Judging Paradigmatic Relations:
a New Collection of English Ratings

Gabriella Lapesa[1,2], Sabine Schulte im Walde[2], Stefan Evert[3]
University of Osnabrück[1], University of Stuttgart[2], FAU Erlangen-Nürnberg[3]
glapesa@uos.de

The Collection

Semantically related English word pairs, rated for the strength of the semantic relation holding between them

Part of a **larger project**, whose goal is to characterize paradigmatic relations **cross-linguistically**

- ➡ **German:** (Scheible and Schulte im Walde, 2014), IMS Stuttgart
- ➡ **Italian**: Computational Linguistics Lab, University of Pisa (collection ongoing)

paradigmatic relations: synonymy, antonymy, hypernymy

3 parts of speech: nouns, verbs, adjectives

degrees of relatedness
Target: *artist*
Relation: synonymy
painter (strongly related)
creator (weakly related)
scientist (negatively related).

+

directionality
For every ⟨*target,relation,relatum*⟩ triple, we collected forward and backward ratings (e.g., *artist-synonym-painter* vs. *painter-synonym-artist*)

+

target selection based on a two-step process

Step 1: Generation Experiment

On Amazon Mechanical Turk, native speakers have been asked to **generate related words for 99 English targets per part-of-speech**

Random selection of targets from WordNet (Miller, 1995) with a **stratified sampling technique** (Scheible and Schulte im Walde, 2014). Criteria:

- **polysemy class**: I) one sense; II) two senses; III) > 3 senses
- **frequency classes**: I) low (200–2,999); II) mid (3,000–9,999); III) high (≥10,000)
- **size of the WordNet semantic class**

Some Examples

- **Target: "aircraft"**
  - **SYN:** plane (4), airplane (3), airship (1), balloon (1) helicopter (1)
  - **ANT:** car (3), watercraft (2), submarine (1), ship (1), landcraft (1), boat (1), _ (1)
  - **HYP:** vehicle (4), machine (2), transportation (2), plane (1), flyer (1)
- **Target: "uplift"**
  - **SYN:** raise (3), encourage (3), inspire (2), rise (1), support (1), elevate (1)
  - **ANT:** depress (4), put down (2), bring down (1), sink (1), defile (1), discourage (1)
  - **HYP:** raise (5), help (3), move (1), encourage (1) movement (1)
- **Target: "able"**
  - **SYN:** capable (5), competent (2), skilled (1), deft (1), apt (1)
  - **ANT:** unable (8), incapable (2)
  - **HYP:** capable (3), can (3), competent (1), functional (1), willing (1), qualifications (1)

Experiment conducted by Giulia Benotto and Alessandro Lenci (Computational Linguistics Lab, University of Pisa).

Step 2: Rating Experiment

⟨**Target, relation**⟩ pairs selected from the generated data
Goal: find pairs for which a full ⟨**target,weakly related, strongly related, not related**⟩ **tuple** was available. Criteria:

- at least 2 different relata had been generated
- a **strongly related** word (e.g, *painter*) was produced at least 4 times
- a **weakly related** word (e.g., *creator*) was produced twice or once
- a **negatively related** word was produced at least twice for the opposing relation: ANT for SYN and HYP, SYN for ANT (e.g., ⟨*painter, antonym, scientist*⟩)

286 ⟨target,relation⟩ pairs, 1,716 target / relation / related word / direction combinations ➡

	NOUN	ADJ	VERB
SYNONYMS	40	22	40
ANTONYMS	21	40	40
HYPERNYMS	27	20	36

Ratings collected with AMT:

- ✓ **"Do you think that the following two words are synonyms?"**
- ✓ 6 points scale (0-5)
- ✓ 10 workers per ⟨target,relation,relatum⟩ triple, per each direction

The Resource

Target	P.Cl	Freq	WN.Class	Relatum	Rel	Degree	Fw	Bw
goodbye	1	mid	communication	farewell	SYN	STRONG	4.6	4.9
goodbye	1	mid	communication	departure	SYN	WEAK	3.0	3.6
goodbye	1	mid	communication	hello	SYN	NOT	0.0	0.0
humble	3	max	all	proud	ANT	STRONG	5.0	4.3
humble	3	max	all	boastful	ANT	WEAK	4.9	4.9
humble	3	max	all	modest	ANT	NOT	0.6	0.4
to bill	3	mid	possession	to charge	HYP	STRONG	4.3	3.5
to bill	3	mid	possession	to notify	HYP	WEAK	3.1	2.9
to bill	3	mid	possession	to pay	HYP	NOT	0.8	0.8

Further information available with the resource:
- Z-score transformed ratings
- Full data per subject (e.g., for linear mixed effect analysis)
- Work Time in Seconds from AMT. Future work: work time as RT

Case Study: Directionality

Are some relations/parts-of-speech more asymmetric than others?
Method: item-based prediction with linear regression models

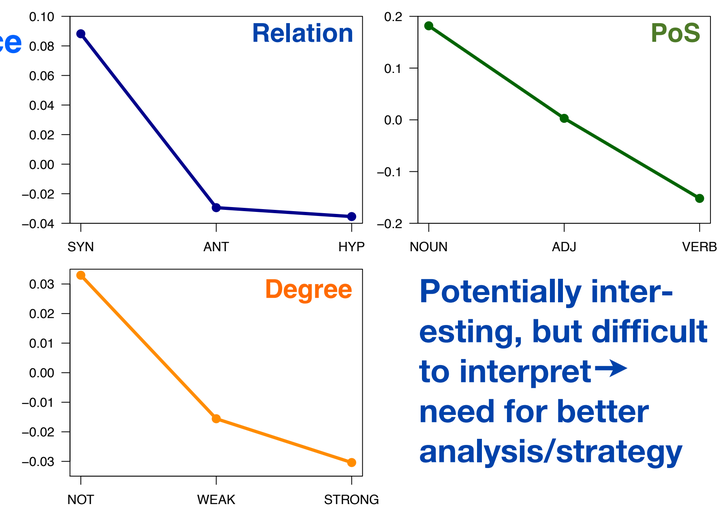
Dependent variable
Difference in Mean Ratings (Forward - Backward)

Predictors
Relation, Part of Speech, Degree of Relatedness (+ 2-way interactions), Target

Model 1: signed difference

Parameter	R^2	p
Relation	1.53	**
PoS	1.49	**
Degree	0.35	
Relation * Degree	1.61	**
Relation * PoS	0.58	
Degree * PoS	0.66	
Target	16.82	

Main effects & int., R^2 (23%)

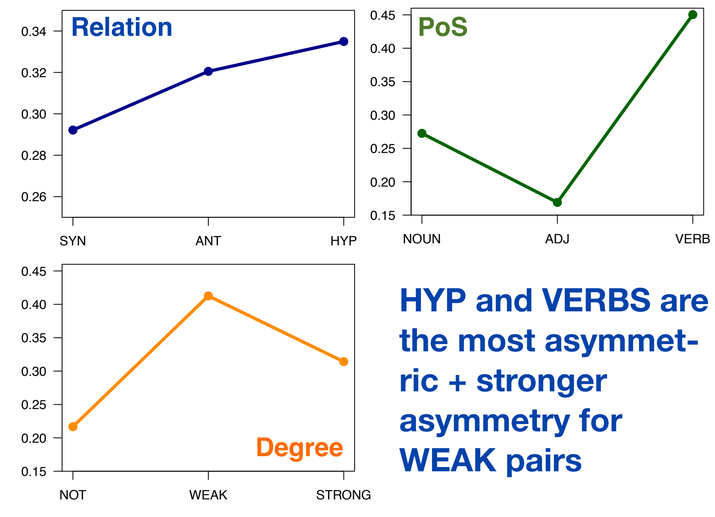


Potentially interesting, but difficult to interpret➔ need for better analysis/strategy

Model 2: abs. difference

Parameter	R^2	p
Relation	0.68	*
PoS	0.79	*
Degree	5.55	***
Relation * Degree	0.94	
Relation * PoS	0.35	
Degree * PoS	0.11	.
Target	21.29	

Main effects & int., R^2 (29.7%)



HYP and VERBS are the most asymmetric + stronger asymmetry for WEAK pairs

References & Acknowledgments

Silke Scheible and Sabine Schulte im Walde (2014). A Database of Paradigmatic Semantic Relation Pairs for German Nouns, Verbs, and Adjectives In: Proceedings of th COLING Workshop on Lexical and Grammatical Resources for Language Processing.
George A. Miller (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41.
We thank Giulia Benotto and Alessandro Lenci for providing us the generation data. This research was funded by the DFG Collaborative Research Centre SFB 732 (Gabriella Lapesa) and the DFG Heisenberg Fellowship SCHU-2580/1-1 (Sabine Schulte im Walde).