Predictability in inflection and word formation

Olivier Bonami

Université Paris Diderot

ParadigMo Conference, Toulouse, June 2017

Derivational paradigms: the view from inflection

- Central question: can the tools and concepts of Word and Paradigm morphology be used to make sense of derivational paradigms?
- Central intuition: paradigms are about predictability.
  - The location of a word in a paradigmatic system predicts (more or less reliably) its form and content.
- In this talk I will deploy quantitative methods to determine whether/to what extent this can be said of structured derivational families.
- I will conclude that predictability of form definitely holds in derivational paradigms, but that inflection and derivation differ in terms of predictability of content.

Structure of the talk

Background: paradigmatic systems

Predictability of form

I. Implicative relations

Joint work with Jana Strnadová (Google, Inc.)

II. Diving into rivalry

Joint work with Juliette Thuilier (Université Toulouse Jean Jaurès)

Predictability of content

Joint work with Denis Paperno (CNRS - Loria)

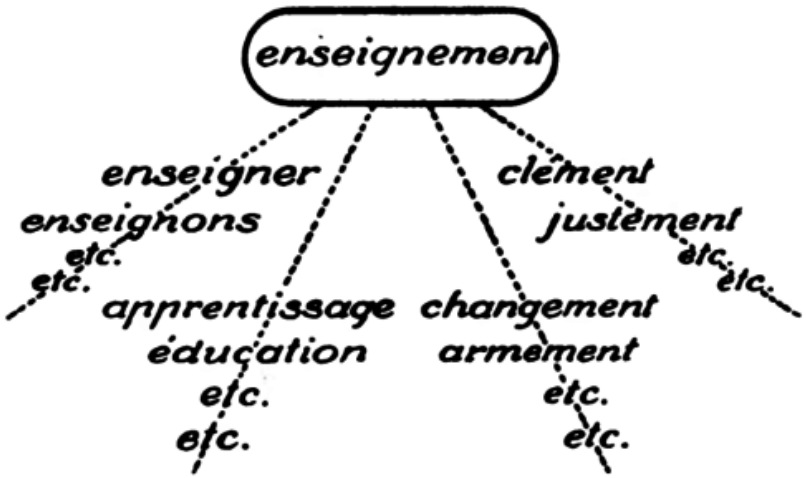
Conclusions

Background: paradigmatic systems

Two notions of paradigm I

1. *"[...] a set of linguistic elements with a common property"*
   (Booij, 2007, p. 8)

   - Here a paradigm corresponds to what Saussure called an associative series:



(Saussure, 1916, p. 175)

   - See among many others: van Marle (1984), Becker (1993), Booij (1997), Hay and Baayen (2005), Roché et al. (2011).

Two notions of paradigm II

2. An inflectional paradigm:

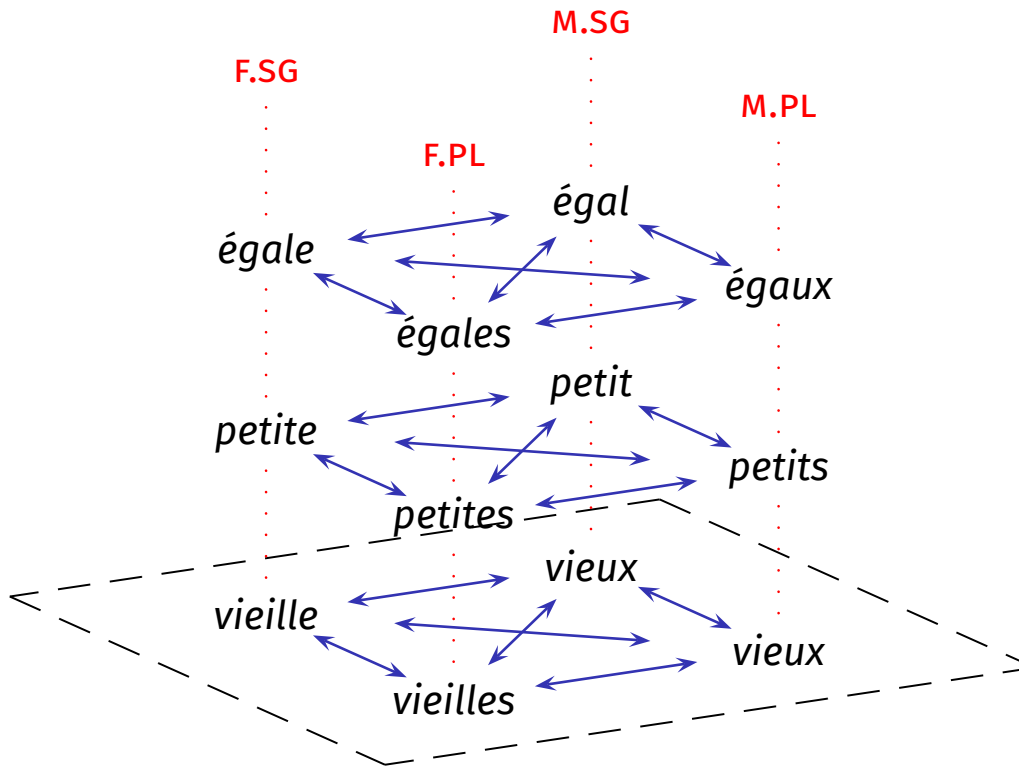
> *"[...] we define the paradigm of a lexeme L as a complete set of cells for L, where each cell is the pairing of L with a complete and coherent morphosyntactic property set (MPS) for which L is inflectable."*
>
> *(Stump and Finkel, 2013, p. 9)*

- Can such a definition be extended so as to encompass aspects of the structure of word formation systems?
  - ☞ Bauer (1997), Blevins (2001), Boyé and Schalchli (2016).
- Leading idea:
  - Inflectional paradigms are structured by contrasts in content (Štekauer, 2014).
  - If we are to make useful parallels with between inflection and derivation, then, "derivational paradigms" should also be structured in that way.

