Paraphrase to Explicate:
Revealing Implicit Noun-Compound Relations

Vered Shwartz

Natural Language Processing Lab, Bar-Ilan University

May 2018

Noun Compounds

- Noun-compounds hold an implicit semantic relation between the head and its modifier(s).

Noun Compounds

- Noun-compounds hold an implicit semantic relation between the head and its modifier(s).
  - *apple cake*: cake *made of* apples

Noun Compounds

- Noun-compounds hold an implicit semantic relation between the head and its modifier(s).
  - *apple cake*: cake *made of* apples
  - *birthday cake*: cake *eaten on a* birthday

Noun Compounds

- Noun-compounds hold an implicit semantic relation between the head and its modifier(s).
  - *apple cake*: cake *made of* apples
  - *birthday cake*: cake *eaten on a* birthday

- They are like "text compression devices" [Nakov, 2013]

Noun Compounds

- Noun-compounds hold an implicit semantic relation between the head and its modifier(s).
  - *apple cake*: cake *made of* apples
  - *birthday cake*: cake *eaten on a* birthday

- They are like “text compression devices” [Nakov, 2013]
- We’re pretty good in decompressing them!

We are good at Interpreting Noun-Compounds

We are good at Interpreting Noun-Compounds

Interpreting new Noun Compounds

- Noun-compounds are prevalent in English, but most are rare

Interpreting new Noun Compounds

- Noun-compounds are prevalent in English, but most are rare
- We easily interpret noun-compounds we've never seen before

Interpreting new Noun Compounds

- Noun-compounds are prevalent in English, but most are rare
- We easily interpret noun-compounds we've never seen before
- What is a "*parsley cake*"?

Interpreting new Noun Compounds

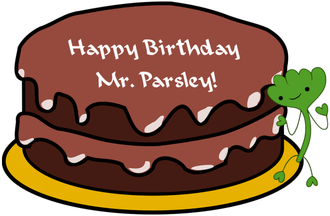
- Noun-compounds are prevalent in English, but most are rare
- We easily interpret noun-compounds we've never seen before
- What is a "*parsley cake*"?

1.

cake with/from parsley

(from `http://www.bazekalim.com`)

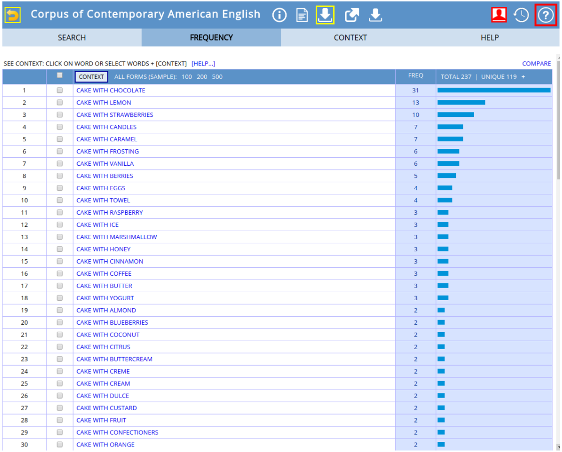
2.



cake for parsley

Interpreting new Noun Compounds

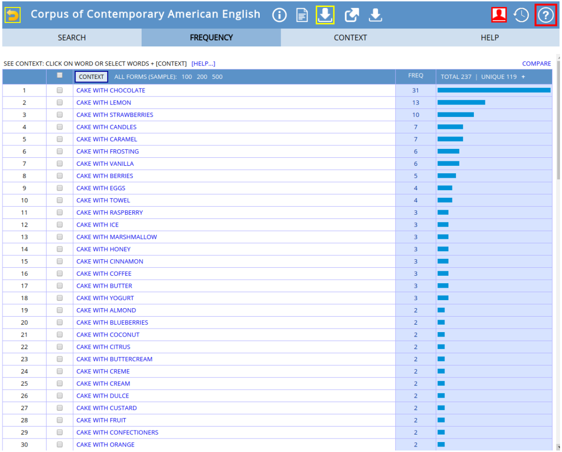
- What can cake be made of?



