International Symposium of Morphology - 2021

What French eventive nominalizations without verbal bases tell us about the salience of paradigmatic networks

Alice Missud[1,2], Florence Villoing[1]

[1]MoDyCo, UMR 7114 (CNRS & Université Paris Nanterre)
[2]LaTTiCe, UMR 8094 (CNRS, Paris Sorbonne Nouvelle & ENS)

1. Introduction: Data (1)

- French *-ion*, *-age* and *-ment* nominalizations can derive from constructed verbs (1)

(1) a. *-iser: créole* 'Creole' → *créoliser* 'to creolize' → *créolisation* 'creolization'
b. *-ifier*: *plan* 'plan' → *planifier* 'to plan' → *planification* 'planning'
c. *a-*: *jour* 'gap/chink' → *ajourer* 'to perforate' → *ajouration* 'perforation'
d. *en-*: *cadre* 'frame' → *encadrer* 'to frame' → *encadrement* 'framing'
e. *é-*: *goutte* 'drop' → *égoutter* 'to drain' → *égouttage* 'draining'
f. *dé-*: *amiante* 'asbestos' → *désamianter* 'to remove asbestos' → *désamiantation* 'asbestos removal'
g. conversion : *frein* 'brake' → *freiner* 'to brake' → *freinage* 'braking'

1. Introduction: Data (1)

- French *-ion*, *-age* and *-ment* nominalizations can derive from constructed verbs (1)

(1) a. *-iser: créole* 'Creole' → *créoliser* 'to creolize' → *créolisation* 'creolization'
b. *-ifier*: *plan* 'plan' → *planifier* 'to plan' → *planification* 'planning'
c. *a-*: *jour* 'gap/chink' → *ajourer* 'to perforate' → *ajouration* 'perforation'
d. *en-*: *cadre* 'frame' → *encadrer* 'to frame' → *encadrement* 'framing'
e. *é-*: *goutte* 'drop' → *égoutter* 'to drain' → *égouttage* 'draining'
f. *dé-*: *amiante* 'asbestos' → *désamianter* 'to remove asbestos' → *désamiantation* 'asbestos removal'
g. conversion : *frein* 'brake' → *freiner* 'to brake' → *freinage* 'braking'

- They constitue **morphological families** of triplets (Hathout 2009, Lignon et al. 2014, Bonami & Strnadová 2019, Namer & Hathout 2020)

(2) $X_{N/Adj}$ → X_V derived → X_V *-ion/-age/-ment* $_N$

  - Represent 26.4 % of French *-ion/-age/-ment* nominalizations in VerNom (Missud et al. 2020)

1. Introduction : Data (2)

However, we found cases where the verb is missing :

(3) a. ***-ion*** : *macbeth*$_N$ 'Macbeth' → *macbethisation*$_N$ 'macbethization' (°*macbethiser*$_V$ 'to macbethize'),
translucide$_{Adj}$ 'translucent'→ *translucidation*$_N$ 'translucentation' (*translucider*$_V$ 'to make translucent'),
disneyland$_N$ 'Disneyland'→ *disneylandification*$_N$ 'disneylandification' (*disneylandifier*$_V$ 'to disneylandify')

b. ***-age*** : *bandelette*$_N$ 'strip'→ *bandelettage*$_N$ 'strip application' (*bandeletter*$_V$ 'to apply strips'),
bestof$_N$ 'best-of'→ *bestofage*$_N$ 'making a best-of' (°*bestofer*$_V$ 'to make a best-of'),
dofus$_N$ 'Dofus' → *dofusage*$_N$ 'playing Dofus' (°*dofuser*$_V$ 'to play Dofus')

c. ***-ment*** : *stupide*$_{Adj}$ 'stupid' → *enstupidement*$_N$ 'making stupid' (°*enstupider*$_V$ 'to make stupid'),
bruyère$_N$ 'heather' → *embruyèrement*$_N$ 'proliferation of heather' (*embruyérer*$_V$ 'to proliferate heather')

1. Introduction: Hypothesis

- The intermediate derivation (the verb) is possible but not lexicalized or attested in corpora (or with a lower frequency than its nominal or adjectival base).

- These nominalizations are formed directly on the base noun or adjective.

(4) $X_{N/Adj}$ ($\rightarrow X_V$ derived) $\rightarrow X_V$ *-ion/-age/-ment* $_N$

- The missing verb (i.e. potential word) is perfectly identifiable and can be easily interpreted.

1. Introduction: Previous work

- Nominalizations on nominal/adjectival bases have already been observed,
  - ➢ In English
    - ♦ *-ation* nominalizations on nominal/adjectival bases (Marchand (1969: 260), Plag (2003: 226))

nominal base	derived verb	action/process/result noun
sediment	not attested	*sedimentation*
epoxide	not attested	*epoxidation*

    - ♦ *-ment* nominalizations on nominal/adjectival bases (Bauer, Lieber, Plag 2013: 198)
      - *-ment* on noms: *illusionment*
      - *-ment* on adjectives : *insensiblement, oddment, surement*
  - ➢ In French
    - ♦ *-isation* on nominal bases: (Dal 2004, Roché 2009, Namer 2009, Lignon et al. 2014, Dal & Namer 2015, Missud & Villoing 2020)

