Baseless derivation: the behavioural reality of derivational paradigms
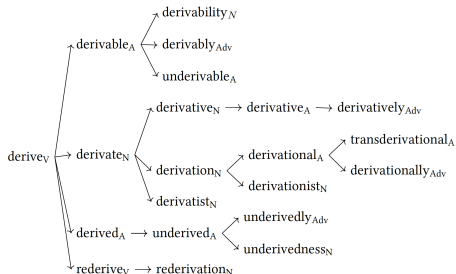
Maria Copot   Olivier Bonami

Université Paris Cité

- Two theoretical approaches to **derivation** and **morphological families**
  - Rooted tree vs paradigmatic
- Why picking the right one matters - different **predictions**
- **Testing** the predictions - what do speakers do?

- Rooted in a **base**
- **Monodirected** links from the base outwards only
- Only **one incoming edge** per word

Multimotivation Multiple candidates for the base

- *Rederivation ← re+derivation? rederive+ation?*

Multimotivation Multiple candidates for the base
  - *Rederivation* ← *re+derivation*? *rederive+ation*?

Back-formation Morphologically simpler words that are obtained from morphologically more complex words
  - *resurrect*←*resurrection* (**to resurge*, ≠*apply*∼*application*)

Uprooting the tree

Multimotivation Multiple candidates for the base
- *Rederivation ← re+derivation? rederive+ation?*

Back-formation Morphologically simpler words that are obtained from morphologically more complex words
- *resurrect←resurrection (*to resurge, ≠apply∼application)*

Cross-formation Leaf nodes having a closer relationship to each other than to their base
- Nouns of pattern *Xism∼Xist - optimism∼optimist*

Multimotivation Multiple candidates for the base
- *Rederivation ← re+derivation? rederive+ation?*

**Back-formation** Morphologically simpler words that are obtained from morphologically more complex words
- *resurrect←resurrection (*to resurge, ≠apply∼application)*

Cross-formation Leaf nodes having a closer relationship to each other than to their base
- Nouns of pattern *Xism∼Xist - optimism∼optimist*

Seen as **peripheral** by proponents of the rooted tree, but to others they represent a need to **reconceptualise** how we think of morphological relationships.
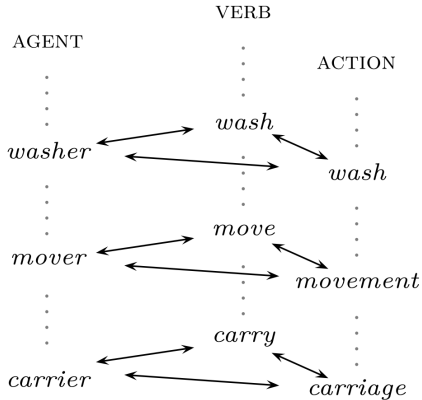
- **Bidirectional** relationships

- **Multiple** incoming edges

- No status of **base**

Variation on a gradient

The two views outlined are **extremes** on a gradient

100% rooted tree Lexeme-based morphology from Aronoff (1976) onwards

Rooted tree + paradigmatic relationships where necessary Construction Morphology (Booij, 2010) and Relational Morphology (Jackendoff & Audring, 2020)

100% paradigmatic Word-and-paradigm approaches to word formation

- Much empirical comparison has been in the form of **case studies**

Comparing frameworks

- Much empirical comparison has been in the form of **case studies**
- Larger scale **quantitative studies** show the information-theoretic need for the paradigmatic alternative
  - e.g. Bonami & Strnadová (2019), Bonami & Guzmàn-Naranjo (2022)

Comparing frameworks

- Much empirical comparison has been in the form of **case studies**
- Larger scale **quantitative studies** show the information-theoretic need for the paradigmatic alternative
    - e.g. Bonami & Strnadová (2019), Bonami & Guzmàn-Naranjo (2022)
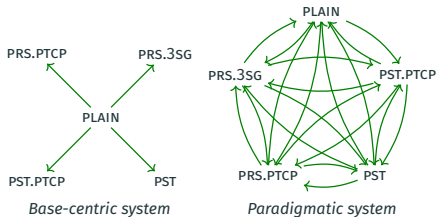- **Speaker behaviour** is not really part of the discussion

The two frameworks make different **predictions** about which relationships between word forms are **accessible** to speakers

Paradigmatic all relationships are available, speakers exploit all generalisations they can

Rooted tree only relationships from a stem to its derived words are tracked by speakers.