		CONTEXT ALL FORMS (SAMPLE): 100 200 500	FREQ	TOTAL 237 | UNIQUE 119 +
1		CAKE WITH CHOCOLATE	31	
2		CAKE WITH LEMON	13	
3		CAKE WITH STRAWBERRIES	10	
4		CAKE WITH CANDLES	7	
5		CAKE WITH CARAMEL	7	
6		CAKE WITH FROSTING	6	
7		CAKE WITH VANILLA	6	
8		CAKE WITH BERRIES	5	
9		CAKE WITH EGGS	4	
10		CAKE WITH TOWEL	4	
11		CAKE WITH RASPBERRY	3	
12		CAKE WITH ICE	3	
13		CAKE WITH MARSHMALLOW	3	
14		CAKE WITH HONEY	3	
15		CAKE WITH CINNAMON	3	
16		CAKE WITH COFFEE	3	
17		CAKE WITH BUTTER	3	
18		CAKE WITH YOGURT	3	
19		CAKE WITH ALMOND	2	
20		CAKE WITH BLUEBERRIES	2	
21		CAKE WITH COCONUT	2	
22		CAKE WITH CITRUS	2	
23		CAKE WITH BUTTERCREAM	2	
24		CAKE WITH CREME	2	
25		CAKE WITH CREAM	2	
26		CAKE WITH DULCE	2	
27		CAKE WITH CUSTARD	2	
28		CAKE WITH FRUIT	2	
29		CAKE WITH CONFECTIONERS	2	
30		CAKE WITH ORANGE	2	

Interpreting new Noun Compounds

- What can cake be made of?



	CONTEXT ALL FORMS (SAMPLE): 100 200 500	FREQ
1	CAKE WITH CHOCOLATE	31
2	CAKE WITH LEMON	13
3	CAKE WITH STRAWBERRIES	10
4	CAKE WITH CANDLES	7
5	CAKE WITH CARAMEL	7
6	CAKE WITH FROSTING	6
7	CAKE WITH VANILLA	6
8	CAKE WITH BERRIES	5
9	CAKE WITH EGGS	4
10	CAKE WITH TOWEL	4
11	CAKE WITH RASPBERRY	3
12	CAKE WITH ICE	3
13	CAKE WITH MARSHMALLOW	3
14	CAKE WITH HONEY	3
15	CAKE WITH CINNAMON	3
16	CAKE WITH COFFEE	3
17	CAKE WITH BUTTER	3
18	CAKE WITH YOGURT	3
19	CAKE WITH ALMOND	2
20	CAKE WITH BLUEBERRIES	2
21	CAKE WITH COCONUT	2
22	CAKE WITH CITRUS	2
23	CAKE WITH BUTTERCREAM	2
24	CAKE WITH CREME	2
25	CAKE WITH CREAM	2
26	CAKE WITH DULCE	2
27	CAKE WITH CUSTARD	2
28	CAKE WITH FRUIT	2
29	CAKE WITH CONFECTIONERS	2
30	CAKE WITH ORANGE	2

TOTAL 237 | UNIQUE 119 +

- Parsley (sort of) fits into this distribution

Interpreting new Noun Compounds

- What can cake be made of?