Some definitions

- Morphological family
  Set of words that are morphologically related.
  - ⇒ sets of words, not lexemes
  - ⇒ not necessarily exhaustive sets
- Paradigmatic system
  Collection of morphological families structured by the same system of oppositions of content charaterized by morphosyntacic property sets.

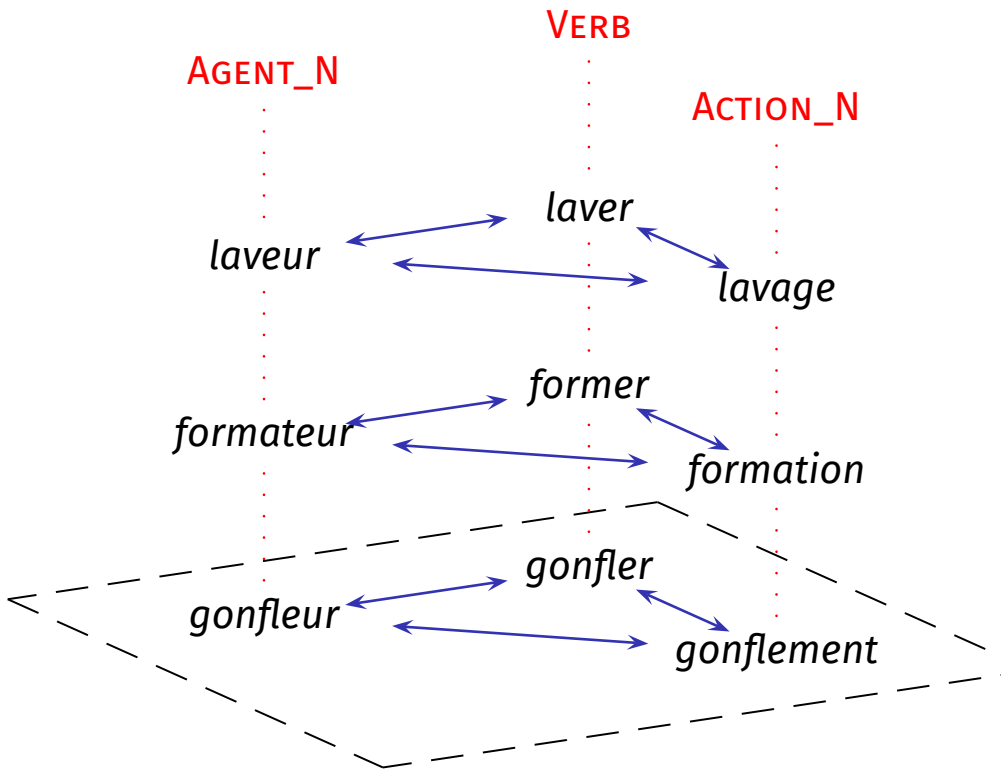
Inflectional example:



Some definitions

- Morphological family
  Set of words that are morphologically related.
  - ⇒ sets of words, not lexemes
  - ⇒ not necessarily exhaustive sets
- Paradigmatic system
  Collection of morphological families structured by the same system of oppositions of content charaterized by morphosemantic relations.

Derivational example:



Discussion

1. I take paradigmatic systems to be collections of partial morphological families. No attempt at exhaustivity.
   - ☞ Presumably, inflectional paradigms are finite, derivational are not.
   - ☞ One may focus on different (partial) paradigmatic systems for different research questions.
2. I do not take organization into orthogonal dimensions to be a defining feature of paradigms, contra e.g. Wunderlich and Fabri (1995).
   - ☞ Not obvious that this is a general property of inflectional paradigms anyway.
3. Defectiveness and overabundance require adjustments.
   - ☞ Higher order notion of paradigmatic system, where cells in the paradigm are (possibly empty) sets of words.
     (Bonami and Stump, 2016; Stump, 2016).
4. I assume that relations of content in derivation are stable enough that paradigmatic systems can be identified.
   - ☞ But see final section.

Predictability of form

I. Implicative relations

Joint work with Jana Strnadová (Google, Inc.)

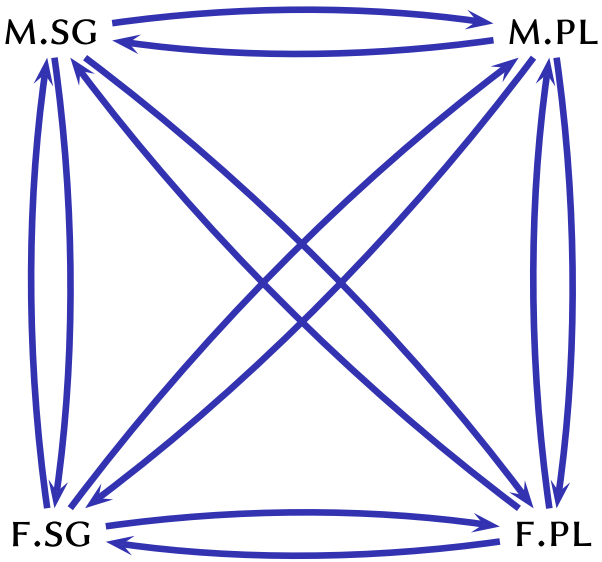
Implicative structure in inflectional paradigms

When a speaker knows only one form of a lexeme, how hard is it to predict the others?

