Baseless derivation: the behavioural reality of derivational paradigms

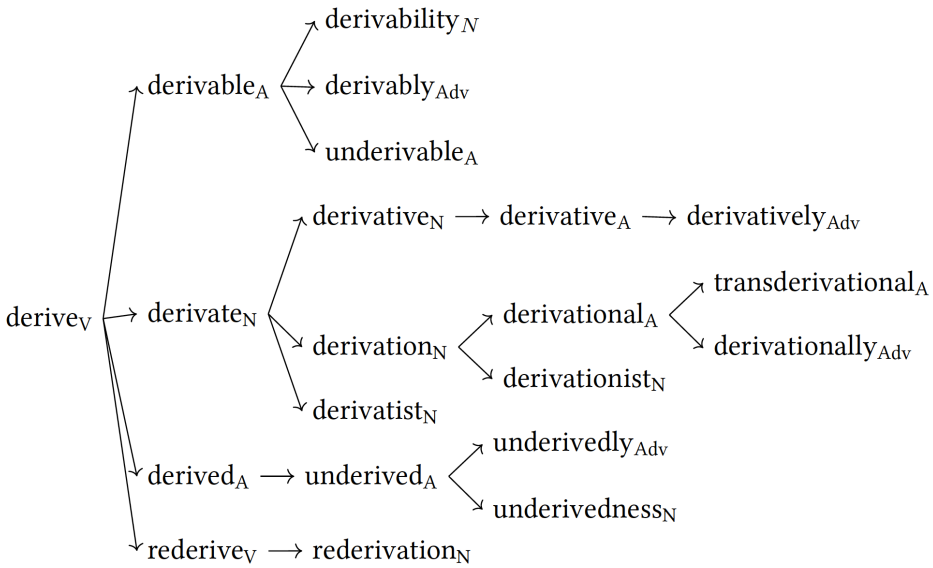
Maria Copot Olivier Bonami

Université Paris Cité

The next 20 minutes of your life

- 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 trees



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

Uprooting the tree

Multimotivation Multiple candidates for the base

- *Ĝāăāʬ̍0̍˙ hʋ¯0̍ωɧ* *ʬā˧ăāʬ̍0̍˙ hʋ¯0̍ωɧ? ʬāăāʬ̍0̍˙ ā˧hʋ¯0̍ωɧ?*

Uprooting the tree

Multimotivation Multiple candidates for the base

- *Ĝāăāẅ̊·hɯ̄̊ωɦ* *ẅāɬăāẅ̊·hɯ̄̊ωɦ?* *ẅāăāẅ̊·āɬhɯ̄̊ωɦ?*

Back-formation Morphologically simpler words that are obtained from morphologically more complex words

- *ẅā'ɷẅẅā!⁻* *ẅā'ɷẅẅā!⁻̊ωɦ* (¸⁻ω ẅā'ɷẅÔā, ɕhɻʕʕɣ̀ *hɻʕʕɣ̊!hɯ⁻̊ωɦ*)

Uprooting the tree

Multimotivation Multiple candidates for the base

- *Ĝāăāẅ̊̇hɯ̄̊ɷɦ* *ẅāɬăāẅ̊̇hɯ̄̊ɷɦ?* *ẅāăāẅ̊̇āɬhɯ̄̊ɷɦ?*

Back-formation Morphologically simpler words that are obtained from morphologically more complex words

- *ẅā'ˬẅẅā!⁻* *ẅā'ˬẅẅā!̄̊ɷɦ* (*¸̄ɷ ẅā'ˬẅÔā, ɕhɯʕʕɣ̀* *hɯʕʕɣ̊!hɯ̄̊ɷɦ*)

Cross-formation Leaf nodes having a closer relationship to each other than to their base

- Nouns of pattern *tˍ̊'g* *tˍ̊'⁻-ɷʕ̄̊gˍ̊'g* *ɷʕ̄̊gˍ̊'⁻*

Uprooting the tree

Multimotivation Multiple candidates for the base

- *Ĝāăāẅ0̇˙ hʋ¯0̇ωɧ ẅā˧ăāẅ0̇˙ hʋ¯0̇ωɧ? ẅāăāẅ0̇˙ ā˧hʋ¯0̇ωɧ?*

Back-formation Morphologically simpler words that are obtained from morphologically more complex words