Corpus of Contemporary American English

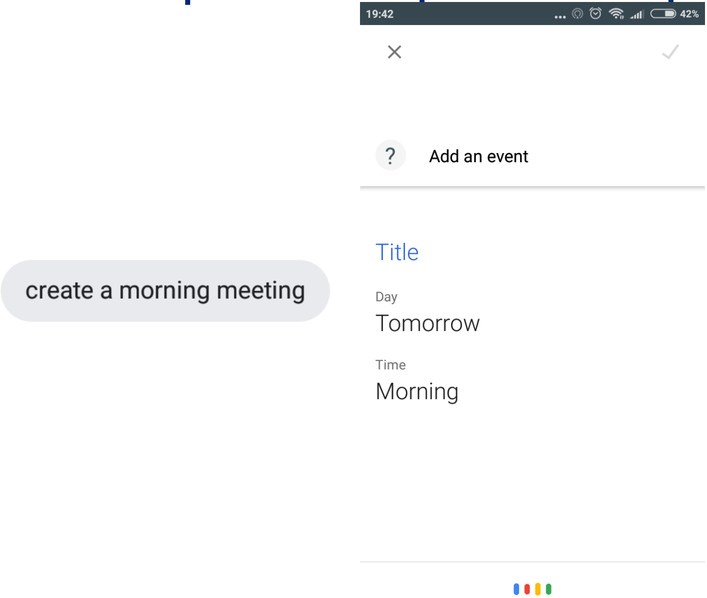
SEARCH | FREQUENCY | CONTEXT | HELP

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...] COMPARE

	CONTEXT ALL FORMS (SAMPLE): 100 200 500	FREQ	TOTAL 237 | UNIQUE 119 +
1	CAKE WITH CHOCOLATE	31	
2	CAKE WITH LEMON	13	
3	CAKE WITH STRAWBERRIES	10	
4	CAKE WITH CANDLES	7	
5	CAKE WITH CARAMEL	7	
6	CAKE WITH FROSTING	6	
7	CAKE WITH VANILLA	6	
8	CAKE WITH BERRIES	5	
9	CAKE WITH EGGS	4	
10	CAKE WITH TOWEL	4	
11	CAKE WITH RASPBERRY	3	
12	CAKE WITH ICE	3	
13	CAKE WITH MARSHMALLOW	3	
14	CAKE WITH HONEY	3	
15	CAKE WITH CINNAMON	3	
16	CAKE WITH COFFEE	3	
17	CAKE WITH BUTTER	3	
18	CAKE WITH YOGURT	3	
19	CAKE WITH ALMOND	2	
20	CAKE WITH BLUEBERRIES	2	
21	CAKE WITH COCONUT	2	
22	CAKE WITH CITRUS	2	
23	CAKE WITH BUTTERCREAM	2	
24	CAKE WITH CREME	2	
25	CAKE WITH CREAM	2	
26	CAKE WITH DULCE	2	
27	CAKE WITH CUSTARD	2	
28	CAKE WITH FRUIT	2	
29	CAKE WITH CONFECTIONERS	2	
30	CAKE WITH ORANGE	2	

- Parsley (sort of) fits into this distribution
- Similar to “selectional preferences” [Pantel et al., 2007]

We need Computers to Interpret Noun-Compounds



Noun-Compound Interpretation Tasks

- Compositionality Prediction
- **Noun-compound Paraphrasing**
- Noun-compound Classification

Compositionality Prediction

Given a noun-compound $w_1 w_2$, assign a score that measures to what extent its meaning is derived from the meanings of w_1 and w_2

- high: *apple cake*
- low: *spelling bee*

Noun-Compound Classification

Given a noun-compound $w_1 w_2$, classify the relation between the head w_2 and the modifier w_1 to one of a set of pre-defined relations

source

ground attack

olive oil

apple cake

part of

sea bass

boat whistle

rotor head

non-compositional

horse radish

baby sitting

hot dog

purpose

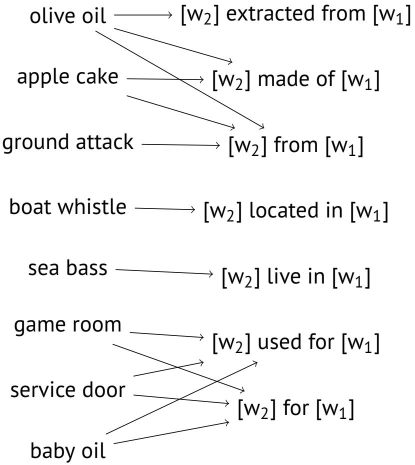
baby oil

game room

service door

Noun-Compound Paraphrasing

Given a noun-compound $w_1 w_2$, express the relation between the head w_2 and the modifier w_1 with multiple prepositional and verbal paraphrases



Noun-Compound Paraphrasing: Background

Task Definition

- Paraphrasing noun-compounds to multiple prepositional and verbal paraphrases [Nakov and Hearst, 2006]

Task Definition

- Paraphrasing noun-compounds to multiple prepositional and verbal paraphrases [Nakov and Hearst, 2006]
- SemEval 2013 task 4 [Hendrickx et al., 2013]:

Task Definition

- Paraphrasing noun-compounds to multiple prepositional and verbal paraphrases [Nakov and Hearst, 2006]

- SemEval 2013 task 4 [Hendrickx et al., 2013]:
  - Systems get a list of noun compounds

Task Definition

- Paraphrasing noun-compounds to multiple prepositional and verbal paraphrases [Nakov and Hearst, 2006]

- SemEval 2013 task 4 [Hendrickx et al., 2013]:
  - Systems get a list of noun compounds
  - Extract paraphrases from free text

Task Definition

- Paraphrasing noun-compounds to multiple prepositional and verbal paraphrases [Nakov and Hearst, 2006]
- SemEval 2013 task 4 [Hendrickx et al., 2013]:
  - Systems get a list of noun compounds
  - Extract paraphrases from free text
  - Rank them

