Mapping Phonology to Semantics: A computational model of cross-lingual spoken word recognition

I. Zaitova, B.M. Abdullah and D. Klakow, 2022

Presented by
Emma Angela Montecchiari

Paper Review
Human Language Technologies
Università di Trento
Cognitive Science Master's
09-02-2024

General framework

Closely related languages exhibit intelligibility =
Metric of cross-language similarity

Inter-comprehension =
Ability of intelligible languages speakers of comprehending each other's speech without explicitly learning the second language

Key influential aspects:

- Lexical distance [cognates]
- Phonological distance (Levenshtein distance)

General framework

Spoken word recognition theories and models

Mapping between words acoustic-phonetic representation (*acoustic realization*) and semantic representation in memory (*lexical knowledge*)

Specific framework

General objective: Exploring mutual intelligibility

Method: Neural model of spoken-word recognition.

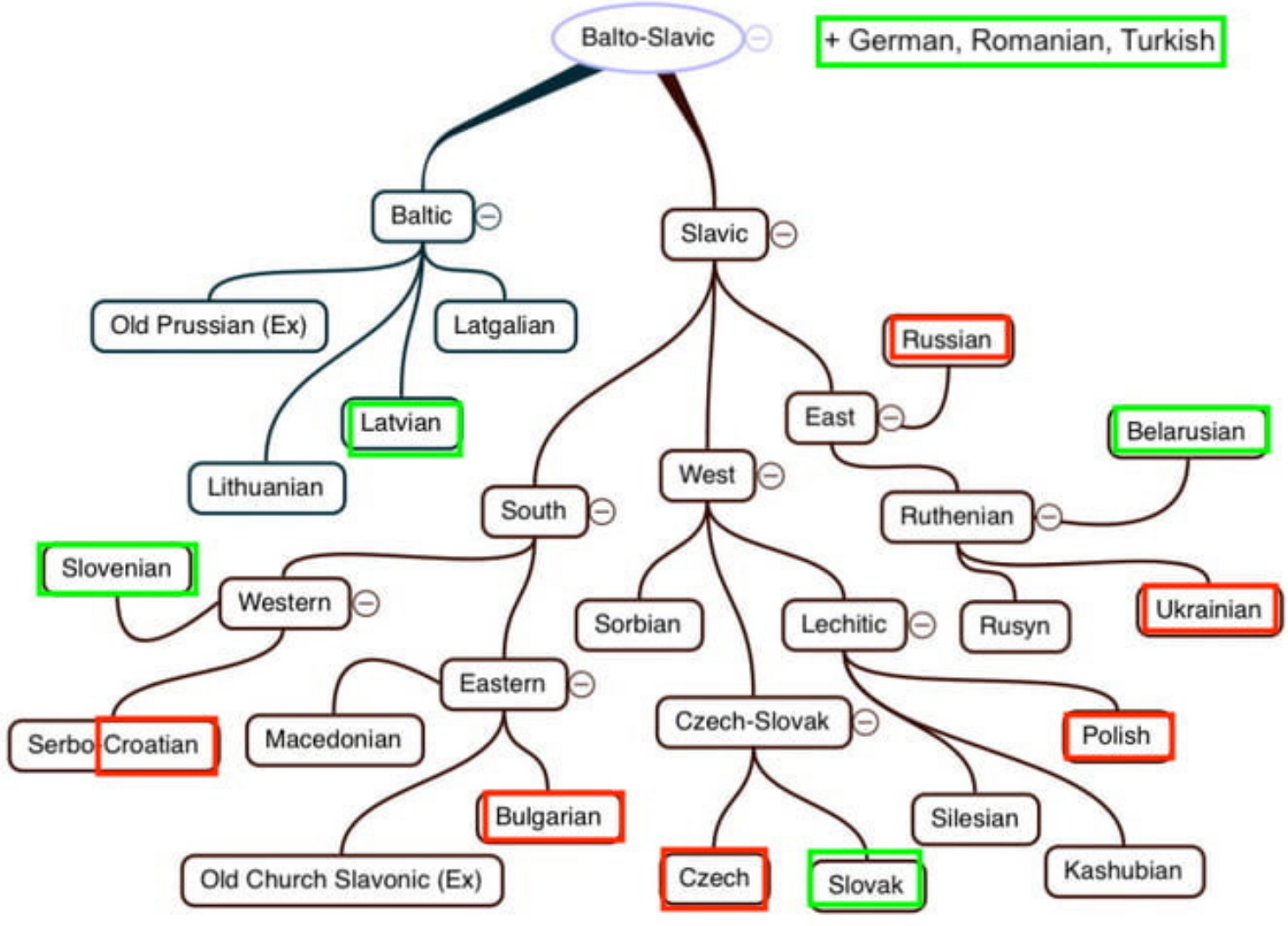
Procedure: Evaluating monolingual models on cross-lingual performance

