



Why are paradigms not just systems of feature oppositions?

Lukáš Kyjánek & Olivier Bonami

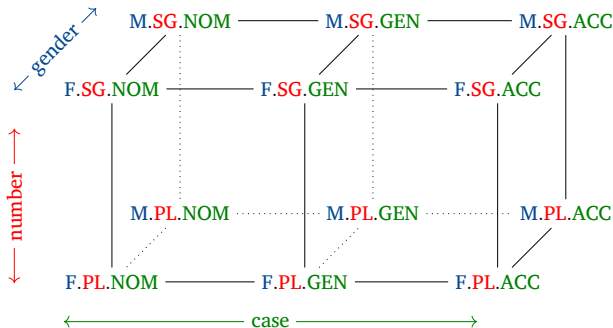
CNRS, Université Paris Cité, Laboratoire de linguistique formelle

October 8 – 10, 2025; Fifth International Symposium of Morphology (ISMo)

Paradigms as systems of (orthogonal) feature oppositions

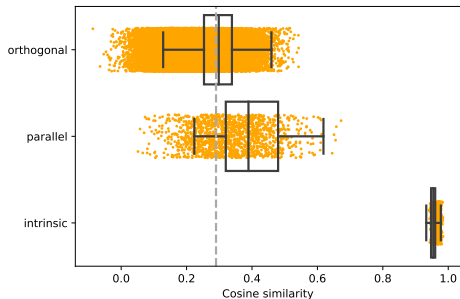
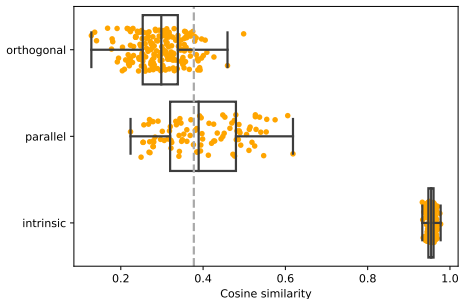
- Many authors define inflectional paradigms in terms of their organization into orthogonal features, cf. Wunderlich and Fabri [1995, p. 266]:

“A paradigm is an n -dimensional space whose dimensions are the attributes (or features) used for the classification of word forms. In order to be a dimension, an attribute must have at least two values. The cells of this space can be occupied by word forms of appropriate categories.”



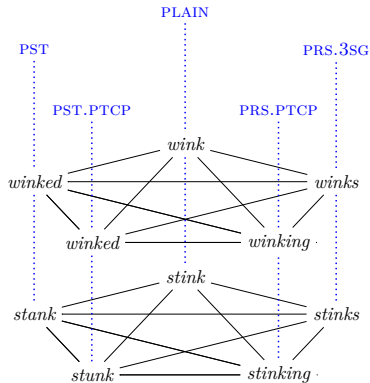
Bonami et al.'s (2023) experiment and observation

- Bonami et al. [2023] focused on Czech nouns and adjectives:
 - (training) vector prediction trained for individual contrasts,
 - (expectation) pairs of contrasts divided into types: orthogonal *vs.* parallel prediction of orthogonal contrasts should have worse performance than parallel ones,
 - (evaluation) vectors predicted for both (extrinsic) types and for intrinsic prediction
- Observation: Some feature contrasts led to a failure of parallelism.



Paradigms are not just systems of feature oppositions

- Bonami et al. [2023], based on distributional data:
Paradigm cells have properties that cannot be explained solely by their morphosyntactic features.
- This suggests Bonami et Strnadová's [2019] characterization of paradigms without reference to feature is on the right track (cf. Štekauer [2014], Blevins [2016]):
 - Paradigms are defined abstractively in terms of aligned pairwise contrasts.
 - Analysis into orthogonal features is a further step of abstraction that is neither necessary nor always insightful.



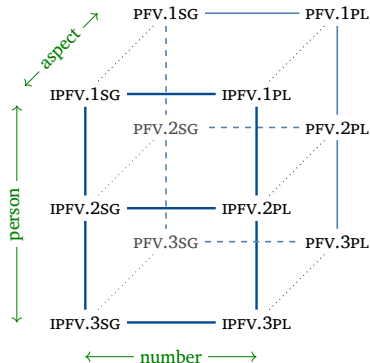
Why are some contrasts more parallel than others?

