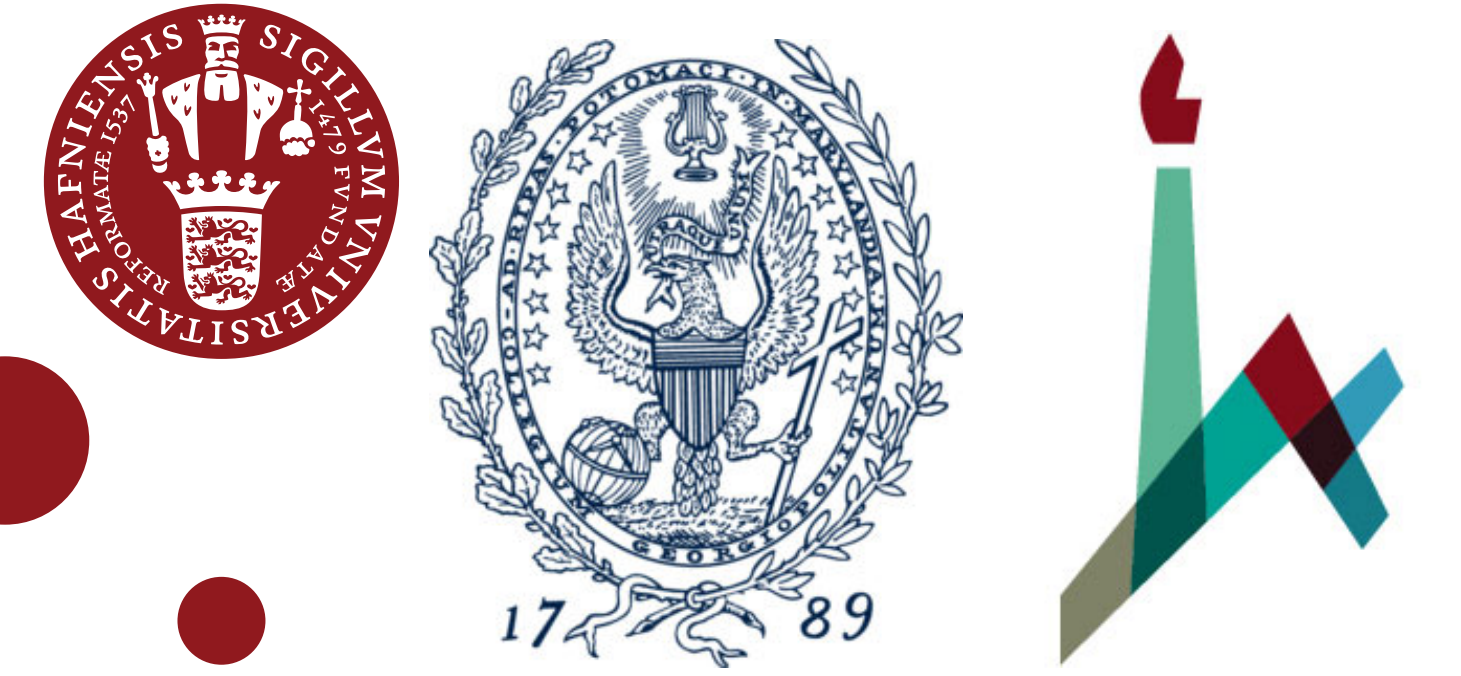


Comparison by Conversion

Reverse-Engineering UCCA from Syntax and Lexical Semantics

