Empirical evidence for paradigmatic organization across morphology

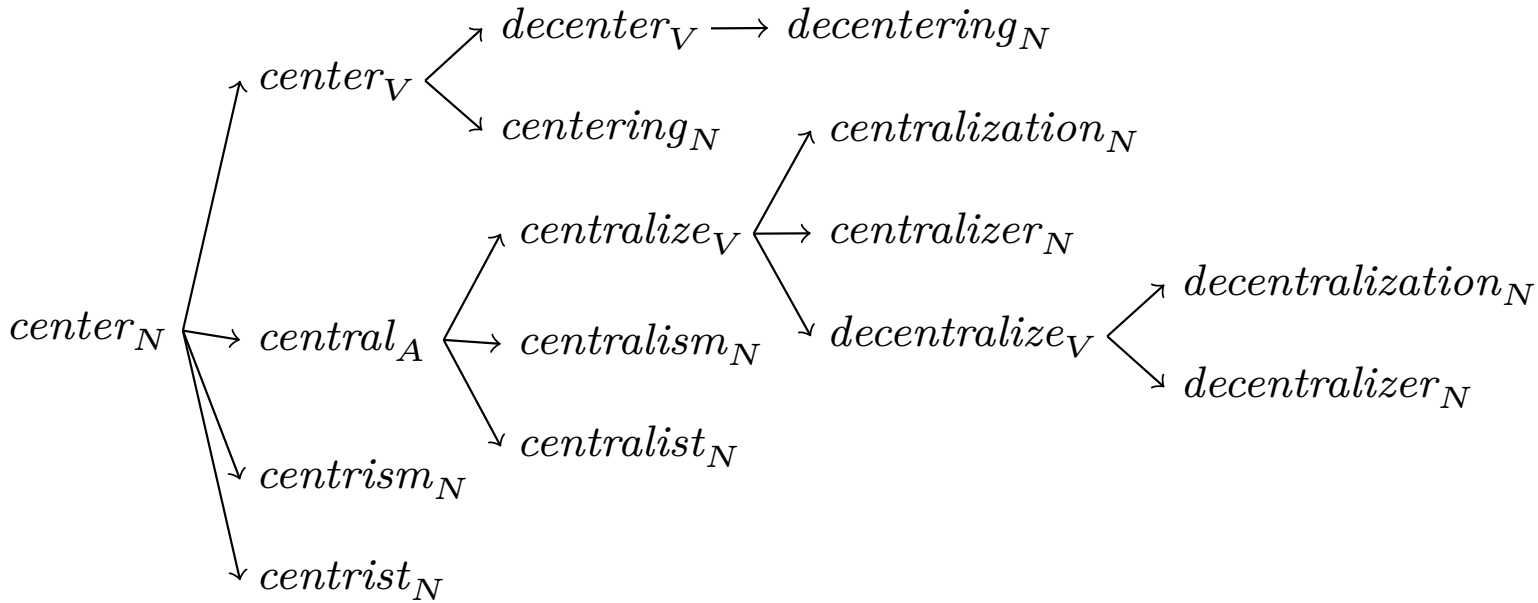
Olivier Bonami

Université Paris Cité, CNRS, Laboratoire de linguistique formelle

Initially presented at: Fred Jelinek Seminar, Prague — February 27, 2023

Derivational families as rooted trees

- Common view of the structure of word formation: every lexeme is either simplex or derives from unique a base.
- I.e., derivational families are structured as rooted trees base.



- Any similarity among items that do not stand in a (base, derivative) relation has to follow from their sharing a common base.
  - View held across many approaches to word formation, e.g. those stemming from Aronoff (1976).
  - Central to morphological resource development efforts such as Derinet (Ševčíková & Žabokrtský, 2014; Vidra et al., 2019) and Universal Derivations (Kyjánek et al., 2020).

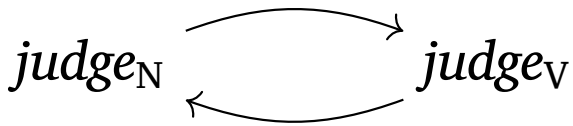
Limitations of rooted trees I

- Many well-identified situations where the rooted-tree approach fails to capture important morphological insights:

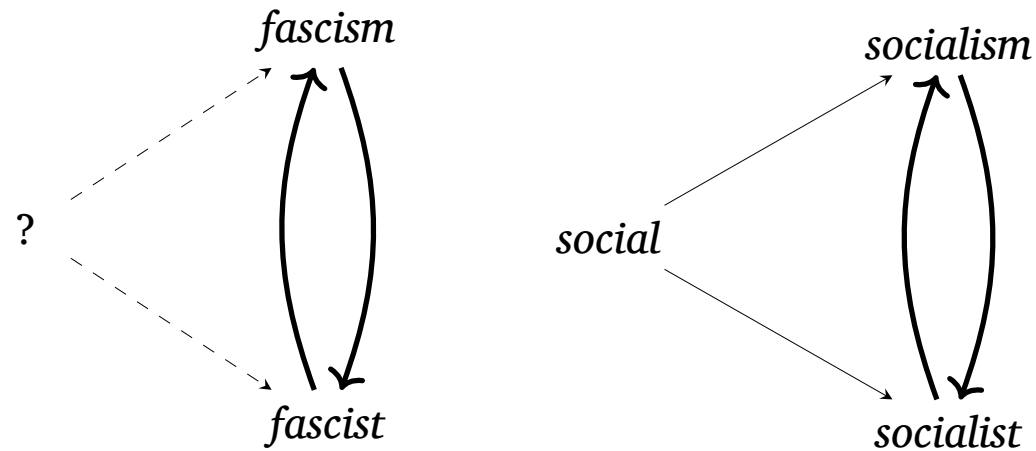
  Back-formation What looks like an output arises first and motivates what looks like the corresponding input

  bartend ⟵ *bartender*

  Conversion The orientation of conversion pairs can be undecidable (Marchand, 1963; Tribout, 2020):

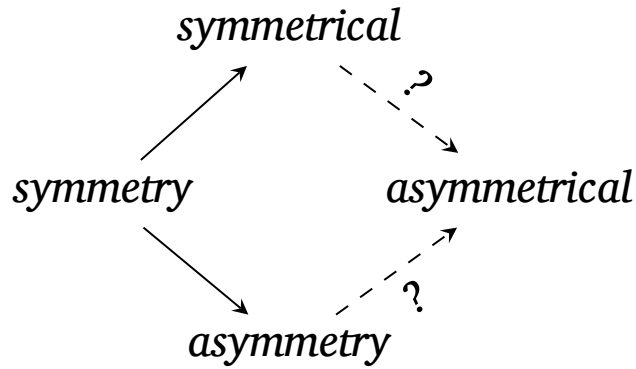
  

  Cross-formation Morphological relation between two complex items, despite absent or poorly motivated base (Becker, 1993).

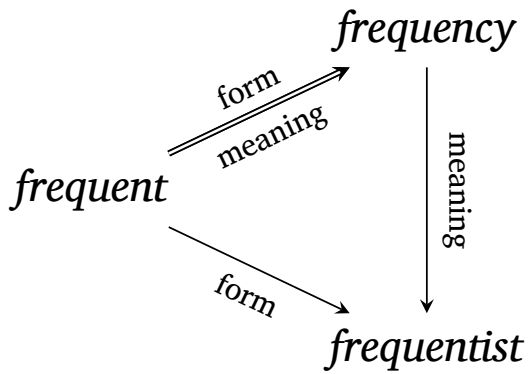
  

Limitations of rooted trees II

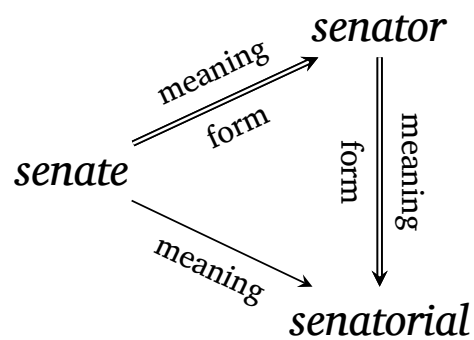
Multimotivation A derived item is equally well motivated by two derivation paths, leading one to be unwilling to choose one over the other (Corbin, 1976).



Form-content mismatches The formal base of a derived item seems different from its semantic base (Hathout & Namer, 2014b).



Family support Multiple members of its family contribute to shaping the semantics of an item (Strnadová, 2014).



Typical reactions

- Three types of reactions by theoretical morphologists:
  - Dismissal: such phenomena are too rare to receive a place in the architecture of morphology.
    - Rooted trees are enough
  - Reform: rooted trees should be supplemented by higher order paradigmatic relations where relevant.
    - secondary analogical coinings (Marle, 1984)
    - second-order schemas in Construction Morphology (Booij, 2010)
    - sister schemas in Relational Morphology (Jackendoff & Audring, 2020)
  - Revolution: such phenomena warrant rethinking the architecture of morphology, using derivational paradigms.
    - Robins (1959), Becker (1993), Bochner (1993), Štekauer (2014), and Bonami & Strnadová (2019), etc.
- The appeal of each option depends on how central and systemic the phenomena under examination turn out to be.

Goals of the talk

- Provide empirical arguments for the systemic role of paradigmatic relations.
- Two steps:
  1. Clarify what we mean by paradigm structure (Bonami & Strnadová, 2019)
  2. Present a series of empirical arguments for the role of paradigm structure in derivation.
     - 2.1 Form predictability: derivational families exhibit omnidirectional form predictability, just as inflectional paradigms do.
     - 2.2 Behavioral confirmations: speakers exhibit awareness of omnidirectional form predictabililty, both in inflectional paradigms and in derivational families.
     - 2.3 Semantic predictability: in some parts of the system, derived items are strongly predictive of each other, while their formal base is not.

What this talk is not about I

- Are inflection and derivation…
  - …irreductibly different components of grammar?
    (e.g. Anderson 1982, 1992; Perlmutter 1998)
  - …extreme points in a large typological space of morphological relatedness?
    (e.g. Dressler 1989; Booij 1996; Bauer 2004; Corbett 2010; Spencer 2013)
  - …inherently the same thing, the difference being "merely a way of speaking"?
    (e.g. Bochner 1993; Ford et al. 1997; Haspelmath forthcoming)
- This is an important issue, in need of better empirical study.
- Distributional methods can help.
  (Bonami & Paperno, 2018; Rosa & Žabokrtský, 2019; Copot et al., 2022)
- However this issue is orthogonal to our concerns.

What this talk is not about II

- What role does segmentation into morph(eme)s play in (derivational) morphology?
  - Decades-old debate between morpheme-based and word-based approaches.
  - Recent research in this area (e.g. Baayen et al. 2019; Bonami & Beniamine 2021 supports the view that
    - Different morphological questions support different segmentations.
    - The signalling values of subword sequences does not support full discretization: various parts of a word are partially informative of various aspects of its content.
- However, again, this is orthogonal to the issue at hand: paradigm structure may be relevant whether morphemic analysis is warranted or not.

Paradigms: a reconceptualization

Two key insights I

- We'd like to take inspiration from fruitful work in the Word and Paradigm tradition that explores "horizontal relations" within inflectional paradigms.
  - See e.g. Robins (1959), Matthews (1972), Wurzel (1984), Zwicky (1985), Stump (1993), Aronoff (1994), Baerman et al. (2005), Blevins (2006), Ackerman et al. (2009), Ackerman & Malouf (2013), Stump & Finkel (2013), Blevins (2016), Bonami & Beniamine (2016), and Sims & Parker (2016).
- Two important insights:
  1. Paradigms are structured by contrasts of content

     (Štekauer, 2014; Stump, 2016)

     - i.e., semantics and/or morphosyntax

PRS	PST
wink	*winked*
stink	*stank*
sting	*stung*
hit	*hit*

Two key insights II

2. Paradigms need not be about combinations of orthogonal features

> *"A paradigm is an n-dimensional space whose dimensions are the attributes (or features) used for the classification of word forms"*
>
> *(Wunderlich & Fabri, 1995, p. 266)*

Orderly paradigms: Italian adjectives

	SG	*PL*
SG	buono	buona
PL	buoni	buone

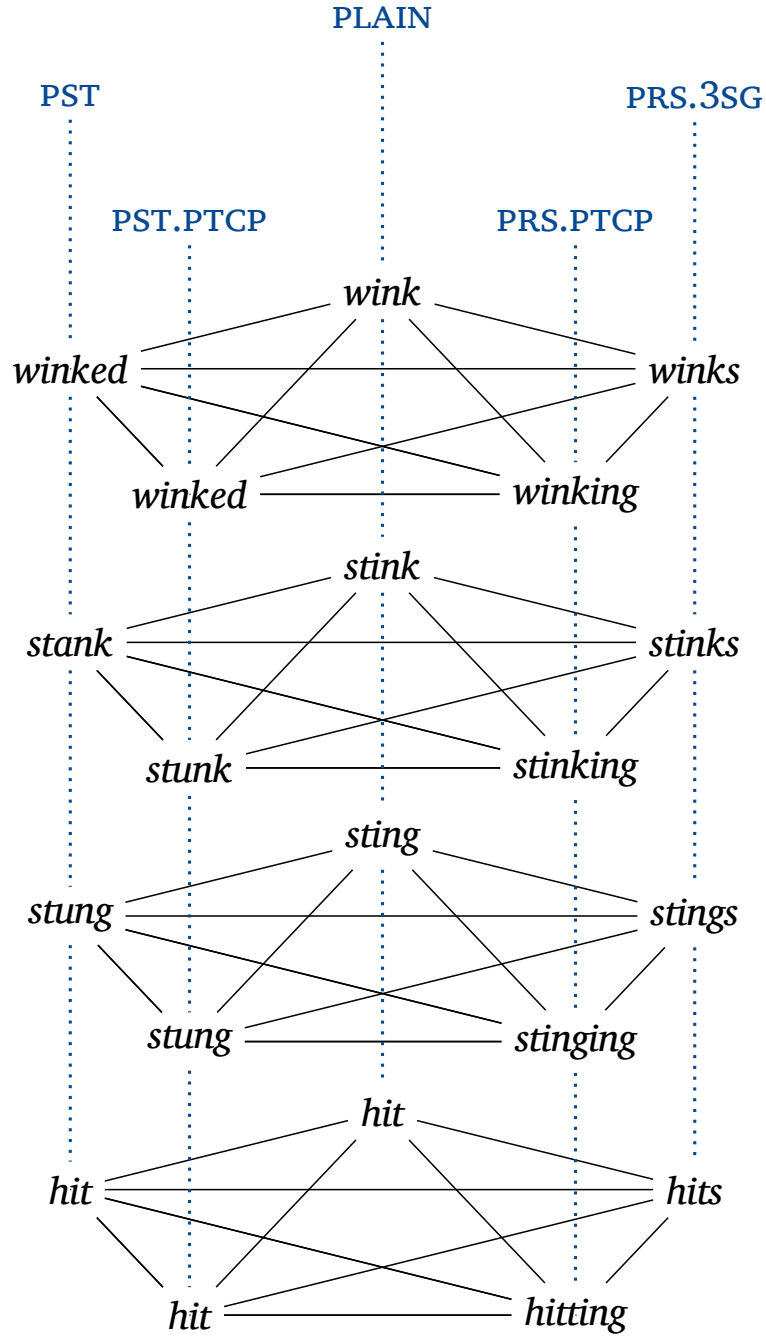
Italian BUONO *'good'.*

Disorderly paradigms: English verbs

			IND PRS	IND PST	IMP
FINITE	SG	1	*eat*	*ate*	
		2	*eat*	*ate*	*eat*
		3	*eats*	*ate*	
	PL	1	*eat*	*ate*	
		2	*eat*	*ate*	*eat*
		3	*eat*	*ate*	
NFIN	PART		*eating*	*eaten*	
	INF		*eat*		

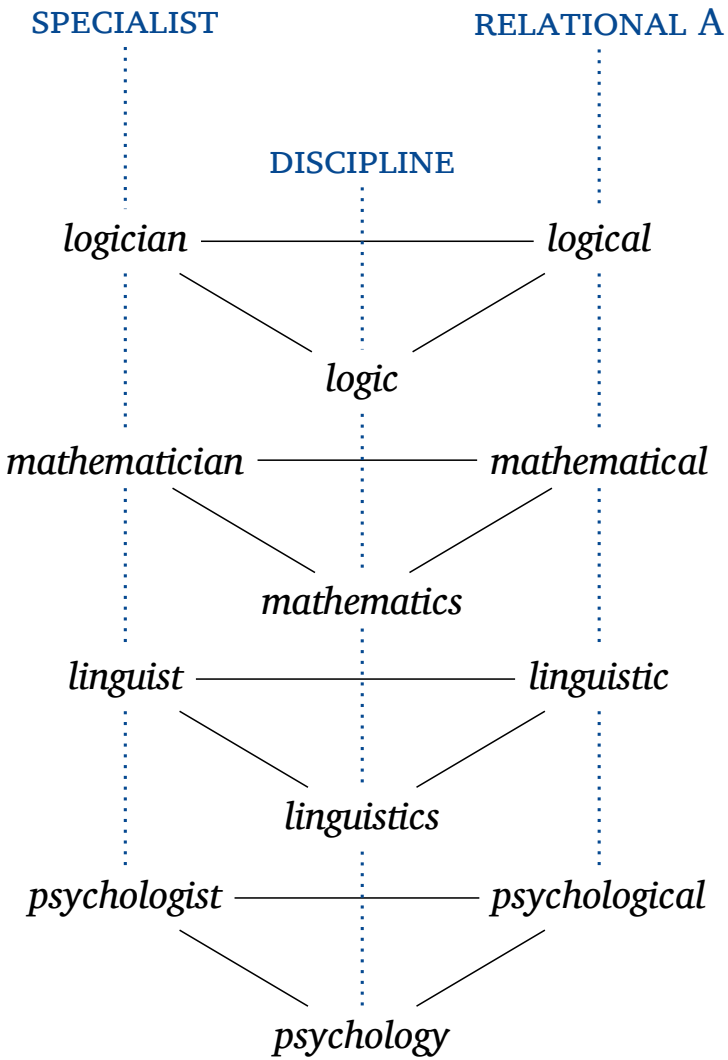
An agnostic definition (Bonami & Strnadová, 2019)

- (Partial) morphological family
  Any set of morphologically related words.
- (Partial) paradigmatic system
  Collection of morphological families exhibiting the same set of contrasts in content.
- Paradigm
  One member of a paradigmatic system.
- Cell
  Set of words that enter the same set of contrasts in their respective families.