(Ackerman et al. (2009)'s Paradigm Cell Filling Problem)

☞ See also a. o. Wurzel (1989); Ackerman and Malouf (2013); Stump and Finkel (2013); Sims (2015)

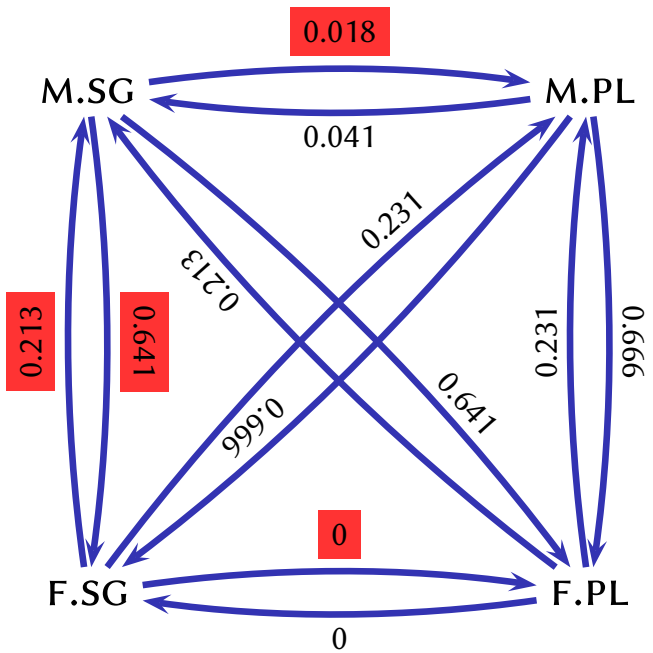
Consider French adjectives:



- F.SG⇒F.PL is trivial
- M.SG⇒M.PL is easy but not trivial, see /lokal/∼/loko/ vs. /banal/∼/banal/
- F.SG⇒M.SG is harder, see /lɛd/∼/lɛ/ vs. /ʁɛd/∼/ʁɛd/
- M.SG⇒F.SG is hardest, see /gɛ/∼/gɛ/ vs. /lɛ/∼/lɛd/ vs. /njɛ/∼/njɛz/ vs. ...

Implicative entropy

- Implicative entropy evaluates how hard it is to guess the pattern relating two words given knowledge of the shape of one word.
  - See Bonami and Beniamine (2016) for discussion of similarities and differences with Ackerman et al.'s use of conditional entropy, and Bonami and Boyé (2014); Bonami and Luís (2014) for more empirical applications.
- Among other things, implicative entropy allows one to quantify differential opacity:



Joint predictiveness

- Bonami and Beniamine (2016) generalize implicative entropy to prediction from multiple paradigm cells.

> When a speaker knows only 2, 3, ..., *n* forms of a lexeme, how hard is it to predict the remaining ones?

- On Romance conjugation, we show that on average, knowing multiple forms of the same lexeme makes the PCFP a lot easier.
- For French adjectives:

1 predictor	0.2966
2 predictors	0.1443
3 predictors	0.0044

- This provides a strong argument for paradigms as first class citizens of the morphological universe: there is useful knowledge on the system that can only be attained by attending to (sub)paradigms.

The dataset I

- We use data from Démonette (Hathout and Namer, 2014), a database of 20,493 derivational relations between 22,570 French lexemes.

…	abandonner	@	abandon	@ACT	…
…	abandonner	@	abandonneur	@AGM	…
…	abandon	@AGT	abandonneur	@AGM	…
…	abandonner	@	abandonnement	@ACT	…
…	…		…	…	

- From Démonette we tabulate 5,414 paradigms for triples (Verb, Action noun, Masculine agent noun)

@	@ACT	@AGM
abaisser	abaissement	abaisseur
abandonner	abandon;abandonnement	abandonneur;abandonnateur
abattre	abattement;abattage	abatteur
affamer		affammeur
	agriculture	agriculteur
…	…	…

The dataset II

- Since we want to deal neither with overabundance nor with defectivity:
  1. We drop all paradigms with an unfilled cell.
  2. In cases of overabundant cells, if one cell-mate makes up $\frac{2}{3}$ or more of the distribution, we drop the other cell-mates; otherwise, we drop the whole paradigm.

@	@ACT	@AGM
abaisser	abaissement	abaisseur
abandonner	abandon~~;abandonnement~~	abandonneur~~;abandonnateur~~
~~abattre~~	~~abattement;abattage~~	~~abatteur~~
~~affamer~~		~~affammeur~~
	~~agriculture~~	~~agriculteur~~
…	…	…

⇒ 1,331 remaining canonical paradigms.

The dataset III

- To assess predictibility on the basis of phonological forms, we use transcription from the GLÀFF, a lexicon derived from French Wiktionary (Hathout et al., 2014)

@	@ACT	@AGM
a.bɛ.se	a.bɛ.smɑ̃;a.bɛs.mɑ̃	a.be.sœʁ
a.bɑ̃.dɔ.ne	a.bɑ̃.dɔ̃	a.bɑ̃.dɔ.nœʁ
…	…	…

⇒ 913 paradigms for which all transcriptions are available.

Results, 1: Differential opacity

	Verb	Action_N	Agent_N
Verb	—	1.115	0.709
Action_N	0.101	—	0.269
Agent_N	0.264	1.114	—

Unary implicative entropy
for (Verb, Action_N, Agent_N) triples

Differential opacity (continued)

	Verb	Action_N	Agent_N
Verb	—	1.115	0.709
Action_N	0.101	—	0.269
Agent_N	0.264	1.114	—

Unary implicative entropy
for (Verb, Action_N, Agent_N) triples

Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
écrire	écriture	scripteur
'write'	'writing'	'writer'
gonfler	gonflement	gonfleur
'inflate'	'inflating'	'inflater'

Sample triples

- Action nouns are hardest to predict, because of the diversity of marking strategies (*-age*, *-ment*, *-ion*, *-ure*, conversion, etc.)