Task Definition

- Paraphrasing noun-compounds to multiple prepositional and verbal paraphrases [Nakov and Hearst, 2006]

- SemEval 2013 task 4 [Hendrickx et al., 2013]:
  - Systems get a list of noun compounds
  - Extract paraphrases from free text
  - Rank them
  - Evaluated for correlation with human judgments

Prior Methods

- Based on corpus occurrences of the constituents:
  "cake made of apples"

Prior Methods

- Based on corpus occurrences of the constituents:
  "cake made of apples"
- SemEval task participants extracted them from Google N-grams

Prior Methods

- Based on corpus occurrences of the constituents:
  "cake made of apples"
- SemEval task participants extracted them from Google N-grams

- **Problems:**
  1. Many unseen NCs, no paraphrases in the corpus

Prior Methods

- Based on corpus occurrences of the constituents:
  "cake made of apples"
- SemEval task participants extracted them from Google N-grams

- **Problems:**
  1. Many unseen NCs, no paraphrases in the corpus
  2. Many NCs with just a few paraphrases

Prior Methods

- Based on corpus occurrences of the constituents:
  "cake made of apples"
- SemEval task participants extracted them from Google N-grams

- **Problems:**
  1. Many unseen NCs, no paraphrases in the corpus
  2. Many NCs with just a few paraphrases
- **Partial solutions:**
  1. [Van de Cruys et al., 2013]: generalize for unseen NCs with similar NCs, e.g. *pear tart* is similar to *apple cake*

Prior Methods

- Based on corpus occurrences of the constituents:

  "cake made of apples"
- SemEval task participants extracted them from Google N-grams

- **Problems:**
  1. Many unseen NCs, no paraphrases in the corpus
  2. Many NCs with just a few paraphrases
- **Partial solutions:**
  1. [Van de Cruys et al., 2013]: generalize for unseen NCs with similar NCs, e.g. *pear tart* is similar to *apple cake*
  2. [Surtani et al., 2013]: learn "is-a" relations between paraphrases: e.g. "[w_2] extracted from [w_1]" $\subset$ "[w_2] made of [w_1]"

Prior Methods

- Based on corpus occurrences of the constituents:
  "cake made of apples"
- SemEval task participants extracted them from Google N-grams

- **Problems:**
  1. Many unseen NCs, no paraphrases in the corpus
  2. Many NCs with just a few paraphrases
- **Partial solutions:**
  1. [Van de Cruys et al., 2013]: generalize for unseen NCs with similar NCs, e.g. *pear tart* is similar to *apple cake*
  2. [Surtani et al., 2013]: learn "is-a" relations between paraphrases: e.g. "[w_2] extracted from [w_1]" $\subset$ "[w_2] made of [w_1]"

- Our solution: multi-task learning to address both problems

Model

Multi-task Reformulation

- Previous approaches: predict a paraphrase for a given NC

Multi-task Reformulation

- Previous approaches: predict a paraphrase for a given NC
- Our model: multi-task learning problem

Multi-task Reformulation

- Previous approaches: predict a paraphrase for a given NC
- Our model: multi-task learning problem
- Training example {w_1 = apple, w_2 = cake, p = "[w_2] made of [w_1]"}

Multi-task Reformulation

- Previous approaches: predict a paraphrase for a given NC
- Our model: multi-task learning problem
- Training example {w_1 = apple, w_2 = cake, p = "[w$_2$] made of [w$_1$]"}
    1. Predict a paraphrase p for a given NC $w_1 w_2$:
       What is the relation between *apple* and *cake*?

Multi-task Reformulation

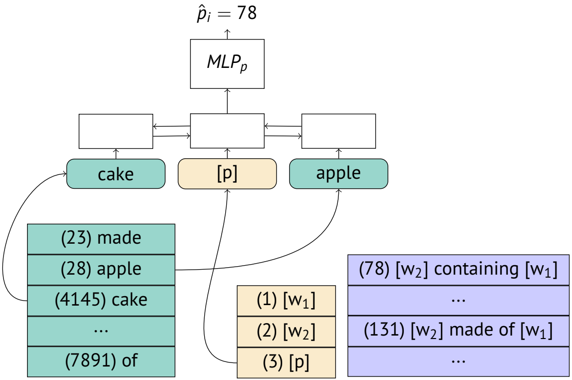
- Previous approaches: predict a paraphrase for a given NC
- Our model: multi-task learning problem
- Training example {w_1 = apple, w_2 = cake, p = "[w$_2$] made of [w$_1$]"}
    1. Predict a paraphrase p for a given NC $w_1 w_2$:
       What is the relation between *apple* and *cake*?
    2. Predict w_1 given a paraphrase p and w_2:
       What can *cake* be made of?