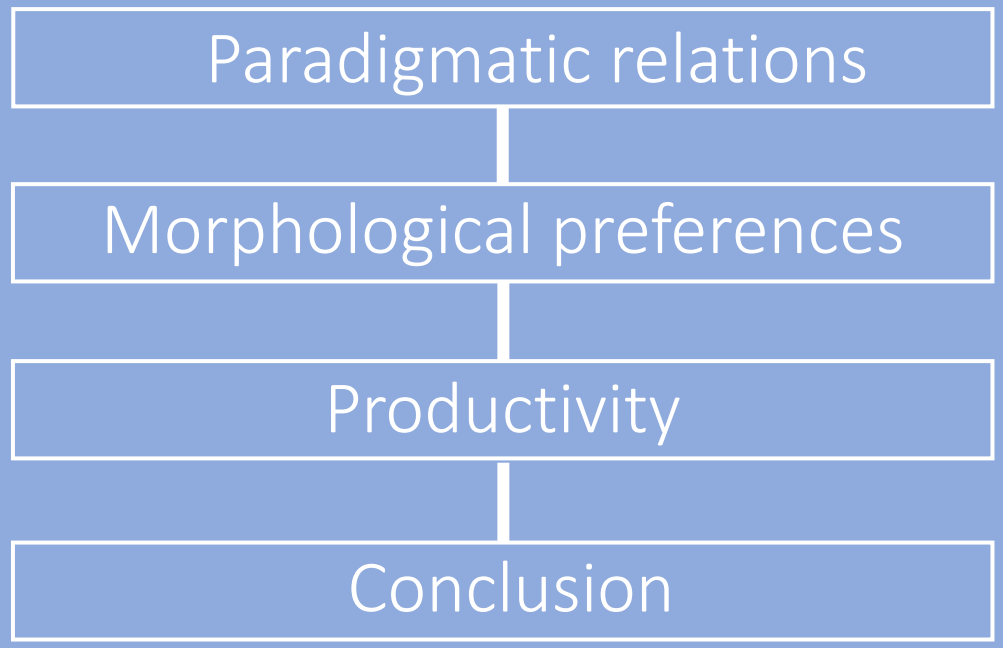
1. Introduction: aims of the talk

- Show that French nominalizations on nominal/adjectival bases do not only concern *–ion* suffixation but also *–age* and *–ment* suffixation;

- Identify the conditions that allow these deadjectival/ denominal nominalizations:
  - Are all potential verbs concerned ?
  - Is there a preference of certain suffixes for certain potential verbs

- Conclude that
  - They are part of Morphological Families structured in Morphological Paradigms
  - The Morphological Paradigms with missing intermediate verbs are quantitatively represented (occurirng more frequently than expected)

ITINERARY


Paradigmatic relations
Morphological preferences
Productivity
Conclusion

I T I N E R A R Y

Paradigmatic relations

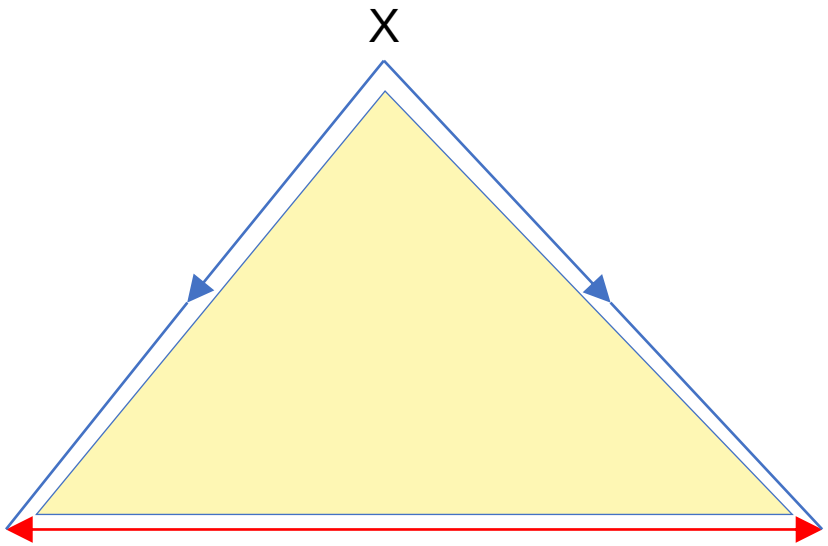
Morphological preferences

Productivity

Conclusion

2.1. Morphological Families structured as Morphological Paradigms

- X*isation* nominalization = triangular structures relating to a paradigmatic relationship (Lignon et al 2014) , of the same type as those mentioned by Roché 2008, Strnadová 2014, for French or Booij 2010 for Dutch.
- Paradigmatic relations: Xisation linked to X without the derivative verb.



Fr: *Italie*$_N$, *italien*$_N$, *italien*$_A$
Fr: *Balkan*$_N$, *balkaniser*$_N$, *balkanisation*$_N$
Nl: *Provence*$_N$, *provençaal*$_N$, *provençaals*$_A$
Nl: *arbeid*$_V$, *arbeider*$_{Nmas}$, *arbeidster*$_{Nfem}$

Fr: Xien N
: **Xiser V**
Nl: Xaal N
Xer N

Fr: Xien A
: **Xisation N**
Nl: Xaals A
Xster N

(Adapted from Lignon et al. 2014)

2.1. Morphological Families structured in Morphological Paradigms

- This paradigmatic organization concerns other French nominalizations taken from families of triplets.
  - $X_{N/Adj}$ - X_V derived - X_V *-ion/-age/-ment* $_N$

- The pairs of each morphological family in (1) constitue a **paradigmatic system** in the sense of Bonami & Strnadová (2019):

  ☞ pairs are aligned because they contrast in content in the same way (« fine-grained » or «coarse-grained» classification)

2.1. Morphological Families structured in Morphological Paradigms

- This paradigmatic organization concerns other French nominalizations taken from families of triplets.
  - $X_{N/Adj}$ - X_V derived - X_V *-ion/-age/-ment* $_N$

- The pairs of each morphological family in (1) constitue a **paradigmatic system** in the sense of Bonami & Strnadová (2019):
  ☞ pairs are aligned because they contrast in content in the same way (« fine-grained » or «coarse-grained» classification)

  - The $X_{N/Adj}$ – [X_V derived - X_V *-ion/-age/-ment*$_N$] pairs are semantically aligned through the **Action relation,**
  - The [$X_{N/Adj}$ - X_V derived] - X_V *-ion/-age/-ment* $_N$ pairs are (roughly) semantically aligned through the **Causative relation,**