- We rely on the same data resources as [Bonami et al. \[2023\]](#) and observe which and why some contrasts are more parallel than others.
- In this talk, we will look into specific contrasts in Czech conjugation where a lack of distributional parallelism is expected on conceptual grounds.
 - We expect distributional parallelism to fail when comparing number contrasts in different persons or person contrasts in different numbers.
 - We do not expect person and number to be fully orthogonal features
 - [Benveniste \[1946\]](#): $1\text{PL} \neq 1 + 1 + \dots + 1$
 - Polite plurals, authorial plural, etc.
 - [Martinet \[1960\]](#), [Wechsler \[2004\]](#) reorganize the person-number feature space on that basis

Czech conjugation dataset

- Perfective and imperfective verbs define a 3-dimensional morphosyntactic space of parallel person and number contrasts. Note that we have parallelism of person and number **across aspects**, but not **parallelism of aspect**, as aspect is a lexical feature.

		FUT	PRS	PST	IMP
SG	1	<i>periph.</i>	<i>kupuji</i>	<i>periph.</i>	—
	2	<i>periph.</i>	<i>kupuješ</i>	<i>periph.</i>	<i>kupuj</i>
	3	<i>periph.</i>	<i>kupuje</i>	<i>kupoval(-/a/o)</i>	—
PL	1	<i>periph.</i>	<i>kupujeme</i>	<i>periph.</i>	—
	2	<i>periph.</i>	<i>kupujete</i>	<i>periph.</i>	<i>kupujte</i>
	3	<i>periph.</i>	<i>kupují</i>	<i>kupoval(i/y/a)</i>	—
<i>kupovat</i> [en. <i>buy</i>] (imperfective)					
		FUT	PRS	PST	IMP
SG	1	<i>koupím</i>	—	<i>periph.</i>	—
	2	<i>koupíš</i>	—	<i>periph.</i>	<i>kup</i>
	3	<i>koupí</i>	—	<i>koupil(-/a/o)</i>	—
PL	1	<i>koupíme</i>	—	<i>periph.</i>	—
	2	<i>koupíte</i>	—	<i>periph.</i>	<i>kupte</i>
	3	<i>koupí</i>	—	<i>koupil(i/y/a)</i>	—
<i>koupit</i> [en. <i>buy</i>] (perfective)					



Source of language data

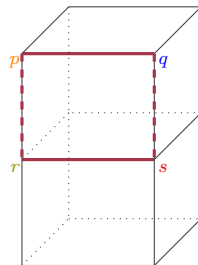
- Following [Bonami et al. \[2023\]](#), we used **Distributional vector models** [[Kyjánek and Bonami, 2022](#)] and inflectional lexicon **MorfFlexCZ 2.0** [[Hajič et al., 2020](#)].
 - **SYN v9 corpus** [[Křen et al., 2021](#)], 5.7 billion tokens from published texts
 - **MorphoDiTa** [[Straková et al., 2014](#)] for tagging and lemmatization in the corpus
 - **Word2vec** [[Mikolov et al., 2013](#)] applied to (token, tag) pairs to train the distributional vector model; syncretic forms are disambiguated
 - **MorfFlexCZ 2.0**, source of morphological data, 125.3M (lemma, tag, wordform) triples
- This results in **highly reliable vectors**, compared to the standards of the field.

Example from MorfFlexCZ:
inflection of 'buy' (perfective).

Lemma	Tag	Word form
koupit	Vf-----A-P--	koupit
koupit	Vf-----A-P-2	koupiti
koupit	VB-S---1P-AAP--	koupím
koupit	VB-S---1P-AAP-6	koupim
koupit	VB-S---2P-AAP--	koupíš
koupit	VB-S---3P-AAP--	koupí
koupit	VB-P---1P-AAP--	koupíme
koupit	VB-P---2P-AAP--	koupíte
koupit	VB-P---3P-AAP--	koupí
koupit	VB-P---3P-AAP-6	koupěj
koupit	VB-P---3P-AAP-3	koupějí
koupit	Vi-S---2--A-P--	kup
koupit	Vi-P---1--A-P--	kupme
koupit	Vi-P---2--A-P--	kupte

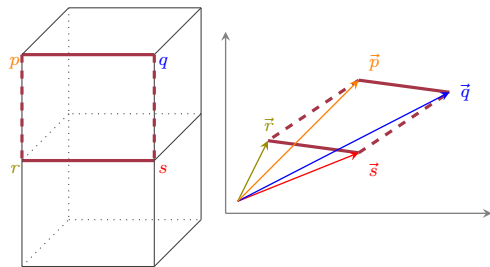
Formalisation of relevant contrasts in vector space