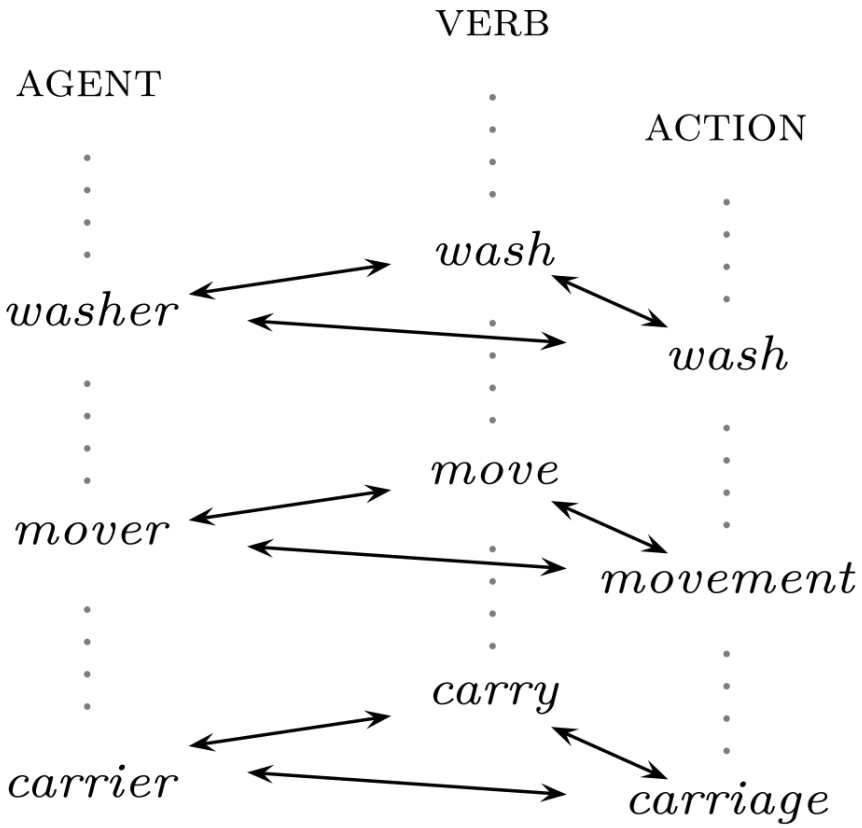
- *ẅā’˷ẅẅā!¯ ẅā’˷ẅẅā!¯0̇ωɧ (¸¯ω ẅā’˷ẅÔā, ϵhʋꞔꞔɏ hʋꞔꞔɏ0̇!hʋ¯0̇ωɧ)*

Cross-formation Leaf nodes having a closer relationship to each other than to their base

- Nouns of pattern *ɬ0̇’g ɬ0̇’¯ - ωꞔ¯0̇g0̇’g ωꞔ¯0̇g0̇’¯*

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

The paradigmatic alternative



- **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

Comparing frameworks

- 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

Framework - behavioural predictions

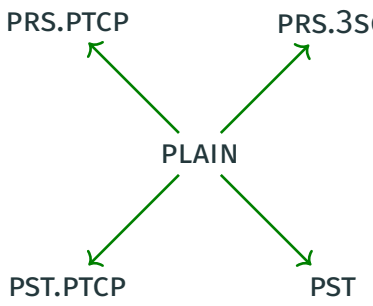
The two frameworks make different **predictions** about which relationships between word forms are **accessible** to speakers

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

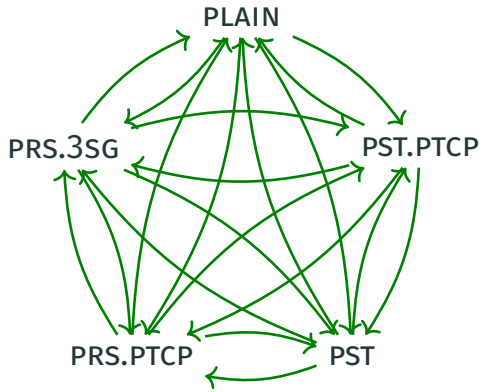
Rooted tree only relationships from a stem to its derived words are tracked by speakers.

The parallels with inflection

- **Inflection** has a similar framework debate



@hu’ā˥!āɧ¯ʬ̊! ’‘ ’¯ā g



ěhuʬhuǎ̊Ô g hu¯̊! ’‘ ’¯ā g

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.”

