

Frame-Based Continuous Lexical Semantics through Exponential Family Tensor Factorization and Semantic Proto-Roles

Francis Ferraro
Adam Poliak
Ryan Cotterell
Benjamin Van Durme

Understanding Participating in Actions

Word embeddings robustly capture word similarity & associativity. What does it mean for “the paper” to “reflect the truth?” And how do we represent it?

The **paper** reflected the **truth**.

Frame Semantics Provide Structured Word Meaning

Understanding through defined and structured concepts (Minsky '74; Fillmore '76, '82)

“reflect” triggers COMMUNICATION frame

Argument	ROLE	Meaning
The paper	COMMUNICATOR	The conveyer of MESSAGE
the truth	MESSAGE	A set of beliefs to convey
---	MEDIUM	How MESSAGE is conveyed

Examples: FrameNet and PropBank

Semantic Proto-Roles Decompose Categorical Roles

Dowty '91, Reisinger et al. '15 : Describe semantic arguments as properties of participating in an action

Property, re: “reflect”	The paper	the truth.
AWARENESS	+	-
MANIPULATED	-	+
SENTIENT	-	-
VOLITIONAL	+	-

Frame Embeddings from Tensor Factorization

Further generalize Cotterell et al. '17's 3-tensor factorization

$$p(t_i, w_j, c_1, \dots, c_K) \propto \exp(\mathbf{1}^T (t_i \odot w_j \odot c_1 \odot \dots \odot c_K))$$