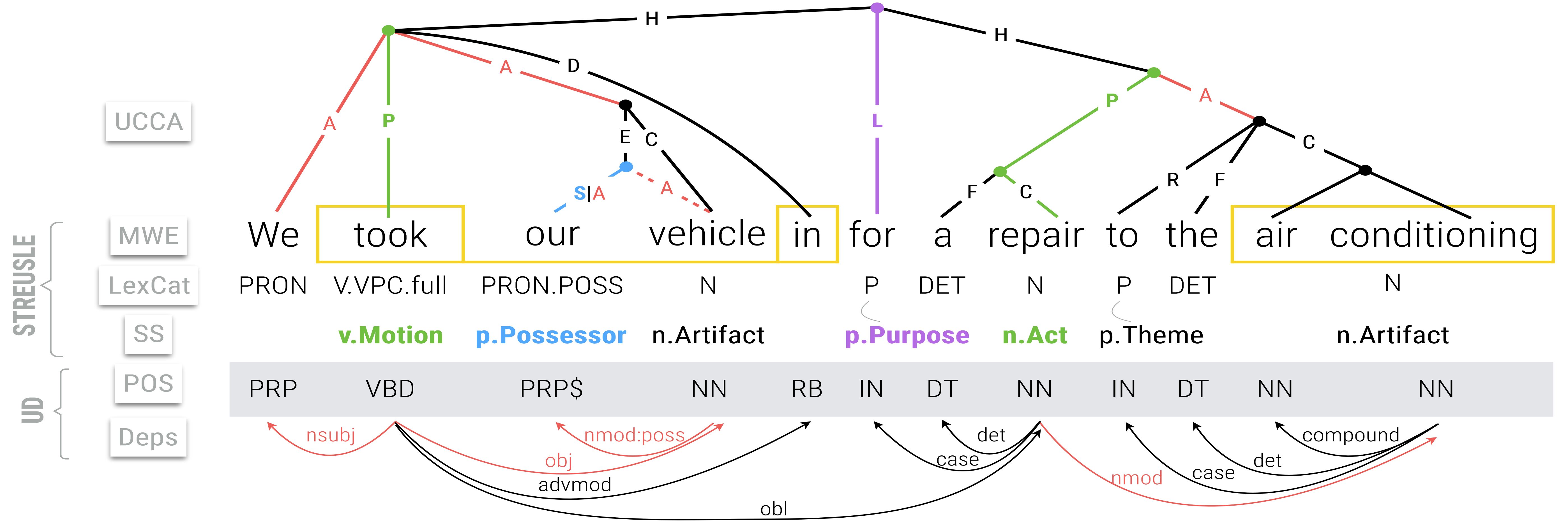


Daniel Hershcovich [◇] Nathan Schneider [♣] Dotan Dvir [♡]
 Jakob Prange [♣] Miryam de Lhoneux [◇] Omri Abend [♡]