Cross-lingual performance: **How are they able to understand spoken words meaning of related languages?**

Specific Framework

Monolingual Models trained on 6 Slavic languages

[exemplifying high mutual intelligibilty]



Cross-Lingual evaluation on Slavic and non-Slavic languages

Specific Framework

Specific objectives:

a) How performance predicts languages mutual intelligibility?

B) Do results reflect languages genetic relation?

C) Identifying linguistic measures affecting performance.

Procedure

6 Monolingual models → training on 6 Slavic languages

[exemplifying high mutual intelligibilty]

↓

Common **cross-language** concepts sample of word forms selected from FastText (linguistic uniformity) (502 concepts)

↓

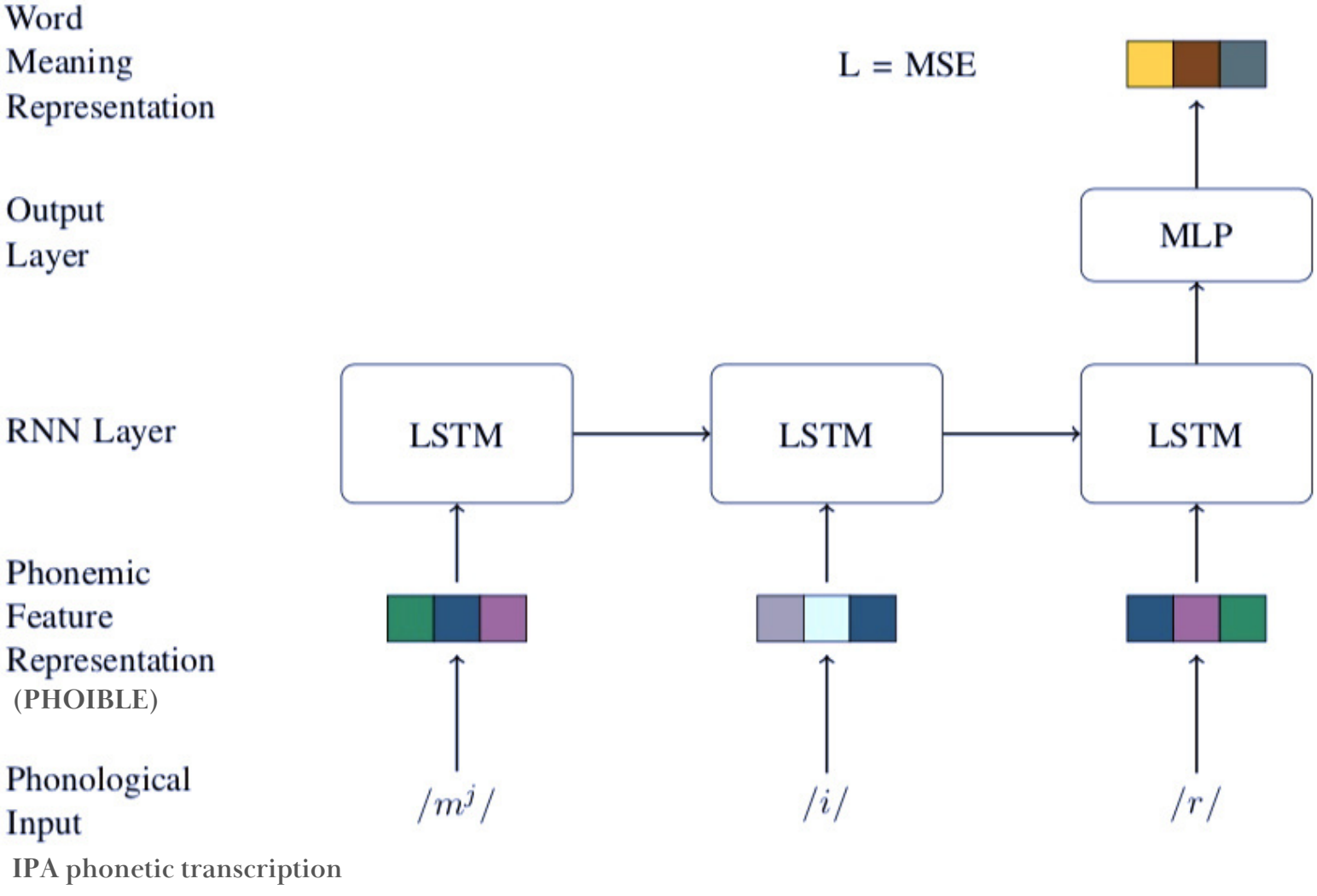
Recurrent Neural Net
(ADAM + MSE)