- **Inflection** has a similar framework debate

Base-centric system   *Paradigmatic system*

The parallels with inflection

- **Inflection** has a similar framework debate
- Longer-standing involvement of **cognitive** predictions (Jun & Albright, 2016 - Single Base Hypothesis)
- Copot & Bonami (2022) tested the predictions controlling for cell frequency and found results suggesting **speakers were aware of and used implicative relationships** in inflection.
  - bidirectionally
  - giving the base no special status
- Is the same true for **derivation**?

Why is testing frameworks important?

- Important for **morphological theory** - stating the obvious
- Important for any fields that **rely on morphological theory**
  - much experimental and psycholinguistic work on morphology assumes a cognitively untested idea.
  - e.g. experiments on "complex words", design reling on a base

Methodology

Acceptability judgement task

"J'aime le monde de la **catonisation**. Je veux être **catoniseur** quand je serai grand."

I love the world of ACTION NOUN. I want to be AGENT NOUN when I grow up.

I love the world of ACTION NOUN. I want to be AGENT NOUN when I grow up.

Does the second word sound good in this context?

Sonne mal ●━━━━━━━ Sonne bien

J'adore le monde de la catonisation.

Je veux être { catonisateur ▬▬▬

catoniseur quand je serai grand. ▪

catonisier ▪

I love the world of ACTION NOUN.

I want to be { AGENT-1

AGENT-2 when I grow up.

AGENT-3

- **Six directed cell pairs**, based on work by Bonami & Strnadová (2019)'s work identifying French derivational families

(a) Rooted tree   (b) Paradigmatic

Predictor $\rightarrow$ Target
VERB $\rightarrow$ AGENT NOUN
AGENT NOUN $\rightarrow$ VERB
VERB $\rightarrow$ ACTION NOUN
ACTION NOUN $\rightarrow$ VERB
AGENT NOUN $\rightarrow$ ACTION NOUN
ACTION NOUN $\rightarrow$ AGENT NOUN

- **Three morphological patterns** chosen per directed cell pair, maximally **differing** in type frequency.
- **Nine items per directed cell pair**, three for each level of type frequency. 54 crucial items.
- Distractors: pseudolexemes in inflectional relationships.

- **Pseudolexemes** based on French derivational families (Bonami & Strnadová, 2019)
  - made with Wuggy (Keuleers & Brysbaert, 2010), to **match phonology** of items belonging to each morphological pattern of interest

If speakers are at all **aware of implicative relationships**

- the **more expected** the second form is from the first, the **better it will be rated.**

If speakers are at all **aware of implicative relationships**

- the **more expected** the second form is from the first, the **better it will be rated.**

If speakers use the distributional information inherent in the **implicative relationships** set up by the paradigm, this will hold true...

- For **all directions of prediction**
- For **all cell pairs**

Quantifying predictability

- To quantify the **expectedness** of the second form conditional on the first, we use the **Minimal Generalisation Learner** (MGL) (Albright & Hayes, 2003) scores.
  - Quantifies **how probable is an output form given an input form**
  - Both **quantitative** and **behavioural** evidence has been gathered thanks to it (Albright & Hayes, 2003; Albright & Hayes, 2002; Albright, 2003; Jun & Albright, 2016)

The Minimal Generalisation Learner

- Method to obtain **mappings** between the two cells of interest.
- **Input:** pairs of forms in the two cells.

VERB	ACTION NOUN
laver	laveur
bouder	boudeur
finir	finisseur
. . .	

- The method extracts **generalisations** mapping the first cell to the second, taking into account the **phonology** of the stem, eg

VERB	ACTION NOUN
Xer →	Xeur
. . .	