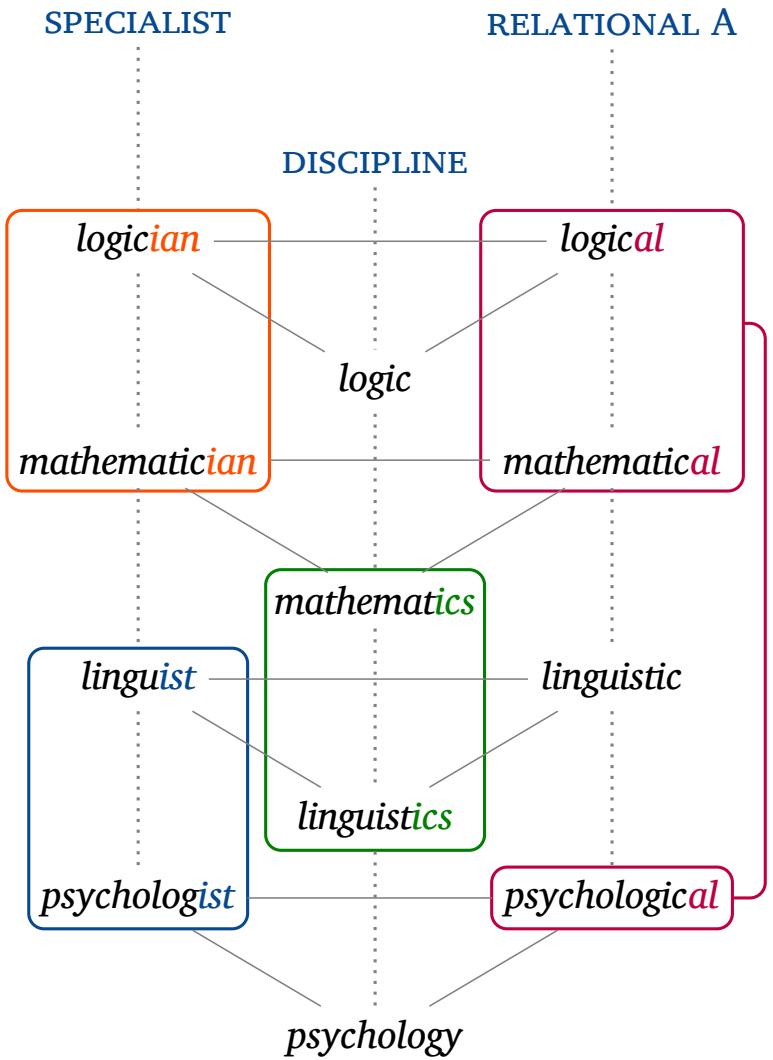
An agnostic definition (Bonami & Strnadová, 2019)

- (Partial) morphological family
  Any set of morphologically related words.
- (Partial) paradigmatic system
  Collection of morphological families exhibiting the same set of contrasts in content.
- Paradigm
  One member of a paradigmatic system.
- Cell
  Set of words that enter the same set of contrasts in their respective families.



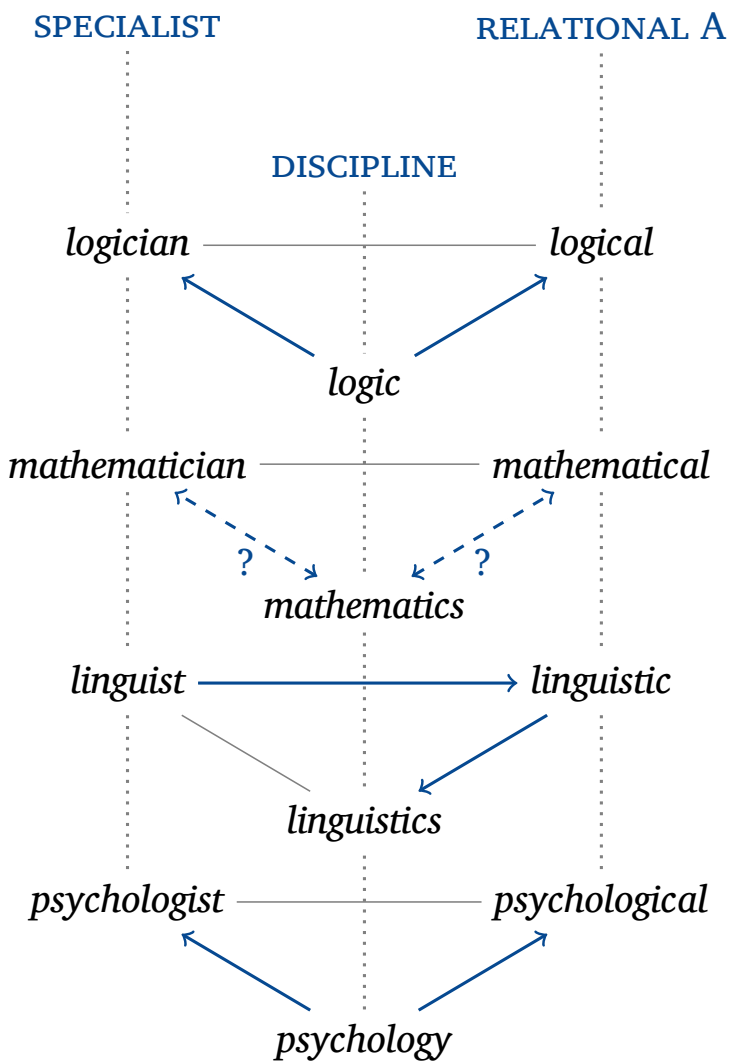
An agnostic definition (Bonami & Strnadová, 2019)

- Under these definitions:
  - Paradigm structure is agnostic to sharing of affixes.



An agnostic definition (Bonami & Strnadová, 2019)

- Under these definitions:
  - Paradigm structure is agnostic to sharing of affixes.
  - Paradigm structure is agnostic to ‘direction of derivation’.



Implicative structure in the extant lexicon

Olivier Bonami & S. Beniamine (2016). “Joint predictiveness in inflectional paradigms.” In: *Word Structure* 9.2, pp. 156–182

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

Implicative structure

- The content of paradigms is (partially) predictable: the wordform filling one cell is predictive of what wordform could fill some other cell.

PLAIN		PRS.PTCP
sing		*singing*
dance		*dancing*
wug	→	?

PLAIN		PST.PTCP
sing		*sang*
sting		*stung*
wug	→	?

- Wurzel (1989) coined the term implicative structure to describe this aspect of the structure of paradigms.

> *The inflectional paradigms are, as it were, kept together by implications. There are no paradigms […] that are not based on implications valid beyond the individual word, so that we are quite justified in saying that inflectional paradigms generally have an implicative structure, regardless of deviations in the individual cases.* Wurzel (1989, p. 114)

- Field of study emerging in the 2000s: *what* are the implications that structure paradigms? How are they organized and set up?
  - Albright (2002), Albright & Hayes (2003), and Jun & Albright (2016)
  - Finkel & Stump (2007, 2009) and Stump & Finkel (2013)
  - Ackerman et al. (2009), Ackerman & Malouf (2013), Bonami & Beniamine (2016), Sims & Parker (2016), Beniamine & Guzmán Naranjo (2021), Pellegrini (2021), and Wilmoth & Mansfield (2021),…

A simple information-theoretic measure

- Bonami & Beniamine (2016), building on Ackerman et al. (2009)

M.SG		M.PL		Alternation	Type frequency
Shape	Example	Shape	Example		
*X*al	/legal/	*X*aux	/lego/	*X*al ∼ *X*o	457
*X*al	/banal/	*X*al	/banal/	*X* ∼ *X*	39
X ≠ *Y*al	/vjø/	*X*	/vjø/	*X* ∼ *X*	10756

Shape alternations in French adjectives

- To assess predictability from cell c_1 to cell c_2:
  1. Identify relevant alternations between pairs of forms

     (This is the hard part, for reasons that I will not get into)
  2. Use this to evaluate the conditional probability of the shape in c_2 given the shape in c_1, e.g.:

$$\begin{aligned} P(\text{M.PL} = X\text{o} \mid \text{M.SG} = X\text{al}) &= P(\text{M.SG} \sim \text{M.PL} = X\text{al} \sim X\text{o} \mid \text{M.SG} = X\text{al}) \\ &= \tfrac{457}{457+39} \approx 0.92 \end{aligned}$$

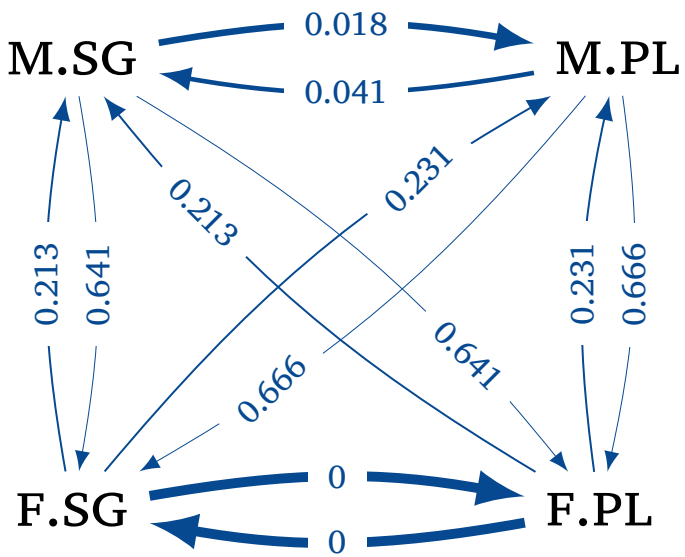
  3. Use conditional entropy as a useful summary of the distribution.

Results on inflection

- High variability in the predictability of one form from another, for a given pair of cells.

Lexeme	M.SG	M.PL	Cond. Prob.
LÉGAL 'legal'	legal	lego	0.92
BANAL 'trivial'	banal	banal	0.08
VIEUX 'old'	vjø	vjø	1

- High variability in the average predictability across pairs of paradigm cells.



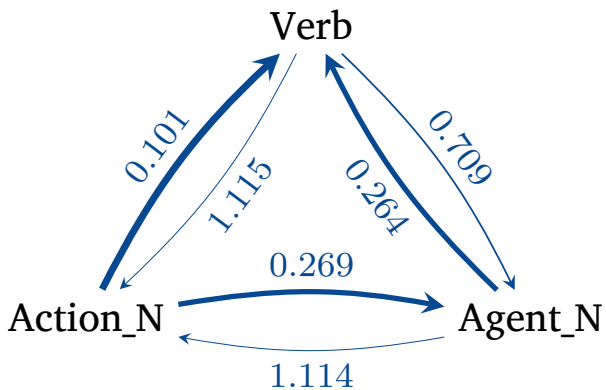
- Resulting insights on the structure of various inflection systems, and the typology of inflection.

Differential predictability in derivation

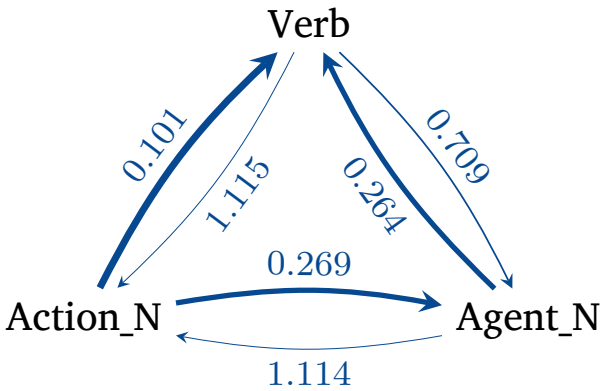
- We apply the same method to a dataset of 913 triples ⟨Verb, Action noun, Masculine agent noun⟩ from French.
  - Derivational relations from the *Démonette* database (Hathout & Namer, 2014a), phonemic transcriptions from the *GLÀFF* lexicon (Hathout et al., 2014).

Family	Verb	Action noun	Agent noun
abaisser ‘lower’	a.bɛ.se	a.bɛs.mɑ̃	a.be.sœʁ
abandonner ‘abandon’	a.bɑ̃.dɔ.ne	a.bɑ̃.dɔ̃	a.bɑ̃.dɔ.nœʁ
…	…	…	…

- Results:



Differential predictability in derivation

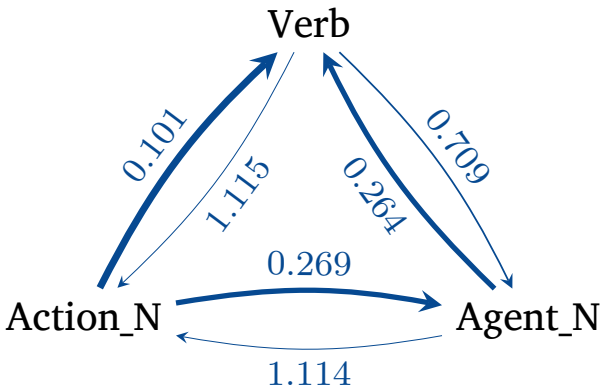


Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
couvrir	couverture	couvreur
'write'	'writing'	'writer'
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 predictability in derivation

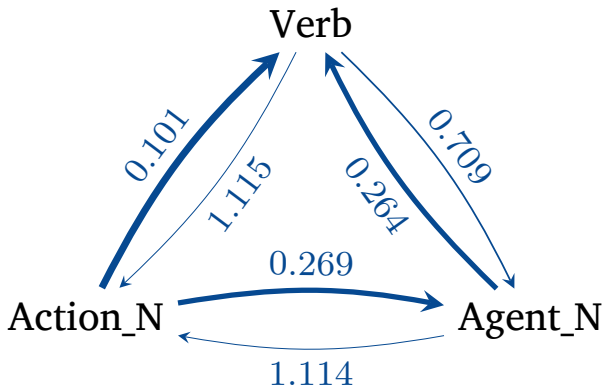


Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
couvrir	couverture	couvreur
'write'	'writing'	'writer'
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 predictability in derivation

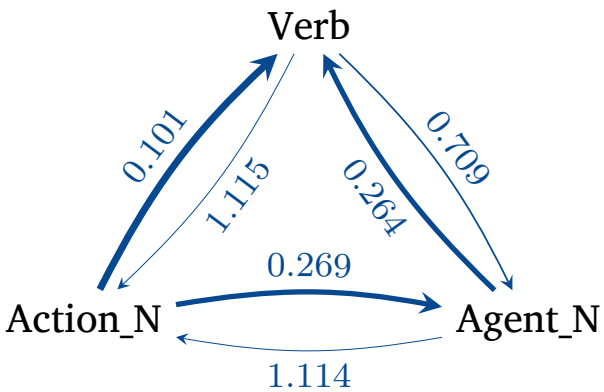


Verb	Action_N	Agent_N
laver	lavage	laveur
'wash'	'washing'	'washer'
contrôler	contrôle	contrôleur
'control'	'control'	'controller'
corriger	correction	correcteur
'correct'	'correction'	'corrector'
former	formation	formateur
'train'	'training'	'trainer'
couvrir	couverture	couvreur
'write'	'writing'	'writer'
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 predictability in derivation



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'
couvrir	couverture	couvreur
'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 the *-at-* augment should be used.

Behavioral evidence for implicative structure

Maria Copot & Olivier Bonami (accepted). “Behavioural evidence for implicative paradigmatic relations.” In: *The Mental Lexicon*

Maria Copot & Olivier Bonami (submitted). “Baseless derivation: the behavioural reality of derivational paradigms.”

Are speakers aware of paradigmatic predictability?

- The previous section has shown the existence of nontrivial implicative structure in both inflection and derivation.
  - Predictability is variable across (predictor cell, target cell) pairs.
  - For a given (predictor, target) pair, predictability is variable across lexemes.
- I now proceed to present evidence that speakers are actually aware of these predictability differentials.

The inflection experiment I

- Participants are shown a video of an utterance containing two forms of a pseudolexeme.

Nous édrilons le quiz de culture générale presque tous les ans. C'est Pierre qui l'a édrili l'anné dernière.

We PRS.1PL the pop culture quiz almost every year. It's Pierre who has PST.PTCP.M.SG it this year

- They are then asked to judge how good the second form sounds.

Est-ce que le deuxième mot sonne bien en tant que mot inventé dans ce contexte ?

Does the second word sound good in this context?

Sonne mal — Sonne bien

The inflection experiment II

- We manipulate:
  1. The pair of cells under examination:

Predictor → Target
INFINITIVE → IMPERFECT INDICATIVE 2PL
IMPERFECT INDICATIVE 2PL → INFINITIVE
PAST PARTICIPLE MASC. SING. → PRESENT INDICATIVE 1PL
PRESENT INDICATIVE 1PL → PAST PARTICIPLE MASC. SING.

