Inflectional and derivational predictions in morphological families.

A case study on Latin

Matteo Pellegrini, Federica Iurescia, Eleonora Litta and Marco Passarotti
{matteo.pellegrini,federica.iurescia,eleonoramaria.litta,marco.passarotti}@unicatt.it

Data-based Research in Word Formation
Biennial of Czech Linguistics 2024
Prague, 19 September 2024

Background

Predictability in inflection

- In the last decades, a large amount of work has been conducted on the mutual **predictability** of different **inflected forms** of lexemes – the Paradigm Cell Filling Problem
- Different methodologies have been devised for a quantitative investigation of this issue
  - Set-theoretic approach: Principal Part Analysis (Stump and Finkel 2013)
  - Graph-theoretic approach (Sims 2020)
  - Information-theoretic approach: implicative entropy (Bonami and Boyé 2014,Beniamine 2018)
- Typological investigations (Ackerman and Malouf 2013, Beniamine 2018)
- Empirical assessments on individual languages
  - Romance: French (Bonami and Boyé 2014), Portuguese (Beniamine, Bonami, and Luís 2021), Romanian (Herce and Pricop 2024)
  - … and beyond: Pitjantjatjara and Yankunytjatjara (Wilmoth and Mansfield 2021), Asama (Lévêque and Pellard 2023)

Derivational paradigms

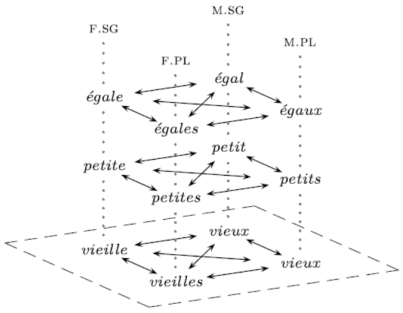
Derivational paradigms are another commonplace topic of recent Word-and-Paradigm approaches

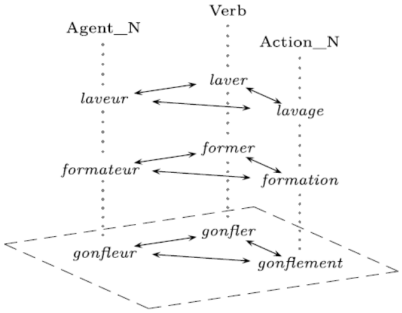
- Štekauer 2014
- Bauer 2019
- Hathout and Namer 2019
- Fernández-Domínguez, Bagasheva, and Lara-Clares 2020
- Ruz, Fernández-Alcaina, and Lara-Clares 2022

→ If the notion of paradigm is extended from inflection to derivation, then also techniques used to investigate inflectional predictions can be extended to derivational predictions

Entropy and predictability in derivation

- Bonami and Strnadová 2019 extend the use of implicative entropy to **derivation** – i.e., predicting (the citation form of) an unknown lexeme from a derivationally related one
- They define a paradigmatic system as a set of morphological families aligned by the same content relation, be it inflectional or derivational





- For a derivational paradigmatic system inlcuding verbs and related agent and action nouns, they show that key predictability properties are shared with inflectional systems

The next step

> *Note that we used the citation form of lexemes for predictions. A more thorough study would have examined the formal relations between all inflected forms of all lexemes, and determined whether different inflected forms contrast in predictiveness of derivationally related words. We doubt that this would have led to significantly different results in this particular instance.*
>
> (Bonami and Strnadová 2019, p. 189)

- In this study, we take a step in this direction, focusing on verbs and agent and action nouns in **Latin**, where this is likely to have an impact
- This allows us to quantify the role of forms other than the citation form in predicting the form of derivatives (cf. Bonami, Boyé, and Kerleroux 2009)