- After training, an **unseen pair** of input and output forms can be submitted
    - how likely is the output conditional on the input…
    - in light of the patterns found in the lexicon and their type frequency?
- For each item, the model calculates its **confidence score** $\propto P(\mathit{target}|\mathit{input})$

Phonological well-formedness judgements

- A different set of participants was asked to provide **phonological well-formedness judgements** on the target forms.
- 20 well-formedness judgements for each target form, averaged into a phonological well-formedness score for the word

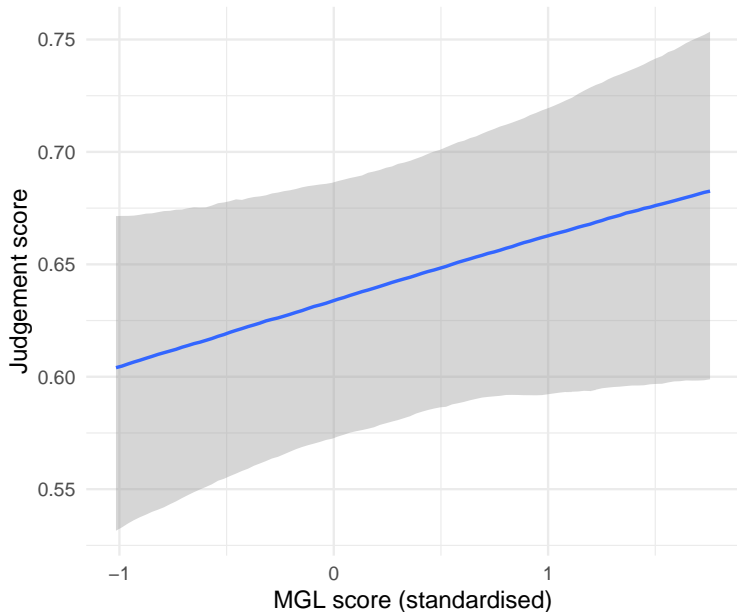Predict **acceptability judgement** of the target form from...

- **MGL form predictability score** of the target form given the predictor
- **well-formedness judgement**
- directed **cell** pair

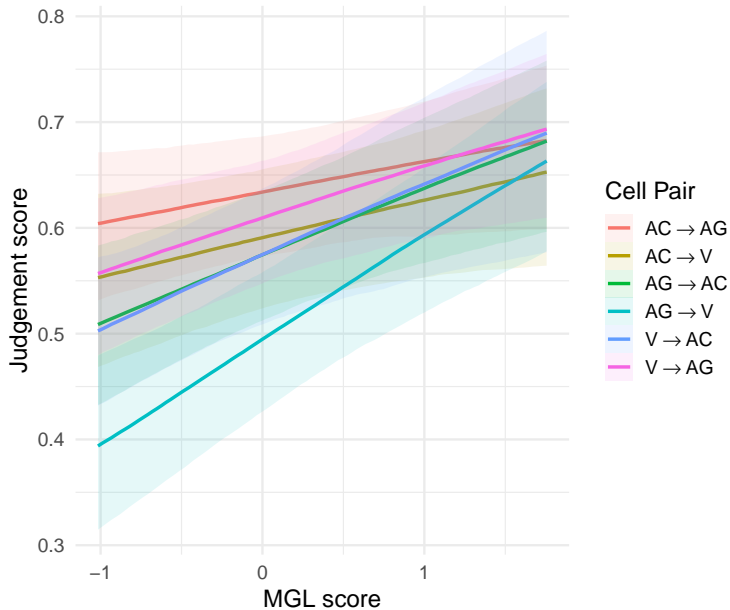Random intercepts for item and participant fitting a beta distribution.

$$judgment \sim MGL\ score * cell + wellformedness +$$
$$(1|participant) + (1|item)$$

60 participants (Prolific.co) * 54 judgements = 3240 datapoints.

- The crucial data point: what happens when speakers are asked to **predict towards the base?**

- The crucial data point: what happens when speakers are asked to **predict towards the base?**
- Model predictions towards the base only (X→BASE)
  Paradigmatic prediction X→BASE scores fit best
  Rooted tree prediction BASE→X scores fit best

- The crucial data point: what happens when speakers are asked to **predict towards the base?**
- Model predictions towards the base only (X→BASE)
  Paradigmatic prediction X→BASE scores fit best
  Rooted tree prediction BASE→X scores fit best
- LOO-CV between models with X→BASE and BASE→X scores: **X→BASE** is a better fit, fulfilling **paradigmatic** prediction.

- Speakers are **aware of implicative relationships** in derivational word families
- Morphological **theories** that wish to claim cognitive relevance should have **mechanisms** that resemble implicative relationships
- **Applications** that are based on morphological theories should apply a paradigmatic filter to the methodology and results interpretation.

Thank you!

maria.copot.s@gmail.com
olivier.bonami@u-paris.fr