²ɣɷ˙ā ¯ɋā ˝ ɷʷɣă ɷì **¯EwɬΠB BΠΨB**¡ ² ˝ ɧɧ¯ ¯ɷ ɛā **¯ώÏBw BΠΨB** ˝ ɋāɧ ² Ôʷɷ˝ ˛ Ɛ¡

Thanks to Cassandre Despujols and Clara Hirst for the videos

Acceptability judgement task

²ɣɷ˙ā ¯ɋā ˝ ɷʷɣă ɷì ¯EwϟΠB BΠΨBⁱ ² ˝ ɯɦ¯ ¯ɷ ɛā ¯ώÏBw BΠΨB ˝ ɋāɦ ² Ôʷɷ˝ ˷ £ⁱ

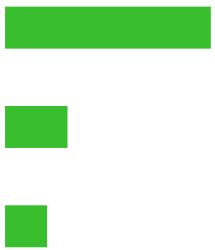
Pɷā’ ¯ɋā ’ā!ɷɦă ˝ ɷʷă ’ɷ˷ ɦă Ôɷɷă ̊ɦ ¯ɋ̊’ !ɷɦ¯ā’ ¯ˈ

Sonne mal Sonne bien

Items - sentences

J'adore le monde de la catonisation.

Je veux être { *catonisateur* / *catoniseur* / *catonisier* } quand je serai grand.

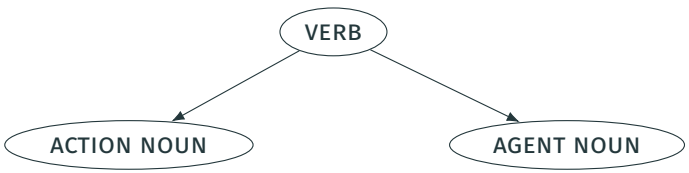


I love the world of ACTION NOUN.

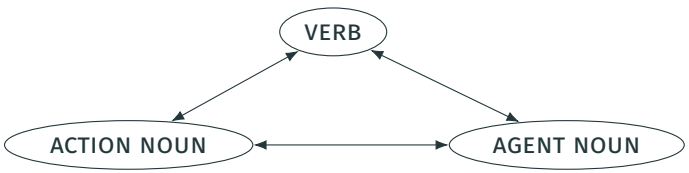
I want to be { AGENT-1 / AGENT-2 / AGENT-3 } when I grow up.

Items - cells

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



(a) Rooted tree



(b) Paradigmatic

Items - cells

Predictor → Target
VERB → AGENT NOUN
AGENT NOUN → VERB
VERB → ACTION NOUN
ACTION NOUN → VERB
AGENT NOUN → ACTION NOUN
ACTION NOUN → 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.

Items - pseudolexemes

- **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

Hypothesis

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.**

Hypothesis

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

The Minimal Generalisation Learner

- 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** / ĕ(¯ƕẘÔā¯j̊0ɧꞨ˞ ¯)

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

Analysis

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

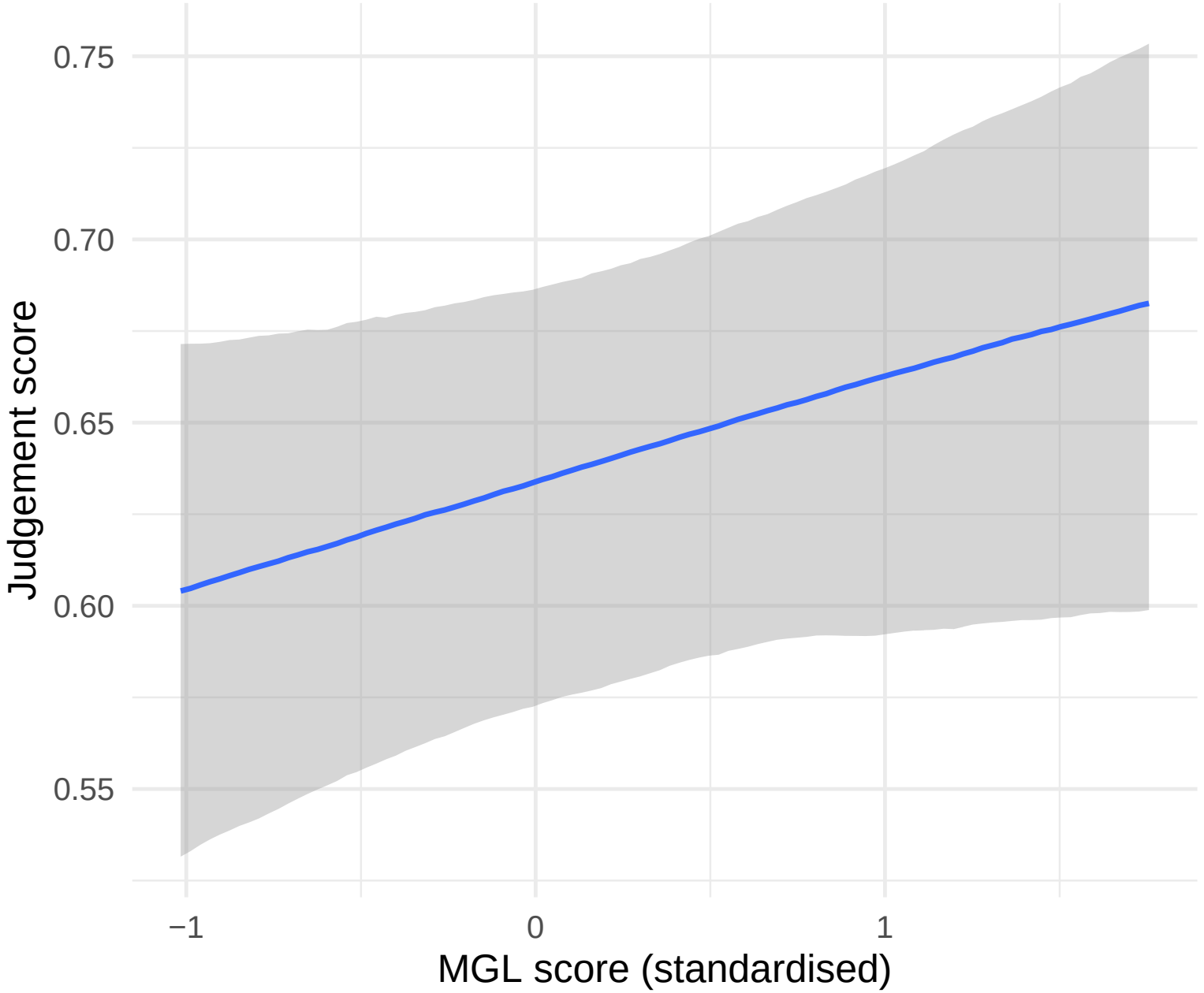
- **MGL form predictability score** of the target form given the predictor
- **well-formedness judgement**
- directed **cell** pair

Random intercepts for item and participant fitting a beta distribution.