- We are interested in **rectangles in morphosyntactic space**.
- To assess whether that is the case:
 1. We pick two ordered pairs of cells corresponding to opposite edges of the morphosyntactic rectangle. $(p, q), (r, s)$



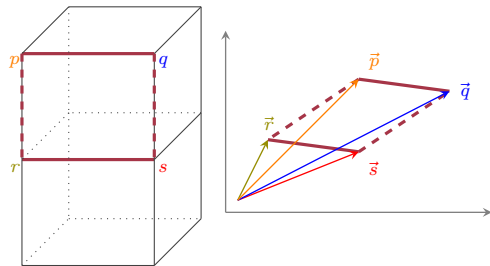
Formalisation of relevant contrasts in vector space

- We are interested in **rectangles in morphosyntactic space**.
- Truly parallel feature contrasts should map to **parallelograms in distrib. space**.
- To assess whether that is the case:
 1. We pick two ordered pairs of cells corresponding to opposite edges of the morphosyntactic rectangle. (p, q) , (r, s)
 2. We associate each cell c with average of vectors of words filling that cell $\vec{c}$.

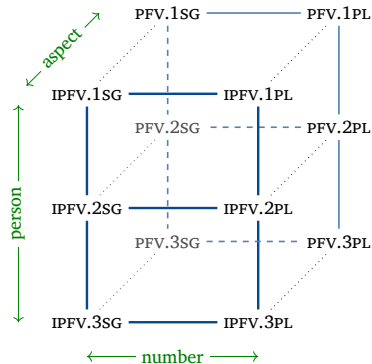


Formalisation of relevant contrasts in vector space

- We are interested in **rectangles in morphosyntactic space**.
- Truly parallel feature contrasts should map to **parallelograms in distrib. space**.
- To assess whether that is the case:
 1. We pick two ordered pairs of cells corresponding to opposite edges of the morphosyntactic rectangle. $(p, q), (r, s)$
 2. We associate each cell c with average of vectors of words filling that cell $\vec{c}$.
 3. We compute the difference vector for each of the two pairs of cells. $\vec{p} - \vec{q}, \vec{r} - \vec{s}$
 4. We compute the **Euclidean distance** between these two vectors. $d(\vec{p} - \vec{q}, \vec{r} - \vec{s})$
- If the cells form a parallelogram, this should be 0. How close to 0 is it?

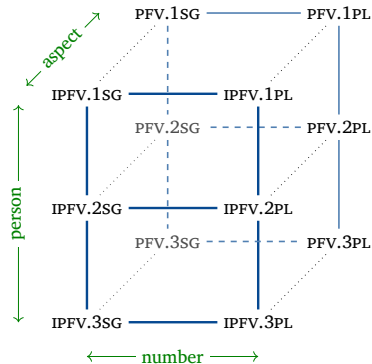
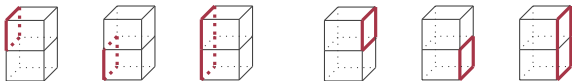


Contrasts (3): rectangles with constant person



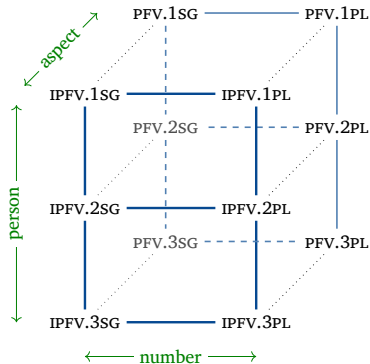
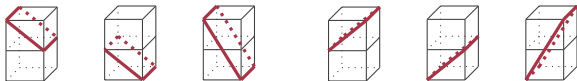
- expectation: parallelograms in distributional space ($ED \approx 0$)