Latin data

Verb inflection

- 3(/4) **stems**
  - Present stem → forms of the "present system" (*infectum*, imperfective meaning)
  - Perfect stem → forms of the "perfect system" (*perfectum*, perfective meaning)
  - Third stem (Aronoff 1994) → perfect participle and supine
  - Fourth stem (Pellegrini 2023b) → future participle (different from the third stem only in a handful of cases)
- 4(/5) **conjugations** (only relevant in the present system)
  - 1st conj. → theme vowel *ā*
  - 2nd conj. → theme vowel *ē*
  - 3rd conj. → theme vowel *e*
  - 4th conj. → theme vowel *ī*
  - Mixed conj. (Dressler 2002) → heteroclitic 3rd/4th (Dressler)
- 4 **principal parts** (Stump and Finkel 2013) are sufficient to encapsulate (most of) this information

amō ~āre ~āuī ~ātum, **moneō** ~ēre ~uī ~itum, **legō**[2] ~ere lēgī lectum. **audiō** ~īre ~īuī or ~iī ~ītum,

Noun inflection

- 5 **declensions**
  - 1st decl. → *a* stems
  - 2nd decl. → *o* stems
  - 3rd decl. → *i* and consonant stems
  - 4th decl. → *u* stems
  - 5th decl. → *e* stems
- In the 3rd declension (and in a few nouns of the 2nd), 2 different **stems** are displayed in the NOM.SG and VOC.SG (and ACC.SG of neuter nouns) vs. other cells
- 2 **principal parts** are sufficient to identify the noun declension (and stem allomorphs if needed)

rosa ~ae, *f.* **lupus** ~ī, *m.*, **corpus** ~oris, *n.* **manus**[1] ~ūs, *f.* **spēs** ~eī, *f.*

Agent nouns

Deverbal agent nouns in Latin are overwhelmingly formed by adding to the third stem (Aronoff 1994, pp. 37 f.) a suffix *-or* that assigns the noun to (a subclass of) the 3rd declension

- ‘plough’: PRS.ACT.INF ***arare***, PRF.PASS.PTCP ***aratum*** → AgentN_NOM.SG ***arator***, AgentN_GEN.SG ***aratoris***
- ‘win’: PRS.ACT.INF ***uincere***, PRF.PASS.PTCP ***uictum*** → AgentN_NOM.SG ***uictor***, AgentN_GEN.SG ***uictoris***

However, there are a few cases where a different stem is displayed in the derivative

- ‘scatter’: PRS.ACT.INF ***spargere***, PRF.PASS.PTCP ***sparsum*** → AgentN_NOM.SG ***spartor***
- ‘tie’: PRS.ACT.INF ***ligare***, PRF.PASS.PTCP ***ligatum*** → AgentN_NOM.SG ***lictor***

Action nouns

For deverbal action nouns, a variety of processes are available and in (partial) competition

- Some of them operate on the third stem (like the agent-noun-forming suffix *-or*)
  - Suffix *-io* that assigns the noun to the 3rd declension (by far the most frequent strategy)
    - 'lead': PRS.ACT.INF ***duc****ere*, PRF.PASS.PTCP ***duct****um* → ActionN_NOM.SG ***duct****io*, ActionN_GEN.SG ***duct****ionis*
  - Suffix *-ura*, that assigns the noun to the 1st declension
    - 'burn': PRS.ACT.INF ***ard****ere*, PRF.PASS.PTCP ***ars****um* → ActionN_NOM.SG ***ars****ura*, ActionN_GEN.SG ***ars****urae*
- Other processes, however, operate on the present stem
  - Suffix *-ium*, that assigns the noun to the 2nd declension
    - 'strike': PRS.ACT.INF ***confligere***, PRF.PASS.PTCP ***conflictum*** → ActionN_NOM.SG ***confligium***, ActionN_GEN.SG ***confligi***
  - Suffix *-ido*, that assigns the noun to the 3rd declension
    - 'desire': PRS.ACT.INF ***cupire***, PRF.PASS.PTCP ***cupitum*** → ActionN_NOM.SG ***cupido***, ActionN_GEN.SG ***cupidinis***
- For all processes, there are exceptional cases with stems different from the one expected
  - 'grind': PRS.ACT.INF ***mol****are*, PRF.PASS.PTCP ***molitum*** → ActionN_NOM.SG ***molat****io*

Expectations

Given this state of affairs, we expect:

- Agent nouns in *-or* to be highly predictable from forms that display the third stem and less predictable from forms that display other stems
- No single form to be a reliable predictor of action nouns, for which several processes based on different stems are available
- Knowledge of multiple forms to be helpful to reduce uncertainty on the forms of derivatives
- Action nouns in *-io* and *-ura* to be good predictors of agent nouns in *-or*

Methodology, data, tools

Methodology: implicative entropy

How difficult is it to predict the PRS.ACT.IND.2PL from the gerund in Italian? (data from LeFFI)

verb (conj.)	GER	PRS.IND.2PL	pattern/context (GER ↔ PRS.IND.2PL)	applicable patterns	n. verbs
'love' (I)	aˈmando	ama:te	1 (_ando ↔ _a:te / C_#)	A (1)	2,081
'see' (II)	veˈdendo	veˈde:te	2 (_endo ↔ _e:te / C_#)	B (2,3)	355
'hear' (III)	senˈtendo	senˈti:te	3 (_endo ↔ _i:te / C_#)	B (2,3)	287

α	$P(\alpha)$	β	$P(\beta)$
A	$\frac{2,081}{2,723}$	1	$\frac{2,081}{2,081}$
B	$\frac{642}{2,723}$	2	$\frac{355}{642}$
		3	$\frac{287}{642}$

$$H(\beta|\alpha) = -\left(\frac{2,081}{2,723} \times \left(\frac{2,081}{2,081} \times \log_2 \frac{2,081}{2,081}\right) + \frac{642}{2,723} \times \left(\frac{355}{642} \times \log_2 \frac{355}{642} + \frac{287}{642} \times \log_2 \frac{287}{642}\right)\right) = 0,234$$

Data extraction

- We start from the principal parts of Latin verbs listed in **PrinParLat** (Pellegrini 2023a)
  - PRS.ACT.INF
  - FUT.ACT.IND.3.SG (more informative than PRS.ACT.IND.1SG)
  - PRF.ACT.IND.1SG
  - PRF.PASS.PTCP.NOM.N.SG
  - FUT.ACT.PTCP.NOM.N.SG (to account for the fourth stem)
- For each verb, we extract the NOM.SG of derivationally related nouns from **WFL** (Litta and Passarotti 2019) and automatically generate the GEN.SG
  - Agent nouns in *-or*
  - Action nouns in *-io*, *-ura*, *-ium*, *-ido* (we exclude deverbal nouns in *-mentum* and *-men* because they seem to have a more specific meaning in Classical Latin, cf. Palmer 1954)
- This yields a paradigmatic system with 9 cells for about 2,500 families

V_PRS. ACT.INF	V_FUT.ACT. IND.3.SG	V_PRF.ACT. IND.1.SG	V_PRF.PASS. PTCP	V_FUT.ACT. PTCP	ActionN_ NOM.SG	ActionN_ GEN.SG	AgentN_ NOM.SG	AgentN_ GEN.SG
abdicare profundere cupere	*abdicabit profundet cupiet*	*abdicaui profudi cupiui/cupii*	*abdicatum profusum cupitum*	*abdicaturum profusurum cupiturum*	*abdicatio profusio cupido*	*abdicationis profusionis cupidinis*	*abdicator profusor cupitor*	*abdicatoris profusoris cupitoris*

Dealing with variants

- Cases of **overabundance** ⇒ need for principled ways to select variants

V_prs. act.inf	V_fut.act. ind.3.sg	V_prf.pass. ptcp.nom.n.sg	ActionN_ nom.sg	AgentN_ nom.sg
arcessere/accersere	*arcesset/accerset*	*arcessitum/accersitum*	*arcessio*	*arcessitor*
aduertere/aduortere	*aduertet/aduortet*	*aduersum/aduorsum*	*aduersio*	*aduersitor*
alere	*alet*	*alitum/altum*	*alitura*	*altor*

- We follow Fradin and Kerleroux 2003 in taking form-based **flexemes** (rather than meaning-based lexemes) as the input of word formation processes, and select the appropriate flexeme

V_prs. act.inf	V_fut.act. ind.3.sg	V_prf.pass. ptcp.nom.n.sg	ActionN_ nom.sg	AgentN_ nom.sg
arcessere	*arcesset*	*arcessitum*	*arcessio*	*arcessitor*
aduertere	*aduertet*	*aduersum*	*aduersio*	*aduersitor*
alere	*alet*	*alitum*	*alitura*	
alere	*alet*	*altum*		*altor*