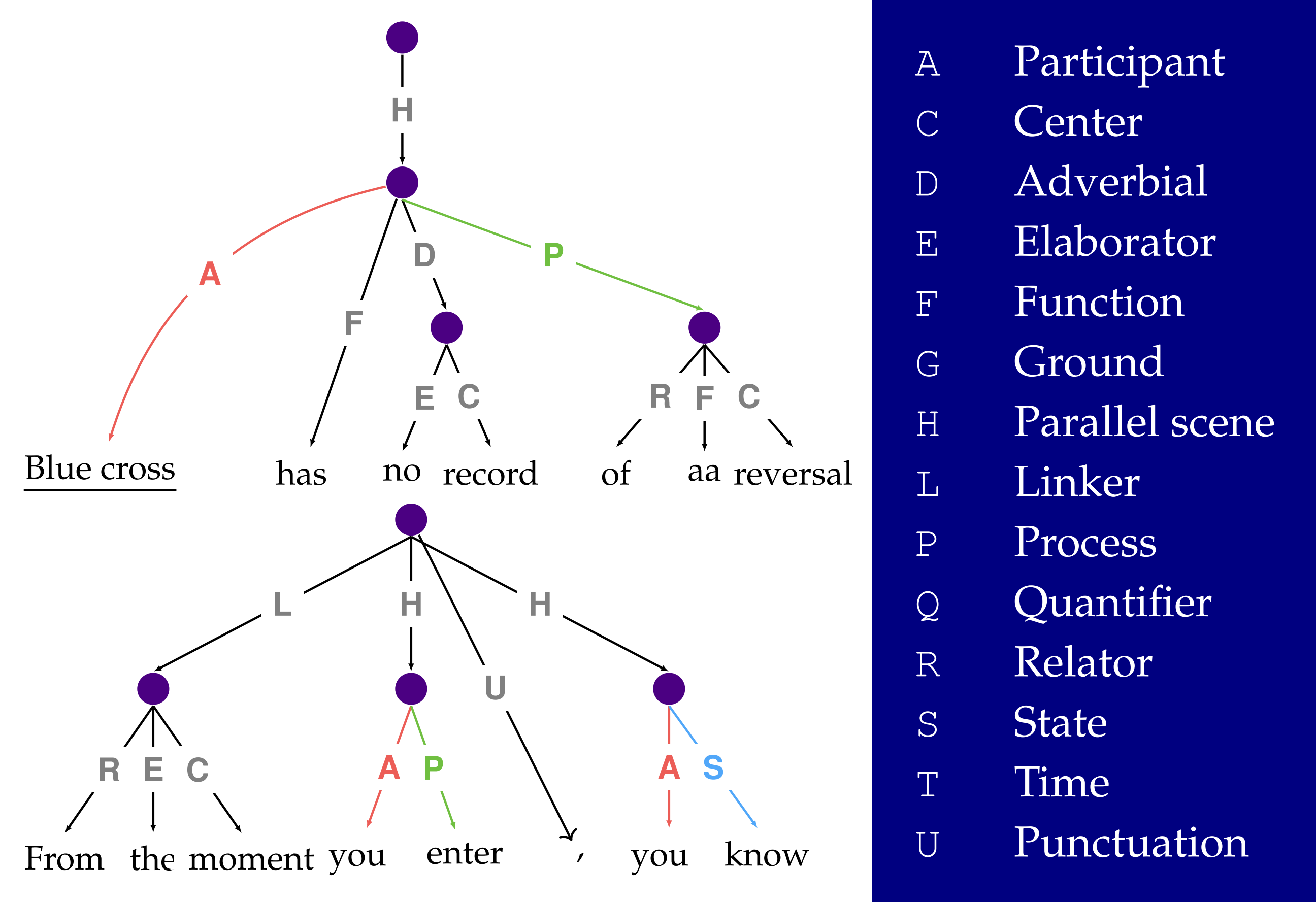
dh@di.ku.dk

◇ University of Copenhagen
 ♣ Georgetown University
 ♡ Hebrew University of Jerusalem

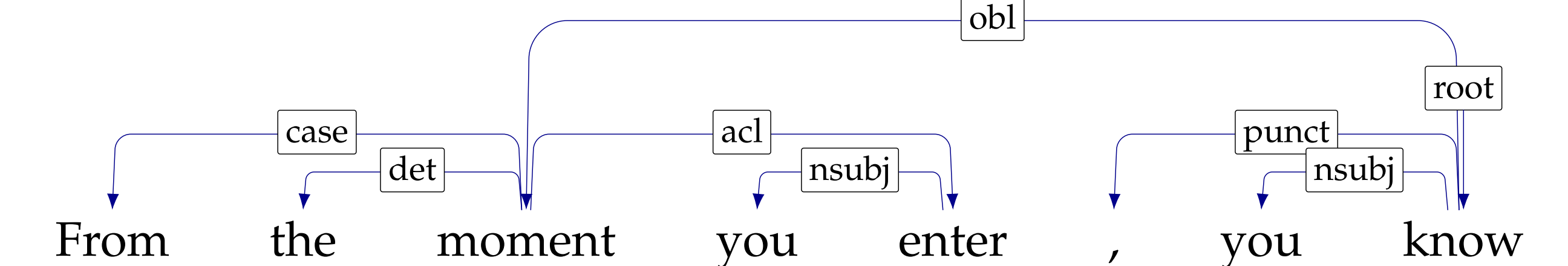
Linguistic and supervised conversion enable deep comparison between meaning representations.