2.1. Morphological Families structured in Morphological Paradigms

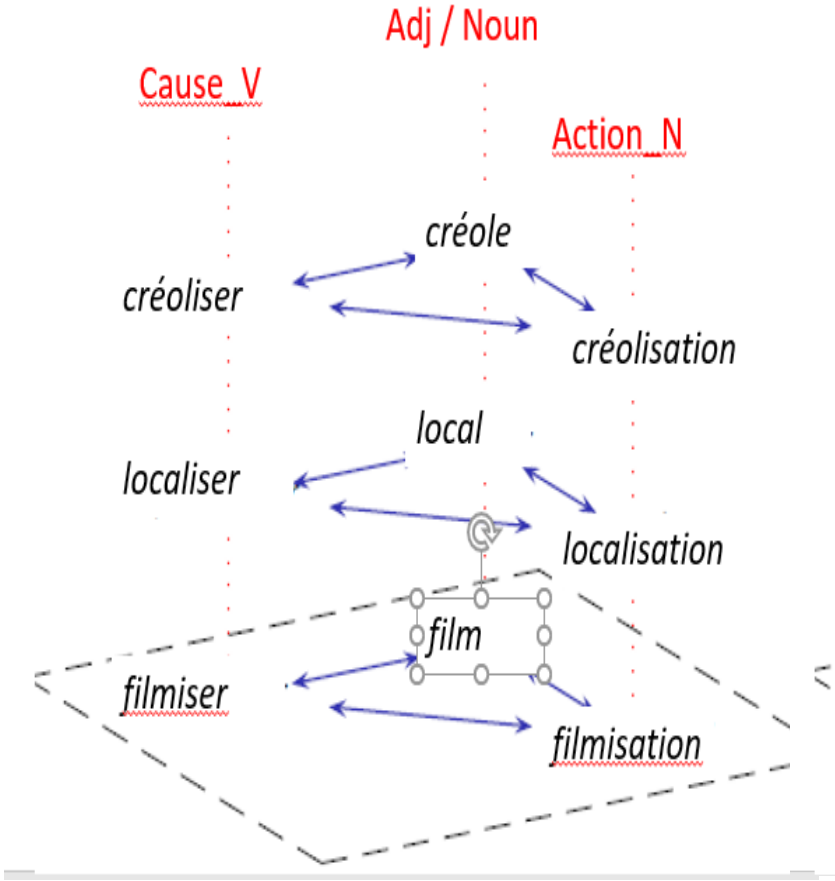


Fig. 1 : French morphological paradigm
$X_{N/Adj}$ - X_V -*iser* - X_V -*ion*$_N$