Multi-task Reformulation

- Previous approaches: predict a paraphrase for a given NC
- Our model: multi-task learning problem
- Training example {w_1 = apple, w_2 = cake, p = "[w_2] made of [w_1]"}
  1. Predict a paraphrase p for a given NC $w_1 w_2$:
     What is the relation between *apple* and *cake*?
  2. Predict w_1 given a paraphrase p and w_2:
     What can *cake* be made of?
  3. Predict w_2 given a paraphrase p and w_1:
     What can be made of *apple*?

Main Task (1): Predicting Paraphrases

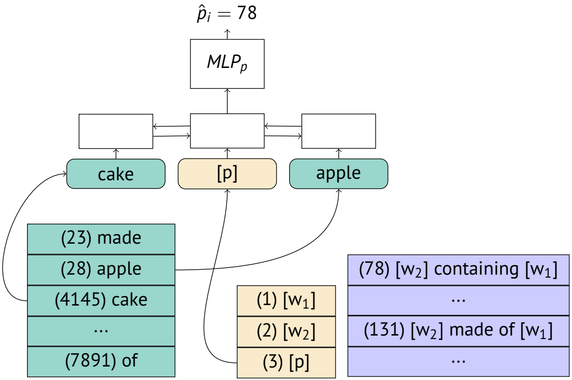
What is the relation between *apple* and *cake*?



- Encode placeholder [p] in “cake [p] apple” using biLSTM

Main Task (1): Predicting Paraphrases

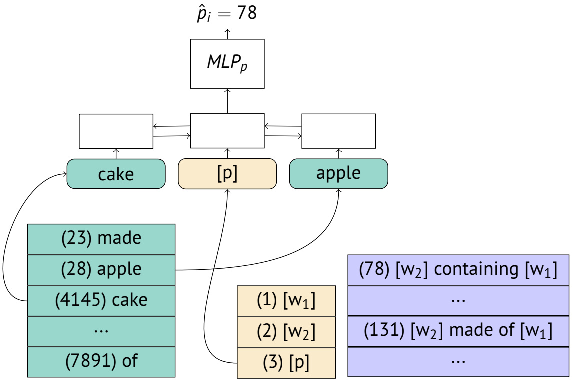
What is the relation between *apple* and *cake*?



- Encode placeholder [p] in "cake [p] apple" using biLSTM
- Predict an index in the paraphrase vocabulary

Main Task (1): Predicting Paraphrases

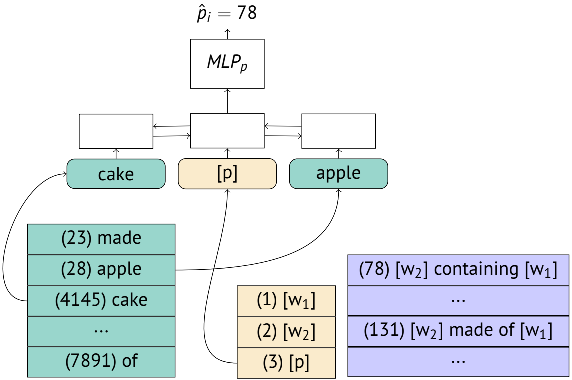
What is the relation between *apple* and *cake*?



- Encode placeholder [p] in "cake [p] apple" using biLSTM
- Predict an index in the paraphrase vocabulary
- Fixed word embeddings, learned placeholder embeddings

Main Task (1): Predicting Paraphrases

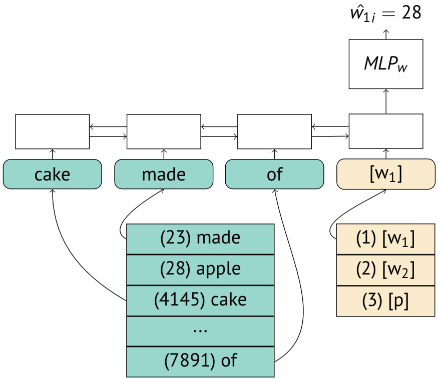
What is the relation between *apple* and *cake*?