UCCA [1]: cross-linguistic meaning representation.



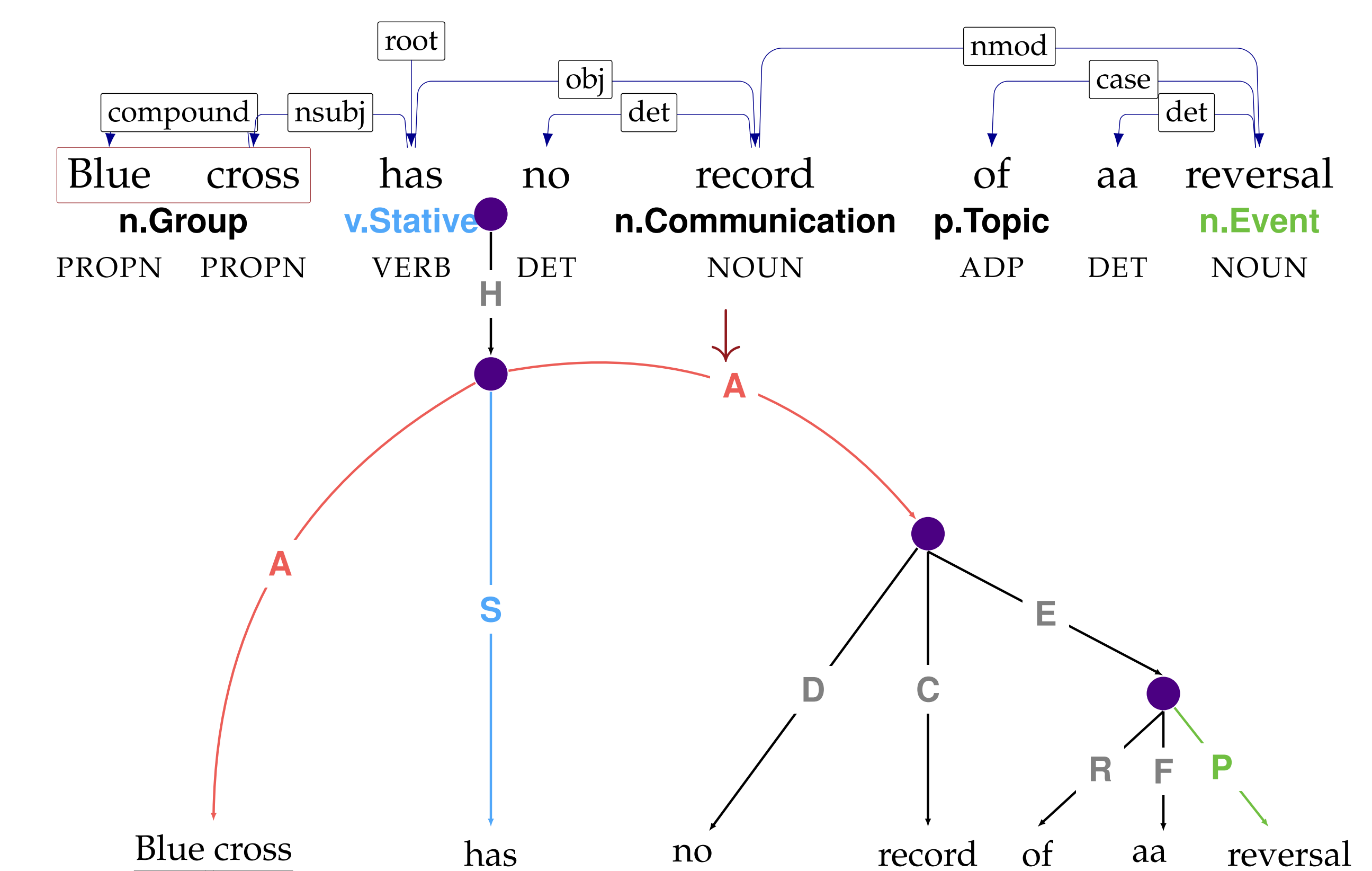
UD [4]: cross-linguistic syntactic representation.