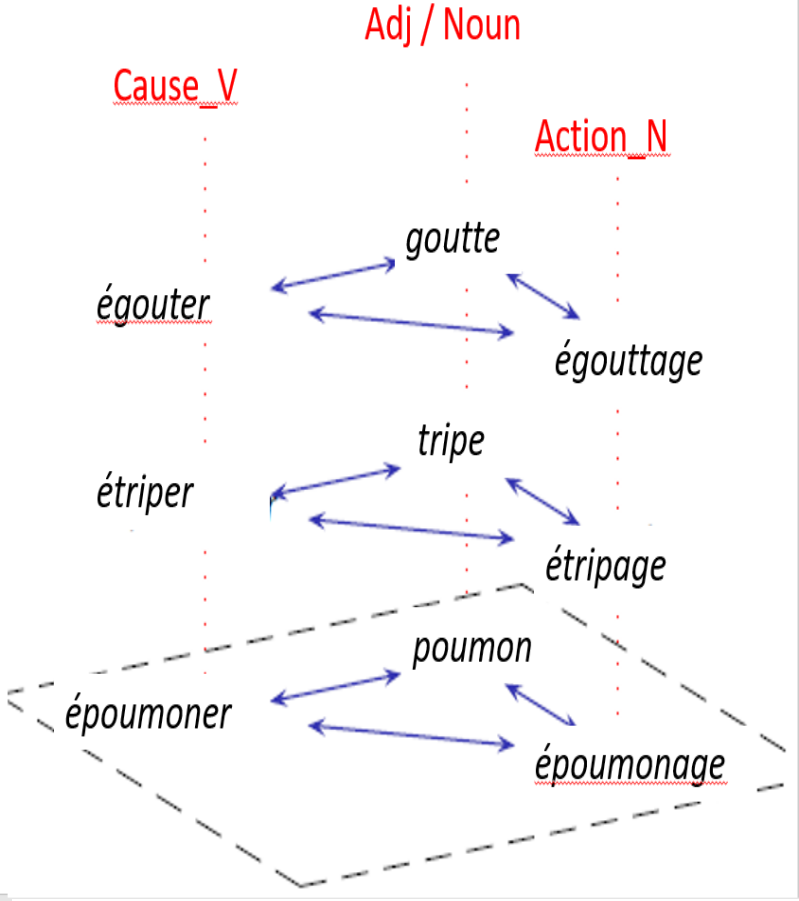


Fig.2: French morphological paradigm
$X_{N/Adj}$ - *é*-X_V - X_V -*age*$_N$

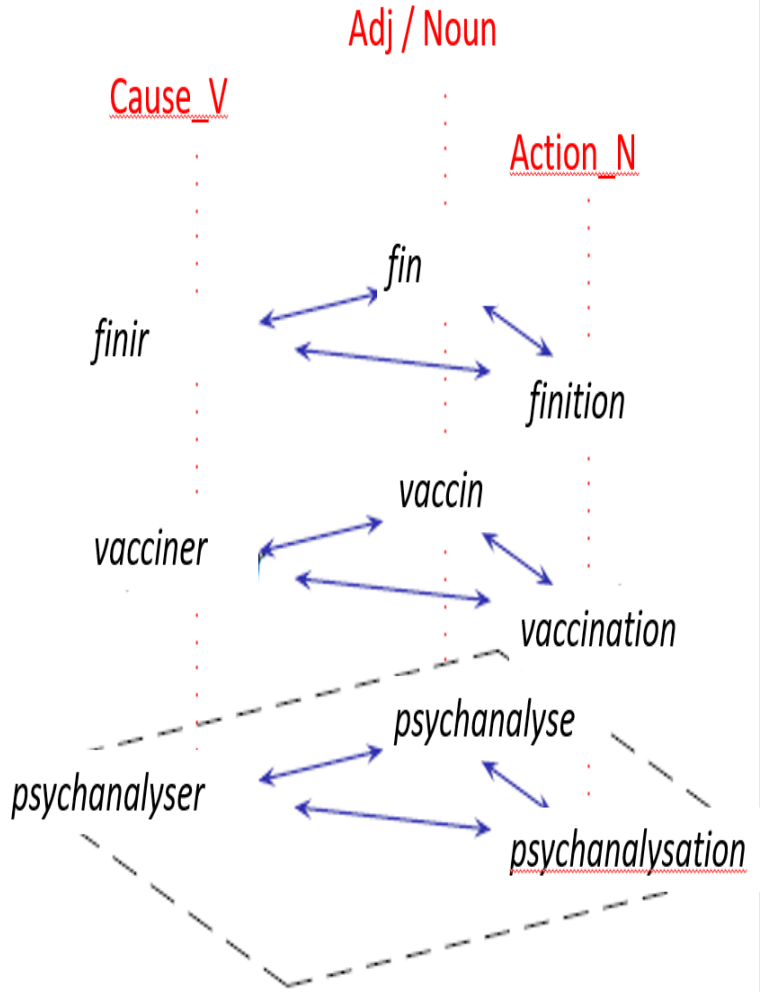


Fig. 3: French morphological paradigm
$X_{N/Adj}$ - X_V - X_V -*ion*$_N$

2.1. Morphological Families structured in Morphological Paradigms

- 7 morphological paradigms for 1 semantic paradigm («coarse-grained») :

- $X_{N/Adj}$ – X_{V} derived [causative] – X_{V} *-ion/-age/-ment* $_{N}$ [Action-N]

(1) a. **X-Xiser-Xisation**'s morphological paradigm *créolisation* 'creolization'
b. **X-Xifier-Xification**'s morphological paradigm *planification* 'planning'
c. **X-aX-aXation**'s morphological paradigm *ajouration* 'perforation'
d. **X-enX-enXment**'s morphological paradigm *encadrement* 'framing'
e. **X-éX-éXage**'s morphological paradigm *égouttage* 'draining'
f. **X-déX-déXation**'s morphological paradigm *désamiantation* 'asbestos removal'
g. **X-X-Xage**'s morphological paradigm *freinage* 'braking'

2.2. Results: belonging to a Morphological Paradigms

- The missing verb is possible only in the context of these morphological paradigms.

- $X_{N/Adj}$ – X_{V} [causative] – X_{V} *-ion/-age/-ment* $_{N}$ [Action-N]

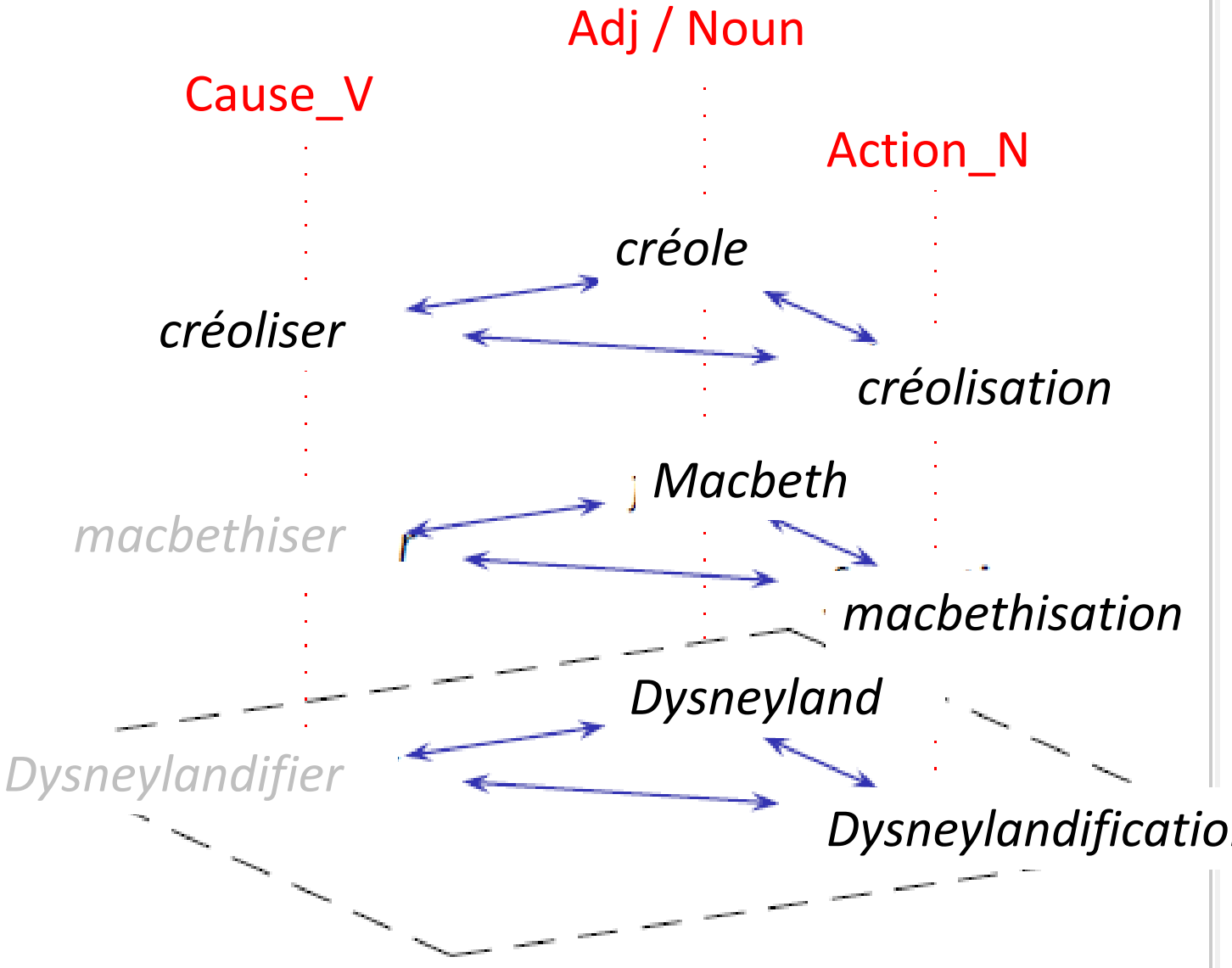
(2) a. X-Xiser-Xisation's morphological paradigm
☞ **X-Xiser-Xisation** *macbethisation*$_{N}$ 'macbethization'

b. X-Xifier-Xification's morphological paradigm
☞ **X-Xifier-Xification** *disneylandification*$_{N}$ 'disneylandification'