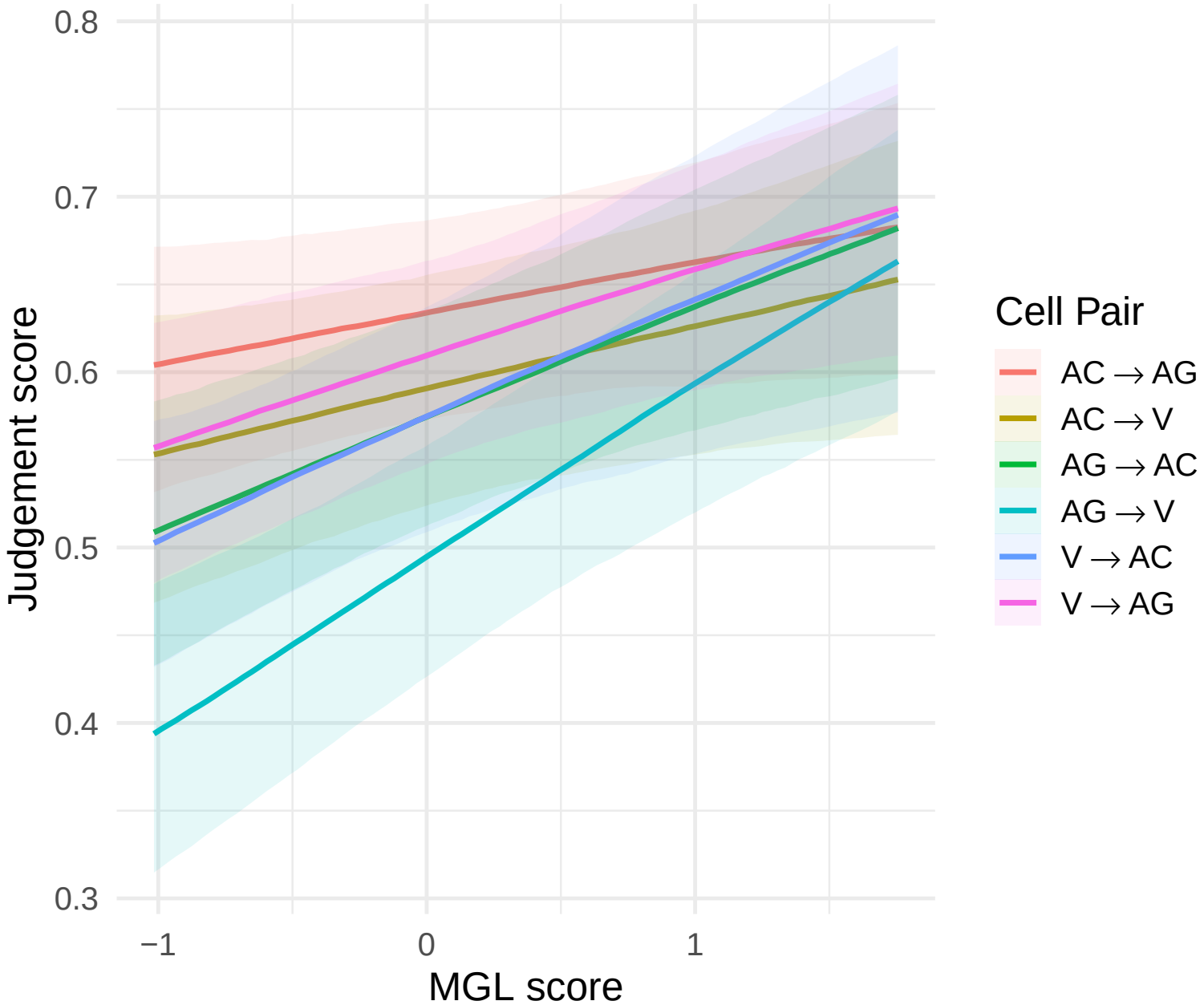
$$ʔ\sim ăÔ\,g\,ā\hbar^{-} \quad Ù\sim Í\ '!\omega^{w}\bar{a}\ ¸\ !\bar{a}\text{ÿÿ}\ \text{ㄐ}\ ''\ \bar{a}\text{ÿÿ}\grave{\imath}\omega^{w}\,g\,\bar{a}\check{a}\hbar\bar{a}''\ \text{ㄐ}$$
$$\text{ㄋ}\ddot{\text{o}}\text{Ł}h\upsilon^{w}\bar{\text{o}}!\bar{\text{o}}\text{Ł}h\upsilon\hbar^{-}\text{◻}\ \text{ㄐ}\ \text{ㄋ}\ddot{\text{o}}\bar{\text{o}}^{-}\bar{a}\,g\,\text{◻}$$

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

Results - word form predictability



Results - cells



Focusing on the base

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

Focusing on 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

Focusing on 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
- LOO-CV between models with X *!* BASE and BASE *!* X scores: **X *!* BASE** is a better fit, fulfilling **paradigmatic** prediction.

Conclusion

- 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

Appendix

The Minimal Generalisation Learner

1. Trained on **pairs of forms** belonging to two paradigm cells. MGL yields all possible **mappings from the first form to the second**
   [hæk] *!* [hækt] ø *!* t/hæk_
   [dis] *!* [dist] ø *!* t/dis_

The Minimal Generalisation Learner

1. Trained on **pairs of forms** belonging to two paradigm cells. MGL yields all possible **mappings from the first form to the second**
   [hæk] *!* [hækt] ø *!* t/hæk_
   [dis] *!* [dist] ø *!* t/dis_
2. The obtained rules are compared to each other and are subsumed under **generalised versions** where possible
   ø *!* t/hæk_
   ø *!* t/dis_
   ø *!* t/[-voiced]_

The Minimal Generalisation Learner

1. Trained on **pairs of forms** belonging to two paradigm cells. MGL yields all possible **mappings from the first form to the second**
   [hæk] *!* [hækt] ø *!* t/hæk_
   [dis] *!* [dist] ø *!* t/dis_
2. The obtained rules are compared to each other and are subsumed under **generalised versions** where possible
   ø *!* t/hæk_
   ø *!* t/dis_
   ø *!* t/[-voiced]_
3. Output: a set of rules with different degrees of specificity – a given input form will usually have more than one applicable rule.