Differential opacity (continued)

	Verb	Action_N	Agent_N
Verb	—	1.115	0.709
Action_N	0.101	—	0.269
Agent_N	0.264	1.114	—

Unary implicative entropy
for (Verb, Action_N, Agent_N) triples

Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
écrire	écriture	scripteur
'write'	'writing'	'writer'
gonfler	gonflement	gonfleur
'inflate'	'inflating'	'inflater'

Sample triples

- Verbs are easiest to predict: the only challenging cases are stem suppletion and non-first conjugation.

Differential opacity (continued)

	Verb	Action_N	Agent_N
Verb	—	1.115	0.709
Action_N	0.101	—	0.269
Agent_N	0.264	1.114	—

Unary implicative entropy
for (Verb, Action_N, Agent_N) triples

Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
écrire	écriture	scripteur
'write'	'writing'	'writer'
gonfler	gonflement	gonfleur
'inflate'	'inflating'	'inflater'

Sample triples

- Action nouns are good predictors of agent nouns, since they almost always use the same stem.

Differential opacity (continued)

	Verb	Action_N	Agent_N
Verb	—	1.115	0.709
Action_N	0.101	—	0.269
Agent_N	0.264	1.114	—

Unary implicative entropy
for (Verb, Action_N, Agent_N) triples

Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
écrire	écriture	scripteur
'write'	'writing'	'writer'
gonfler	gonflement	gonfleur
'inflate'	'inflating'	'inflater'

Sample triples

- On the other hand, verbs are not so good predictors of agent nouns, because, even in the absence of suppletion, one has to guess whether a learned stem (typically in *-at*) should be used.

Results, 2: Joint predictiveness

- Predicting from two members of a morphological family is a lot easier than predicting from just one.

1 predictor	0.595
2 predictors	0.196

Average implicative entropy

Joint predictiveness (continued)

- Predicting from two members of a morphological family is a lot easier than predicting from just one.
- In particular, predicting the form of verbs from knowledge of the two nouns is trivial.

Predictors	Predicted	Entropy
Verb, Agent_N	Action_N	0.444
Verb, Action_N	Agent_N	0.138
Agent_N, Action_N	Verb	0.006

  - All the remaining uncertainty is caused by a handful of *-ionner* verbs (Lignon and Namer, 2010).

(Action_N , Agent_N) ⇒ Verb
(percussion , percuteur) ⇒ percuter (inspection , inspecteur) ⇒ inspecter
(perquisition , perquisiteur) ⇒ perquisitionner (fonction , foncteur) ⇒ fonctionner

Sample triples

Taking stock

- We have established that joint knowledge of two members of a derivational family is much more predictive of the rest of the family than knowledge of just one form.
- Thus predictability in derivation cannot just be a matter of a relation between a derivative and its base.
- Note that this is congruent with literature on the role of morphological family size in morphological processing (e.g. Schreuder and Baayen, 1997)
- While it gives a good overall picture, implicative entropy has inherent limitations when addressing the fine structure of predictability
  - Focus on one type of predictive variable (phonological shape).
  - Cannot deal with gaps or doublets.
- For a finer view of predictability, we turn to statistical modelling.

Predictability of form

II. Diving into rivalry

Joint work with Juliette Thuilier (Université Toulouse Jean Jaurès)

The issue

- We focus on rivalry between *-iser* and *-ifier* suffixation in French.
- Both suffixes form verbs from nouns or adjectives; it is often undecidable which of the two should be considered the base.

	Noun	Adjective	Derived verb
(i)	—	*banal*	*banaliser*
		‘trivial’	‘trivialize’
(ii)	*aval*	—	*avaliser*
	‘approval’		‘approve’
(iii)	*république*	*républicain*	*républicaniser*
	‘republic’	‘republican’	‘make republican’
(iv)	*Staline*	*stalinien*	*staliniser*
	‘Stalin’	‘stalinist’	‘make stalinist’
(v)	*morale*	*moral*	*moraliser*
	‘morality’	‘moral’	‘make ethical’
(vi)	*hôpital*	*hospitalier*	*hospitaliser*
	‘hospital’	‘of hospital’	‘hospitalize’

- We will establish that the overall makeup of the morphological family has predictive value as to which suffix is used.

The dataset

- Starting point: 1263 verbs with infinitives in *-ifier* or *-iser* that are:
  1. documented in French Wiktionary (Hathout et al., 2014), and
  2. attested in Google Ngrams (Michel et al., 2010).
- Manual filtering of underived verbs (e.g. *miser* 'bet') prefixed verbs (e.g. *décoloniser*), borrowings (e.g. *randomiser*), and verbs based on suppletive stems (e.g. *pacifier*), leading to a set of 791 lexemes.
- Annotation for age of the lexeme and stem phonology deduced from the resources above.
- Hand annotation for the structure of the morphological family by the authors.

The variables

- We used the following predictor variables, building on Lignon (2013) for inspiration on what to look at:
  1. Date of first attestation in Google Books
  2. Length of the derivational stem
  3. Last Consonant of the derivational stem
  4. Complexity of the final consonant Cluster of the stem
  5. Makeup of the Ascending Morphological Family (AMF) of the verb
     - adjective only?
     - noun only?
     - both?
  6. Morphological Class of the Adjective (MCA)
     - suffixed denominal?
     - conversion?
     - other?
  7. If the morphological family contains a denominal adjective, does it have Relational readings?
- Continuous variables normalized to a standard deviation of 1.

Main results

- We ran logistic regression models in R, and shrinked the models by backward stepwise selection.
- The first model takes into account all datapoints but leaves out variables that presuppose the existence of both a noun and an adjective in the family.