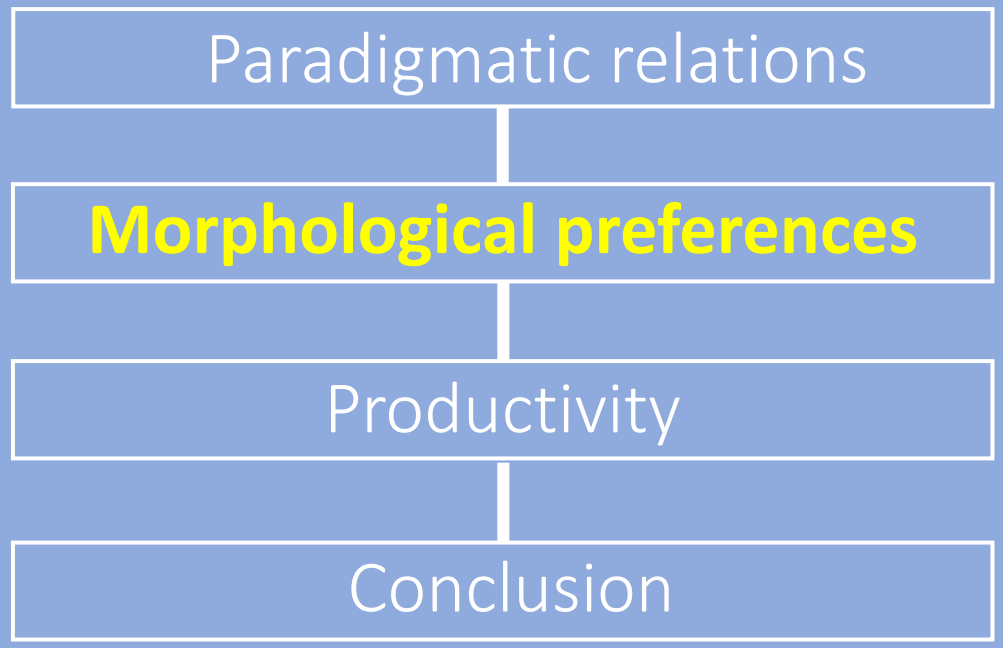
c. X-enX-enXment's morphological paradigm
☞ **X-enX-enXment** *enstupidement*$_{N}$ 'making stupid'

d. X-X-Xage's morphological paradigm
☞ **X-X-Xage** *bandelettage*$_{N}$ 'strip application'

2.2. Results: belonging to a Morphological Paradigms



I T I N E R A R Y


Paradigmatic relations
Morphological preferences
Productivity
Conclusion

3.1. Data collection

- Extracting N, Adj, Nion, Nage and Nment from frCOW (Schäfer & Bildhauer 2012; Schäfer 2015), 9 billion words
- Matching N / Adj → Nion / Nage / Nment with no verb
- Random sampling: 30% N/Adj → *-ion*, *-age* and *–ment* pairs
- Manual selection: first attestation of nouns and presumed verbs on Google Search (+ Twitter)

  → attested since 1950, <= 10 pages of results, looking at spelling variants / inflection

	-ion	*-age*	*-ment*	Total
Total number of N-N pairs	918	593	667	2 178
Total number of selected pairs	**186**	**80**	**8**	**274**

3.2. Morphological preferences when the verb is missing

Inferred schemas that construct verbs amongst derivatives without a verbal base:

	-iser	*-ifier*	*conversion*	*a-*	*en-*	*é-*	Total
-ion	**0.51**	**0.08**	0.08	-	-	-	0.67
-age	-	-	**0.27**	-	0.01	-	0.29
-ment	-	-	< 0.01	-	**0.02**	-	0.02
Total	0.51	0.08	0.36	-	0.04	-	1.0

3.3. Morphological preferences in the general case

Preferred denominal or deadjectival verbal bases:

	-iser	*-ifier*	*conversion*	*a-*	*en-*	*é-*	*dé-*	Total
-ion	**0.49**	**0.05**	0.04	< 0.01	< 0.01	< 0.01	0.01	0.62
-age	0.01	< 0.01	**0.11**	< 0.01	0.03	**0.01**	**0.04**	0.22
-ment	< 0.01	< 0.01	0.05	**0.01**	**0.05**	< 0.01	0.01	0.15
Total	0.51	0.07	0.2	0.03	0.09	0.03	0.06	1.0

3.3. Morphological preferences in the general case

Preferred denominal or deadjectival verbal bases (derived event nouns of < 10 occurrences) :

	-iser	*-ifier*	*conversion*	*a-*	*en-*	*é-*	*dé-*	Total
-ion	**0.5**	**0.05**	0.05	< 0.01	< 0.01	< 0.01	0.1	0.63
-age	0.01	< 0.01	**0.09**	< 0.01	0.03	**0.01**	**0.05**	0.22
-ment	< 0.01	0.05	< 0.01	**0.01**	**0.04**	< 0.01	0.01	0.14
Total	0.53	0.05	0.2	0.01	0.07	0.02	0.08	1.0