- When this does not help, we keep the most frequent variant (data from the LASLA corpus)

V_prs. act.inf	V_prf.act. ind.1sg	V_prf.pass. ptcp.nom.n.sg	ActionN_ nom.sg
intellegere (71)/*intelligere* (1)	*intellexi* (20)/***intellegi*** (76)	*intellectum*	*intellectio*
aedificare	*aedificaui*	*aedificatum*	*aedificatio* (5)/***aedificium*** (89)

Qumin-ready dataset

- Automatic conversion to an underspecified IPA transcription
  - No distinction [i]-[j] and [v]-[w]
  - Vowel length is only marked on endings
- Characterization of each segment in terms of phonological features
  - The sounds file of LatInfLexi (Pellegrini and Passarotti 2018) 2.0 is used (on its turn based on Cser 2020)

sound_id	tier	value	labial_vocalic	dorsal_vocalic	high	back
a	segmental			1	0	1
e	segmental			1	0	0
i	segmental			1	1	0
o	segmental		1	1	0	1
u	segmental		1	1	1	1
ˈ	stress	stressed				
ː	length	long				

- After these last steps, we can use tools developed for the Quantitative Modelling of Inflection (Qumin, Beniamine 2018) to automatically compute implicative entropy on our dataset

Results

Predictability of inflectionally related verb forms

	V_PRS. ACT.INF	V_FUT.ACT. IND.3.SG	V_PRF.ACT. IND.1.SG	V_PRF. PASS.PTCP. NOM.N.SG	V_FUT. ACT.PTCP. NOM.N.SG	ActionN_ NOM.SG	ActionN_ GEN.SG	AgentN_ NOM.SG	AgentN_ GEN.SG
V_PRS.ACT.INF		0.15	0.35	0.23	0.24	0.41	0.42	0.31	0.31
V_FUT.ACT.IND.3.SG	0.05		0.35	0.22	0.23	0.41	0.41	0.3	0.3
V_PRF.ACT.IND.1.SG	0.4	0.42		0.39	0.38	0.48	0.47	0.36	0.36
V_PRF.PASS.PTCP.NOM.N.SG	0.51	0.53	0.76		0.003	0.29	0.29	0.01	0.01
V_FUT.ACT.PTCP.NOM.N.SG	0.51	0.53	0.75	0.01		0.30	0.30	0.02	0.02
ActionN_NOM.SG	0.39	0.4	0.63	0.01	0.02		0	0	0
ActionN_GEN.SG	0.39	0.41	0.62	0.01	0.02	0		0	0
AgentN_NOM.SG	0.46	0.44	0.76	0.01	0.03	0.3	0.3		0
AgentN_GEN.SG	0.45	0.44	0.74	0.01	0.03	0.3	0.3	0	

- A much more in-depth investigation is provided by Pellegrini 2023b
- Results are not identical, due to different data and analytical choices (e.g. regarding granularity of IPA transcription and phonological features), but they are comparable
- The most important difference is the relatively lower predictiveness of perfect and future participles in the results of the current study, due to the fact that vowel lenght distinctions are not coded on stems

Predictability of agent nouns from verb forms

	V_PRS. ACT.INF	V_FUT.ACT. IND.3.SG	V_PRF.ACT. IND.1.SG	V_PRF. PASS.PTCP. NOM.N.SG	V_FUT. ACT.PTCP. NOM.N.SG	ActionN_ NOM.SG	ActionN_ GEN.SG	AgentN_ NOM.SG	AgentN_ GEN.SG
V_PRS.ACT.INF		0.15	0.35	0.23	0.24	0.41	0.42	0.31	0.31
V_FUT.ACT.IND.3.SG	0.05		0.35	0.22	0.23	0.41	0.41	0.3	0.3
V_PRF.ACT.IND.1.SG	0.4	0.42		0.39	0.38	0.48	0.47	0.36	0.36
V_PRF.PASS.PTCP.NOM.N.SG	0.51	0.53	0.76		0.003	0.29	0.29	0.01	0.01
V_FUT.ACT.PTCP.NOM.N.SG	0.51	0.53	0.75	0.01		0.30	0.30	0.02	0.02
ActionN_NOM.SG	0.39	0.4	0.63	0.01	0.02		0	0	0
ActionN_GEN.SG	0.39	0.41	0.62	0.01	0.02	0		0	0
AgentN_NOM.SG	0.46	0.44	0.76	0.01	0.03	0.3	0.3		0
AgentN_GEN.SG	0.45	0.44	0.74	0.01	0.03	0.3	0.3	0	