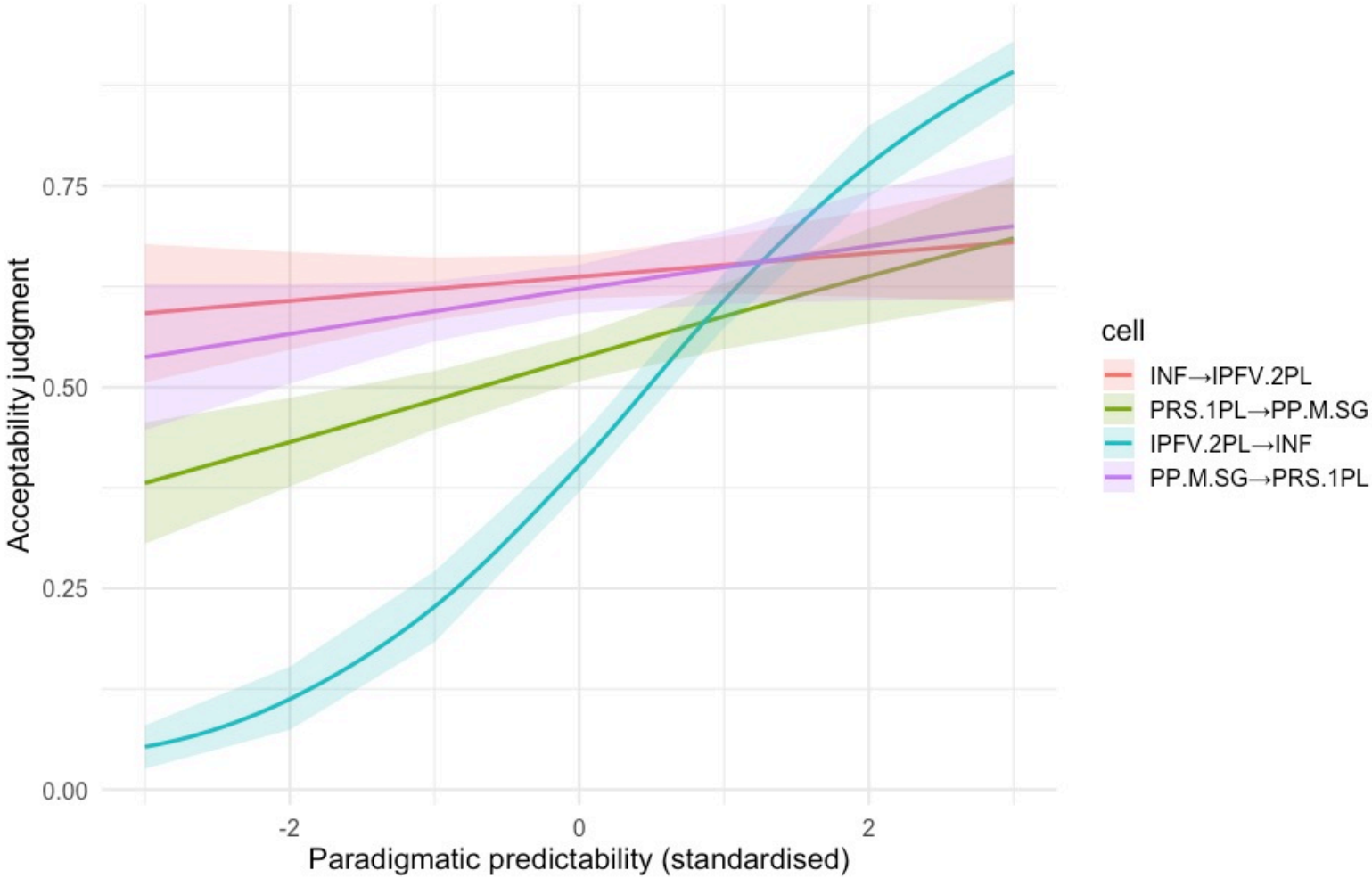
  2. The predictability of the alternation:

Nous édrilons le quiz de culture générale presque tous les ans.

C'est Pierre qui l'a { *édrilé* ■■■■■ / *édrili* ■■ / *édrilu* ■ } l'année dernière.

We IND.PRS.1PL *the general culture quiz almost every year. Pierre has* { PST.PTCP.M.SG-1 / PST.PTCP.M.SG-2 / PST.PTCP.M.SG-3 } *it last year.*

The inflection experiment: results



- Manifest awareness of paradigmatic predictability across pairs of cells.
- No privileged status for the citation form.
- In fact, speakers are most sensitive to predictability scores when predicting the infinitive, the opposite of what we would expect.

Modeling details

- Maximal GLMM with random intercepts for item and participant fitting a beta distribution.

```
judgment ~ predictability * cell + wellformedness +
           (predictability * cell + wellformedness|participant) +
           (predictability|item)
```

- Contrast coding of (predictor,target) pairs

	2	3	4
INF→IPFV.2PL	-0.25	-0.25	-0.25
PRS.1PL→PST.PTCP	0.75	-0.25	-0.25
IPFV.2PL→INF	-0.25	0.75	-0.25
PTS.PTCP→PRS.1PL	-0.25	-0.25	0.75

- Model coefficients:

	Estimate	*Std.Error*	$z-value$	$p-value$	
(Intercept)	0.21	0.09	2.33	0.02	*
Predictability	0.31	0.06	4.83	$<.001$	***
Condition2	-0.42	0.13	-3.33	0.001	***
Condition3	-0.96	0.15	-6.48	$<.001$	***
Condition4	-0.06	0.13	-0.49	0.623	
Well-formedness	0.21	0.11	1.97	0.049	*
Predictability:Condition2	0.15	0.15	0.97	0.332	
Predictability:Condition3	0.77	0.19	4.13	$<.001$	***
Predictability:Condition4	0.05	0.17	0.32	0.751	

The derivation experiment I

- Our goal is now to show that speakers exhibit the same kind of awareness of paradigmatic predictability within derivational families.
- Participants are shown a video of an utterance containing two pseudolexemes that are derivationally related.

> Un rancibateur est un professionnel du rancibatage du linge.
>
> *An AGENT_NOUN is a professional of laundry ACTION_NOUN*

- They are then asked to judge how good the second form sounds.

> Est-ce que le deuxième mot sonne bien en tant que mot inventé dans ce contexte ?
>
> *Does the second word sound good in this context?*
>
> 
>

The derivation experiment II

- We manipulate:
  1. The pair of cells under examination:

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

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

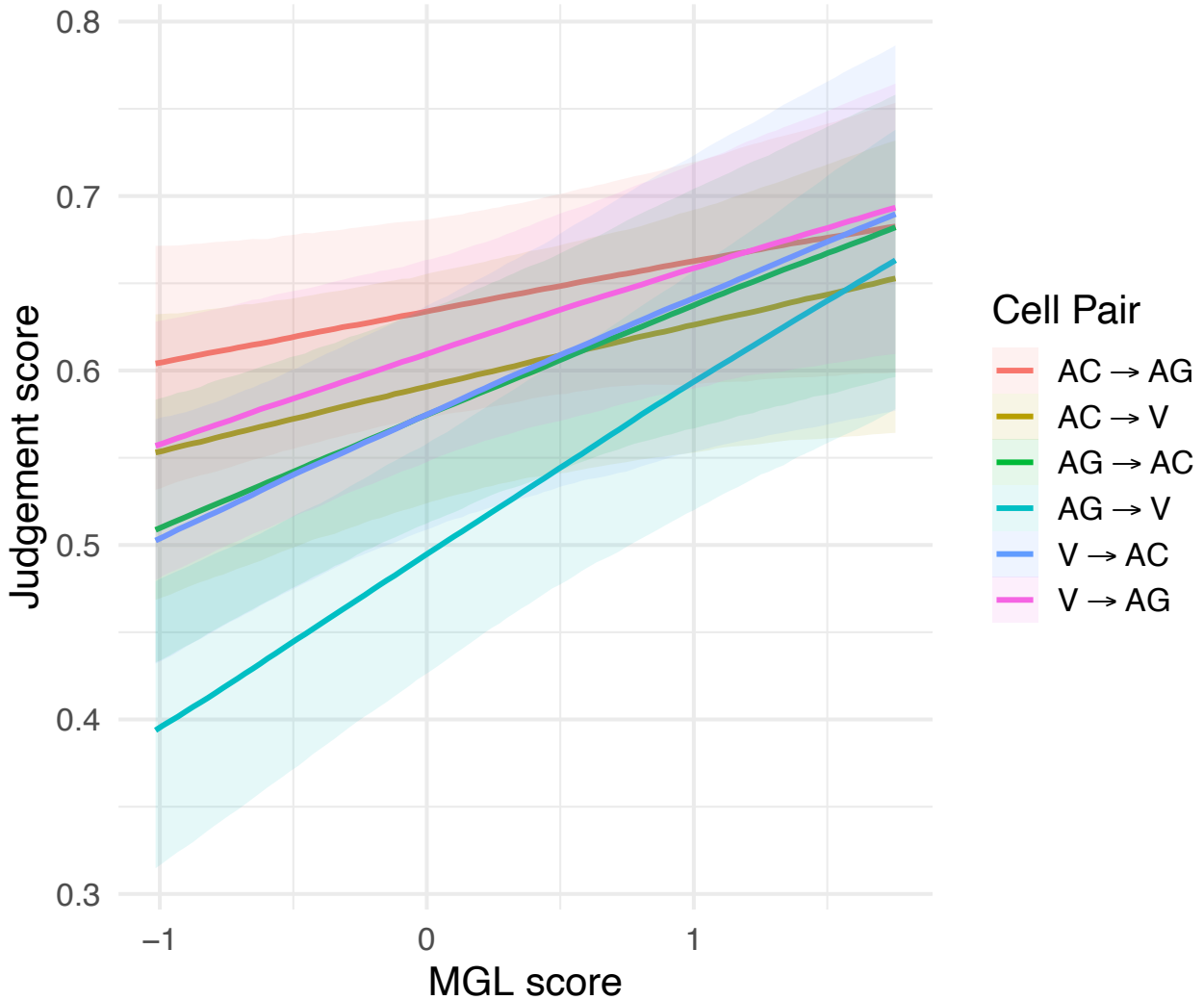
  2. The predictability of the alternation:

Un rancibateur est un professionnel { de la rancibation  / du rancibatage  / du rancibage 

An AGENT_NOUN is a professional of { *ACTION_NOUN-1* / *ACTION_NOUN-2* / *ACTION_NOUN-3*

The derivation experiment: results

- Clear effect of predictability in all 6 conditions.
- Prediction from the "base" (the verb) has no privileged status.
- Hence speakers are aware of paradigmatic predictability between all cells in the paradigm, in both directions.



Modeling details I

- Mixed effects zero-and-one inflated Bayesian beta regression.

```
judgment ~ predictability * cell + wellformedness +
           (predictability * cell + wellformedness|participant) +
           (predictability|item)
```

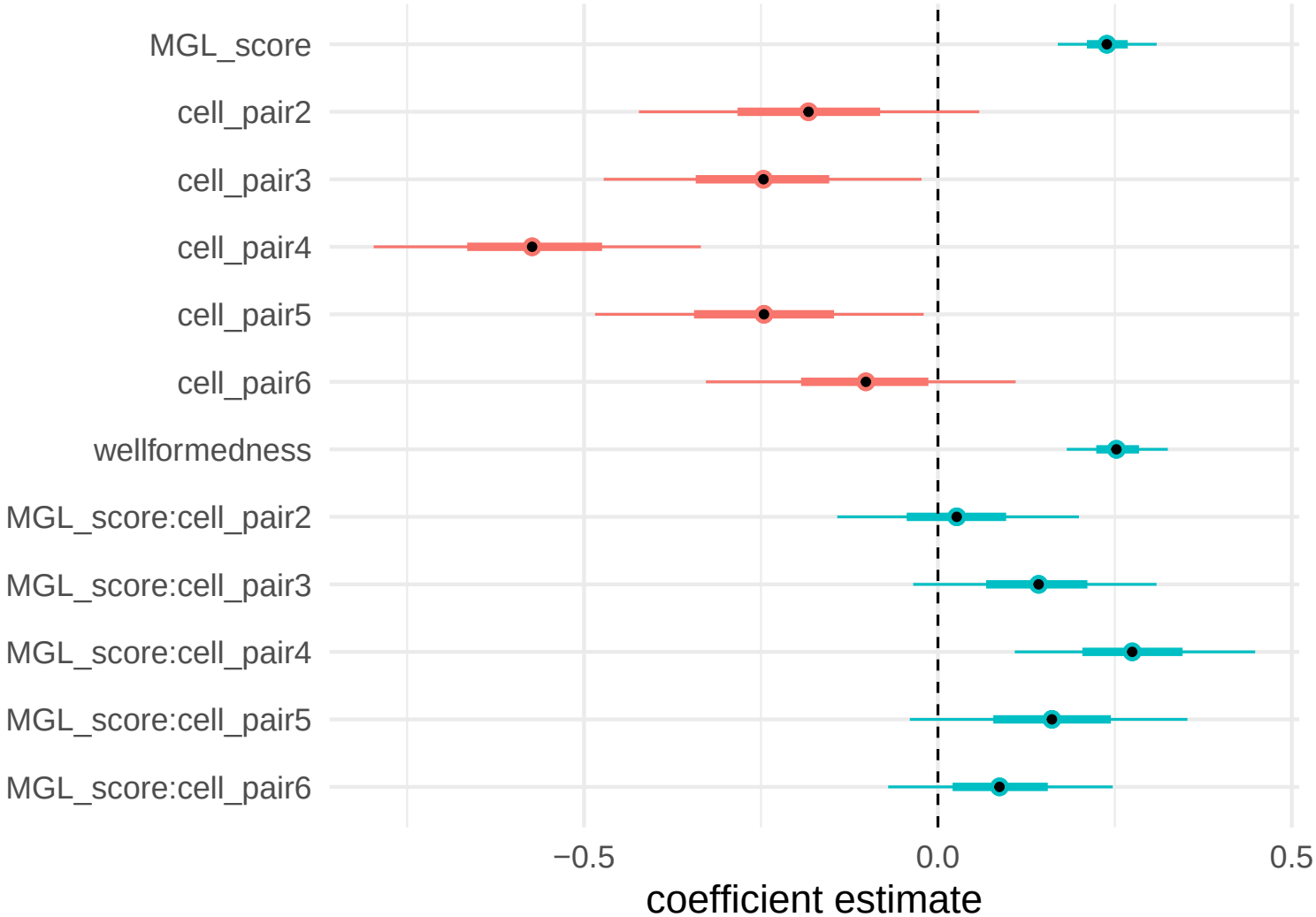
- Contrast coding of (predictor,target) pairs

	2	3	4	5	6
AC→AG	-0.17	-0.17	-0.17	-0.17	-0.17
AC→V	0.83	-0.17	-0.17	-0.17	-0.17
AG→AC	-0.17	0.83	-0.17	-0.17	-0.17
AG→V	-0.17	-0.17	0.83	-0.17	-0.17
V→AC	-0.17	-0.17	-0.17	0.83	-0.17
V→AG	-0.17	-0.17	-0.17	-0.17	0.83

Modeling details II

- Model coefficients:

Whiskers = 95% CrI



Distributional evidence for derivational paradigms

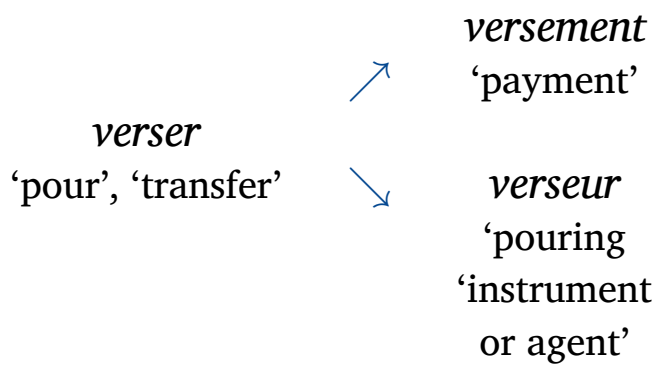
Olivier Bonami & Matías Guzman Naranjo (in press). "Distributional evidence for derivational paradigms." In: *The semantics of derivational morphology: theory, methods, evidence*. Ed. by Sven Kotowski & Ingo Plag. Berlin: De Gruyter

The plan

- Up to now, we have gathered evidence on the importance of paradigmatic relations based on form predictability
  - Within a given derivational family, knowing the form of the the lexeme meaning M_1, what is the form of the lexeme meaning M_2?
  - E.g., what is the action noun corresponding to the agent noun *directeur* 'director'?
- We now turn to the converse question:
  - Within a given derivational family, knowing the meaning of the lexeme of morphological category C_1, what is the meaning of the lexeme of morphological category C_2?
  - E.g., what does the meaning of *directeur* tell us about the meaning of *direction*?

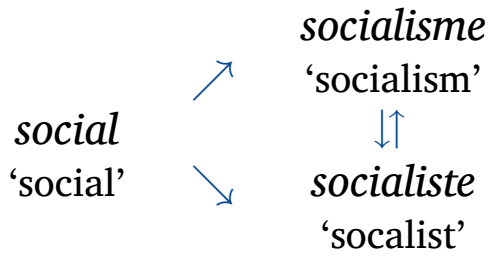
Semantic predictions I

- If derivation is organized as a rooted tree:
  - Derived forms should be typically predictable from their base
  - Derived forms are not expected to be predictable from other members of the derivational family



Semantic predictions II

- If derivation is organized in terms of paradigms, then there can be situations where this does not hold:



Witness definitions in the Oxford English Dictionary:

Socialism	A theory or system of social organization based on state or collective ownership and regulation of the means of production, distribution, and exchange for the common benefit of all members of society; advocacy or practice of such a system, esp. as a political movement. Now also: any of various systems of liberal social democracy which retain a commitment to social justice and social reform, or feature some degree of state intervention in the running of the economy.
Socialist	An advocate or supporter of socialism.

- We are trying to find cases where this holds not only for individual triples, but in a systematic fashion.

Parenthesis:
Morphology and distributional vector spaces

The distributional hypothesis

- We start from the observation that:

  Similarity of meaning results in similarity of linguistic distribution. (Boleda, 2020, p. 214)

  Hence we can hypothesize that

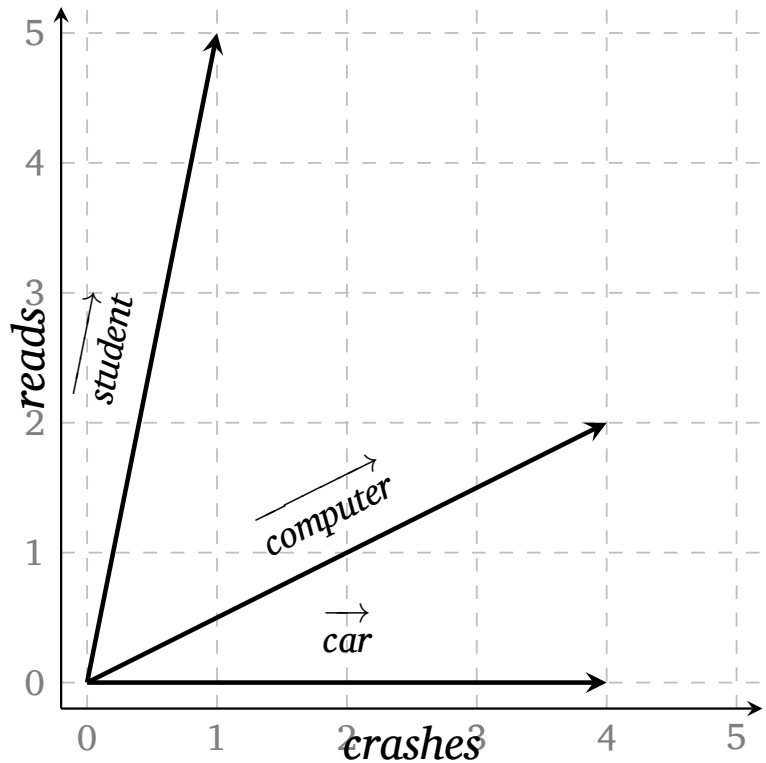
  The degree of semantic similarity between two linguistic expressions A and B is a function of the similarity of the linguistic contexts in which A and B can appear. (Lenci, 2008, p. 3)

- Notes:
  - This is an old idea (Harris, 1954; Firth, 1957)
  - Became practical with the development of word embedding technology (Mikolov et al. 2013, Pennington et al. 2014 and many others).
  - Taken litterally, the distributional hypothesis is very likely to be false, but it is still a very useful approximation.
  - For purposes of morphology it is useful to generalize it to not just semantics but also morphosyntactic content: words that share morphosyntactic features will be distributionally similar.