1 LSTM layer
linear-tanh MLP

→ Learning the mapping between (A) **phonological** sequences and (B) **semantics** embeddings (meaning representation)

Procedure - Architecture

FastText cbow

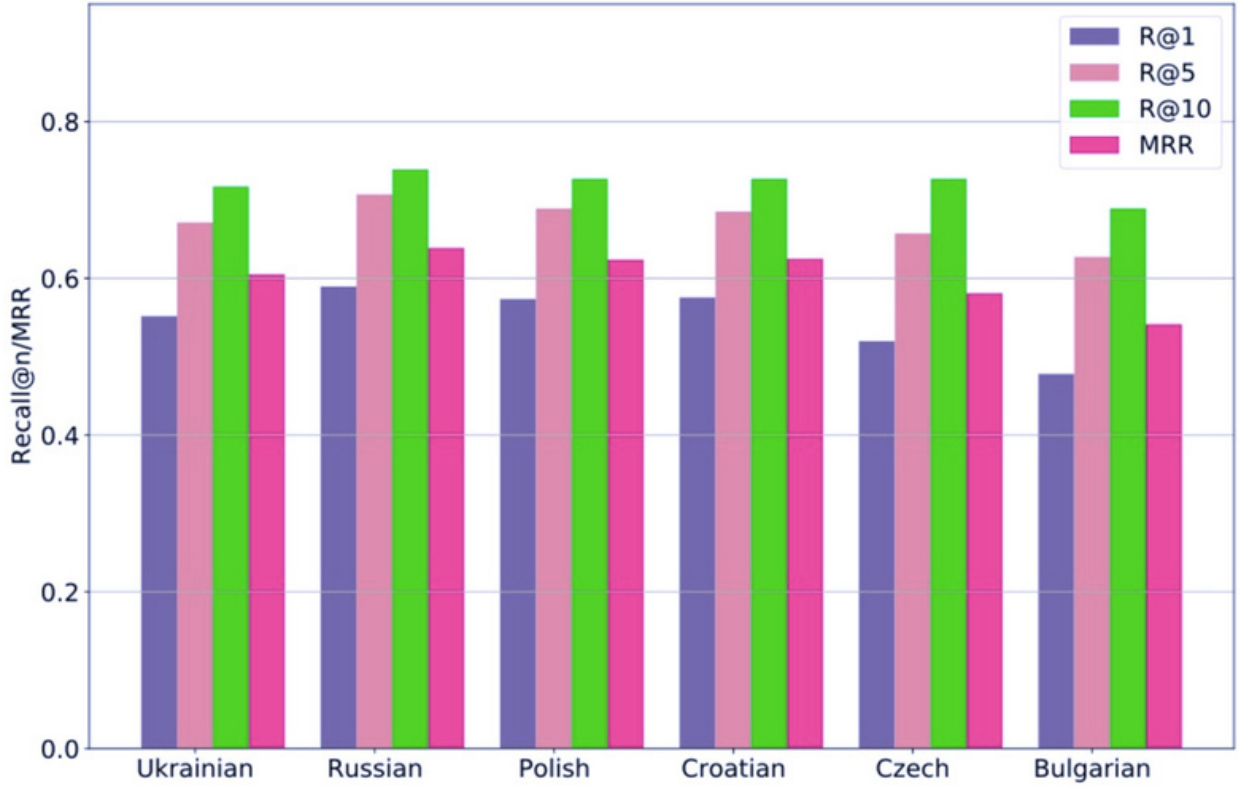


[502 common cross-lingual concepts words]

Procedure - Evaluation

Evaluation: cosine similarity between **embeddings**
Test set retrieval - [output]-[target in training set]

Monolingual [training language=evaluation language]:
[output]-[target in test set].



Procedure - Evaluation

Cross-lingual evaluation [training (L1) language≠evaluation (L2) language]:

Cosine similarity between **embeddings** - [output]-[target in L2 test] + [L2 target]-[L1 target] + [L1 target]-[possible outputs in training set]

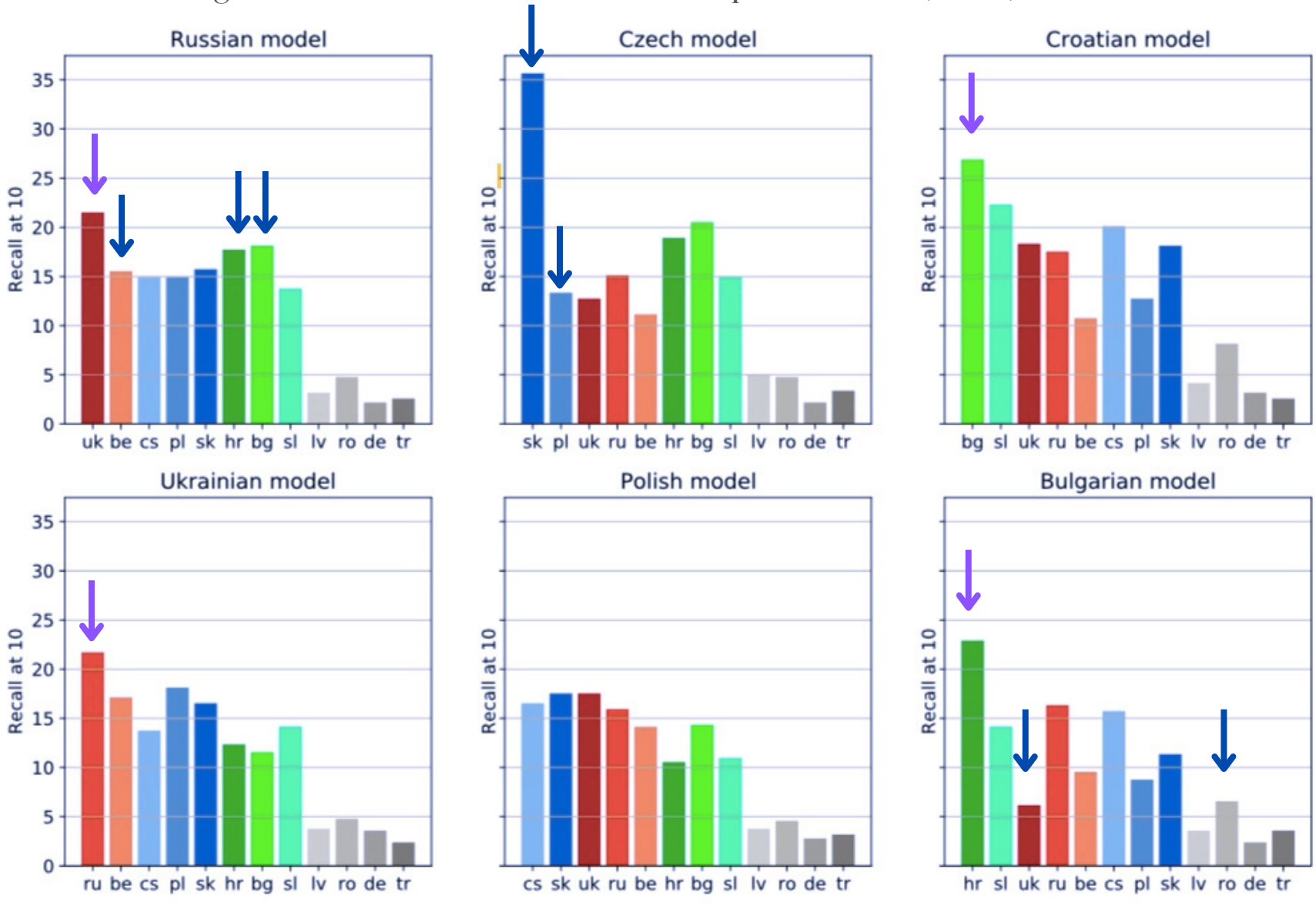
*training: Russian -- e.g. word **люди /ljudi/** 'people',*
testing: Czech concepts.
*evaluation: (1) compute the semantics representation of Czech **lidé /lid@/** 'people' (2) estimate its similarity to test sequences in Russian with the target meaning representation being that of the Russian word **люди /lj u d i/**.*

Concept	Russian		Czech		Bulgarian	
	Orth	IPA	Orth	IPA	Orth	IPA
EAR	ухо	/u x ɑ/	ucho	/u x o/	ухо	/u x ɔ/
NOSE	нос	/n o s/	nos	/n o s/	нос	/n ɔ s/

Performance Results

Performance metrics:

- Average recall at 1, 5 and 10 & Mean Reciprocal Rank (MRR)

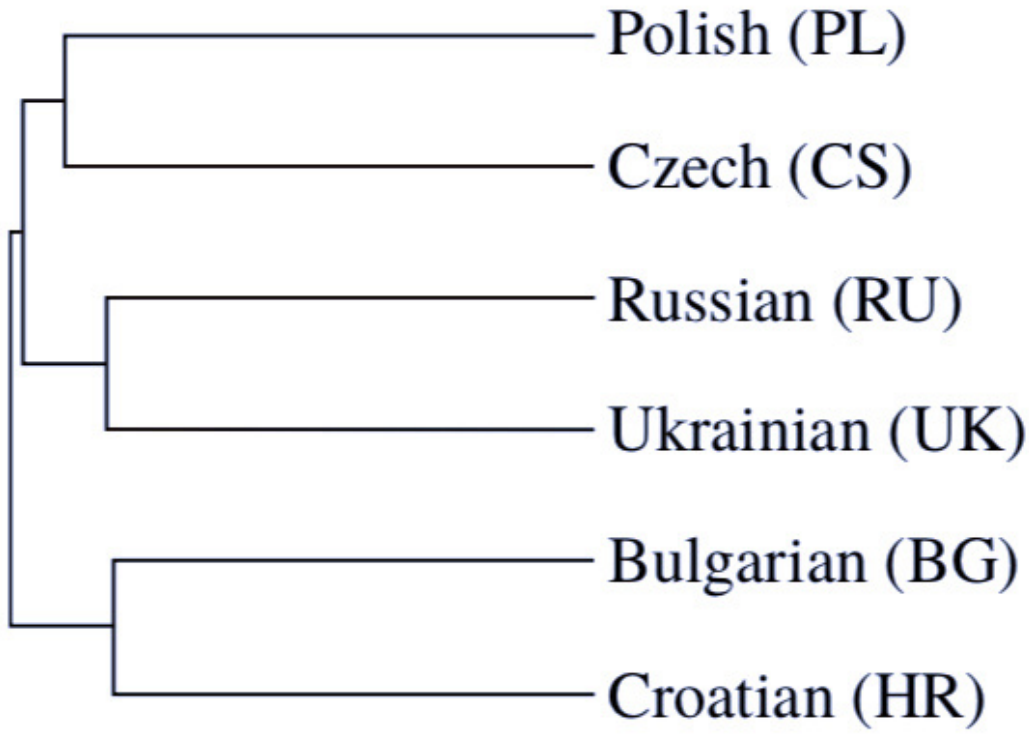


Ukrainian – uk, Russian – ru, Belarusian – be, Czech – cs, Polish – pl, Slovak – sk, Croatian – hr, Bulgarian – bg, Slovene – sl, Latvian – lv, Romanian – ro, German – de, Turkish – tr

Linguistic Predictors Results

Linguistic predictors:

Levenshtein distance and PWLD & Hierarchical clustering among Recall@10



Hierarchical clustering

Overcomes

Additional Positive Overcomes:

- Long-short term memory network (LSTM) and multi-layer perceptron (MLP) are proven for cognitive validity in predicting human behaviour and cognitive features
- Adherence to multiple-trace theory, embracing continuity and coupling between speech perception, production, and memory.
- Ruled system discrete phonological representation, tackling the acoustic-phonetic invariance challenge

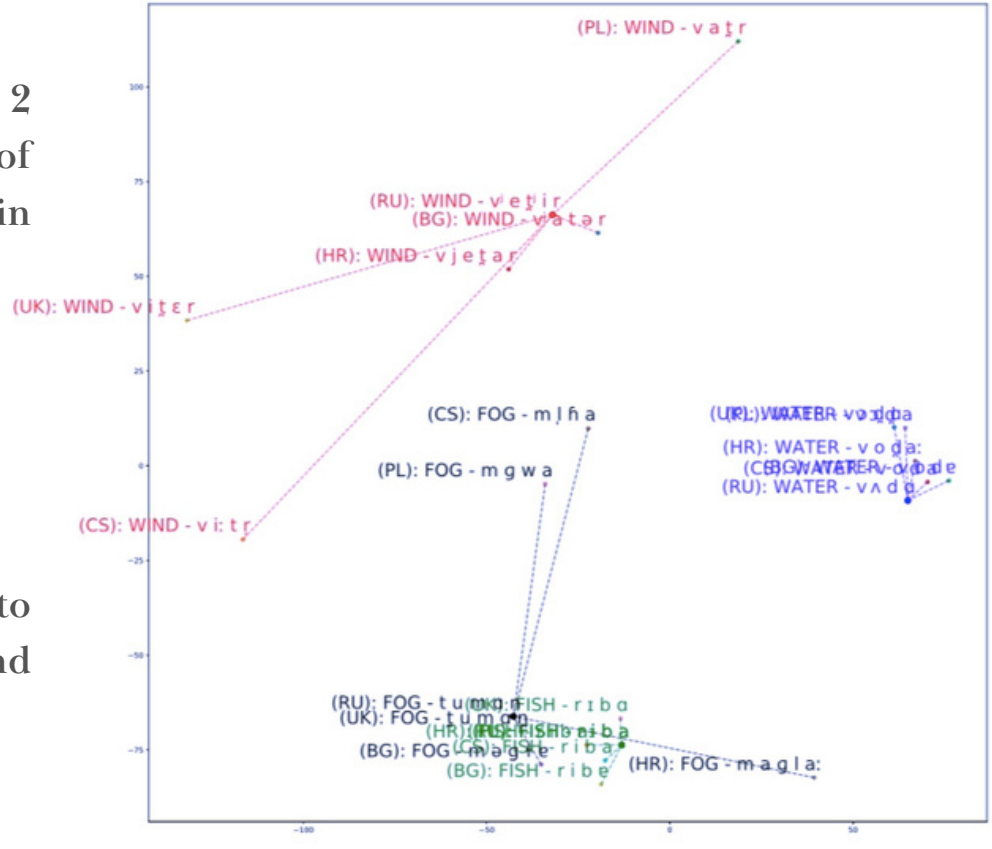
Overcomes

Additional Negative Overcomes:

- PWLD, has a lower correlation with retrieval metrics than LD

	R@10	MRR	cos sim	LD	PWLD
R10		0.98***	0.5***	-0.74***	-0.57***
MRR			0.5***	-0.75***	-0.56***
cos sim				-0.29*	-0.44***
LD					0.8***
PWLD					

- t-SNE clustering results raise 2 concerns about the alignment of similarly sounding words in different languages.



- Lack of accessibility to implementation code and training/testing data

Suggestions and Remarks

Remarks:

- Claim for better explanation of semantic space computations (more examples);
- Claim for a visual representation of Slavic languages family tree.

Suggestions

- Underlie the quality of the model's architecture to enhance validity;
- Enhance a theoretical/historical framework in respect to the used model.

Suggestions and Remarks

Suggestions for future directions:

- Bilingual training set towards an observation whether multilingualism enhance inter-comprehension;
- Explore a behavioural paradigm for model testing (multilingual modelling tasks as such hinting word meaning for L2 languages inputs, computing word similarity assessments);
- Asymmetrical intelligibility (Spanish and Portuguese)

Thank you for your attention!