Contrasts (6): rectangles with constant number



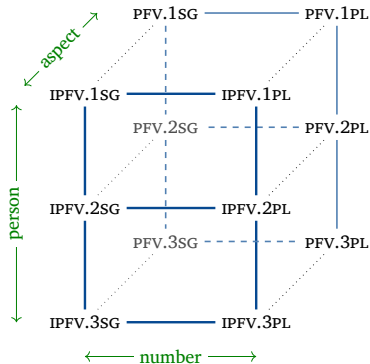
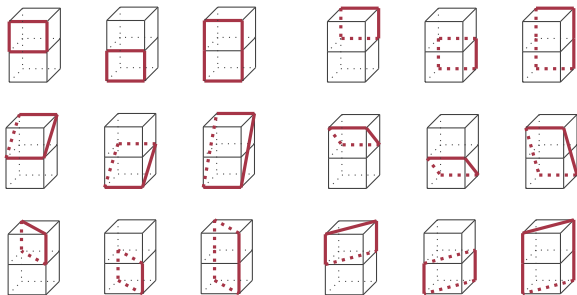
- expectation: parallelograms in distributional space ($ED \approx 0$)

Contrasts (6): rectangles covarying person and number



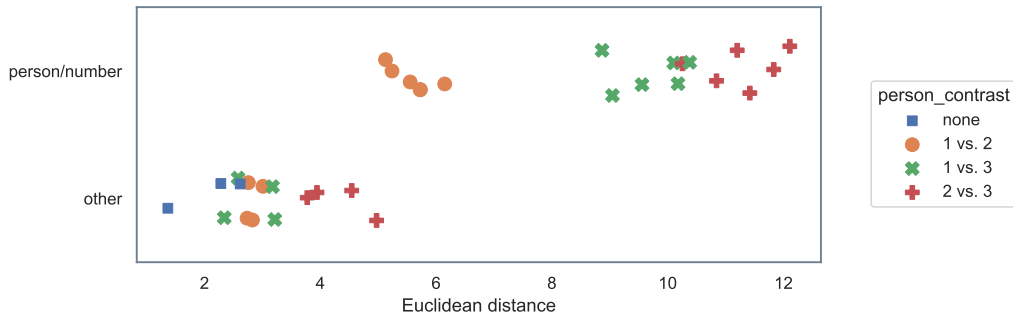
- expectation: parallelograms in distributional space ($ED \approx 0$)

Contrasts (18): rectangles opposing person and number



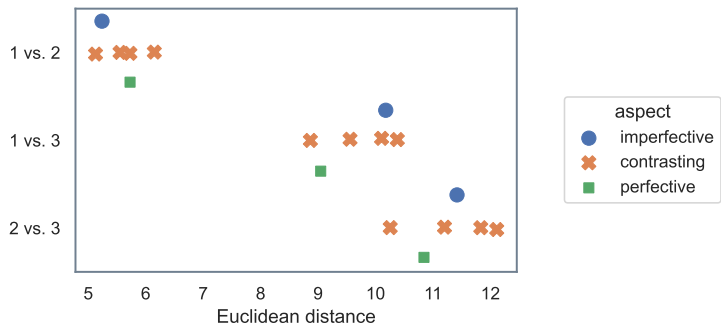
- expectation: **not** parallelograms in distributional space ($ED \gg 0$)

Results: Person/Number *vs.* other rectangles



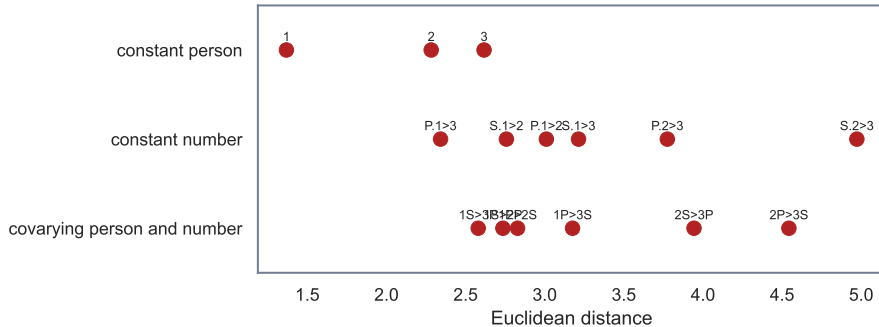
- Rectangles opposing person/number are nearly categorically separated from others.
- The 2vs3 contrast achieves insignificantly higher ED than other person contrasts.

Results: Deeper into person/number rectangles



- In observation of pers/num rectangles, the person contrasts matter.
- There is a difference between the ED of perfective-imperfective person contrasts (1vs2 *vs.* 1vs3, 2vs3).

Results: Deeper into other rectangles



- In observation of other rectangles than pers/num, no obvious organization is present.