Distributional vector spaces in one slide

- Cooccurrence counts are vectors, angles between vectors measure similarity :

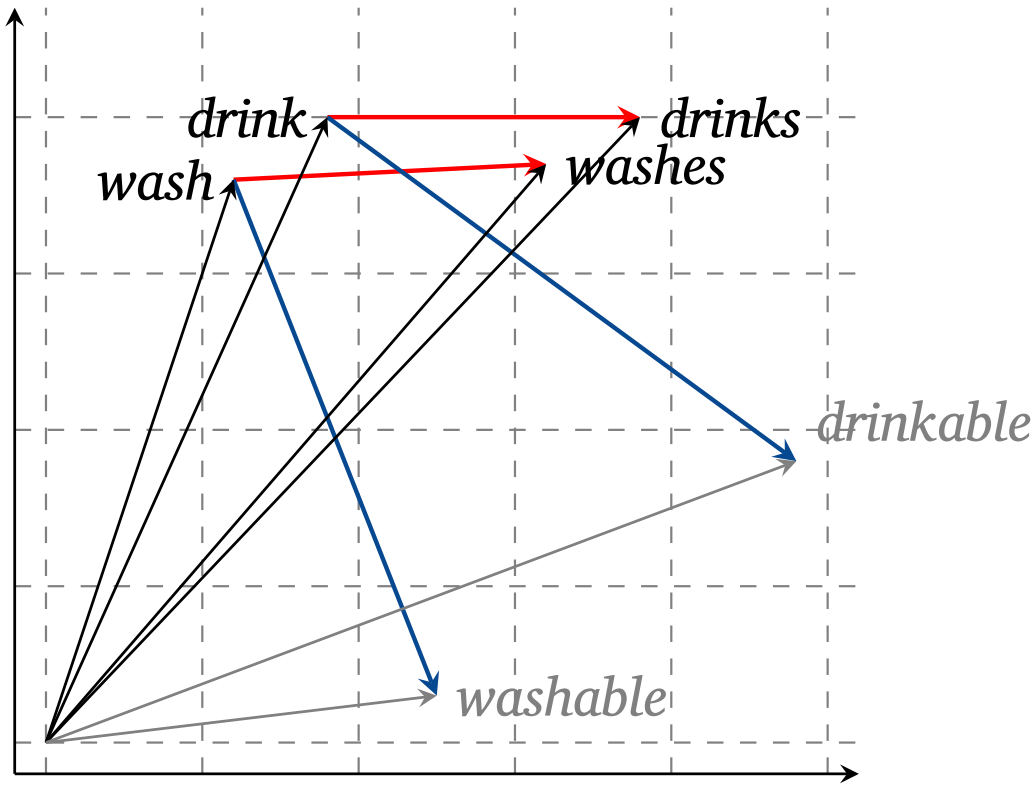
	crashes	reads
student	1	5
computer	4	2
car	4	0



- In practice:
  - Realistic representations rely on cooccurrences with very large lexica in large corpora ⇒ many more dimensions.
  - Most current systems rely on prediction tasks to infer vector representations.

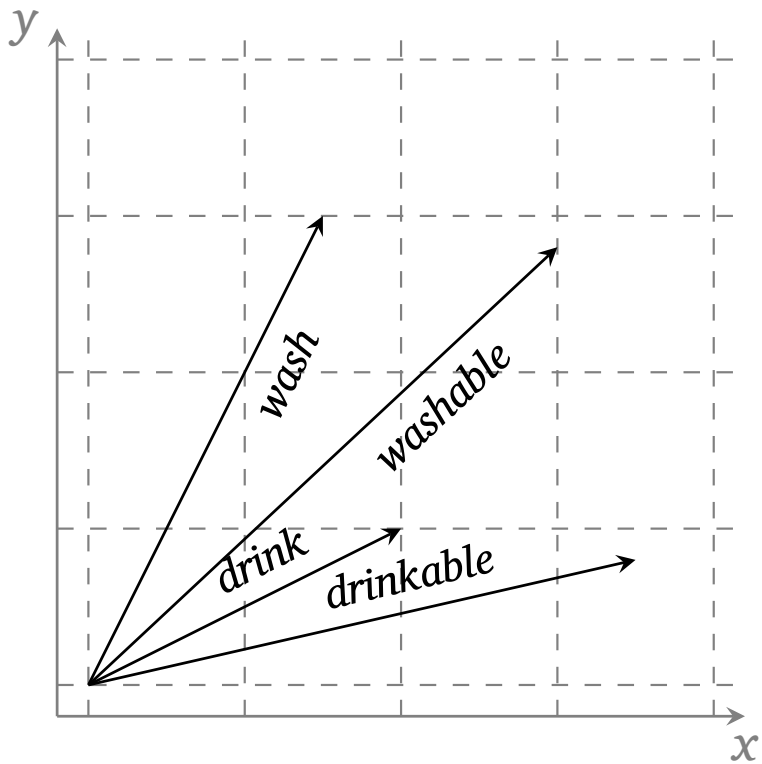
Comparing morphological relations

- Difference vectors capture the distributional relation between words
- Bonami & Paperno (2018): The variability of difference vectors reflects the semantic predictability of the relation between pairs of words.
- Empirically, we found that variability is higher for words related by derivation than for words related by inflection.



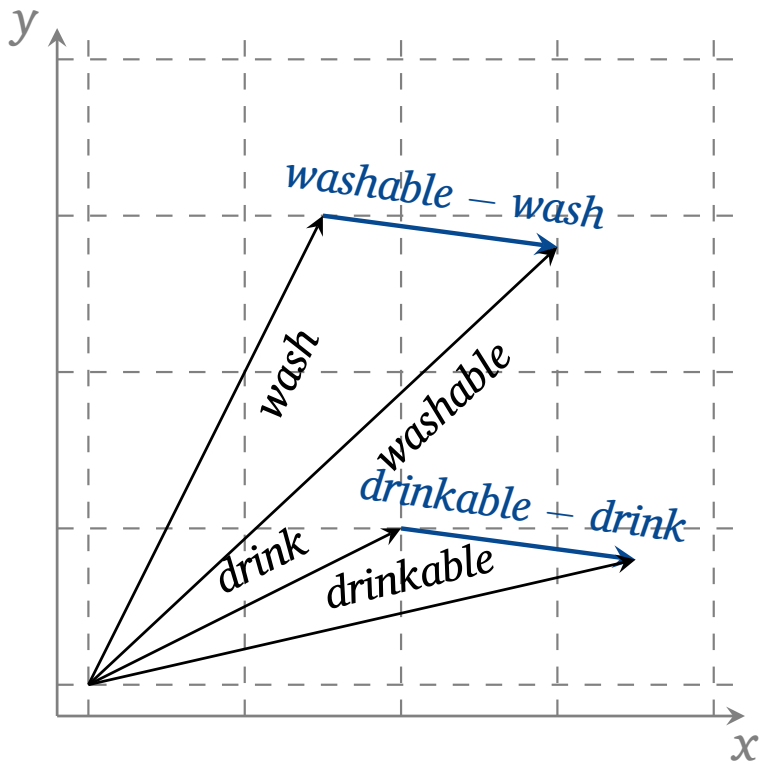
Distributional properties of morphological relations

- Distributional vectors do capture morphological properties of words: all other things being equal, the vectors for words that share some morphology are more similar that those of words that don't.



Distributional properties of morphological relations

- Distributional vectors do capture morphological properties of words: all other things being equal, the vectors for words that share some morphology are more similar that those of words that don't.



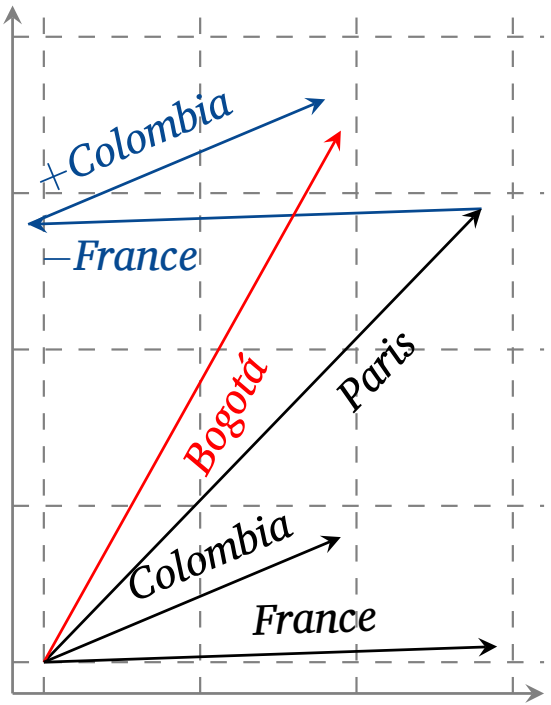
- We now want to compare how pairs of words relate to one another:

$$\frac{drink}{drinkable} \stackrel{?}{\approx} \frac{wash}{washable}$$

- This can be done by comparing difference vectors.

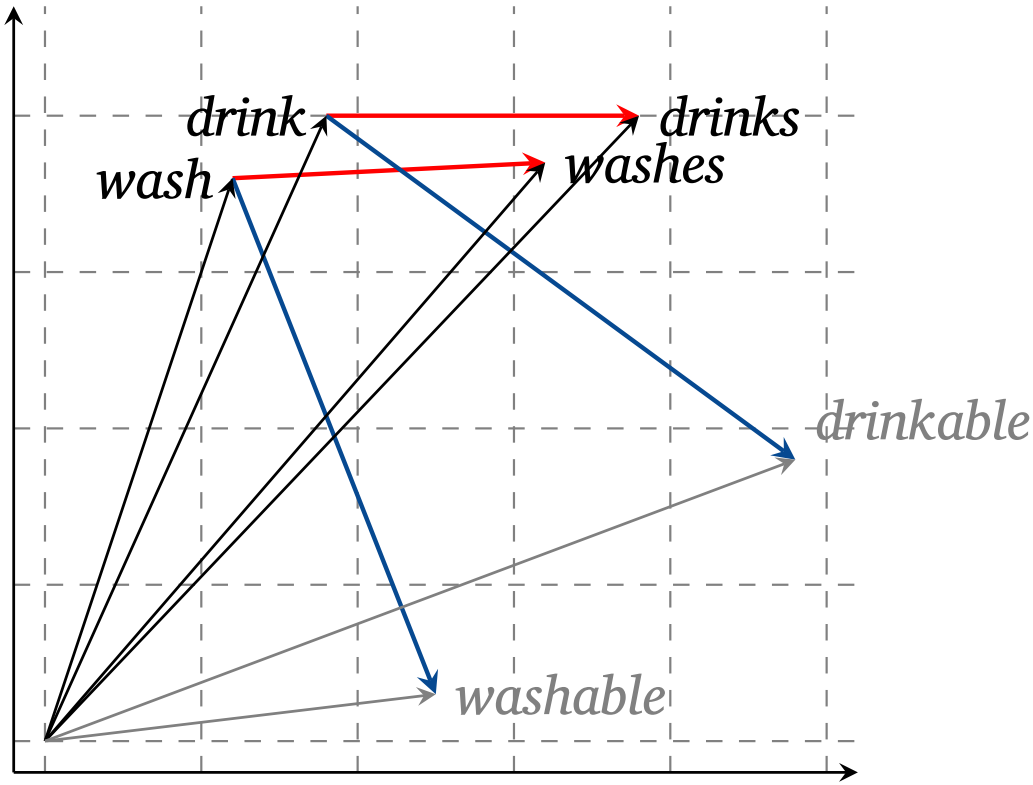
Mikolov et al. (2013)

- A good distributional vector space should solve accurately semantic analogies using difference vectors:



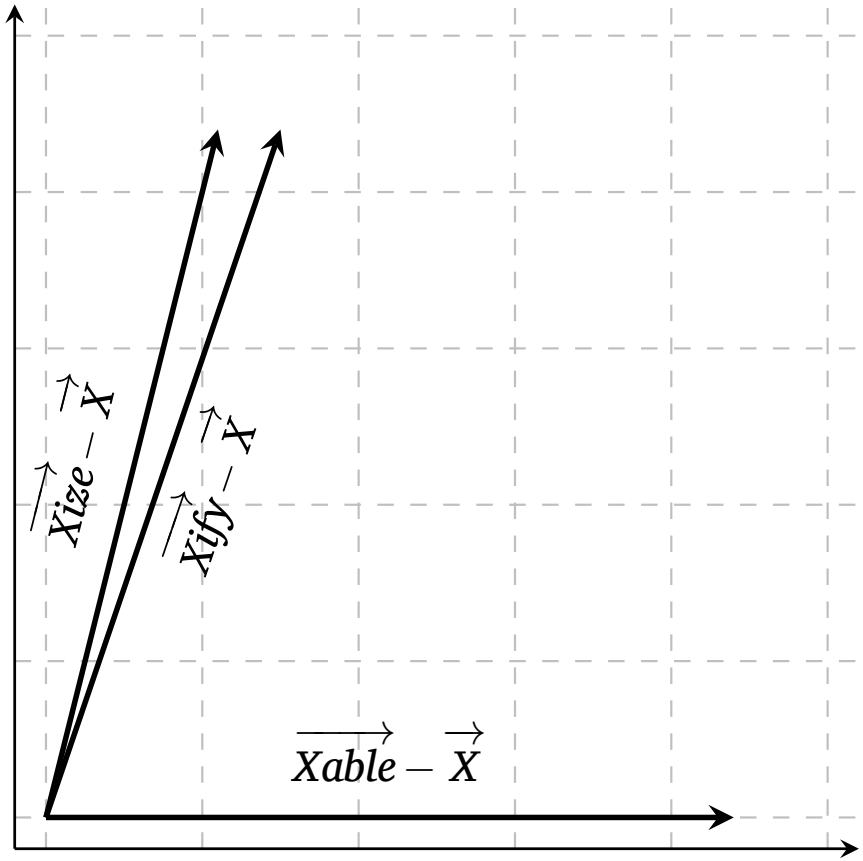
Bonami & Paperno (2018)

- The variability of difference vectors reflects the semantic predictability of the relation between pairs of words.
- Empirically, we found that variability is higher for words related by derivation than for words related by inflection.



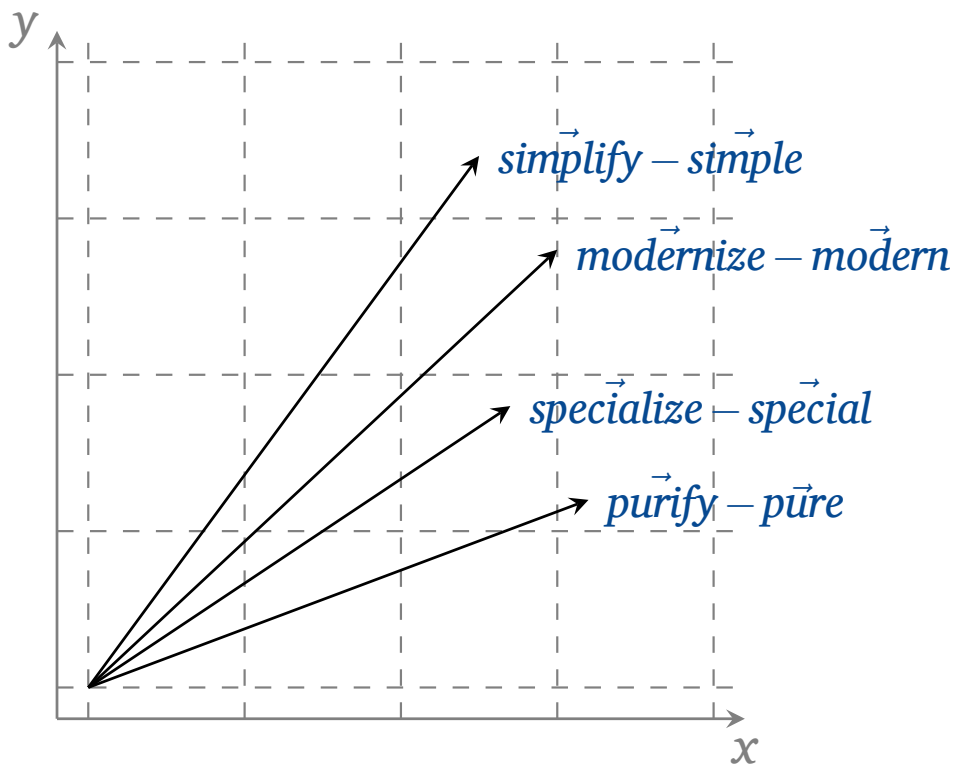
Guzman23 I

- Similarity between derivational processes can be assessed by comparing average difference vectors



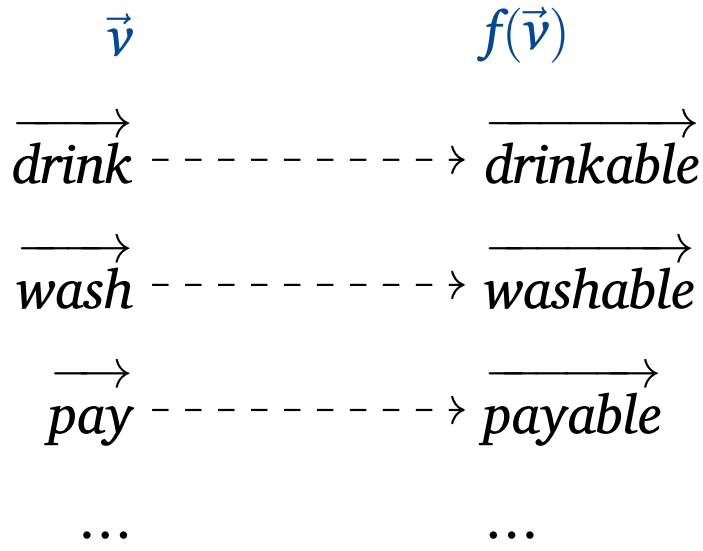
Guzman23 II

- Rival processes are those that are distributionally undistinguishable
  - I.e., a classifier cannot guess from seeing the difference vector which process led to it.



Marelli & Baroni (2015)

- The semantic import of a derivational process can be modeled as a function from base vectors to derivative vectors.