Coefficients	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	5.6290	0.8355	6.737	1.61e-11
DATE	-1.1360	0.4173	-2.722	0.00649
LENGTH	-7.0933	0.7080	-10.018	< 2e-16
CONSONANT==AlvObs	0.2324	0.5009	0.464	0.64267
CONSONANT==Son	-0.8537	0.4926	-1.733	0.08308
AMF==N	-0.7734	0.5400	-1.432	0.15211
AMF==both	-1.2199	0.4506	-2.707	0.00679

(Modelled: $P(Y = \text{-IFIER} \mid X)$; Intercept: CONSONANT==other, AMF==A)

- The model is highly predictive: AUC $= 0.918$.
- The co-presence of a related noun and adjective in the morphological family significantly shift the preference in favor *-iser*.

Main results (continued)

- The second model looks only at cases where the AMF contains both a noun and an adjective, but takes into account the formal and semantic relation between those.

Coefficients	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	5.5029	0.8535	6.447	1.14e-10
DATE	-0.8419	0.5905	-1.426	0.153923
LENGTH	-7.8332	0.9512	-8.235	< 2e-16
CONSONANT==son	-1.2054	0.4174	-2.888	0.003878
MCA==ique	-1.2240	0.5540	-2.210	0.027134
MCA==other_sfx	1.0299	0.4797	2.147	0.031810
RELATIONAL==True	-1.4755	0.4390	-3.361	0.000777

(Intercept: MCA==conversion, CONSONANT==other, RELATIONAL==False)

- The model is highly predictive: AUC = 0.943.
- Both the form (MCA) and the content (RELATIONAL) of the relation between noun and adjective play a significant role in explaining the choice of *-iser* or *ifier*.

Taking stock

- The present study improves our understanding of predictability of form in derivational paradigms:
  1. In this particular instance, how populated the paradigm is plays an important role in predicting which derivational suffix is preferred.
  2. In addition, the exact nature of the relation between paradigm cells, both in terms of form and in terms of content, plays a role.
- It is striking that such a result could be reached, despite a rather coarse-grained and superficial annotation of morphological families and lexical semantics.
  - More detailed studies are expected to uncover better predictors.
- We predicted form from form and form from content. What about predicting content?

Predictability of content

Joint work with Denis Paperno (CNRS - Loria)

The issue I

- Organizing derivational families into paradigmatic systems presupposes that derivational operations can be associated with stable semantic contrasts.
- This goes against a strong intuition that inflection and derivation differ in terms of predictability of content
  (cf. e.g. Robins 1959; Matthews 1974; Wurzel 1989; Stump 1998):
  - Pairs of cells in an inflectional paradigm contrast in the same way.

    table : *tables*
    mouse : *mice*
    committee : *committees*

  - But pairs of words entertaining the "same derivational relation" typically contrast in various ways, because of affix polysemy, lexicalization, and lexical meaning shift.

    barbe 'beard' : *barbier* 'barber'
    épice 'spice' : *épicier* 'grocer'
    pompe 'pump' : *pompier* 'firefighter'

The issue II

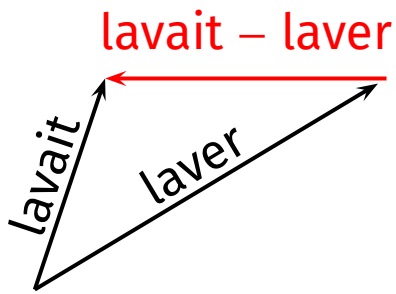
- While this is a commonly held view, it rests on semantic intuitions that are in need of explicit testing.
  - Some derivational relations seem semantically quite regular, e.g. that between place names and demonyms.
  - Some amount of variability will be found in inflection too
    - Because of systematic lexical semantics: e.g. the shift from singular to plural is not the same for count and mass nouns.
    - Because of lexical accidents: e.g. semi-*pluralia tantum*

      menotte 'small hand' : *menottes* 'small hand/handcuffs'
      ciseau 'chisel' : *ciseaux* 'chisel/scissors'
      vacance 'vacancy' : *vacances* 'vacancies/holidays'

- In this final part, we explore means of assessing empirically whether inflection and word formation differ in this respect.

Semantic contrasts as shift vectors I

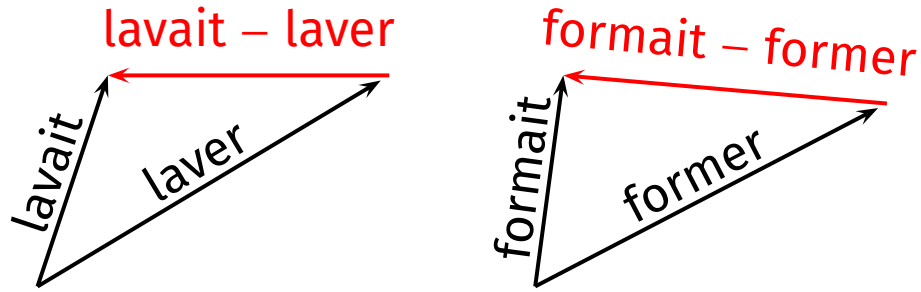
- We rely on distributional semantics: the meaning of a word is approximated by a high-dimensional vector representing its distribution in a corpus.
- Within such a framework, we can examine how vectors representing derivationally-related words relate to each other (Marelli and Baroni, 2015).
- Simple way of doing this: the contrast in meaning between two words is the difference between their two vectors; i.e., the vector representing what it takes to go from the meaning of one word to the meaning of the other.



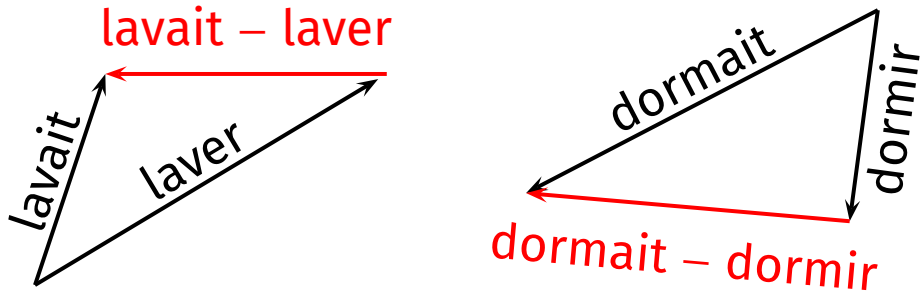
- We will call this vector the shift vector.