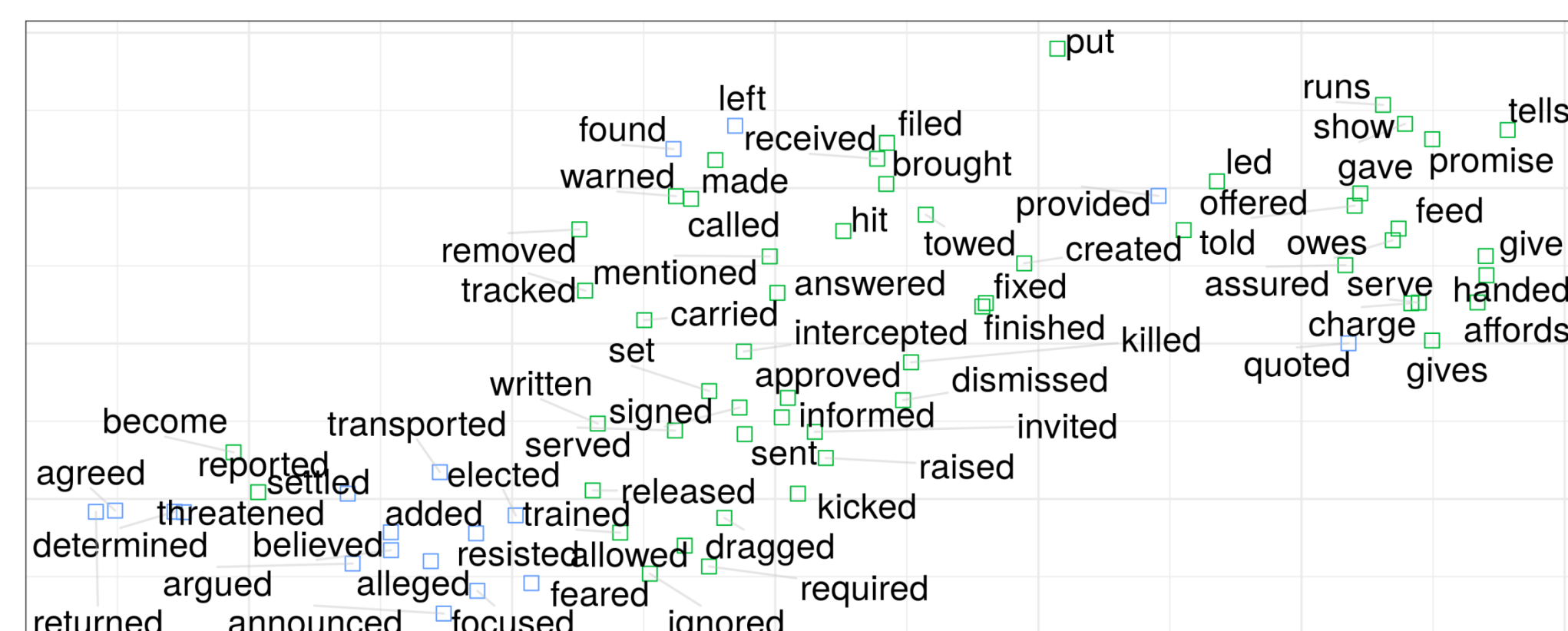
 /fmof/tensor-factorization

Evaluating Attributive Embeddings

QVEC: Correlate learned and oracle ontology vectors (Tsvetkov et al. '15)

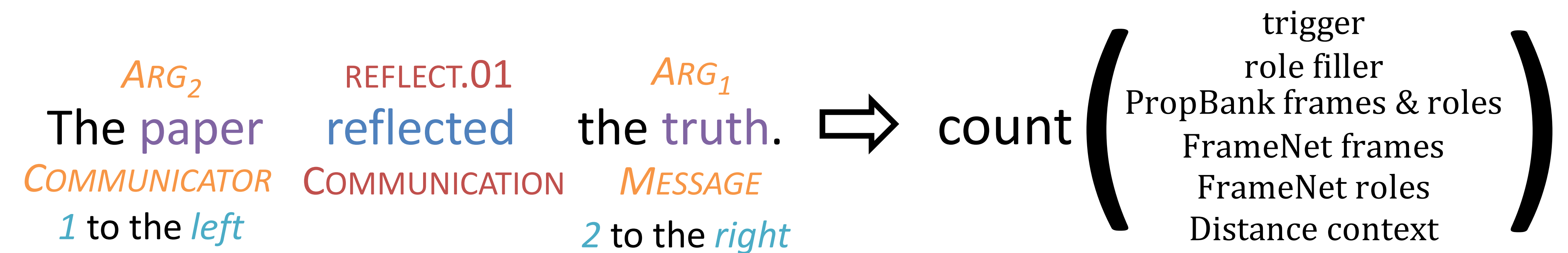
Use SPR annotations as ontology:

- Properties & syntax are coordinates: 80 total
- SPR Likert → 0-5 rating
- Sum & normalize over ratings



Extracting Frame Counts from Large Corpora

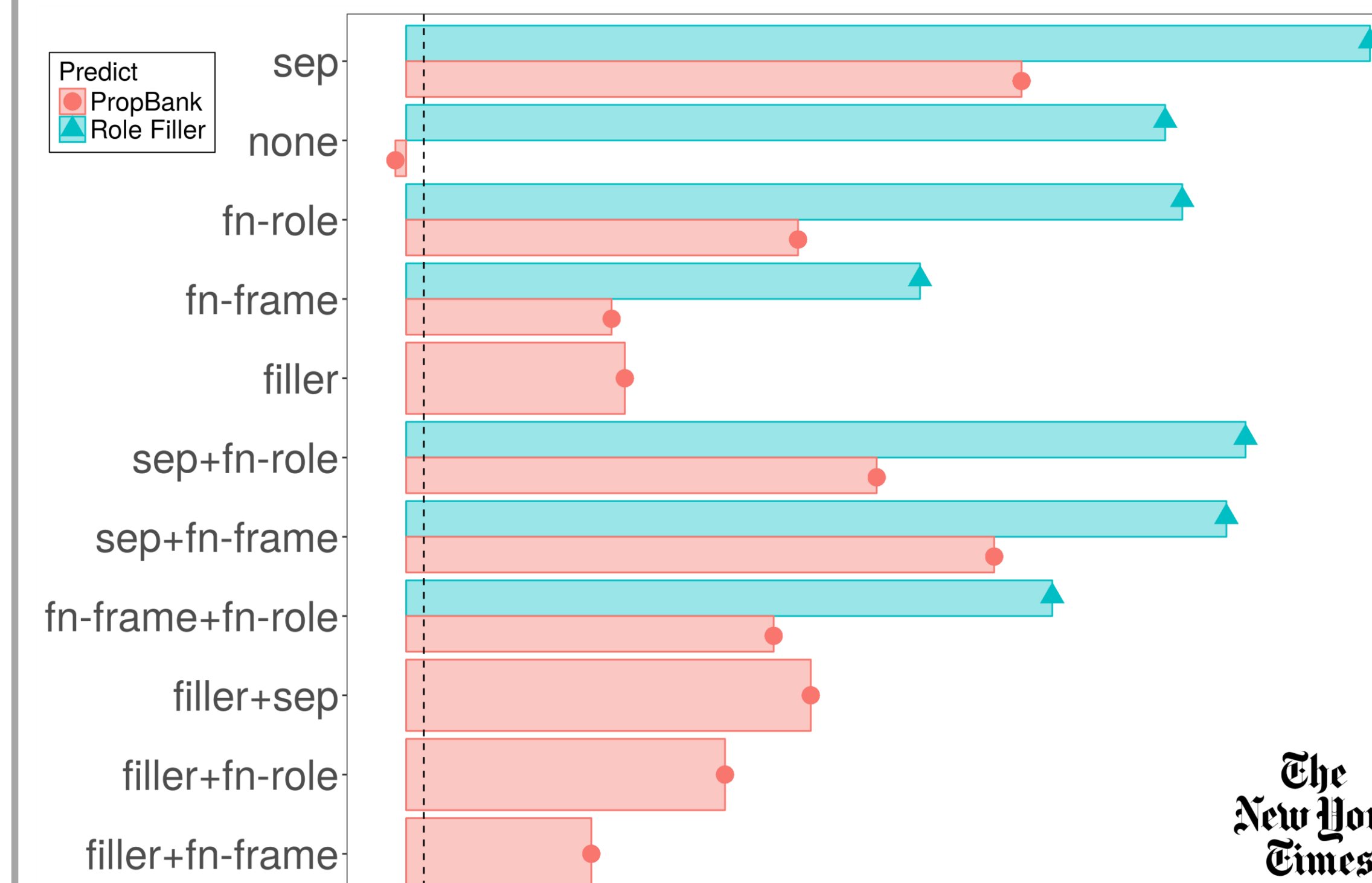
- Record every word triggering FRAMES...
- Each ROLE's fillers...
- FRAMES and ROLES...
- And context



Extract counts from **Concretely Annotated Corpora**: 15M+ documents annotated with three semantic parses (Ferraro et al. '14; <http://hltcoe.github.io/>)