3.3. Morphological preferences in the general case

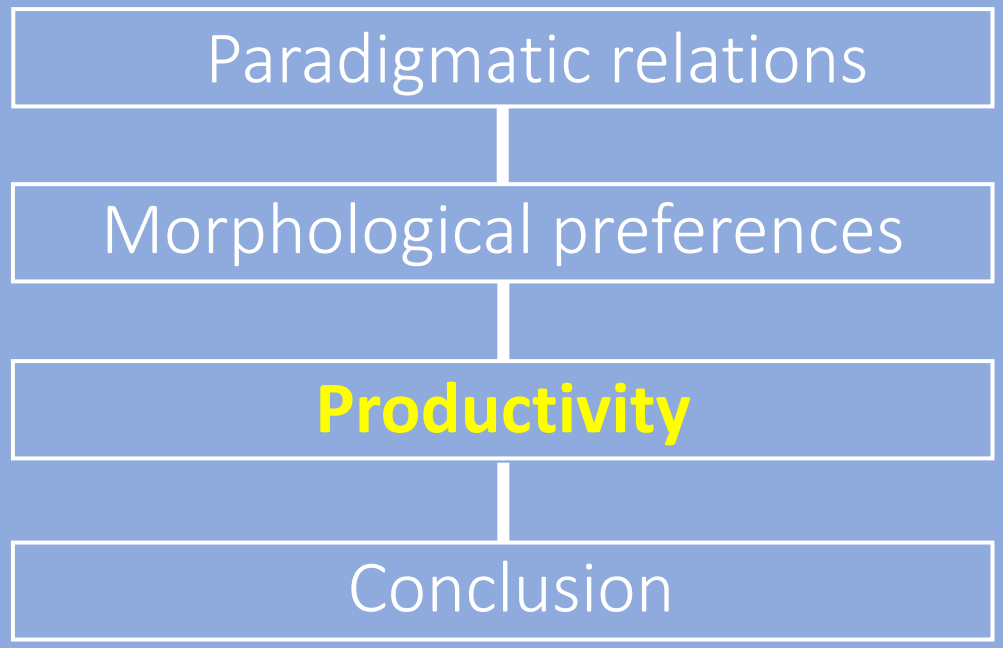
	-ion	***-age***	***-ment***
-iser	*local* 'local' - *localiser* 'to locate'- *localisation* 'localization'	*balise* 'marker' - *baliser* 'to mark [sth] out' - *balisage* 'cone [sth] off'	*voix* 'voice' - *voiser* 'to voice' - *voisement* 'voicing'
-ifier	*plan* 'plan' - *planifier* 'to plan' - *planification* 'planning'	*plastique* 'plastic' - *plastifier* 'to laminate' - *plastifiage* 'laminating'	*mort* 'death' - *mortifier* 'to mortify' - *mortifiement* 'mortification'
conversion	*fin* 'end' - *finir* 'to finish' - *finition* 'finishing stage'	*frein* 'brake' - *freiner* 'to brake' - *freinage* 'bracking'	*fleur* 'flower' - *fleurir* 'to flower' - *fleurissement* 'flowering'
en-	*capsule* 'cap' - *encapsuler* 'to encapsulate' - *encapsulation* 'encapsulation'	*bouteille* 'bottle' - *embouteiller* 'to bottle' - *embouteillage* 'bottling'	*cadre* 'a frame' - *encadrer* 'to frame' - *encadrement* 'framing'
é-	*ventre* 'stomach' - *éventrer* - 'to disembowel' - *éventration* 'disembowelment'	*tripe* 'guts' - *étriper* 'to gut'- *étripage* 'gutting'	*merveille* 'wonder' - *émerveiller* 'to amaze' - *émerveillement* 'wonder'
a-	*bête* 'stupid' - *abêtir* 'to make dumb' - *abêtisation* 'dumbing'	*plat* 'flat' - *aplatir* 'to flatten'- *aplatissage* 'flattening'	*ligne* 'line' - *aligner* 'to align' - *alignement* 'alignment'

- The degree of specialization in the general case varies depending on the schema:

-ion **>** *-age* **>** *-ment*

→ Does it have a link with the global productivity of these schemas?

ITINERARY


Paradigmatic relations
Morphological preferences
Productivity
Conclusion

4.1. Data to measure productivity

VerNom lexical database (Missud, Amsili & Villoing, 2020):

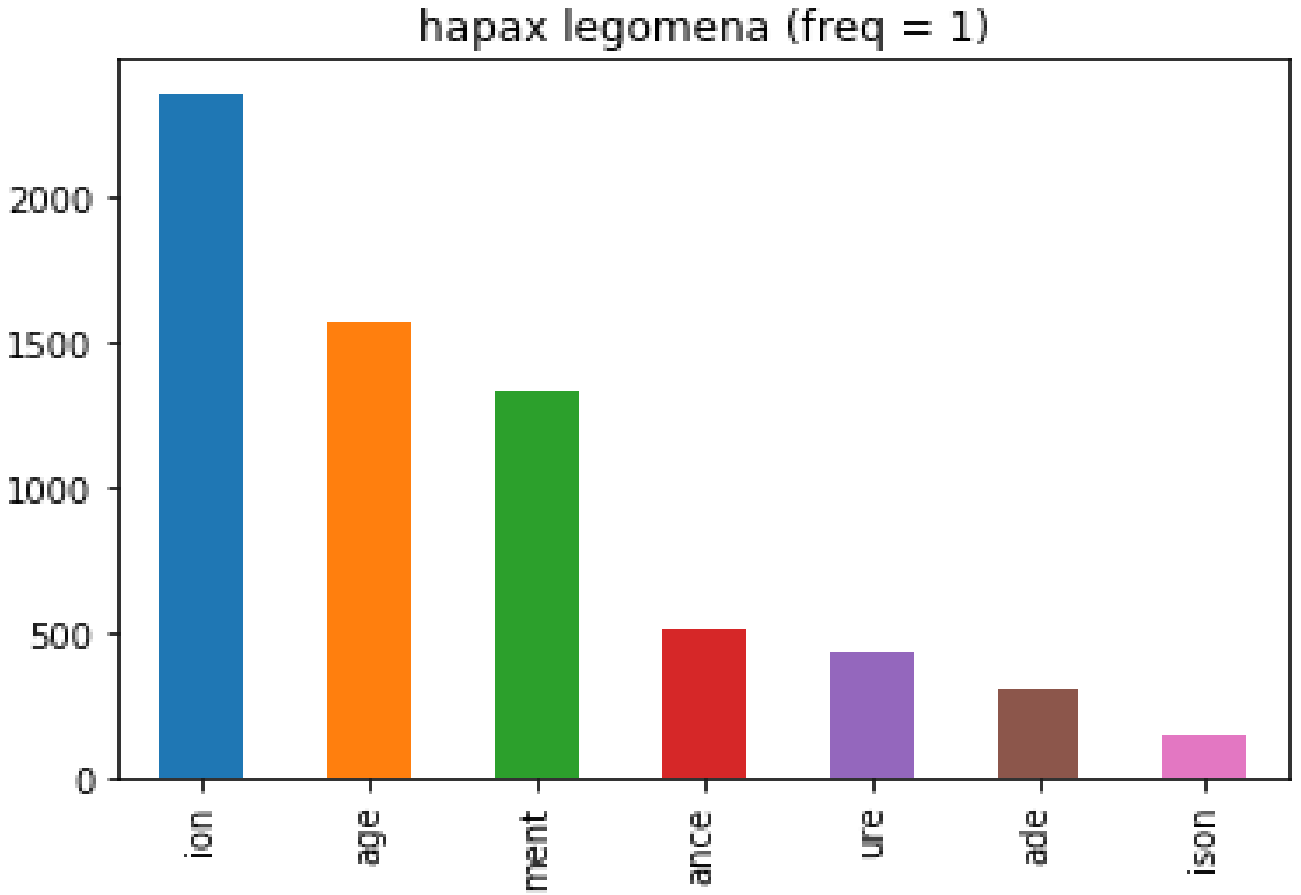
- *-ion*, *-age*, *-ment*, *-ance*, *-ure*, *-ade* and *–aison*
- Based on frCOW (Schäfer & Bildhauer 2012; Schäfer 2015), 9 billion words
- 25 857 verb-noun pairs with their frequencies (automatically)

→ https://www.ortolang.fr/market/lexicons/vernom

4.2. Productivity measure: Baayen

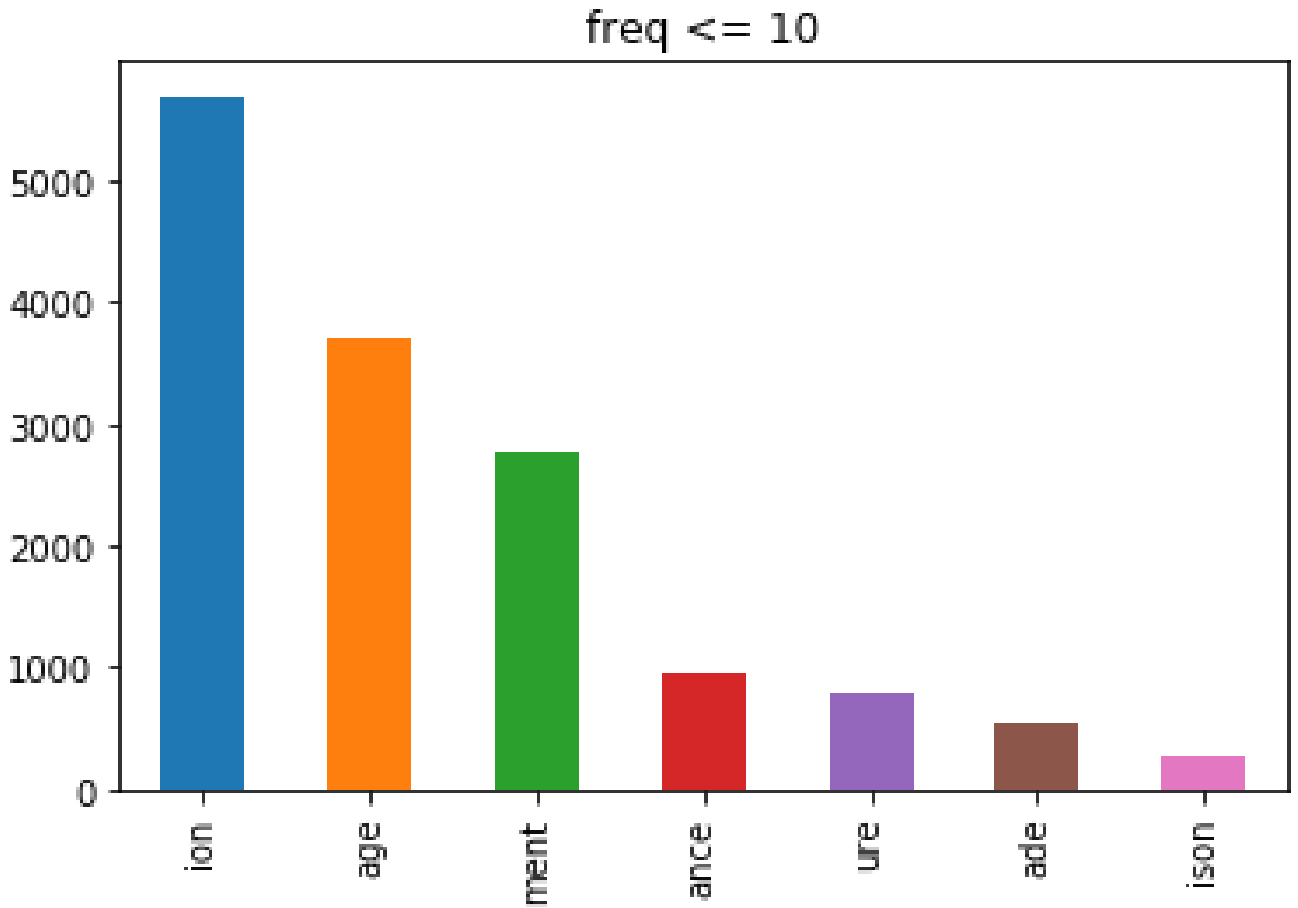
The hapax-conditioned degree of productivity (Baayen, 1994)

From VerNom (Missud, Amsili & Villoing 2020)

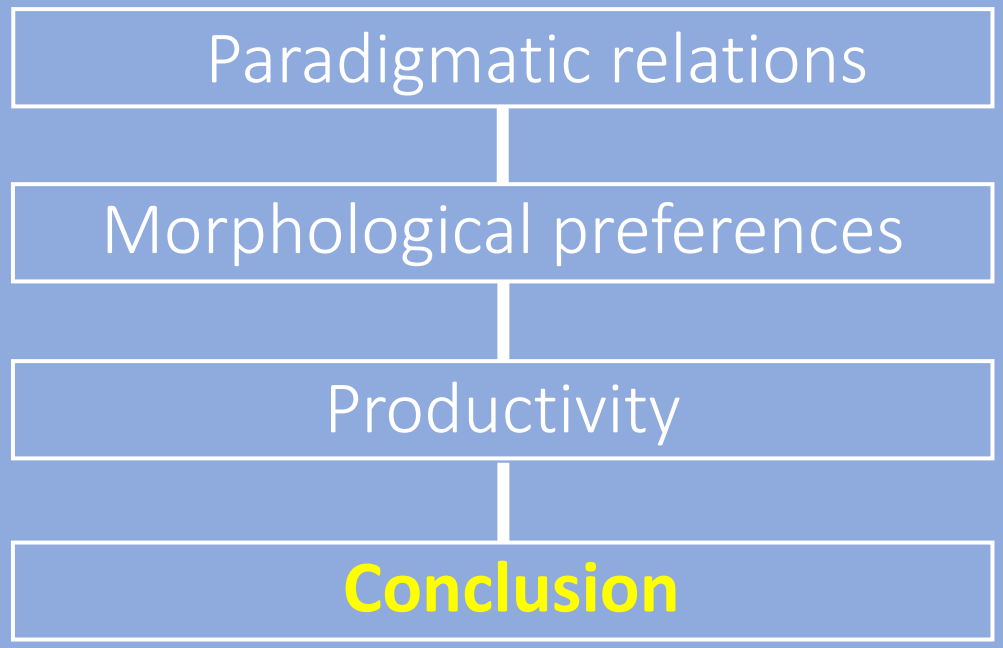


4.2. Productivity measure: Baayen

With different frequency thresholds (<10, <50, <100): **the rank remains identical**.



ITINERARY


Paradigmatic relations
Morphological preferences
Productivity
Conclusion

5.1. Summary

The preferences highlight highly frequent paradigms:

- *-ion*, *-age* and *-ment* display distinct preferences on the morphological level
- The schemas' preferences are reflected even when no verb is attested
- Different degrees of preference: *-ion* > *-age* > *-ment*
- Potential shortcuts in one morphological paradigm: schemas look pluritaleral
- Link between the salience of the schemas' preferences and their global productivity

5.2. Next step: focus on the semantics

X_N [toponym / anthroponym] – X_V *-iser* [causative] – X_V *-ion* $_N$ [Action-N]:

frankenstein – (°frankensteiniser) – frankensteinisation
Sarkhollande – (°sarkhollandiser) – sarkhollandisation
Kaboul – (°kabouliser) – kaboulisation

X_N [instrument / object] – X_V converted [instrumental] – X_V *–age* $_N$ [Action-N]:

caméscope – (°caméscoper) – caméscopage
caravane – (°caravaner) – caravanage
gommette – (°gommetter) – gommettage

Thank you!

References

Baayen, H. (1994). Derivational productivity and text typology. *Journal of Quantitative Linguistics* 1: 16-34.

Bauer, L., Lieber, R., & Plag, I. (2013). The Oxford reference guide to English Morphology. Oxford: Oxford University Press.

BONAMI, O. & STRNADOVA, J. (2019). Paradigm structure and predictability in derivational morphology. *MORPHOLOGY* 28, NO 2 (2019), 167-197.

Booij, G. (2010). Construction morphology. Oxford: Oxford University Press.

Dal, G. (2004). *Vers une morphologie de l'évidence: d'une morphologie de l'input à une morphologie de l'output* (Doctoral dissertation, Université Lille 3).

Dal, G., & Namer, F. (2015). La fréquence en morphologie: pour quels usages?. *Langages*, (1), 47-68.

Hathout, N. (2009). *Contributions à la description de la structure morphologique du lexique et à l'approche extensive en morphologie*. Habilitation à diriger des recherches. Universités de Toulouse II-Le Mirail. Toulouse.

Lignon, S., Namer, F., & Villoing, F. (2014). De l'agglutination à la triangulation ou comment expliquer certaines séries morphologiques. In *SHS Web of Conferences* (Vol. 8, pp. 1813-1835). EDP Sciences.

Marchand, H. (1969). The categories and types of present-day English wordformation. A synchronic-diachronic approach. Munich, Germany: Beck.

Missud, A., Amsili, P. & Villoing, F. (2020). VerNom : une base de paires morphologiques acquise sur très gros corpus. *Actes de la 27e conférence sur le Traitement Automatique des Langues Naturelles* (TALN 2020).

Missud, A. & Villoing, F. (2020). The morphology of rival –ion, –age and –ment selected verbal bases. In Dany Amiot & Delphine Tribout (eds.), *Lexique*, vol. 26, 29–52. Presses Universitaires de Lille.

References

Namer, F. (2009). *Morphologie, lexique et traitement automatique des langues,* pp.444. Hermès-Lavoisier.

Namer, F., & Hathout, N. (2020). ParaDis and Démonette–From Theory to Resources for Derivational Paradigms. *The Prague Bulletin of Mathematical Linguistics*, *114*(1), 5-34.

Plag, I. (2003). *Word-formation in English*. Cambridge University Press.

Roché, M. (2008). Structuration du lexique et principe d'économie: le cas des ethniques. In *Congrès Mondial de Linguistique Française* (p. 145). EDP Sciences.

Roché, M. (2009). Un ou deux suffixes ? Une ou deux suffixations ?. B. Fradin, F. Kerleroux et M. Plénat. *Aperçus de morphologie du français*, Presses Universitaires de Vincennes, pp.143-173.

Schäfer, R. (2015). Processing and querying large web corpora with the COW14 architecture. In P. BAÅ„SKI, H. BIBER, E. BREITENEDER, M. KUPIETZ, H. LÃNGEN & A. WITT (Éds.), Proceedings of Challenges in the Management of Large Corpora 3 (CMLC-3), Lancaster : UCREL IDS.

Schäfer, R. & Bildhauer, F. (2012). Building large corpora from the web using a new efficient tool chain. In N. C. C. CHAIR), K. CHOUKRI, T. DECLERCK, M. U. DOÄŸAN, B. MAEGAARD, J. MARIANI, A. MORENO, J. ODIJK & S. PIPERIDIS (Éds.), Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12), p. 486–493, Istanbul, Turkey : European Language Resources Association (ELRA).

Strnadová, J. (2014). *Les réseaux adjectivaux. Sur la grammaire des adjectifs dénominaux en français* (Doctoral dissertation, Université Paris Diderot (Paris 7) Sorbonne Paris Cité; Univerzita Karlova, Prague).