Semantic contrasts as shift vectors II

- Word vectors corresponding to the same paradigm cell will be similar in some dimensions and different in others.



- The word vectors may be very different but the shift vectors still be very similar.

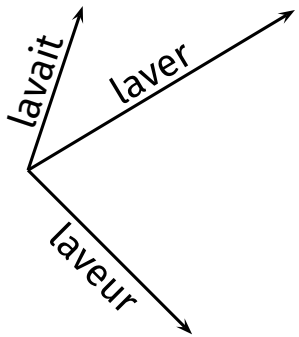


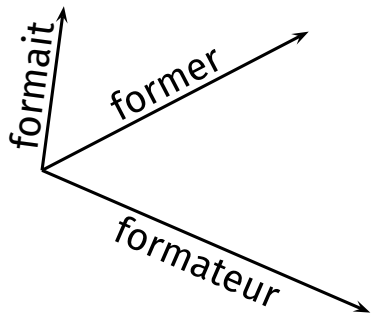
- Stability of semantic contrasts amounts to similarity of shift vectors.

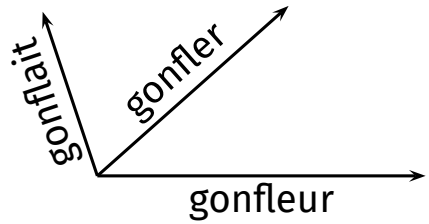
NB: We are not examining distance between word meanings but distance between shifts in meaning (compare Wauquier 2016) .

The hypothesis

- We look at triplets of morphologically-related forms, one of which is used as the pivot for comparison.

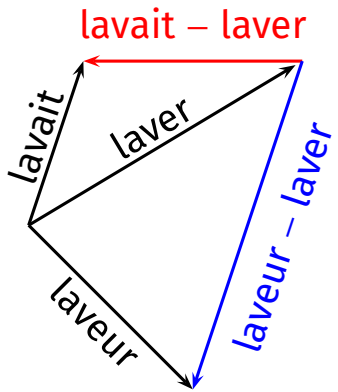


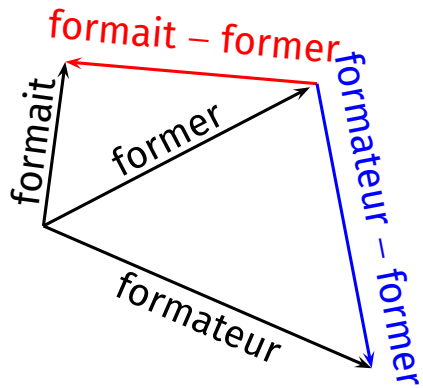


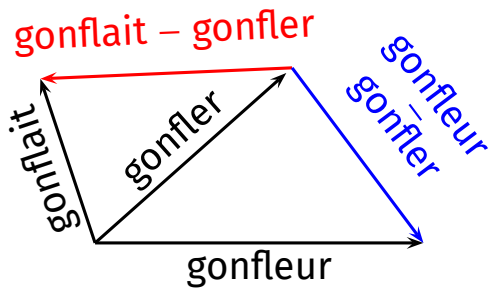


The hypothesis

- We look at triplets of morphologically-related forms, one of which is used as the pivot for comparison.
- We compute shift vectors between the pivot and the other forms.