Discussion/Challenge: Periphrastic verb forms in Czech

- Grammatical expressions created with multiple words rather than a single, inflected word, often using auxiliary verbs, are usual in Czech.

		FUT	PRS	PST	FUT	PRS	PST
SG	1	budu kupovat	<i>single</i>	kupoval(−/a/o) jsem	<i>single</i>	—	koupil(−/a/o) jsem
	2	budeš kupovat	<i>single</i>	kupoval(−/a/o) jsi	<i>single</i>	—	koupil(−/a/o) jsi
	3	bude kupovat	<i>single</i>	<i>single</i>	<i>single</i>	—	<i>single</i>
PL	1	budeme kupovat	<i>single</i>	kupoval(i/y/a) jsme	<i>single</i>	—	koupil(i/y/a) jsme
	2	budete kupovat	<i>single</i>	kupoval(i/y/a) jste	<i>single</i>	—	koupil(i/y/a) jste
	3	budou kupovat	<i>single</i>	<i>single</i>	<i>single</i>	—	<i>single</i>
<i>kupovat</i> [en. <i>buy</i>] (imperfective)					<i>koupit</i> [en. <i>buy</i>] (perfective)		

- Work with vectors is disputable – auxiliaries affect trained vectors. Is it a solution to train vectors based on a pseudo-corpus?
 - excluding auxiliaries
 - labelling past forms with verbs
- Could contextualised word embeddings solve that technical issue?

Conclusions

- Some morphosyntactic contrasts are more parallel than others.
 - We showed this in the obvious place, *person* vs. *number*.
 - Another good candidate for Czech:
 - periphrastic verb forms or
 - GEN.PL in nouns and adjectives.
 - Ultimately, this could lead to a data-driven delineation of feature space.
- Distributional vectors provide a very welcome way of exploring the import of features in an inductive data-driven way.
- The properties of distributional vectors support the view that [paradigms are prior to feature organization](#) [[Blevins, 2016](#), [Bonami and Strnadová, 2019](#)].

References

- Émile Benveniste. Structure des relations de personne dans le verbe. *Bulletin de la Société de Linguistique de Paris*, XLIII, 1946.
- James P. Blevins. *Word and Paradigm Morphology*. Oxford University Press, Oxford, 2016.
- Olivier Bonami and Jana Strnadová. Paradigm structure and predictability in derivational morphology. *Morphology*, 29(2):167–197, 2019.
- Olivier Bonami et al. Assessing the featural organisation of paradigms with distributional methods. In *Proceedings of the Society for Computation in Linguistics (vol 6)*, pages 310–320, 2023.
- Jan Hajič, Jaroslava Hlaváčová, Marie Mikulová, Milan Straka, and Barbora Štěpánková. MorfFlex CZ 2.0, 2020. URL <http://hdl.handle.net/11234/1-3186>. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.
- Michal Křen, Václav Cvrček, Jan Henyš, Milena Hnátková, Tomáš Jelínek, Jan Koček, Dominika Kovářiková, Jan Křivan, Jiří Milička, Vladimír Petkevič, Pavel Procházka, Hana Skoumalová, Jana Šindlerová, and Michal Škrabal. SYN v9: large corpus of written czech, 2021. URL <http://hdl.handle.net/11234/1-4635>. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.
- Lukáš Kyjánek and Olivier Bonami. Package of word embeddings of czech from a large corpus, 2022. URL <http://hdl.handle.net/11234/1-4920>. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.
- André Martinet. *Éléments de linguistique générale*. Armand Colin, Paris, 1960.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- Jana Straková, Milan Straka, and Jan Hajič. Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 13–18, Baltimore, Maryland, June 2014. Association for Computational Linguistics. URL <http://www.aclweb.org/anthology/P/P14/P14-5003.pdf>.
- Pavol Štekauer. Derivational paradigms. In Rochelle Lieber and Pavol Štekauer, editors, *The Oxford Handbook of Derivational Morphology*, pages 354–369. Oxford University Press, Oxford, 2014.
- Stephen Wechsler. Number as person. In Olivier Bonami and Patricia Cabredo Hofherr, editors, *Empirical issues in formal syntax and semantics*, vol. 5, pages 255–274, 2004. URL <http://www.cssp.cnrs.fr/eiss5>.
- Dieter Wunderlich and Ray Fabri. Minimalist Morphology: an approach to inflection. *Zeitschrift für Sprachwissenschaft*, 14(2):236–294, 1995.