- Simplest possible approximation: $f(\vec{v}) = \vec{v} + \vec{a}$, where $\vec{a}$ is the average difference vector for all observed pairs of related words (Mickus et al., 2019).
  - i.e. we add a constant to each dimension.
- More sophisticated take (Marelli & Baroni, 2015): f is approximated by a linear transformation.
  - i.e. we fit a linear model to the prediction of each dimension in the output vectors from all the dimensions in the input vector.

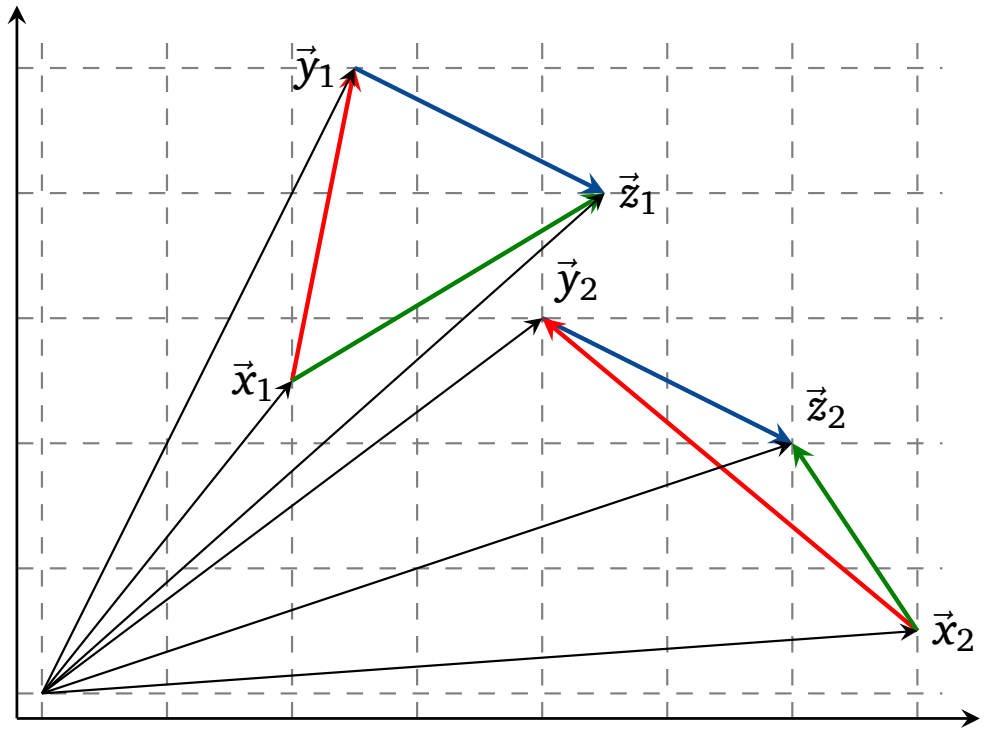
(end of parenthesis)

A distributional reformulation of our hypothesis

- We need to find matched pairs of processes:

Base	Derivative$_1$	Derivative$_2$
x_1	y_1	z_1
x_2	y_2	z_2
x_3	y_3	z_3
...	...	...

where, as a general tendency, $\vec{y}_i$ is a better predictor of $\vec{z}_i$ than $\vec{x}_i$:



The data: vector space

- We computed a vector space on the FRCOW corpus (Schäfer, 2015; Schäfer & Bildhauer, 2012) using the Gensim (Řehůřek, 2010) implementation of word2vec (Mikolov et al., 2013).

  Hyperparameters: 2 training epochs, 5 negative samples, window size 5, vector size 100.

- We need vectors for lexemes rather than wordforms.
- To this end we built a version of the corpus with:
  - Lemmas rather than wordforms.
    - e.g. *dînera* ⇝ `dîner_ver`
  - Tagged lemmas rather than bare lemmas
    - e.g. *un dîner* ⇝ `un_art dîner_nom`
  - Careful gender-neutralization
    - e.g. *du* ⇝ `de_prep le_art`

  …and used that as input for word2vec.

The data: lexicon I

- 10 datasets with at least 150 triples exemplifying two derivatives on the same base, where all words have a frequency of at least 5 in FRCOW.

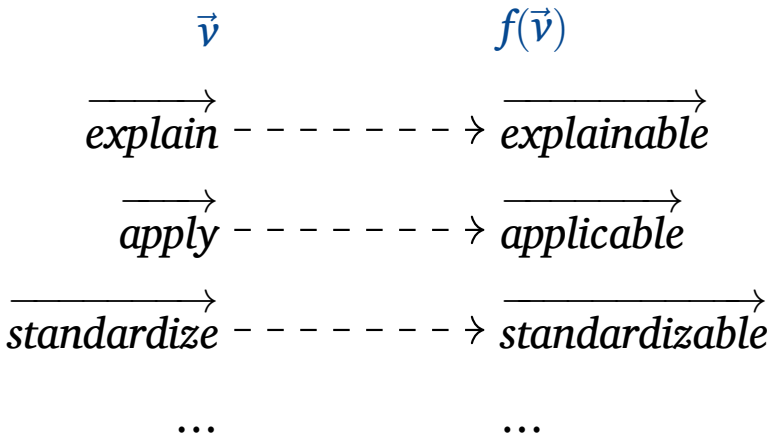
$Process_1$	$Process_2$	Sample size
age:V > N	conversion:V > N	833
age:V > N	eur:V > N	584
age:V > N	ment:V > N	354
ant:V > A	ment:V > N	302
conversion:V > N	eur:V > N	679
conversion:V > N	ment:V > N	377
ier:N > N	erie:N > N	151
eur:V > N	ion:V > N	514
eur:V > N	ment:V > N	342
isme:A/N > N	iste:A/N > N	277

(Data from *Démonette* (Hathout & Namer, 2014a) + ad-hoc data extraction)

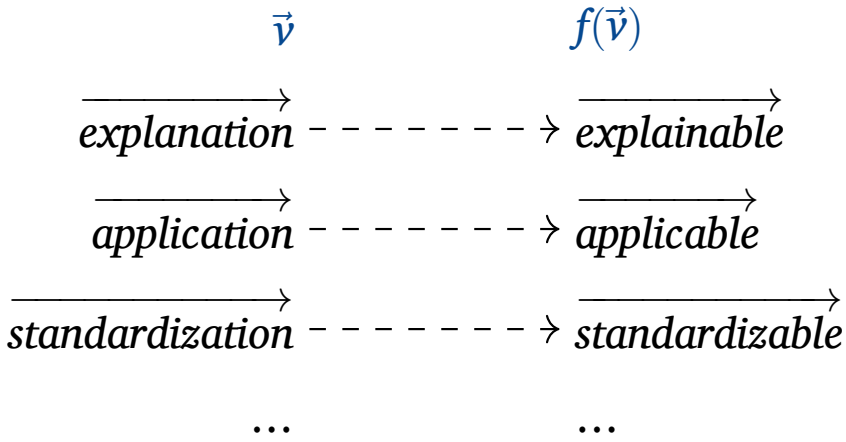
- Note that some pairs of processes (highlighted in gray) are rivals, i.e., processes that convey the same types of meanings.

The method, 1

- Crucial insight from Marelli & Baroni (2015): the semantic import of a derivational process can be modeled as a function from base vectors to derivative vectors.

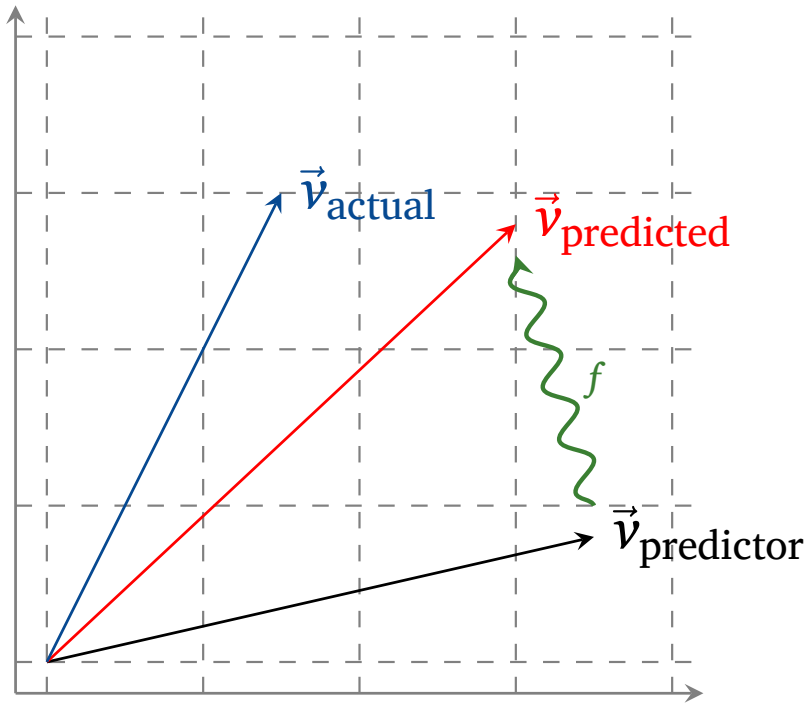


- We want to use exactly that insight but generalize it to any paradigmatic relation across morphological families.



The method, 2

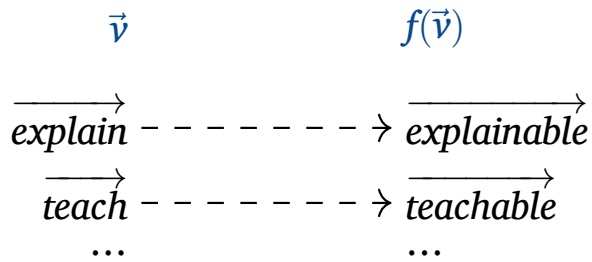
- We then measure how good the function *f* is at capturing the semantics of the morphological relation in particular cases by examining the cosine between the predicted and the actual target vector.



- The average value of $\cos(\vec{v}_{\text{predicted}}, \vec{v}_{\text{actual}})$ is indicative of how predictable the meaning of targets is from that of predictors for that particular morphological relation.

The method, 3

- There are various ways of estimating the function f from data:



- Addition of the average difference vector (Drozd et al., 2016; Mickus et al., 2019)
- Linear transformations (Marelli & Baroni, 2015)

$$\begin{aligned}
\texttt{target_val_1} &\sim \texttt{pred_val_1} + \texttt{pred_val_2} + \cdots + \texttt{pred_val_100} \\
\texttt{target_val_2} &\sim \texttt{pred_val_1} + \texttt{pred_val_2} + \cdots + \texttt{pred_val_100} \\
&\vdots \qquad\qquad\qquad\qquad\qquad \vdots \\
\texttt{target_val_100} &\sim \texttt{pred_val_1} + \texttt{pred_val_2} + \cdots + \texttt{pred_val_100}
\end{aligned}$$

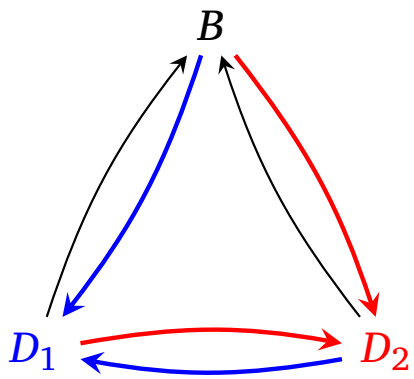
- Here we use a linear model predicting each dimension in the target vector from that dimension in the predictor vector plus 10 principal components of the whole predictor vector.

$$\texttt{target\_val} \sim \texttt{pred\_val} * \texttt{dimension} + \texttt{PC1} + \texttt{PC2} + \cdots + \texttt{PC10}$$

- We perform 10-fold crossvalidation throughout and report the aggregated performance across folds on unseen data.

The method, 4

- For each pair of processes under consideration, we can do this for each of the 6 prediction relations, and compare the averages.



	Predictor	Target
1	B	D_1
2	B	D_2
3	D_2	D_1
4	D_1	D_2
5	D_1	B
6	D_2	B

- If on average:
  - D_1 is better predicted by D_2 than by B, or
  - D_2 is better predicted by D_1 than by B

  then we have found evidence of paradigmatic organization.

The method, 5

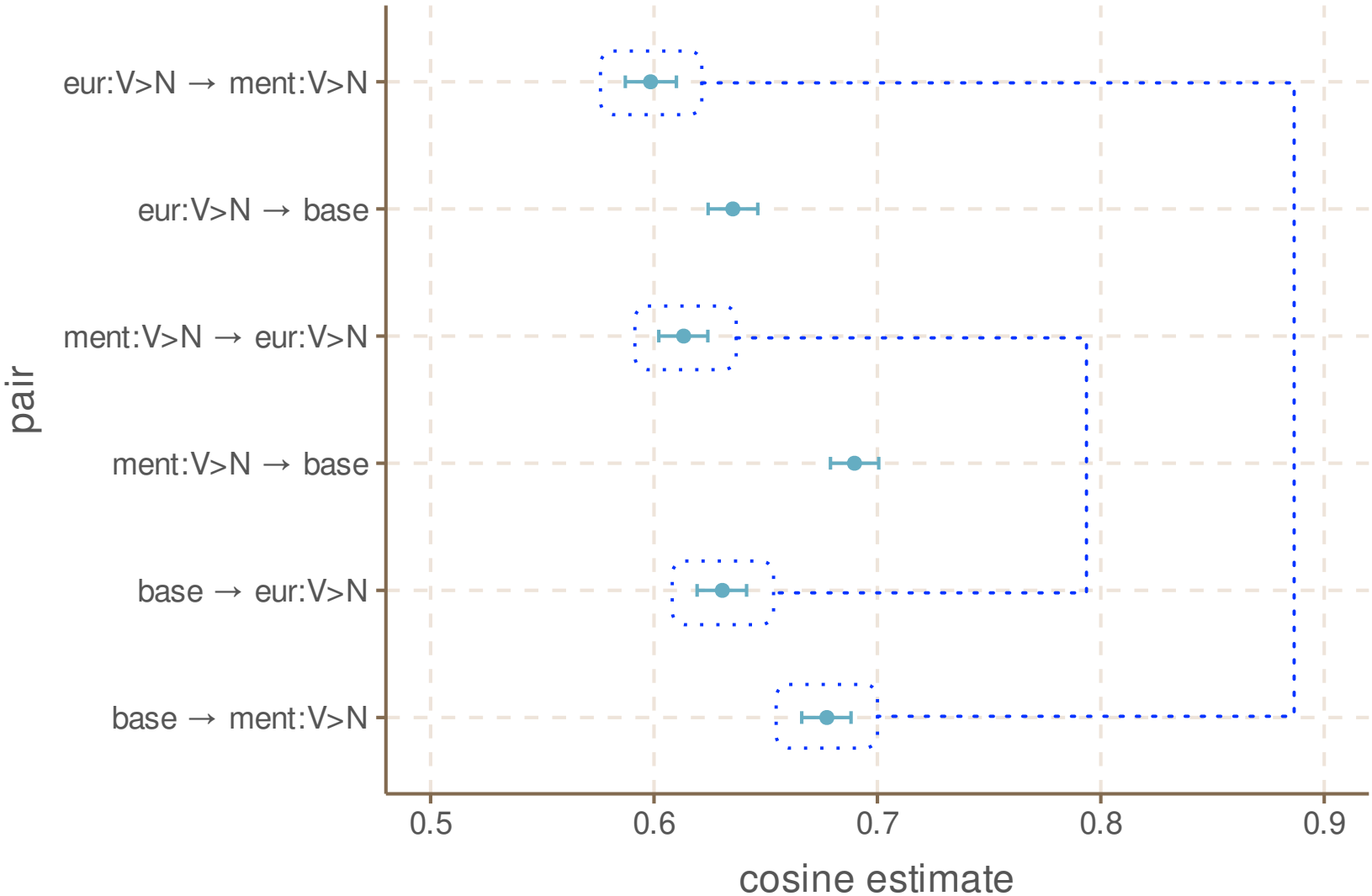
- This gives us raw results of the form:

Prediction relation	Sample predictor	Sample target	Sample performance	Average performance
`base>eur`	*accorder*	*accordeur*	0.640	0.676
`eur>base`	*accordeur*	*accorder*	0.753	0.689
`base>ment`	*accorder*	*accordement*	0.849	0.633
`ment>base`	*accordement*	*accorder*	0.869	0.637
`eur>ment`	*accordeur*	*accordement*	0.712	0.615
`ment>eur`	*accordement*	*accordeur*	0.493	0.600

- We might be tempted to conclude directly from the average performance.
- However, we are not too confident about the quality of our vectors.
- Hence we use a Bayesian Beta regression to estimate credible intervals around average performance, where each pair of words within a set of triples is a data point.

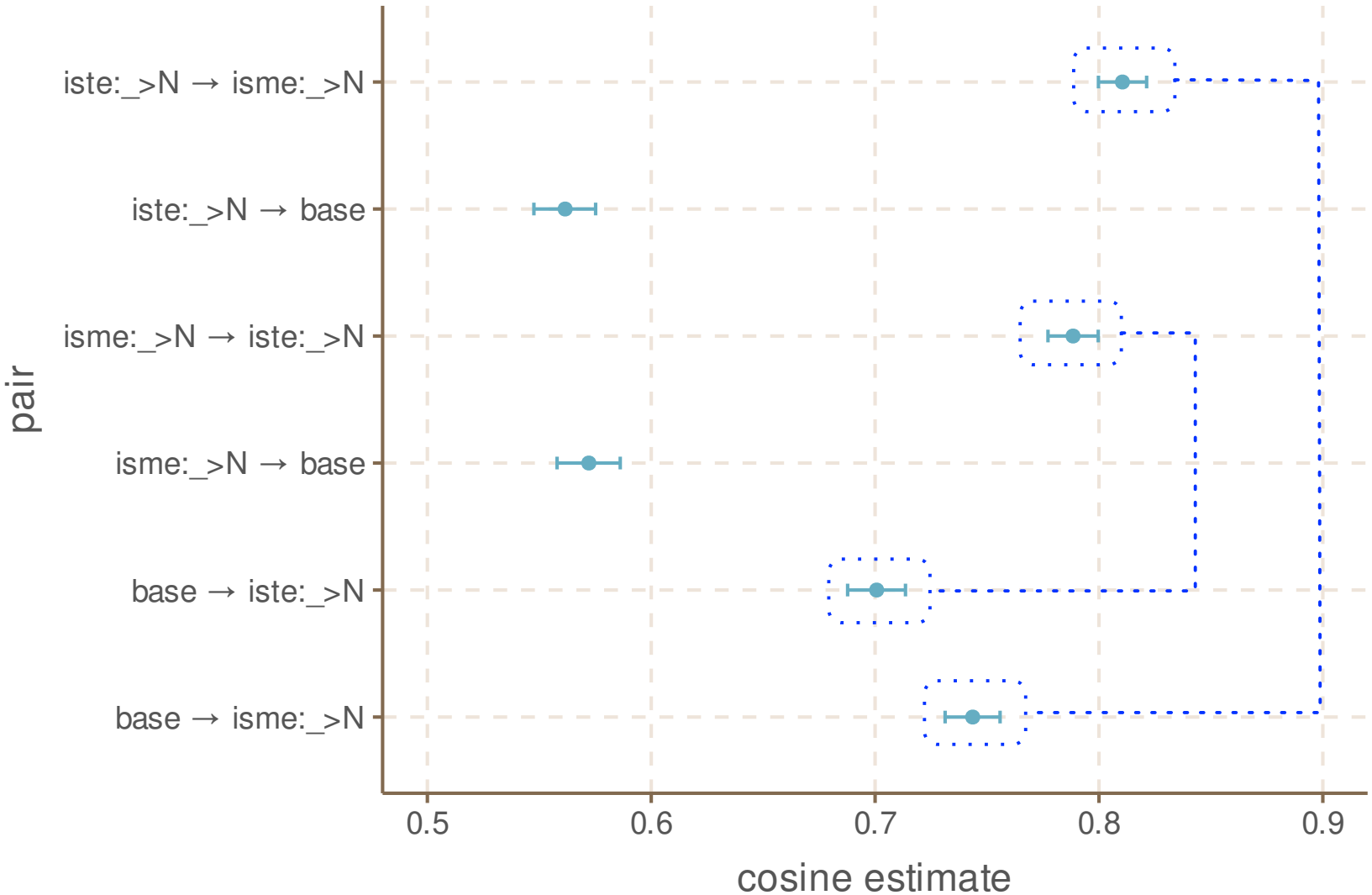
$$\texttt{Cosine similarity} \sim \texttt{Process}$$

Results: *-eur* vs. *-ment*



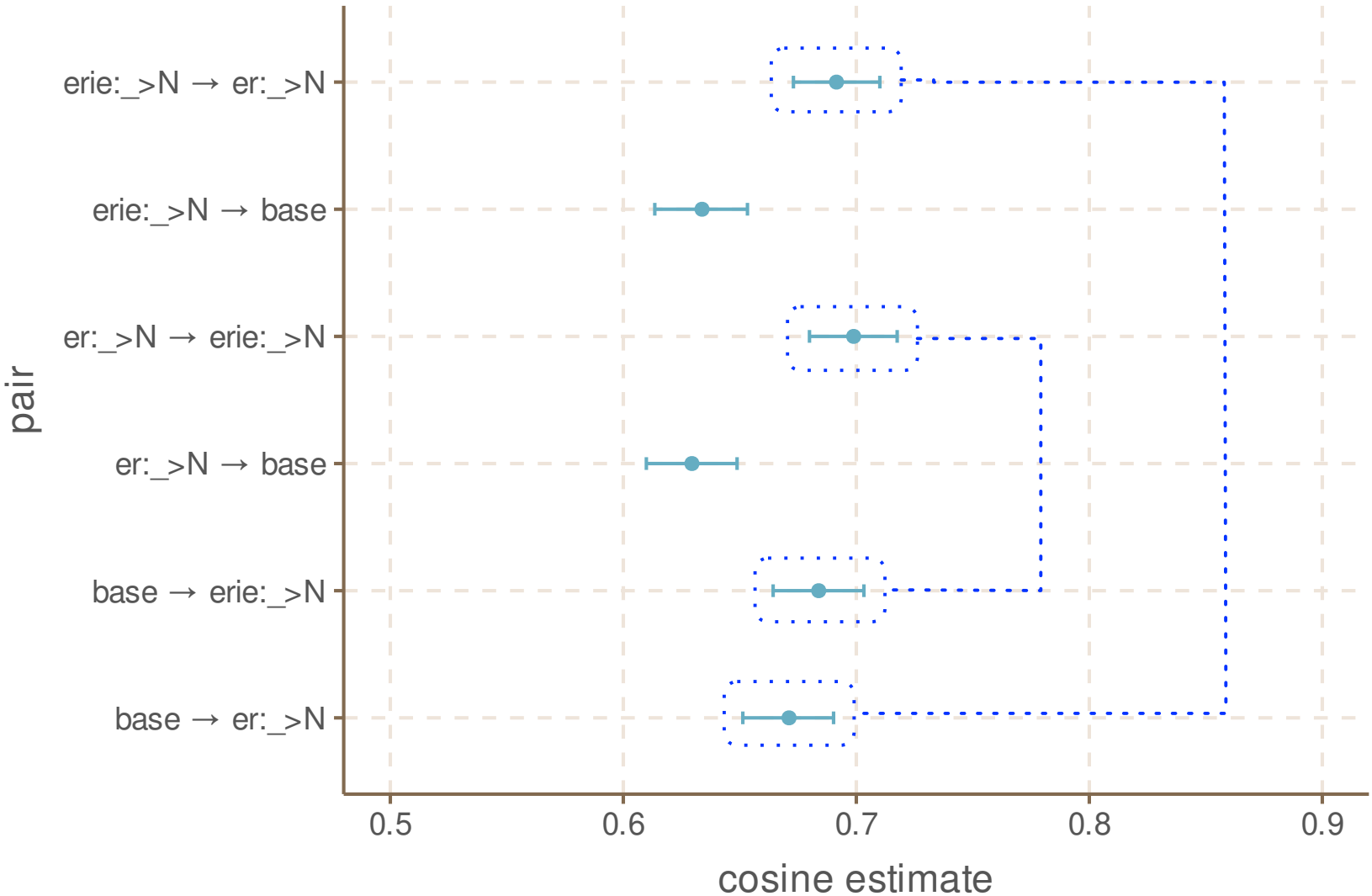
(Whiskers: 95% uncertainty interval obtained by Bayesian Beta regression)

Results: *-isme* vs. *-iste*



(Whiskers: 95% uncertainty interval obtained by Bayesian Beta regression)

Results: *-ier* vs. *-erie*



(Whiskers: 95% uncertainty interval obtained by Bayesian Beta regression)

Discussion

- Clear evidence that for some (but not all) pairs of processes, derivatives are more interpredictable than either is predictable from their base, on average.
- This is contradictory with a rooted-tree model, and entirely compatible with a paradigmatic model.
- (*-isme*, *-iste*) is the one pair of process exhibiting a very clear paradigmatic effect.
  - Not surprising, as this is the poster child for paradigmatic relations in derivation (see e.g. Becker 1993; Bauer 1997; Booij 2010; Roché 2011)
  - However most of the literature focuses on missing bases (e.g. *optimism*, *optimist*): we firmly conclude that a strong paradigmatic bond exists even when a base is present.
- Such a result is all we need to prove our point:
  - The paradigmatic hypothesis predicts that derivatives will sometimes be highly interpredictable, not that they always do.
  - The rooted tree hypothsesis predicts that bases should always be the best predictor.
- On the other hand, one datapoint is not a lot.
  - Replication on other languages where more data is available would be very welcome, e.g. using Derinet (Vidra et al., 2019) and SYN-derived vectors (Kyjánek & Bonami, 2022)!

Conclusions

Taking stock

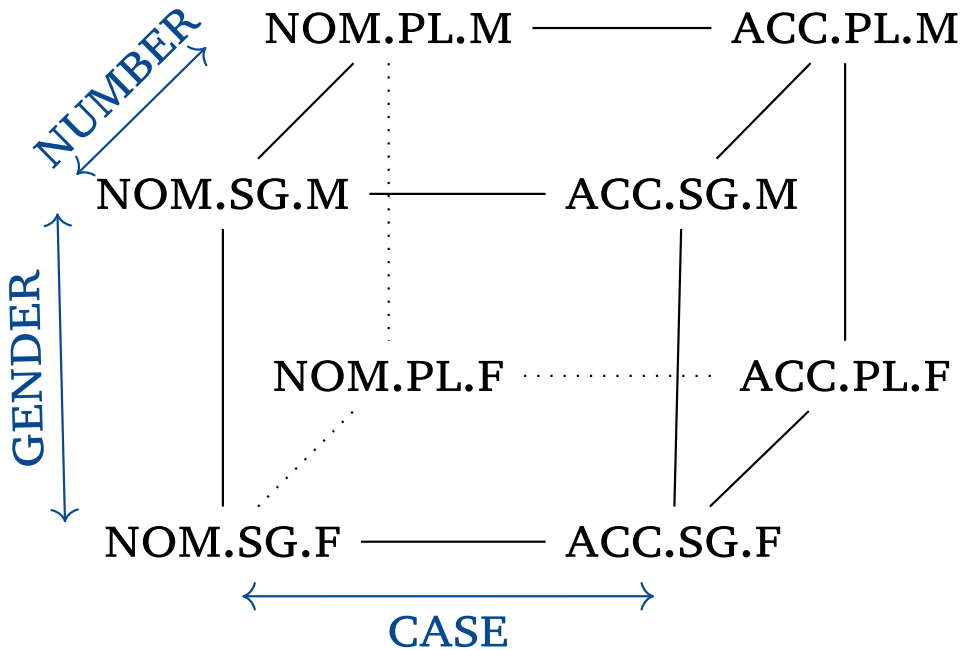
- Substantial evidence of a systemic role of paradigmatic relations in shaping derivational families.
  - Strong evidence from form predictability, promising evidence from predictability of meaning.
  - The argument is not that only paradigmatic relations matter, but that they can't be set aside.
- I have exemplified how computational methods are crucial to answering key theoretical questions in morphology, by allowing one to move from anecdotal to systematic evidence.
- I have explored various facets of a 3-dimensional search space:

Dimension	Methods	Inflection	Derivation
Form	Computational	Bonami & Beniamine (2016)	Bonami & Strnadová (2019)
	Behavioral	Copot & Bonami (accepted)	Copot & Bonami (submitted)
Meaning	Computational	…	Bonami & Guzmán Naranjo (2023)
	Behavioral	…	…

- Where do we go from here?

Next: Bonami, Kyjánek & Wauquier (submitted)

- Study of Czech nouns and adjectives
- Distributional predictability from cell to cell in the paradigm is extremely accurate.
- Morphosyntactic feature systems convey the assumption that some contrasts are parallels: e.g. the SG-PL contrast is the same in the NOM and the ACC.



- We establish empirically that parallel contrasts are somewhat similar but by no means identical.
- Stay tuned for more!

Thanks

- Collaborators: Sacha Beniamine, Matías Guzman Naranjo, Maria Copot, Lukáš Kyjánek, Jana Strnadová, Marine Wauquier
- Institutions:
  - Labex EFL, Strand 2: Experimental grammar
  - ANR Project *Demonext* (PI Fiammetta Namer)
  - Laboratoire de linguistique formelle (Université Paris Cité & CNRS)

References I

Ackerman, Farrell, James P. Blevins, & Robert Malouf (2009). "Parts and wholes: implicative patterns in inflectional paradigms." In: *Analogy in Grammar*. Ed. by James P. Blevins & Juliette Blevins. Oxford: Oxford University Press, pp. 54–82 (cit. on pp. 10, 17, 18).

Ackerman, Farrell & Robert Malouf (2013). "Morphological organization: the low conditional entropy conjecture." In: *Language* 89, pp. 429–464 (cit. on pp. 10, 17).

Albright, Adam C. (2002). "The Identification of Bases in Morphological Paradigms." PhD thesis. University of California, Los Angeles (cit. on p. 17).

Albright, Adam C. & Bruce P. Hayes (2003). "Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study." In: *Cognition* 90, pp. 119–161 (cit. on p. 17).

Anderson, Stephen R. (1982). "Where's Morphology?" In: *Linguistic Inquiry* 13, pp. 571–612 (cit. on p. 7).

— (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press (cit. on p. 7).

Aronoff, Mark (1976). *Word Formation in Generative Grammar*. Cambridge: MIT Press (cit. on p. 2).

— (1994). *Morphology by itself*. Cambridge: MIT Press (cit. on p. 10).

Baayen, R. Harald et al. (2019). "The Discriminative Lexicon: A Unified Computational Model for the Lexicon and Lexical Processing in Comprehension and Production Grounded Not in (De)Composition but in Linear Discriminative Learning." In: *Complexity* 2019, p. 4895891 (cit. on p. 8).

References II

Baerman, Matthew, Dunstan Brown, & Greville G. Corbett (2005). *The Syntax-Morphology Interface: A Study of Syncretism*. Cambridge: Cambridge University Press (cit. on p. 10).

Bauer, Laurie (1997). "Derivational Paradigms." In: *Yearbook of Morphology 1996*. Ed. by Geert Booij & Jaap van Marle. Dordrecht: Kluwer, pp. 243–256.

— (2004). "The function of word-formation and the inflection-derivation distinction." In: *Words in their Places. A Festschrift for J. Lachlan Mackenzie*. Amsterdam: Vrije Universiteit, pp. 283–292 (cit. on p. 7).

Becker, Thomas (1993). "Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology." In: *Yearbook of Morphology 1993*. Ed. by Geert Booij & Jaap van Marle. Dordrecht: Kluwer, pp. 1–25 (cit. on pp. 3, 5).

Beniamine, Sacha & Matías Guzmán Naranjo (2021). "Multiple alignments of inflectional paradigms." In: *Proceedings of the Society for Computation in Linguistics* 4.21. URL: https://scholarworks.umass.edu/scil/vol4/iss1/21 (cit. on p. 17).

Blevins, James P. (2006). "Word-based morphology." In: *Journal of Linguistics* 42, pp. 531–573 (cit. on p. 10).

— (2016). *Word and Paradigm Morphology*. Oxford: Oxford University Press (cit. on p. 10).

Bochner, Harry (1993). *Simplicity in Generative Morphology*. Berlin: Mouton de Gruyter (cit. on pp. 5, 7).

References III

Boleda, Gemma (2020). “Distributional Semantics and Linguistic Theory.” In: *Annual Review of Linguistics* 6.1, pp. 213–234 (cit. on p. 41).

Bonami, Olivier & S. Beniamine (2016). “Joint predictiveness in inflectional paradigms.” In: *Word Structure* 9.2, pp. 156–182 (cit. on pp. 10, 16–18).

Bonami, Olivier & Sacha Beniamine (2021). “Leaving the stem by itself.” In: *All Things Morphology*. Ed. by Marcia Haag et al. Amsterdam: John Benjamins, pp. 82–98 (cit. on p. 8).

Bonami, Olivier & Matías Guzman Naranjo (2023). “Distributional evidence for derivational paradigms.” In: *The semantics of derivational morphology: theory, methods, evidence*. Ed. by Sven Kotowski & Ingo Plag. Berlin: De Gruyter, pp. 219–258.

— (in press). “Distributional evidence for derivational paradigms.” In: *The semantics of derivational morphology: theory, methods, evidence*. Ed. by Sven Kotowski & Ingo Plag. Berlin: De Gruyter (cit. on p. 36).

Bonami, Olivier & Denis Paperno (2018). “Inflection vs. derivation in a distributional vector space.” In: *Lingue e Linguaggio* 17.2, pp. 173–195 (cit. on pp. 7, 43).

Bonami, Olivier & Jana Strnadová (2019). “Paradigm structure and predictability in derivational morphology.” In: *Morphology* 29.2, pp. 167–197 (cit. on pp. 5, 6, 12–16).

Booij, Geert (1996). “Inherent versus contextual inflection and the split morphology hypothesis.” In: *Yearbook of Morphology 1995*. Ed. by Geert Booij & Jaap van Marle. Dordrecht: Kluwer, pp. 1–16 (cit. on p. 7).

— (2010). *Construction morphology*. Oxford: Oxford University Press (cit. on p. 5).

References IV


Copot, Maria & Olivier Bonami (accepted). "Behavioural evidence for implicative paradigmatic relations." In: *The Mental Lexicon* (cit. on p. 25).


— (submitted). "Baseless derivation: the behavioural reality of derivational paradigms." (Cit. on p. 25).


Copot, Maria, Timothee Mickus, & Olivier Bonami (2022). "Idiosyncratic Frequency as a Measure of Derivation vs. Inflection." In: *Journal of Language Modelling* 10.2, pp. 193–240 (cit. on p. 7).


Corbett, Greville G. (2010). "Canonical derivational morphology." In: *Word Structure* 3, pp. 141–155 (cit. on p. 7).


Corbin, Danielle (1976). "Peut-on faire l'hypothèse d'une dérivation en morphologie?" In: *Grammaire transformationnelle : syntaxe et lexique*. Ed. by Jean-Claude Chevalier. Lille: Presses Universitaires de Lille, pp. 47–91 (cit. on p. 4).


Dressler, Wolfgang U. (1989). "Prototypical differences between inflection and derivation." In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42, pp. 3–10 (cit. on p. 7).


Drozd, Aleksandr, Anna Gladkova, & Satoshi Matsuoka (Dec. 2016). "Word Embeddings, Analogies, and Machine Learning: Beyond king - man + woman = queen." In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. Osaka, Japan: The COLING 2016 Organizing Committee, pp. 3519–3530. URL: https://aclanthology.org/C16-1332.

References V

Finkel, Raphael & Gregory T. Stump (2007). “Principal parts and morphological typology.” In: *Morphology* 17, pp. 39–75 (cit. on p. 17).

— (2009). “Principal parts and degrees of paradigmatic transparency.” In: *Analogy in Grammar*. Ed. by James P. Blevins & Juliette Blevins. Cambridge: Cambridge University Press, pp. 13–54 (cit. on p. 17).

Firth, John R. (1957). “Modes of meaning.” In: *Papers in Linguistics, 1934-1951*. Oxford: Oxford University Press, pp. 190–215 (cit. on p. 41).

Ford, Alan, Rajendra Singh, & Gita Martohardjono (1997). *Pace Panini*. Berlin: Peter Lang (cit. on p. 7).

Harris, Zelig S. (1954). “Distributional Structure.” In: *Word* 10, pp. 146–162 (cit. on p. 41).

Haspelmath, Martin (forthcoming). “Inflection and derivation as traditional comparative concepts.” In: *Linguistics* (cit. on p. 7).

Hathout, Nabil & Fiammetta Namer (2014a). “Démonette, a French derivational morpho-semantic network.” In: *Linguistic Issues in Language Technology* 11.5, pp. 125–168 (cit. on p. 20).

— (2014b). “Discrepancy between form and meaning in Word Formation: the case of over- and under-marking in French.” In: *Morphology and meaning (Selected papers from the 15th International Morphology Meeting, Vienna, February 2012)*. Ed. by Franz Rainer et al. Current Issues in Linguistic Theory. Amsterdam: John Benjamins, pp. 177–190 (cit. on p. 4).

References VI


Hathout, Nabil, Franck Sajous, & Basilio Calderone (2014). "GLÀFF, a Large Versatile French Lexicon." In: *Proceedings of LREC 2014* (cit. on p. 20).


Jackendoff, Ray & Jenny Audring (2020). *The texture of the lexicon*. Oxford: Oxford University Press (cit. on p. 5).


Jun, Jongho & Adam C. Albright (2016). *Speakers' Knowledge of Alternations Is Asymmetrical: Evidence from Seoul Korean Verb Paradigms*. Berlin: Cambridge University Press (cit. on p. 17).


Kyjánek, Lukáš & Olivier Bonami (2022). "A Distributional Approach to Inflection *vs.* Derivation in Czech." Paper read at *Word Formation Theories VI*, Košice, 26/06/2022.


Kyjánek, Lukáš et al. (2020). "Universal Derivations 1.0, A Growing Collection of Harmonised Word-Formation Resources." In: *The Prague Bulletin of Mathematical Linguistics* 115, pp. 5–30. ISSN: 0032-6585. URL: https://ufal.mff.cuni.cz/pbml/115/art-kyjanek-et-al.pdf (cit. on p. 2).


Lenci, Alessandro (2008). "Distributional semantics in linguistic and cognitive research." In: *Rivista di Linguistica* 20, pp. 1–31 (cit. on p. 41).


Marchand, Hans (1963). "On a question of contrary analysis with derivationally connected but morphologically uncharacterized words." In: *English Studies* 44, pp. 176–187 (cit. on p. 3).


Marelli, Marco & Marco Baroni (2015). "Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics." In: *Psychological Review* 122, pp. 485–515.

References VII

Marle, Jaap van (1984). *On the Paradigmatic Dimension of Morphological Creativity*. Dordrecht: Foris (cit. on p. 5).

Matthews, P. H. (1972). *Inflectional Morphology. A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press (cit. on p. 10).

Mickus, Timothee, Olivier Bonami, & Denis Paperno (2019). "Distributional Effects of Gender Contrasts Across Categories." In: *Proceedings of the 2nd Meeting of the Society for Computation in Linguistics*.

Mikolov, Tomas et al. (2013). "Efficient Estimation of Word Representations in Vector Space." In: *CoRR* abs/1301.3781. URL: http://arxiv.org/abs/1301.3781 (cit. on p. 41).

Pellegrini, Matteo (2021). "Patterns of interpredictability and principal parts in Latin verb paradigms: an entropy-based approach." In: *Journal of Latin Linguistics* 19.2, pp. 195–229 (cit. on p. 17).

Pennington, Jeffrey, Richard Socher, & Christopher D Manning (2014). "Glove: Global vectors for word representation." In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543 (cit. on p. 41).

Perlmutter, David M. (1998). "Interfaces: explanation of allomorphy and the architecture of grammars." In: *Morphology and its relation to phonology and syntax*. Ed. by Steven G. Lapointe, Diane K. Brentari, & Patrick M. Farrell. Stanford, CA: CSLI Publications, pp. 307–338 (cit. on p. 7).

References VIII

- Řehůřek, Radim (2010). "Software Framework for Topic Modelling with Large Corpora." In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 45–50.
- Robins, R. H. (1959). "In defense of WP." In: *Transactions of the Philological Society* 58, pp. 116–144 (cit. on pp. 5, 10).
- Roché, Michel (2011). "Quel traitement unifié pour les dérivations en *-isme* et en *-iste*?" In: ed. by Michel Roché et al. Hermès Lavoisier, pp. 69–143.
- Rosa, Rudolf & Zdeněk Žabokrtský (Sept. 2019). "Attempting to separate inflection and derivation using vector space representations." In: *Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology*. Prague, Czechia: Charles University, Faculty of Mathematics, Physics, Institute of Formal, and Applied Linguistics, pp. 61–70. URL: https://aclanthology.org/W19-8508 (cit. on p. 7).
- Schäfer, Roland (2015). "Processing and Querying Large Web Corpora with the COW14 Architecture." In: *Proceedings of Challenges in the Management of Large Corpora*, pp. 28–34.
- Schäfer, Roland & Felix Bildhauer (2012). "Building Large Corpora from the Web Using a New Efficient Tool Chain." In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, pp. 486–493.
- Ševčíková, Magda & Zdeněk Žabokrtský (2014). "Word-Formation Network for Czech." In: *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. Reykjavík, Iceland: European Language Resources Association, pp. 1087–1093. ISBN: 978-2-9517408-8-4 (cit. on p. 2).

References IX

Sims, Andrea D. & Jeff Parker (2016). "How inflection class systems work: On the informativity of implicative structure." In: *Word Structure* 9, pp. 215–239 (cit. on pp. 10, 17).

Spencer, Andrew (2013). *Lexical relatedness: a paradigm-based model*. Oxford: Oxford University Press (cit. on p. 7).

Strnadová, Jana (2014). "Les réseaux adjectivaux: Sur la grammaire des adjectifs dénominaux en français." PhD thesis. Université Paris Diderot et Univerzita Karlova V Praze (cit. on p. 4).

Stump, Gregory T. (1993). "On rules of referral." In: *Language* 69, pp. 449–479 (cit. on p. 10).

— (2016). *Inflectional paradigms*. Cambridge: Cambridge University Press (cit. on p. 10).

Stump, Gregory T. & Raphael Finkel (2013). *Morphological Typology: From Word to Paradigm*. Cambridge: Cambridge University Press (cit. on pp. 10, 17).

Tribout, Delphine (2020). "Nominalization, verbalization or both? Insights from the directionality of noun-verb conversion in French." In: *Zeitschrift für Wortbildung/Journal of Word Formation* 4.2, pp. 187–207 (cit. on p. 3).

Vidra, Jonáš et al. (2019). "DeriNet 2.0: Towards an All-in-One Word-Formation Resource." In: *Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology (DeriMo 2019)*. Praha, Czechia: ÚFAL MFF UK, pp. 81–89. ISBN: 978-80-88132-08-0 (cit. on p. 2).

References X

Štekauer, Pavol (2014). "Derivational Paradigms." In: *The Oxford Handbook of Derivational Morphology*. Ed. by Rochelle Lieber & Pavol Štekauer. Oxford: Oxford University Press, pp. 354–369 (cit. on pp. 5, 10).

Wilmoth, Sasha & John Mansfield (2021). "Inflectional predictability and prosodic morphology in Pitjantjatjara and Yankunytjatjara." In: *Morphology* (cit. on p. 17).

Wunderlich, Dieter & Ray Fabri (1995). "Minimalist Morphology: an approach to inflection." In: *Zeitschrift für Sprachwissenschaft* 14.2, pp. 236–294 (cit. on p. 11).

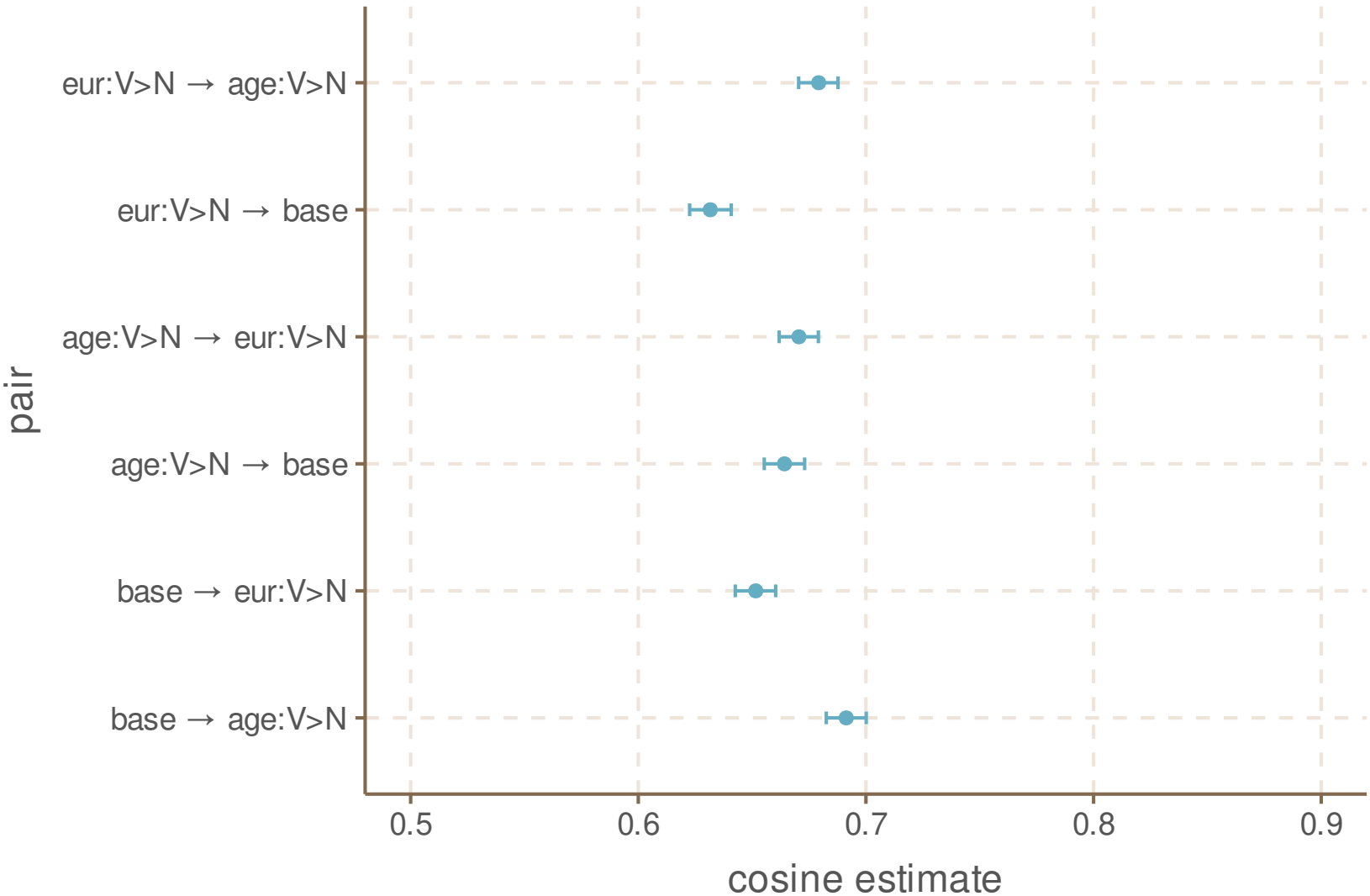
Wurzel, Wolfgang Ulrich (1984). *Flexionsmorphologie und Natürlichkeit. Ein Beitrag zur morphologischen Theoriebildung*. Translated as Wurzel, 1989. Berlin: Akademie-Verlag (cit. on p. 10).

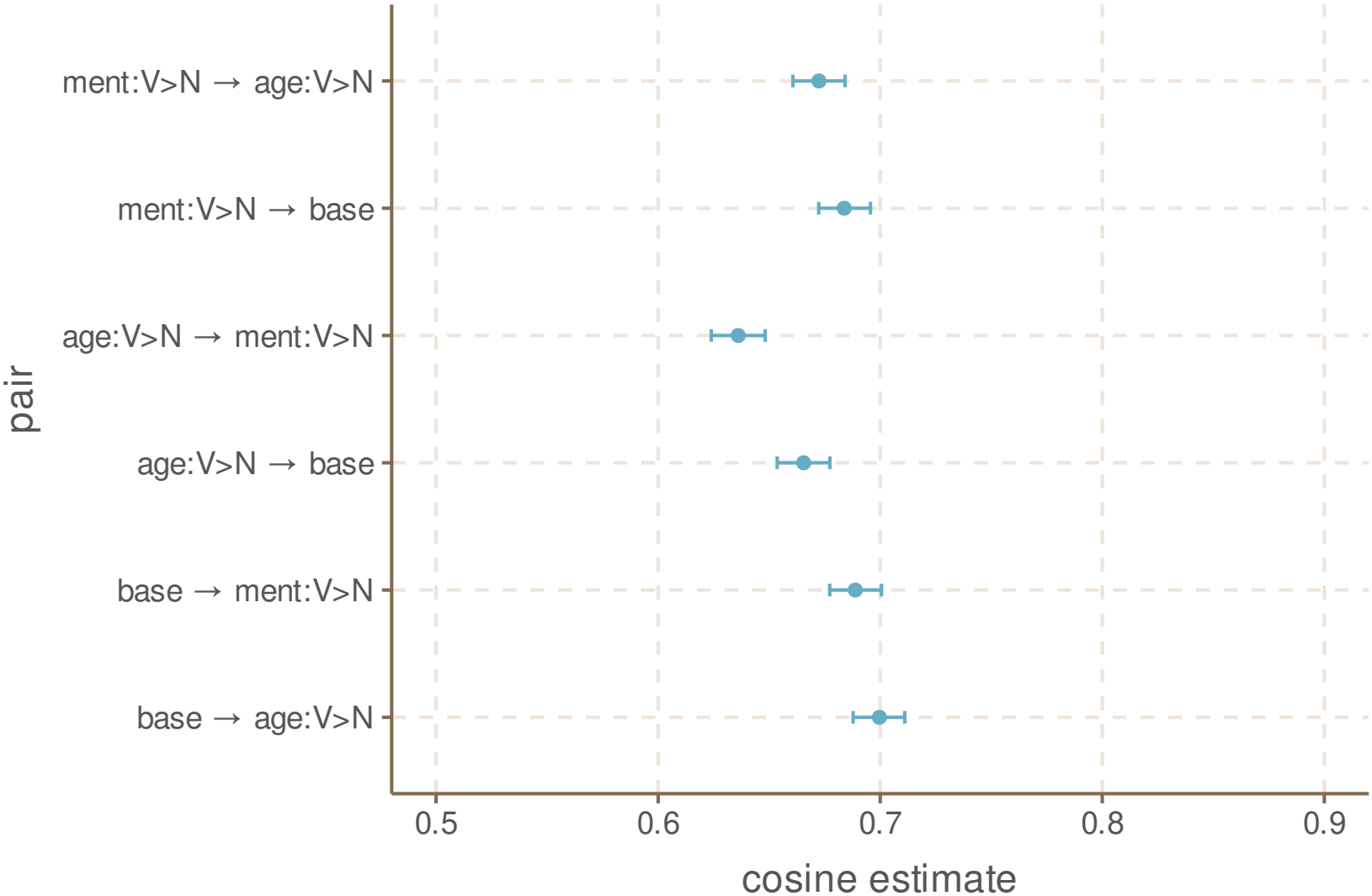
— (1989). *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer (cit. on p. 17).

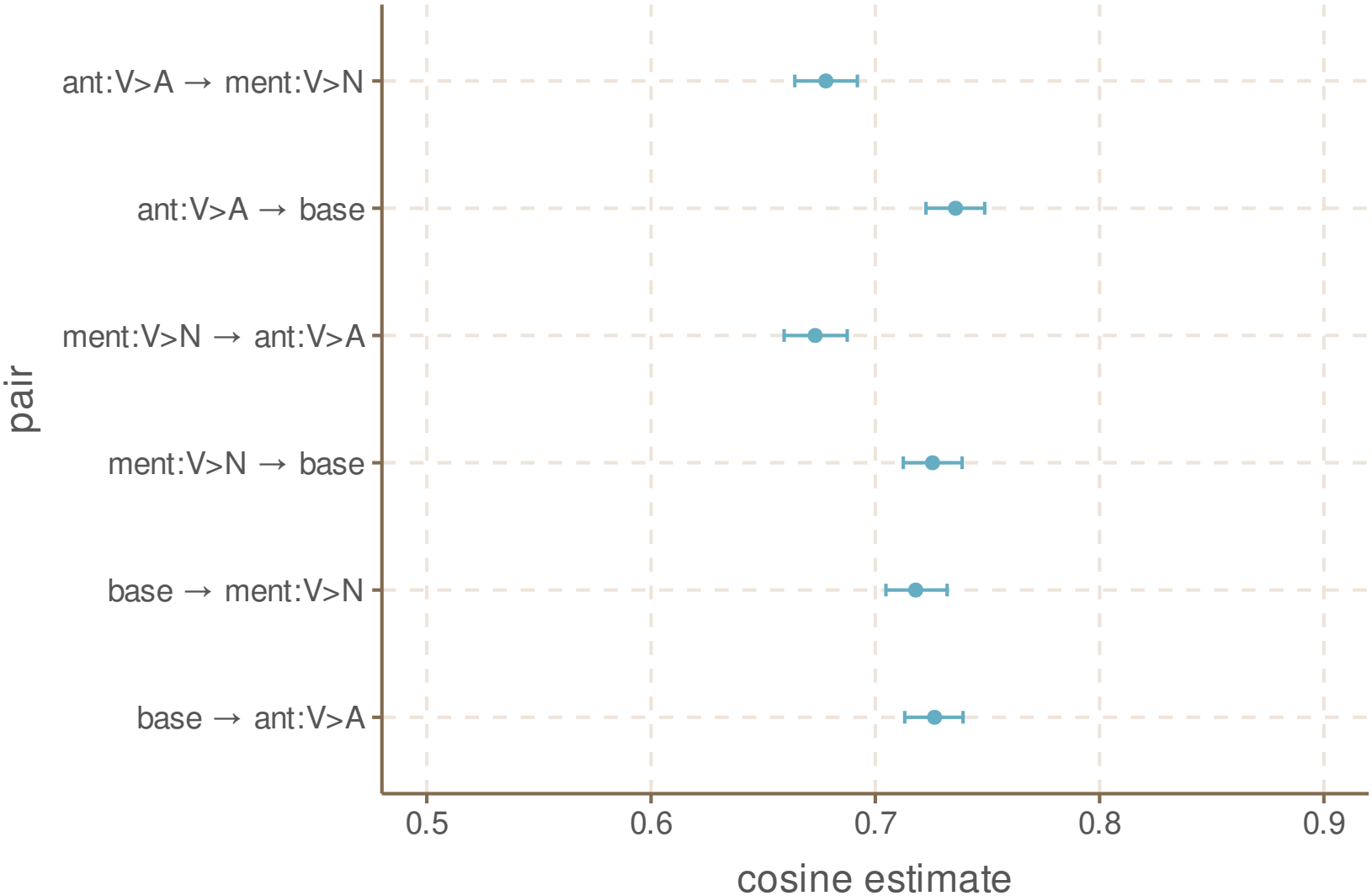
Zwicky, Arnold M. (1985). "How to describe inflection." In: *Proceedings of Berkeley Linguistics Society 11*, pp. 372–386 (cit. on p. 10).

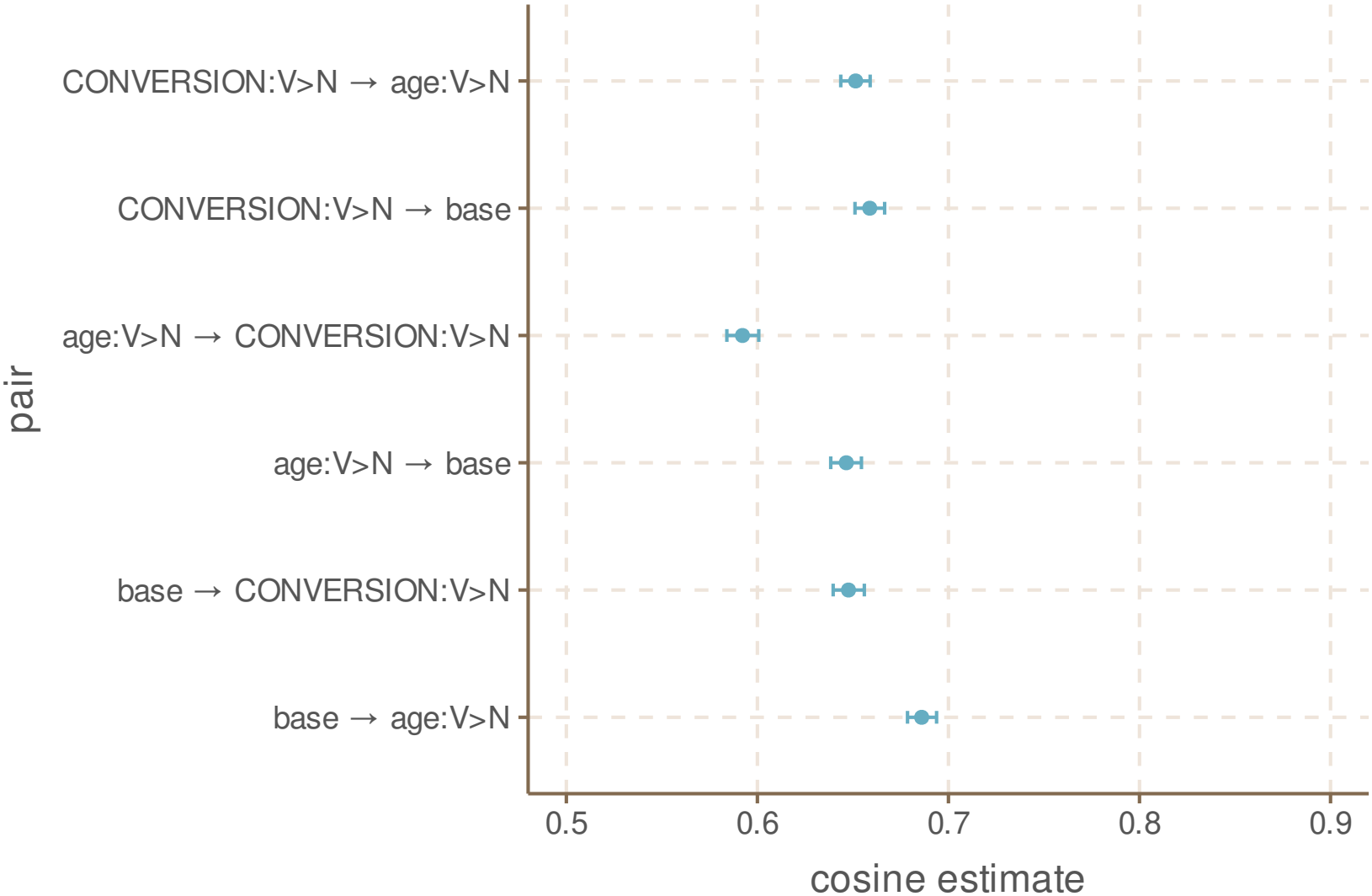
Summary of overall results

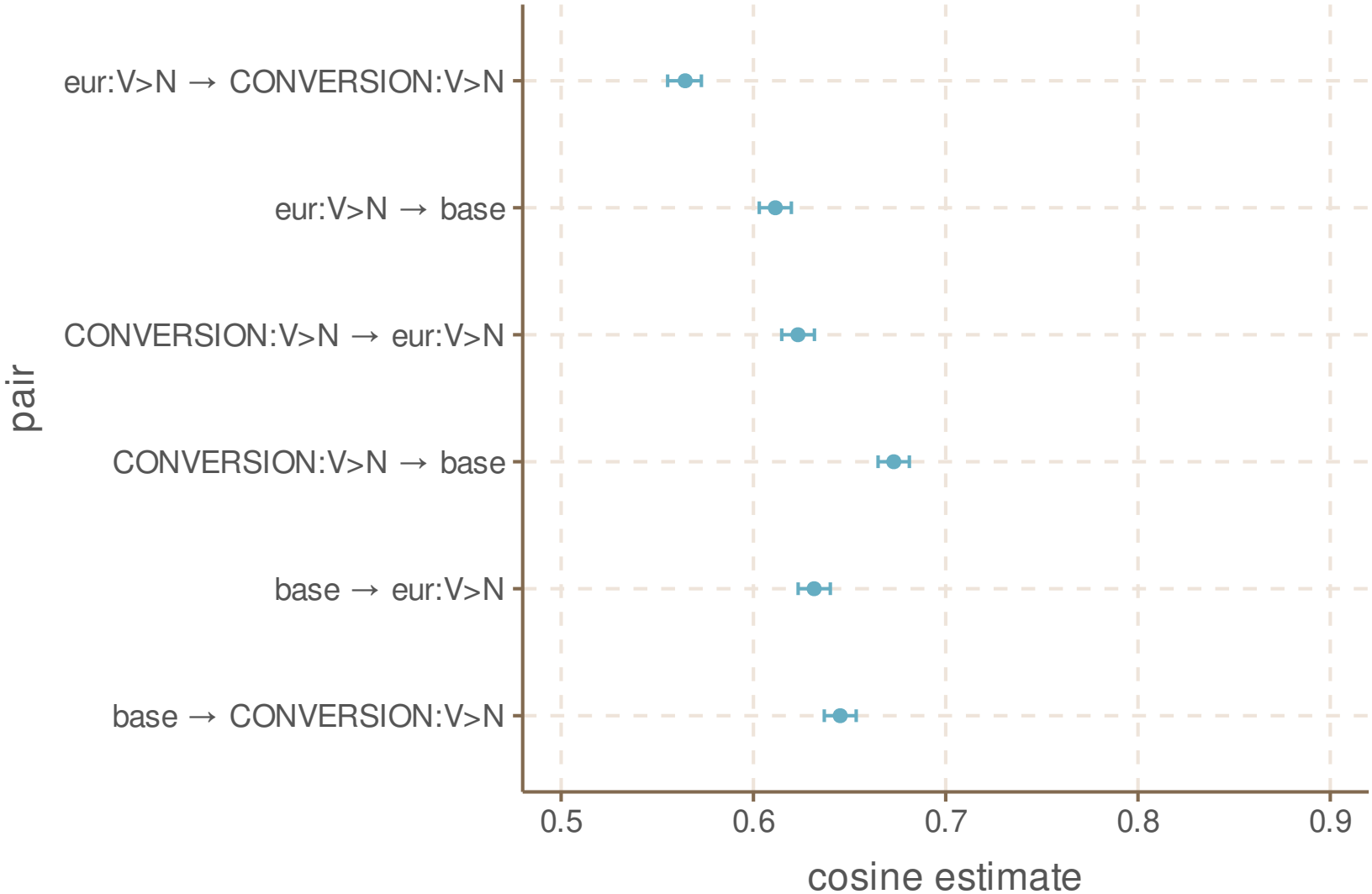
	$D_2 \rightarrow D_1$ vs. $B \rightarrow D_1$	$D_1 \rightarrow D_2$ vs. $B \rightarrow D_2$	$B \rightarrow D_1$ vs. $D_1 \rightarrow B$	$B \rightarrow D_2$ vs. $D_2 \rightarrow B$	
(isme:N > N, iste:N > N)	$\gg$	$\gg$	$\ll$	$\ll$	
(er:N > N, erie:N > N)	$\geq$	$\geq$	$\ll$	$\ll$	
(age:V > N, conversion:V > N)	$\ll$	$\ll$	$\geq$	$\ll$	Rivals
(age:V > N, ment:V > N)	$\ll$	$\ll$	$\ll$	$\ll$	Rivals
(conversion:V > N, ment:V > N)	$\ll$	$\ll$	$\gg$	$\geq$	Rivals
(age:V > N, eur:V > N)	$\leq$	$\gg$	$\ll$	$\leq$	Agt/Act
(conversion:V > N, eur:V > N)	$\ll$	$\leq$	$\gg$	$\ll$	Agt/Act
(ion:V > N, eur:V > N)	$\ll$	$\geq$	$\leq$	$\leq$	Agt/Act
(ment:V > N, eur:V > N)	$\ll$	$\leq$	$\geq$	$\leq$	Agt/Act
(ant:V > A, ment:V > N)	$\ll$	$\ll$	$\geq$	$\geq$	

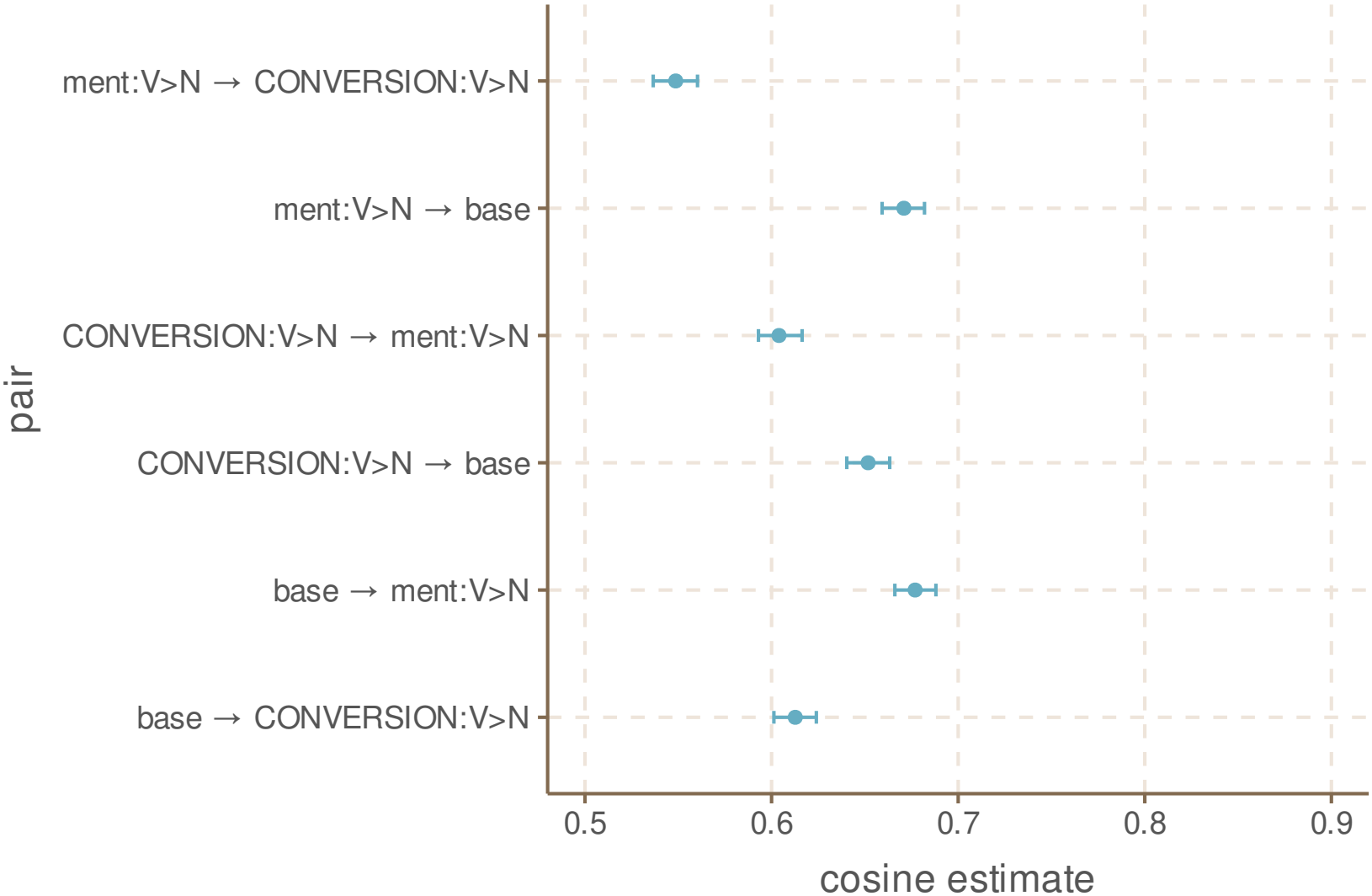

eur:V>N → age:V>N
eur:V>N → base
age:V>N → eur:V>N
age:V>N → base
base → eur:V>N
base → age:V>N
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate

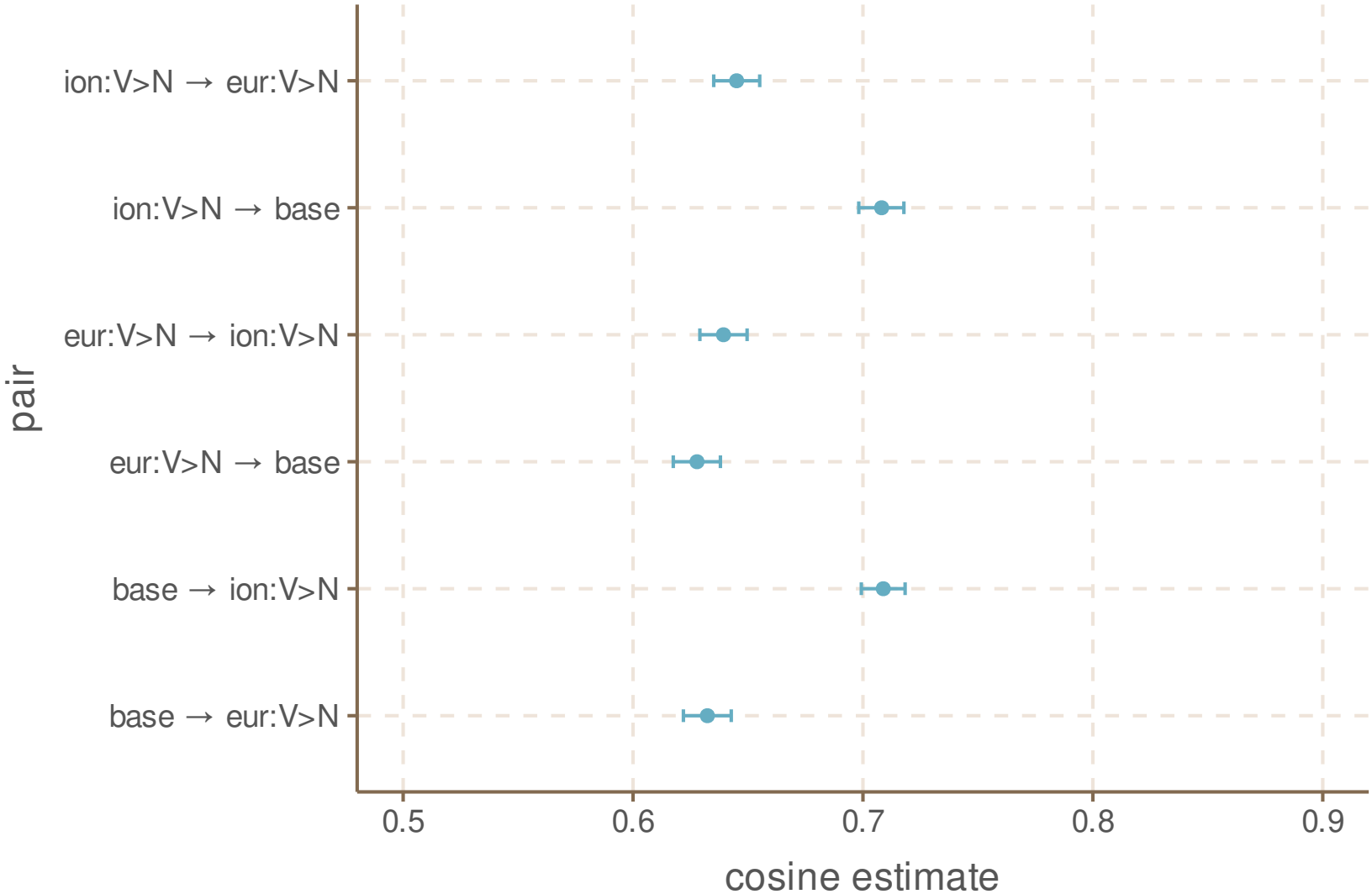

ment:V>N → age:V>N
ment:V>N → base
age:V>N → ment:V>N
age:V>N → base
base → ment:V>N
base → age:V>N
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate


ant:V>A → ment:V>N
ant:V>A → base
ment:V>N → ant:V>A
ment:V>N → base
base → ment:V>N
base → ant:V>A
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate


CONVERSION:V>N → age:V>N
CONVERSION:V>N → base
age:V>N → CONVERSION:V>N
age:V>N → base
base → CONVERSION:V>N
base → age:V>N
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate


eur:V>N → CONVERSION:V>N
eur:V>N → base
CONVERSION:V>N → eur:V>N
CONVERSION:V>N → base
base → eur:V>N
base → CONVERSION:V>N
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate


ment:V>N → CONVERSION:V>N
ment:V>N → base
CONVERSION:V>N → ment:V>N
CONVERSION:V>N → base
base → ment:V>N
base → CONVERSION:V>N
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate


ion:V>N → eur:V>N
ion:V>N → base
eur:V>N → ion:V>N
eur:V>N → base
base → ion:V>N
base → eur:V>N
pair
0.5
0.6
0.7
0.8
0.9
cosine estimate