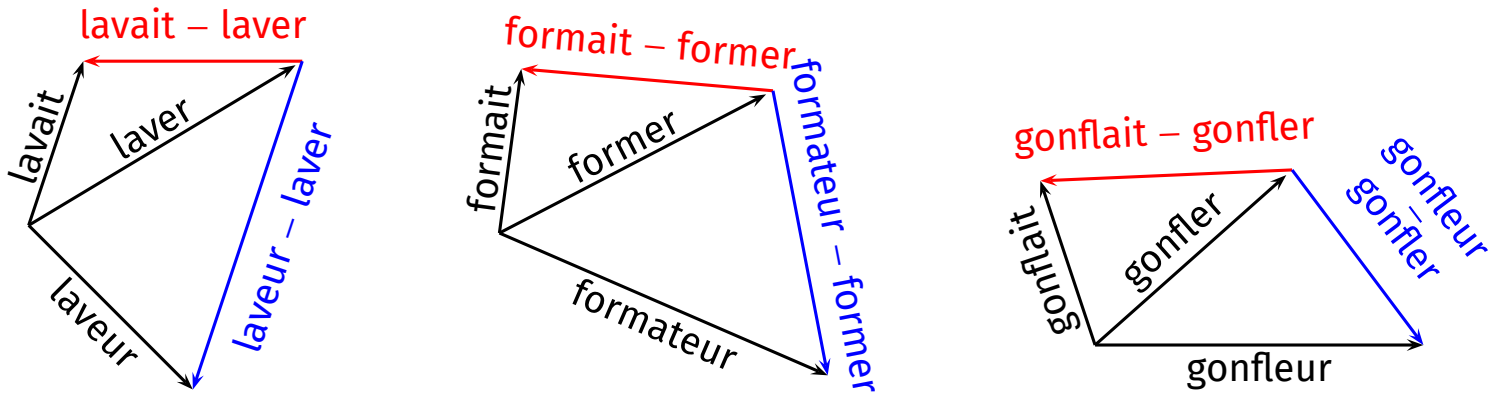




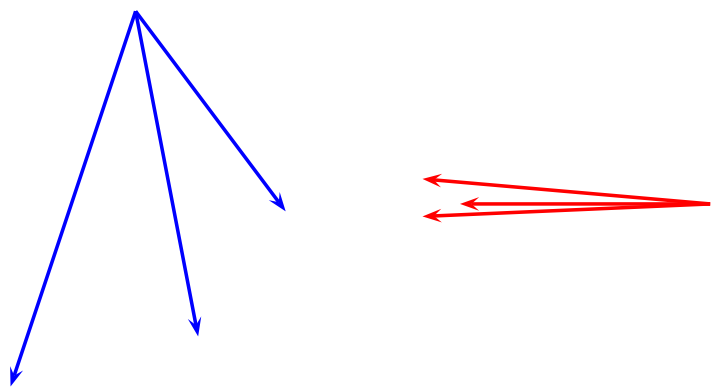


The hypothesis

- We look at triplets of morphologically-related forms, one of which is used as the pivot for comparison.
- We compute shift vectors between the pivot and the other forms.



- We then expect the shift vectors for derivationally-related pairs to be more diverse than those for inflectionally-related pairs.



The execution, I

- Vector space constructed from the FrWac corpus (Baroni et al., 2009) using `word2vec` (Mikolov et al., 2013).
  - CBOW algorithm, window size 5, negative sampling with 10 samples, 400 dimensions
- Paradigmatic system of 6576 (partial) families and 59 cells constructed from:
  1. Derivational relations between verbs, action nouns and agent nouns from Démonette (Hathout and Namer, 2014)
  2. Hand-constructed set of derivational relations between verbs and *-able* adjectives
  3. Inflectional relations from the GLÀFF (Hathout et al., 2014)
- We then look for triplets of cells where:
  1. There is a derivational relation between the first (pivot) and second cell and an inflectional relation between the first and third.
  2. We have enough data to select 100 triplets of words such that
     - 2.1 there is a single word in each cell,
     - 2.2 no word has homonyms,
     - 2.3 all words have a frequency above 50,
     - 2.4 the frequency ratio between the nonpivot cells is between $\frac{1}{5}$ and 5,
     - 2.5 the median frequency ratio is 1 or very close to 1.

The execution, II

- We found 174 partial paradigmatic systems verifying these requirements.
- Note that two different systems may provide evidence on the same derivational relation:

pivot	comparison 1	comparison 2	ratio
changer	*changeur*	*changeait*	0.356
prolonger	*prolongateur*	*prolongeait*	0.380
entendre	*entendeur*	*entendait*	0.389
...	...	...	...

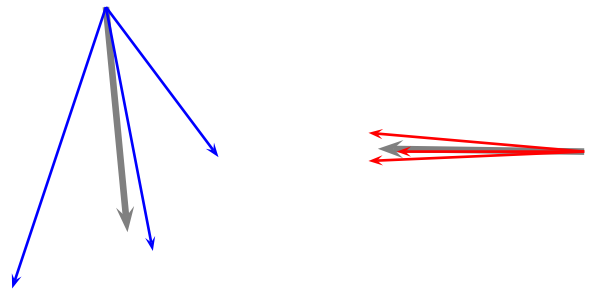
Sample system 1: (V.INF, Agent_N.SG, V.IPFV.3SG)

pivot	comparison 1	comparison 2	ratio
possesseurs	*possesseur*	*possédez*	0.236
finisseurs	*finisseur*	*finissez*	0.244
dégustateurs	*dégustateur*	*dégustez*	0.229
...	...	...	...

Sample system 2: (Agent_N.PL, Agent_N.SG, V.PRS.2PL)

The execution, II

- For each of the 174 systems:
  - We compute the two **shift vector averages**.



  - We compute the Euclidian distance between each individual vector and the average vector.
  - We perform a *t*-test to assess whether there is a significant difference in distance to the average between the shift vectors for the two compared cells.

Results

- Main result:
  - In all 174 situations, there is higher dispersion around the average for shift vectors between derivationally-related words than for shift vectors between inflectionally-related word.
  - This difference is statistically highly significant ($p < 0.001$) in all but 2 cases.
  - Interestingly, these 2 cases straddle the inflection-derivation divide (infinitive-participle-action noun)
- We thus have strong distributional evidence that derivational relations are less stable semantically than inflectional relations.

Discussion

- Our results do not entail that there is a categorical distinction between inflection and derivation in terms of predictability of content.
- In fact, if we do the same exercise with inflectionally-related forms, we find some interesting contrasts, e.g.
  - When using a finite form as a pivot, the shift vectors relating it to the infinitive are significantly more diverse than those relating it to another finite form.
- This suggests that there is a gradient of morphological contrasts in terms of semantic predictability, with derivational contrasts clustering towards the low predictability end and inflectional contrasts clustering towards the high predictability end.
- Also suggests new research questions:
  - Can we rank morphosyntactic features in terms of semantic predictability? Does the ranking vary across languages?
  - Do distinctions of dubious status on the inflection-derivation divide (finiteness, voice, etc.) fall in the middle in terms of semantic predictability?
  - Some contrasts are often said to be derivational in some languages and inflectional in others (aspect, diminutives). Does this have measurable effects in terms of semantic predictability?

Conclusions

Conclusions

- Using various quantitative methods, we established properties of derivational families that require a paradigmatic formulation.
  1. In derivation as in inflection, bidirectional predictability relations of variable reliability can be documented. (section 2)
  2. In derivation as in inflection, prediction from multiple words is vastly easier than prediction from single words. (section 2)
  3. In derivation, the degree of saturation of paradigms is predictive of affix choices. (section 3)
  4. In derivation, the formal and semantic relations between members of a paradigm are predictive of affix choice. (section 3)
- This strongly suggests that a (partial, content-based, dynamic, opportunistic) notion of derivational paradigm is a necessary component of the study of word formation.
- However differences between inflection and derivation should not be neglected.
  - Semantic contrasts between pairs of words are less stable if these are related by derivation than if they are related by inflection. (section 4)

References I

Ackerman, F., Blevins, J. P., and Malouf, R. (2009). 'Parts and wholes: implicative patterns in inflectional paradigms'. In J. P. Blevins and J. Blevins (eds.), *Analogy in Grammar*. Oxford: Oxford University Press, 54–82.