- The perfect participle is by far the best predictor (close-to-perfect reliability) - ‘win’: PRS.ACT.INF ***uincere***, PRF.PASS.PTCP ***uictum*** → AgentN_NOM.SG ***uictor***
- The future participle is very similar (although slightly less reliable)
- Forms of the present system are better predictors than those of the perfect system! - ‘cleanse’: PRS.ACT.INF ***abluere***, PRF.PASS.PTCP ***ablutum*** → AgentN_NOM.SG ***ablutor***

Predictability of action nouns from verb forms

	V_PRS. ACT.INF	V_FUT.ACT. IND.3.SG	V_PRF.ACT. IND.1.SG	V_PRF. PASS.PTCP. NOM.N.SG	V_FUT. ACT.PTCP. NOM.N.SG	ActionN_ NOM.SG	ActionN_ GEN.SG	AgentN_ NOM.SG	AgentN_ GEN.SG
V_PRS.ACT.INF		0.15	0.35	0.23	0.24	0.41	0.42	0.31	0.31
V_FUT.ACT.IND.3.SG	0.05		0.35	0.22	0.23	0.41	0.41	0.3	0.3
V_PRF.ACT.IND.1.SG	0.4	0.42		0.39	0.38	0.48	0.47	0.36	0.36
V_PRF.PASS.PTCP.NOM.N.SG	0.51	0.53	0.76		0.003	0.29	0.29	0.01	0.01
V_FUT.ACT.PTCP.NOM.N.SG	0.51	0.53	0.75	0.01		0.30	0.30	0.02	0.02
ActionN_NOM.SG	0.39	0.4	0.63	0.01	0.02		0	0	0
ActionN_GEN.SG	0.39	0.41	0.62	0.01	0.02	0		0	0
AgentN_NOM.SG	0.46	0.44	0.76	0.01	0.03	0.3	0.3		0
AgentN_GEN.SG	0.45	0.44	0.74	0.01	0.03	0.3	0.3	0	

- The perfect and future participle are still the best predictors (nouns in *-io*, that behave very similarly to agent nouns in *-or*, are by far the most frequent strategy) - 'pour out': PRS.ACT.INF ***profund****ere*, PRF.PASS.PTCP ***profus****um* → ActionN_NOM.SG ***profus****io*
- No single form is a very reliable predictor (due to competition between different processes: *-io*, *-ura*, *-ium*, *-ido*)

Interpredictability of agent and action nouns

	V_PRS. ACT.INF	V_FUT.ACT. IND.3.SG	V_PRF.ACT. IND.1.SG	V_PRF. PASS.PTCP. NOM.N.SG	V_FUT. ACT.PTCP. NOM.N.SG	ActionN_ NOM.SG	ActionN_ GEN.SG	AgentN_ NOM.SG	AgentN_ GEN.SG
V_PRS.ACT.INF		0.15	0.35	0.23	0.24	0.41	0.42	0.31	0.31
V_FUT.ACT.IND.3.SG	0.05		0.35	0.22	0.23	0.41	0.41	0.3	0.3
V_PRF.ACT.IND.1.SG	0.4	0.42		0.39	0.38	0.48	0.47	0.36	0.36
V_PRF.PASS.PTCP.NOM.N.SG	0.51	0.53	0.76		0.003	0.29	0.29	0.01	0.01
V_FUT.ACT.PTCP.NOM.N.SG	0.51	0.53	0.75	0.01		0.30	0.30	0.02	0.02
ActionN_NOM.SG	0.39	0.4	0.63	0.01	0.02		0	0	0
ActionN_GEN.SG	0.39	0.41	0.62	0.01	0.02	0		0	0
AgentN_NOM.SG	0.46	0.44	0.76	0.01	0.03	0.3	0.3		0
AgentN_GEN.SG	0.45	0.44	0.74	0.01	0.03	0.3	0.3	0	

- Action nouns are fully reliable predictors of agent nouns (the implicative relation between action nouns in *-io* and *-ura* and agent nouns is even stronger than the one between the latter and the perfect/future participle!)
- Agent nouns are much less reliable predictors of action nouns (due to the availability, besides *-io*, of other action-noun forming processes, that cannot be predicted from the agent noun in *-or*)

Predicting agent nouns: binary implicative entropy

What happens if, rather than predicting derivatives from a single form, we do that from multiple forms?

	H(AgentN|C1)	H(AgentN|C2)	H(AgentN|C1+C2)
V_PRS.ACT.INF, V_PRF.ACT.IND.1.SG	0.31	0.36	0.04
V_PRS.ACT.INF, V_PRF.PASS.PTCP.NOM.N.SG	0.31	0.01	0.008
V_PRS.ACT.INF, ActionN_NOM.SG	0.31	0	0
V_PRF.ACT.IND.1.SG, V_PRF.PASS.PTCP.NOM.N.SG	0.36	0.01	0.006
PRF.ACT.IND.1.SG, ActionN_NOM.SG	0.36	0	0
PRF.PASS.PTCP.NOM.N.SG, ActionN_NOM.SG	0.01	0	0

- Joint knowledge of more than one form always leaves virtually no uncertainty on the formation of the agent noun in *-or*
- Interestingly, this happens even if both forms are not strongly predictive by themselves (i.e., for the pair PRS.ACT.INF+PRF.ACT.IND.1.SG)

Predicting action nouns: binary implicative entropy

	H(ActionN|C1)	H(ActionN|C2)	H(ActionN|C1+C2)
V_PRS.ACT.INF, V_PRF.ACT.IND.1.SG	0.41	0.48	0.17
V_PRS.ACT.INF, V_PRF.PASS.PTCP.NOM.N.SG	0.41	0.29	0.16
V_PRS.ACT.INF, AgentN_NOM.SG	0.41	0.3	0.14
V_PRF.ACT.IND.1.SG, V_PRF.PASS.PTCP.NOM.N.SG	0.48	0.29	0.18
V_PRF.ACT.IND.1.SG, AgentN_NOM.SG	0.48	0.3	0.18
V_PRF.PASS.PTCP.NOM.N.SG, AgentN_NOM.SG	0.29	0.3	0.28

- Joint knowledge of more than one form – be them inflectionally or derivationally related – always reduces uncertainty remarkably
- The pair composed of the perfect participle and the agent noun is the only exception (the most useful piece of information provided by the latter – i.e., the third stem – is already provided by the former)
- A non-negligible amount of uncertainty remains (due to the availability of different action-noun forming processes)

Conclusions

Conclusions

We have investigated the differential predictability of derivatives from various inflected forms, using the same techniques that have been applied to inflectionally and derivationally related forms

- Confirmation of expectations from qualitative observations:
  - the perfect participle is by far the best predictor of agent nouns;
  - the difference in predictiveness of different verb forms is smaller for action nouns;
  - action nouns predict agent nouns better than the reverse.
- New empirical findings:
  - forms of the present system are better predictors than those of the perfect system for both agent nouns and action nouns;
  - joint knowledge of more than one form is helpful, regardless of whether the forms are inflectionally or derivationally related, and even if the forms are not strongly predictive by themselves.

$\Rightarrow$ Evaluating the predictability of derivational processes on the basis of the citation form alone might mean missing important information!

Thanks!

Get in touch

Matteo Pellegrini, Federica Iurescia, **Eleonora Litta** and Marco Passarotti

{matteo.pellegrini,federica.iurescia,eleonoramaria.litta,marco.passarotti}@unicatt.it

https://github.com/CIRCSE

https://centridiricerca.unicatt.it/circse_index.html

Largo Gemelli 1, 20123 Milan, Italy

Works cited I

Ackerman, Farrell and Robert Malouf (2013). "Morphological organization: the low conditional entropy conjecture". In: *Language* 89, pp. 429–464.

Aronoff, Mark (1994). *Morphology by itself: Stems and inflectional classes*. MIT press.

Bauer, Laurie (2019). "Notions of paradigm and their value in word-formation". In: *Word Structure* 12.2, pp. 153–175.

Beniamine, Sacha (2018). "Classifications flexionnelles. Étude quantitative des structures de paradigmes". PhD thesis. Université Sorbonne Paris Cité-Université Paris Diderot.

Beniamine, Sacha, Olivier Bonami, and Ana R Luís (2021). "The fine implicative structure of European Portuguese conjugation". In: *Isogloss. Open Journal of Romance Linguistics* 7, pp. 1–35.

Bonami, Olivier and Gilles Boyé (2014). "De formes en thèmes". In: *Foisonnements morphologiques. Études en hommage à Françoise Kerleroux*, pp. 17–45.

Bonami, Olivier, Gilles Boyé, and Françoise Kerleroux (2009). "L'allomorphie radicale et la relation flexion-construction". In: *Aperçus de morphologie du français*, pp. 103–125.

Bonami, Olivier and Jana Strnadová (2019). "Paradigm structure and predictability in derivational morphology". In: *Morphology* 29.2, pp. 167–197.

Cser, András (2020). *The phonology of classical Latin*. John Wiley & Sons.

Dressler, Wolfgang U (2002). "Latin inflection classes". In: *Theory and description in Latin linguistics*, pp. 91–110.

Fernández-Domínguez, Jesús, Alexandra Bagasheva, and Cristina Lara-Clares (2020). *Paradigmatic relations in word formation*. Brill.

Fradin, Bernard and Françoise Kerleroux (2003). "Troubles with lexemes". In: *Topics in morphology. Selected Papers from the Third Mediterranean Morphology Meeting*, pp. 177–196.

Hathout, Nabil and Fiammetta Namer (2019). "Paradigms in word formation: what are we up to?" In: *Morphology* 29, pp. 153–165.

Herce, Borja and Bogdan Pricop (2024). "VeLeRo: an inflected verbal lexicon of standard Romanian and a quantitative analysis of morphological predictability". In: *Language Resources and Evaluation*.

Works cited II

Lévêque, Dimitri and Thomas Pellard (2023). "The implicative structure of Asama verb paradigms: A quantitative study of segmental and suprasegmental alternations". In: *Morphology* 33.3, pp. 261–286.

Litta, Eleonora and Marco Passarotti (2019). "(When) inflection needs derivation: a word formation lexicon for Latin". In: *Lemmata Linguistica Latina*, pp. 224–239.

Palmer, Leonard Robert (1954). *The latin language*. Faber and Faber.

Pellegrini, Matteo (2023a). "Flexemes in theory and in practice: Modelling overabundance in Latin verb paradigms". In: *Morphology* 33.3, pp. 361–395.

— (2023b). *Paradigm Structure and Predictability in Latin Inflection*. Springer.

Pellegrini, Matteo and Marco Passarotti (2018). "LatInfLexi: an inflected lexicon of Latin verbs". In: *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*, pp. 324–329.

Ruz, Alba E, Cristina Fernández-Alcaina, and Cristina Lara-Clares (2022). *Paradigms in Word Formation: Theory and Applications*. John Benjamins Publishing Company.

Sims, Andrea D. (2020). "Inflectional networks: Graph-theoretic tools for inflectional typology". In: *Proceedings of the Society for Computation in Linguistics 2020*, pp. 302–312.

Štekauer, Pavol (2014). "Derivational paradigms". In: *The Oxford handbook of derivational morphology*, pp. 354–369.

Stump, Gregory T. and Raphael A. Finkel (2013). *Morphological typology: From word to paradigm*. Cambridge University Press.

Wilmoth, Sasha and John Mansfield (2021). "Inflectional predictability and prosodic morphology in Pitjantjatjara and Yankunytjatjara". In: *Morphology* 31.4, pp. 355–381.