Higher SPR Correlations in Learned Frame Trigger Embeddings

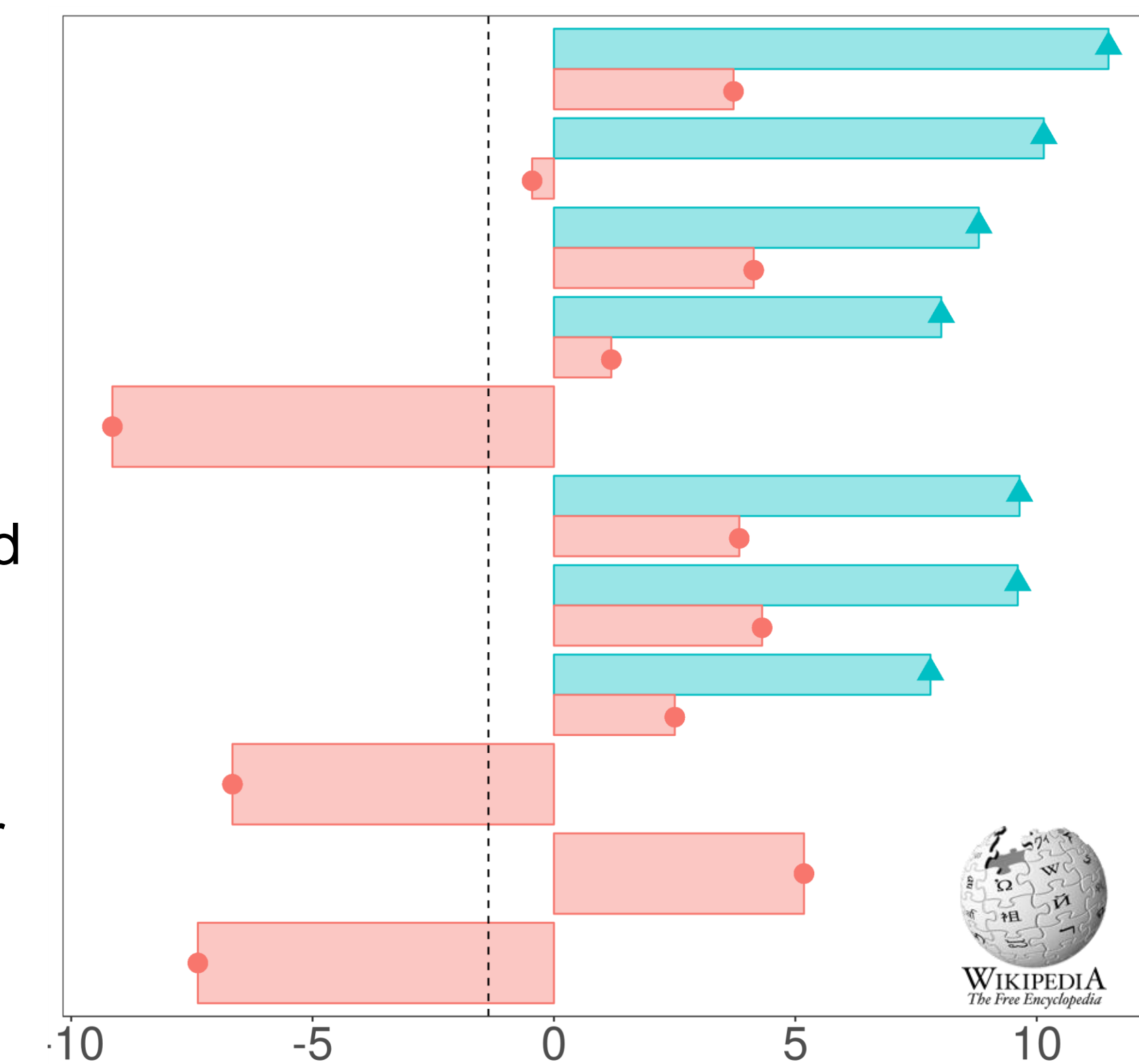


Frames give helpful extended context

Frames rival and outperform strong lexical models

Extended context and frames are complementary

Predict role fillers for higher correlation



Learning Paraphrases and Inflectional Relations

Under three newswire models, what triggers are most similar to **anticipated**?

Filler sep		Filler sep + fn-role		PropBank sep	
1 foresaw	6 pondered	1 aggravate	6 rubbed	1 anticipate	6 intimidated
2 figuring	7 kidded	2 scouts	7 analyzes	2 anticipating	7 separating
3 alleviated	8 constituted	3 jams	8 jailed	3 anticipates	8 separates
4 craved	9 uttering	4 appealing	9 discerned	4 stabbing	9 drag
5 jeopardized	10 forgiven	5 provoking	10 halved	5 separate	10 guarantee

Conclusion

Semantic frames obtained from large, disparate corpora can be used to learn enriched word vectors resulting in higher semantic proto-role based correlations.