Ackerman, F. and Malouf, R. (2013). 'Morphological organization: the low conditional entropy conjecture'. *Language*, 89:429–464.

Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2009). 'The wacky wide web: A collection of very large linguistically processed web-crawled corpora'. In *Language Resources and Evaluation*, vol. 43. 209–226.

Bauer, L. (1997). 'Derivational paradigms'. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 243–256.

Becker, T. (1993). 'Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology'. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 1–25.

Blevins, J. P. (2001). 'Paradigmatic derivation'. *Transactions of the Philological Society*, 99:211–222.

Bonami, O. and Beniamine, S. (2016). 'Joint predictiveness in inflectional paradigms'. *Word Structure*, 9:156–182.

Bonami, O. and Boyé, G. (2014). 'De formes en thèmes'. In F. Villoing, S. Leroy, and S. David (eds.), *Foisonnements morphologiques. Etudes en hommage à Françoise Kerleroux*. Presses Universitaires de Paris Ouest, 17–45.

Bonami, O. and Luís, A. R. (2014). 'Sur la morphologie implicative dans la conjugaison du portugais : une étude quantitative'. In J.-L. Léonard (ed.), *Morphologie flexionnelle et dialectologie romane. Typologie(s) et modélisation(s).*, no. 22 in Mémoires de la Société de Linguistique de Paris. Leuven: Peeters, 111–151.

Bonami, O. and Strnadová, J. (2016). 'Derivational paradigms: pushing the analogy'. In *49th Annual Meeting of the Societas Linguistica Europaea*. Naples.

Bonami, O. and Stump, G. T. (2016). 'Paradigm Function Morphology'. In A. Hippisley and G. T. Stump (eds.), *Cambridge Handbook of Morphology*. Cambridge: Cambridge University Press, 449–481.

Bonami, O. and Thuilier, J. (submitted). 'A statistical approach to affix rivalry: French *-iser* and *-ifier*'. May, 2017.

Booij, G. (1997). 'Autonomous morphology and paradigmatic relations'. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 35–53.

Booij, G. E. (2007). *The Grammar of Words*. Oxford University Press, 2nd edn.

Boyé, G. and Schalchli, G. (2016). 'The status of paradigms'. In A. Hippisley and G. Stump (eds.), *The Cambridge Handbook of Morphology*. Cambridge University Press, 206–234.

References II

Hathout, N. and Namer, F. (2014). 'Démonette, a French derivational morpho-semantic network'. *Linguistic Issues in Language Technology*, 11:125–168.

Hathout, N., Sajous, F., and Calderone, B. (2014). 'GLÀFF, a large versatile French lexicon'. In *Proceedings of LREC 2014*.

Hay, J. B. and Baayen, R. H. (2005). 'Shifting paradigms: gradient structure in morphology'. *TRENDS in Cognitive Science*, 9:342–348.

Lignon, S. (2013). '*-ISER* and *-IFIER* suffixation in French: Verifying data to 'verize' hypotheses'. In N. Hathout, F. Montermini, and J. Tseng (eds.), *Morphology in Toulouse, Selected prodeedings of Décembrettes 7*. Munich: Lincom Europa, 119–132.

Lignon, S. and Namer, F. (2010). 'Comment conversionner les v-ion ? ou la construction de v-ionnerverbe par conversion'. In *Actes du 2eme Congrès Mondial de Linguistique Française*. 1009–1028.

Marelli, M. and Baroni, M. (2015). 'Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics'. *Psychological Review*, 122:485–515.

Matthews, P. H. (1974). *Morphology*. Cambridge: Cambridge University Press.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., andSteven Pinker, J. O., Nowak, M. A., and Aiden, E. L. (2010). 'Quantitative analysis of culture using millions of digitized books'. *Science*, 14:176–182.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). 'Efficient estimation of word representations in vector space'. *CoRR*, abs/1301.3781.

Robins, R. H. (1959). 'In defense of WP'. *Transactions of the Philological Society*:116–144.

Roché, M., Boyé, G., Hathout, N., Lignon, S., and Plénat, M. (eds.) (2011). *Des unités morphologiques au lexique*. Hermès Lavoisier.

Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.

Schreuder, R. and Baayen, R. H. (1997). 'How complex simple words can be'. *Journal of Memory and Language*, 37:118–139.

Sims, A. (2015). *Inflectional defectiveness*. Cambridge: Cambridge University Press.

Stump, G. T. (1998). 'Inflection'. In A. Spencer and A. Zwicky (eds.), *The Handbook of Morphology*. London: Blackwell, 13–43.

——— (2016). *Inflectional paradigms*. Cambridge: Cambridge University Press.

Stump, G. T. and Finkel, R. (2013). *Morphological Typology: From Word to Paradigm*. Cambridge: Cambridge University Press.

van Marle, J. (1984). *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris.

References III

Štekauer, P. (2014). 'Derivational paradigms'. In R. Lieber and P. Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*. Oxford: Oxford University Press, 354–369.

Wauquier, M. (2016). *Indices distributionnels pour la comparaison sémantique de dérivés morphologiques*. Master's thesis, Université Toulouse Jean Jaurès.

Wunderlich, D. and Fabri, R. (1995). 'Minimalist Morphology: an approach to inflection'. *Zeitschrift für Sprachwissenschaft*, 14:236–294.

Wurzel, W. U. (1989). *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer.