F2,1)	/lavœʁ/ (F3,1)
/sal/ (F1,2)	/salaʒ/ (F2,2)	/salœʁ/ (F3,2)

S1	S2	S3
's_1'	'act (of s_1'/performed by s_3')	'he who (s_1'/ performs s_2')
'to wash' (S1,1)	'act (of wash' / performed by a washer') (S2,1)	'he who (washes' / performs washing') (S3,1)
'to salt' (S1,2)	'act of salt' / 'act performed by a salter' (S2,2)	'he who salts' / 'he who performs salting' (S3,2)

M1	M2	M3
LAVER (M1,1)	LAVAGE (M2,1)	LAVEUR (M3,1)
SALER (M1,2)	SALAGE (M2,2)	SALEUR (M3,2)

♢ Concrete paradigms: superposition of families with members in the same relations of contrast

♢ Aligned members form series

♢ Abstract paradigms form graphs of relations btw patterns

- ♢ (S1,S2,S3) forms a complete graph
- ♢ In (F1,F2,F3) F2 and F3 are not connected

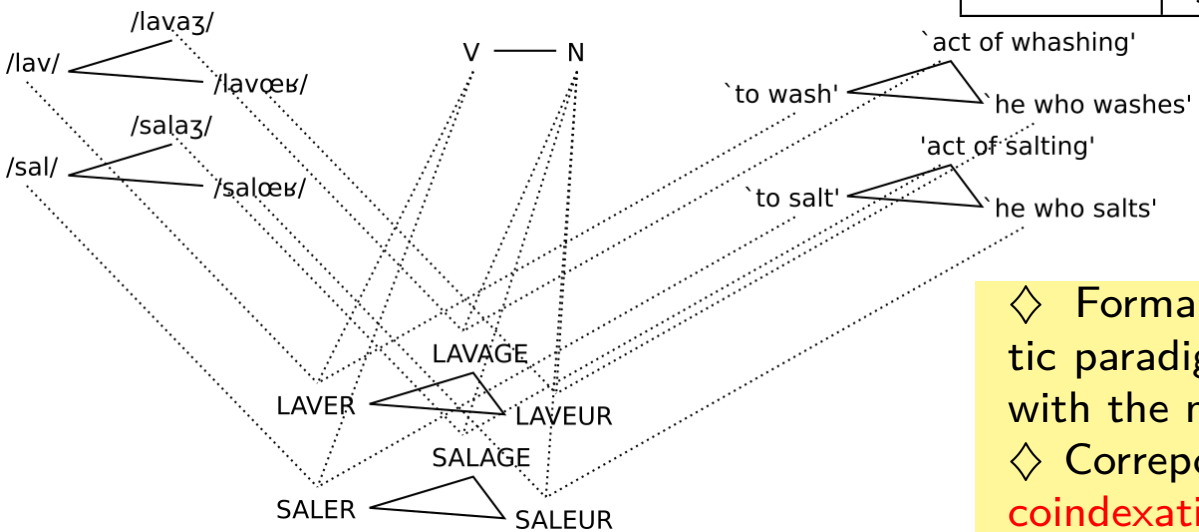
♢ (Categorial families and paradigm omitted here)

ParaDis: Correspondences

F1	F2	F3
/f_1/	/f_1-aʒ/	/f_1-œʁ/
/lav/ (F1,1)	/lavaʒ/ (F2,1)	/lavœʁ/ (F3,1)
/sal/ (F1,2)	/salaʒ/ (F2,2)	/salœʁ/ (F3,2)

C1	C2
$c1$	$c2$
V	N
(C1,1)	(C2,1)

S1	S2	S3
's_1'	'act of s_1'/'act performed by s_3'	'he who s_1'/ 'he who performs s_2'
'to wash' (S1,1)	'act of wash' / 'act performed by a washer' (S2,1)	'he who washes' / 'he who performs washing' (S3,1)
'to salt' (S1,2)	'act of salt' / 'act performed by a salter' (S2,2)	'he who salts' / 'he who performs salting' (S3,2)



♢ Formal, categorial and semantic paradigms are in correspondence with the morphological paradigm
♢ Correpondences are expressed by coindexation

M1	M2	M3
LAVER (M1,1) F1,1/C1,1/S1,1	LAVAGE (M2,1) F2,1/C2,1/S2,1	LAVEUR (M3,1) F3,1/C2,1/S3,1
SALER (M1,2) F1,2/C1,1/S1,2	SALAGE (M2,2) F2,2/C2,1/S2,2	SALEUR (M3,2) F3,2/C2,1/S3,2

ParaDis: Derivational Paradigms

M4	M5	M6
LANCER	LANCEMENT	LANCEUR
RONFLER	RONFLEMENT	RONFLEUR

M1	M2	M3
LAVER	LAVAGE	LAVEUR
SALER	SALAGE	SALEUR

M1	M2	M3
M4	M5	M6

Derivational Paradigm

S1	S2	S3
's_1'	'act of s_1'/'act performed by s_3'	'he who s_1'/ 'he who performs s_2'
'to launch'	'act of launching' / 'act performed by the launcher'	'he who launches' / 'he who performs launching'
'to wash'	'act of washing' / 'act performed by the washer'	'he who washes' / 'he who performs washing'

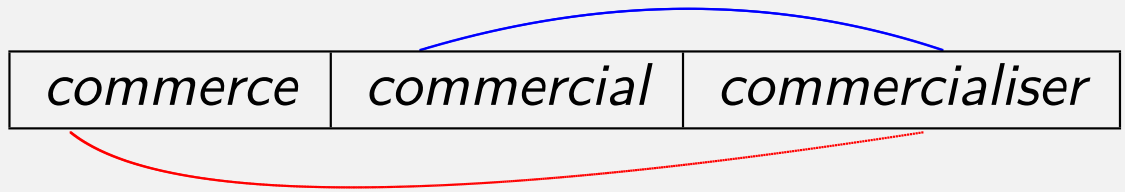
♢ Morphological paradigms sharing the same semantic paradigm can be superposed into a derivational paradigm
→ "paradigm of paradigms"
♢ (M1,M2, M3) and (M4, M5, M6) are in correspondence with (S1,S2,S3)
♢ They form a Derivational Paradigm

Beauty Contest: first result

	CxM	CP	PS	ParaDis
Families	(✓)	✓	✓	✓
Concrete paradigms			✓	✓
Abstract paradigms	✓	✓	?	✓
Generalization of Paradigmatic Organisation				✓

Form-Meaning Discrepancies

Have a separate access to the verbal formal and semantic description



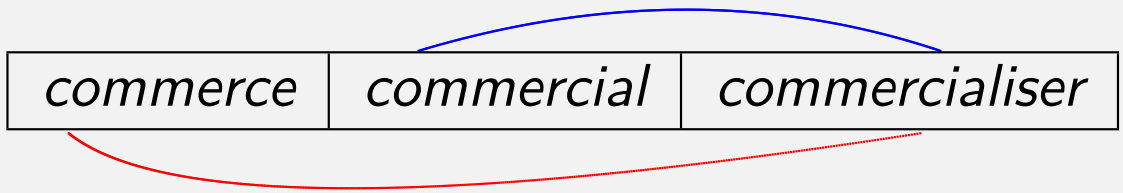
- CxM - multiple coindexation system

CxM

$$< [[[x]_{Ni} \text{ -al}]_{Aj}\text{-iser}]_{Vw} \leftrightarrow [\text{related to } [\text{SEM}]_i]_w >$$

Form-Meaning Discrepancies

Have a separate access to the verbal formal and semantic description



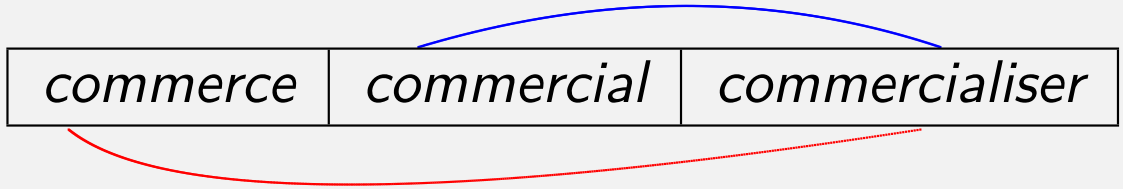
- CP - multiple variable sharing

CP

$$\left\{ \begin{bmatrix} /X/ \\ N \\ `Z' \end{bmatrix}, \begin{bmatrix} /Y/ = /X\text{ -al}/ \\ A \\ `\text{related to Z}' \end{bmatrix}, \begin{bmatrix} /Y\text{ -iser}/ \\ V \\ `\text{put on Z}' \end{bmatrix} \right\}$$

Form-Meaning Discrepancies

Have a separate access to the verbal formal and semantic description



- PS - paradigms are semantic-based, discrepancies are ignored
- ParaDis - form and meaning are represented at independent levels

ParaDis

S9	S11	S12
's_9'	'related to s_9'	'put on s_9'
'market'	'related to market'	'put on the market'

S11 and S12 unrelated

F9	F11	F12
/f_9/	/f_{11}/=/f_9-al/	/f_{12}/ = /f_{11}-iz/
/kɔmɛʁs/	/kɔmɛʁsjal/	kɔmɛʁsjaliz/

complete graph

M9	M11	M12
COMMERCE	COMMERCIAL	COMMERCIALISER

Defectiveness

Represent incomplete families

école *commerce*	*écolier* *commerçant*	*commercial*	*commercialiser*

- CxM - no
- CP - no
- PS - Families are made of sets of lexemes (and not lexemes). A gap in a defective family is an empty set.

{*école*}	{*écolier*}	{ }	{ }

Defectiveness

Represent incomplete families

école *commerce*	*écolier* *commerçant*	*commercial*	*commercialiser*

- ParaDis - Different morphological paradigms, the same semantic paradigm: superposition in the same derivational paradigm. Gaps revealed at paradigm-level.

ParaDis

M13	M14		
ÉCOLE	ÉCOLIER		
M9	M10	M11	M12
COMMERCE	COMMERÇANT	COMMERCIAL	COMMERCIALISER

Suppletion

Represent differences and relatedness

école	*écol*ier	*scol*aire	*scol*ariser

- CxM - independent variables, same semantic content

CxM

$< [x]_{Ni} \leftrightarrow [\text{SEM}]_i > \approx < [[x]_{Ni}\ \text{-ier}]_{Nj} \leftrightarrow [\text{sme in relation with } [\text{SEM}]_i]_j > \approx$
$< [[y]_{Ni}\ \text{-aire}]_{Ak} \leftrightarrow [\text{related to } [\text{SEM}]_i]_k > \approx$
$< [[[y]_{Ni}\ \text{-aire}]_{Ak}\ \text{-iser}]_{Vw} \leftrightarrow [\text{put to } [\text{SEM}]_i]_w >$

Suppletion

Represent differences and relatedness

école	*écol*ier	*scol*aire	*scol*ariser

- CP - independent formal variables, same semantic variable

CP

$$\left\{ \begin{bmatrix} /X/ \\ N \\ \text{`Z'} \end{bmatrix}, \begin{bmatrix} /X \text{ -ier}/ \\ N \\ \text{`sme in relation with Z'} \end{bmatrix}, \begin{bmatrix} W = /Y \text{ -aire}/ \\ A \\ \text{`related to Z'} \end{bmatrix}, \begin{bmatrix} /W \text{ -iser}/ \\ V \\ \text{`put to Z'} \end{bmatrix} \right\}$$

Suppletion

Represent differences and relatedness

école	*écolier*	*scolaire*	*scolariser*

- PS - not relevant (secondariness of formal contrasts)
- ParaDis - different morphological families and paradigms but the same semantic paradigm and even the same semantic family
  - superposed in the same derivational paradigm.

ParaDis

		M15	M16
		SCOLAIRE	SCOLARISER
M13	M14		
ÉCOLE	ÉCOLIER		
M9	M10	M11	M12
COMMERCE	COMMERÇANT	COMMERCIAL	COMMERCIALISER

n-uplets

Represent differences, semantic identity and same cell position in the family

prison	*prisonnier*		*emprisonner*
		carcéral	*incarcérer*

- CxM - second order schema with the same semantic content

CxM

$< [\text{en-}\ [x]_{Ni}]_{Vj} \leftrightarrow [\text{put in}[\text{SEM}]_i]_j > \approx$

$< [\text{in-}\ [y]_{Ni}]_{Vk} \leftrightarrow [\text{put in}[\text{SEM}]_i]_k >$

n-uplets

Represent differences, semantic identity and same cell position in the family

prison	*prisonnier*		*emprisonner*
		carcéral	*incarcérer*

- CP - 5-members CS with the same semantic content

CP

$$\left\{ \begin{bmatrix} /X/ \\ N \\ `Z' \end{bmatrix}, \begin{bmatrix} /X\text{ -ier}/ \\ N \\ `\text{sme in relation with Z}' \end{bmatrix}, \begin{bmatrix} /Y\text{ -al}/ \\ A \\ `\text{related to Z}' \end{bmatrix}, \begin{bmatrix} /\text{in- Y}/ \\ V \\ `\text{put in Z}' \end{bmatrix}, \begin{bmatrix} /\text{en- X}/ \\ V \\ `\text{put in Z}' \end{bmatrix} \right\}$$

n-uplets

Represent differences, semantic identity and same cell position in the family

prison	*prisonnier*	*carcéral*	*emprisonner* *incarcérer*

- PS - in the same two-element set

PS

{prison}	{prisonnier}	{carcéral}	{ emprisonner, incarcérer }

n-uplets

Represent differences, semantic identity and same cell position in the family

prison	*prisonnier*	*carcéral*	*emprisonner* *incarcérer*

- ParaDis - different morphological families and paradigms, but same semantic paradigm, the same family, and verbs in the same semantic cell
  - superposed in the same derivational paradigm

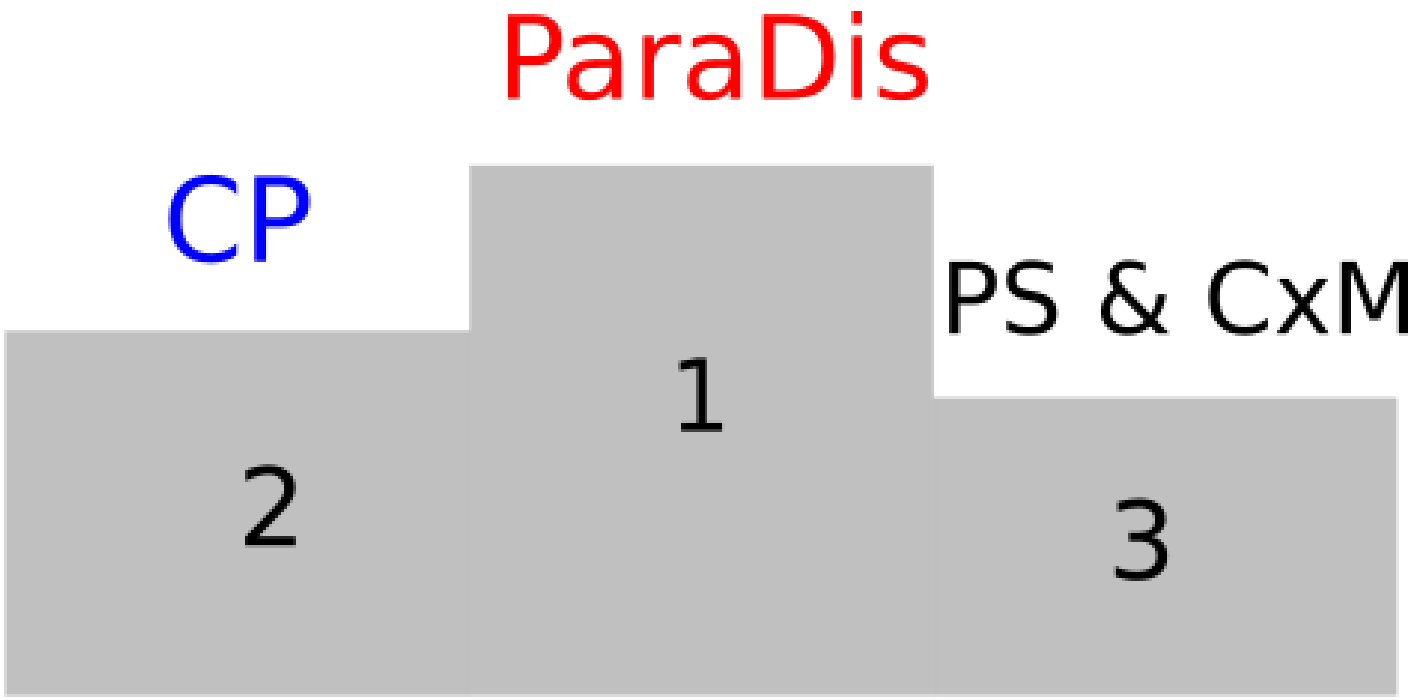
ParaDis

M19	M20		M21
PRISON	PRISONNIER		EMPRISONNER
		M17	M18
		CARCÉRAL	INCARCÉRER
		M15	M16
		SCOLAIRE	SCOLARISER
M13	M14		
ÉCOLE	ÉCOLIER		
M9	M10	M11	M12
COMMERCE	COMMERÇANT	COMMERCIAL	COMMERCIALISER

Beauty Contest: final score

	CxM	CP	PS	ParaDis
Families	(✓)	✓	✓	✓
Concrete paradigms			✓	✓
Abstract paradigms	✓	✓	?	✓
Generalization of Paradigmatic Organization				✓
Form-Meaning Discrepancies	✓	✓		✓
Defectiveness			✓	✓
Suppletion	✓	✓		✓
n-uplets	✓	✓	✓	✓

Beauty Contest: final score



Conclusion

- This is a first attempt to compare derivational paradigm-based models
- We have selected four models different enough from each other
- We have selected (some) phenomena widely used to test capabilities of derivational approaches
  - meaning-form mismatches
  - traditional infringements to canonicity
- In the future, we would like to rely on a more normalized benchmark, in line with [Corbett 2010]'s canonicity principles used to compare derivational descriptions
  - test a wider (standard) range of phenomena, in order to perform new sorts of analyses and obtain a more accurate classification of the contending models

References