STREUSLE [5]: English corpus of web reviews with comprehensive annotation of lexical semantics.

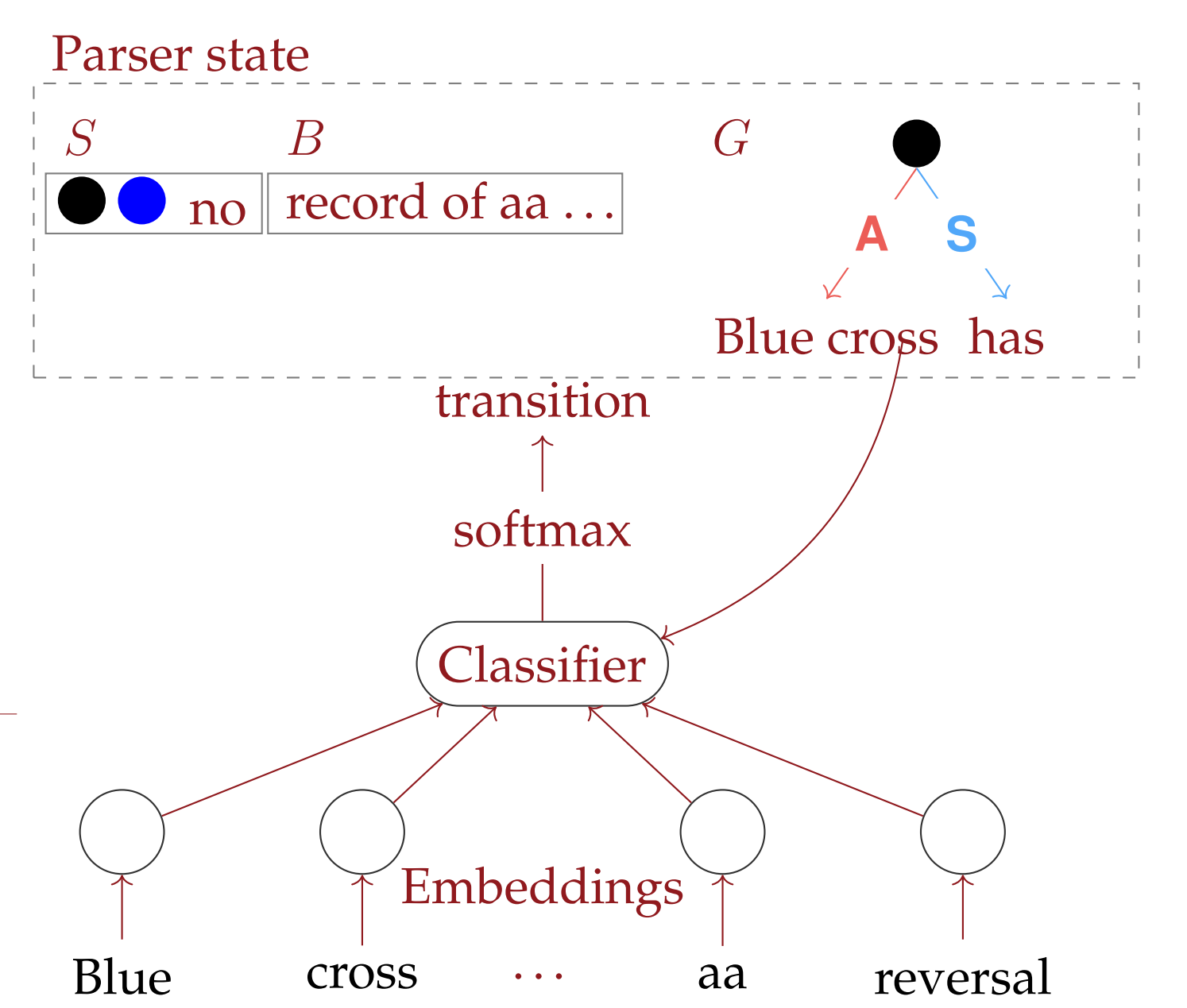
Converted UCCA STREUSLE	Gold UCCA
tap_water (unanalyzable) n.Substance	[E tap] [C water]
[P road_construction] n.Event	[A road] [P construction]
[F a] [C meal] [E [R on]] [F the] n.Food p.Locus	[F a] [C meal] [E [R on]] [F the] [C menu]]
[C menu]]]] n.Communication	
[P answered] [A [Q all]] [A my]] [C questions]] v.Communication p.Originator n.Communication p.Gestalt	[P answered] [A [D all]] [A my] [P questions]]

Rule-Based Conversion



Delexicalized Parsing

Transition-based parsers:
 TUPA [3],
 HIT-SCIR [2].



EWT reviews corpus:

	Train	Dev
# Sentences	2,723	554
# Tokens	44,804	5,394

Evaluation

	F1		Predicted Category																Gold Category		
	Primary	Remote	A	A G	A P	A S	C	D	D T	E	F	G	H	L	N	P	Q	R		S	T
Rule-based	71.7	44.2	758	4	7	12	17	11	9	4	1	6	1	14	1	1	19	150			
TUPA, delex	69.5	46.4	50	7	12	457	27	11	1	1	12	3	31	2	5	12	1	48			
UD only	64.4	35.9	10			12	280	40	8	12	2	2	6	4	1	7	18	20			
STREUSLE only	62.4	27.5	48	1		20	42	1	294	3	1	17	3	7	1	24	4	49			
HIT-SCIR, delex	67.9	41.6	3			3			613				1	1	3	4					
TUPA +UD+GloVe	71.7	47.0	2			2	6	2					2	6	2						
HIT-SCIR (BERT)	71.9	41.8	40	2		1	29	6	13	1	19	1	450	4	22	2	8	265			
HIT-SCIR (GloVe)	67.0	42.4				7			1	19	1		221	14	1	27					
+UD+STREUSLE	72.2	46.9				1	1		1	1			10	31	1						
			3			16	15	1	2	13	12	1	1	345	2	29	32				
			3			8	5	1						40							
			6			6							13	1	211	14	3				
			2			48	49	4	26	6			10	1	251	5					
			4	2	3	4	2	3					1			45	5				
			∅	148	1	3	6	136	60	100	32	124	9	2	65	12	34	23	6		

References

- [1] Omri Abend and Ari Rappoport. Universal Conceptual Cognitive Annotation (UCCA). In *ACL*, pages 228–238, August 2013.
- [2] Wansiang Che, Longxu Dou, Yang Xu, Yuxuan Wang, Yijia Liu, and Ting Liu. HIT-SCIR at MRP 2019: A unified pipeline for meaning representation parsing via efficient training and effective encoding. In *Proceedings of the Shared Task on Cross-Framework Meaning Representation Parsing at the 2019 Conference on Computational Natural Language Learning*, pages 76–85, Hong Kong, China, 2019.
- [3] Daniel Hershcovich, Omri Abend, and Ari Rappoport. A transition-based directed acyclic graph parser for UCCA. In *ACL*, pages 1127–1138, July 2017.
- [4] Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers, and Daniel Zeman. Universal Dependencies v2: An evergrowing multilingual treebank collection. In *LREC*, pages 4034–4043, May 2020.
- [5] Nathan Schneider, Jena D. Hwang, Vivek Srikumar, Jakob Prange, Austin Blodgett, Sarah R. Moeller, Aviram Stern, Adi Bitan, and Omri Abend. Comprehensive supersense disambiguation of English prepositions and possessives. In *ACL*, pages 185–196, 2018.