- Encode placeholder [p] in “cake [p] apple” using biLSTM
- Predict an index in the paraphrase vocabulary
- Fixed word embeddings, learned placeholder embeddings
- (1) Generalizes NCs: *pear tart* expected to yield similar results

Helper Task (2): Predicting Missing Constituents

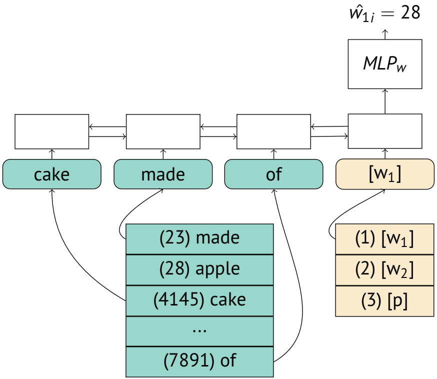
What can *cake* be made of?



- Encode placeholder in "cake made of [w_1]" using biLSTM

Helper Task (2): Predicting Missing Constituents

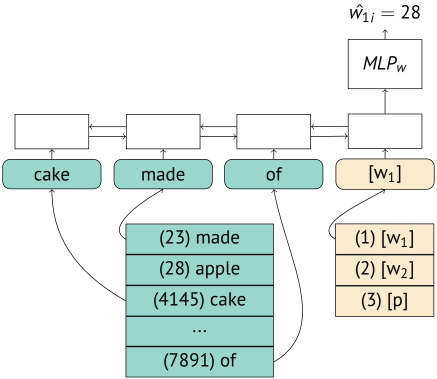
What can *cake* be made of?



- Encode placeholder in "cake made of [w_1]" using biLSTM
- Predict an index in the word vocabulary

Helper Task (2): Predicting Missing Constituents

What can *cake* be made of?



- Encode placeholder in "cake made of $[w_1]$" using biLSTM
- Predict an index in the word vocabulary
- (2) Generalizes paraphrases:

  "$[w_2]$ containing $[w_1]$" expected to yield similar results

Training Data

- Collected from Google N-grams

Training Data

- Collected from Google N-grams
- Input:
  - Set of NCs
  - Templates of POS tags (e.g. “[w_2] verb prep [w_1]”)

Training Data

- Collected from Google N-grams
- Input:
  - Set of NCs
  - Templates of POS tags (e.g. “[w_2] verb prep [w_1]”)
- Weighting by frequency and length

Training Data

- Collected from Google N-grams
- Input:
  - Set of NCs
  - Templates of POS tags (e.g. "[w_2] verb prep [w_1]")
- Weighting by frequency and length
- 136,609 instances

Qualitative Analysis: Predictions

$[w_1]$	$[w_2]$	**Predicted Paraphrases**
software	company	$[w_2]$ of $[w_1]$ $[w_2]$ to develop $[w_1]$ $[w_2]$ in $[w_1]$ industry $[w_2]$ involved in $[w_1]$
$[w_2]$	**Paraphrase**	**Predicted** $[w_1]$
company	$[w_2]$ engaged in $[w_1]$	management production computer business
Paraphrase	$[w_1]$	**Predicted** $[w_2]$
$[w_2]$ engaged in $[w_1]$	software	company firm engineer industry

Produces unseen triplets from similar triplets, e.g.:

firm in software industry

company in insurance industry

$\Rightarrow$ company in software industry

Evaluation 1: Paraphrasing

Model

- Predict top k paraphrases for each noun compound

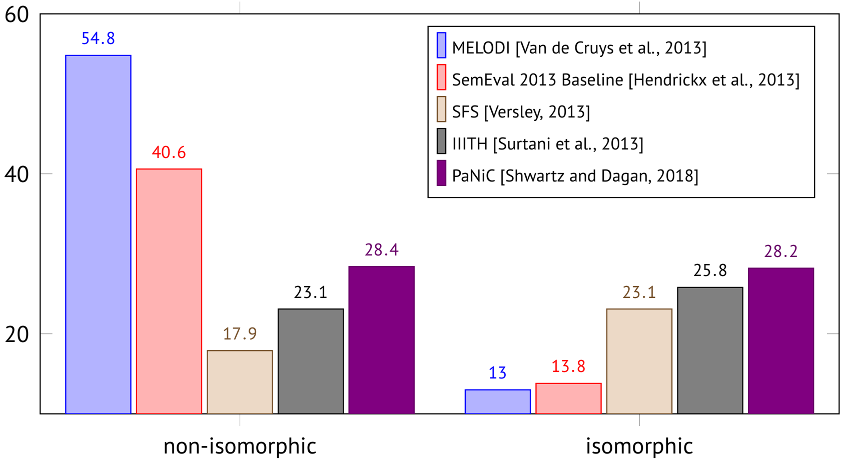
Model

- Predict top *k* paraphrases for each noun compound
- Learn to re-rank the paraphrases
  - to better correlate with human judgments

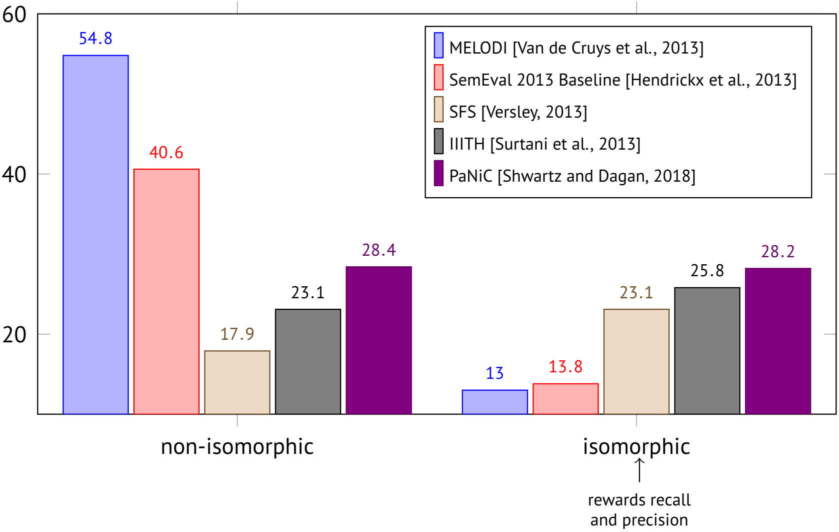
Model

- Predict top *k* paraphrases for each noun compound
- Learn to re-rank the paraphrases
  - to better correlate with human judgments
- SVM pair-wise ranking with the following features:
  - POS tags in the paraphrase
  - Prepositions in the paraphrase
  - Length
  - Special symbols
  - Similarity to predicted paraphrase

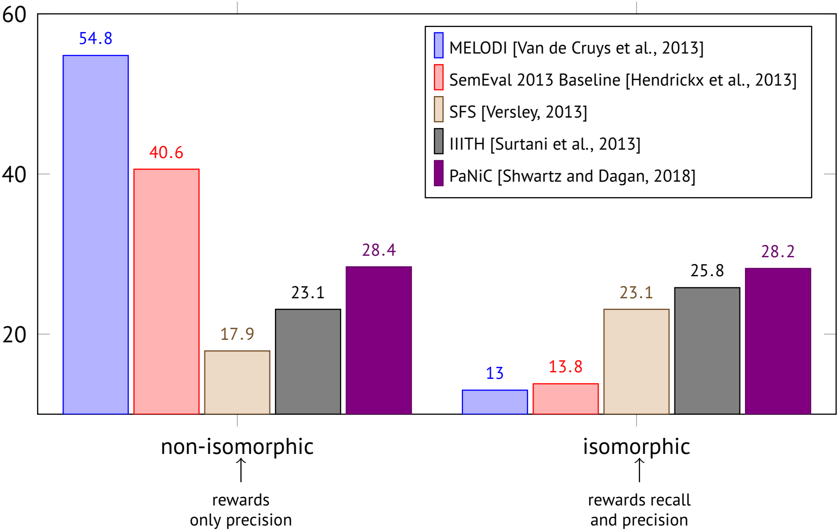
Results


60
40
20
54.8
40.6
17.9
23.1
28.4
13
13.8
23.1
25.8
28.2
MELODI [Van de Cruys et al., 2013]
SemEval 2013 Baseline [Hendrickx et al., 2013]
SFS [Versley, 2013]
IIITH [Surtani et al., 2013]
PaNiC [Shwartz and Dagan, 2018]
non-isomorphic
isomorphic

Results

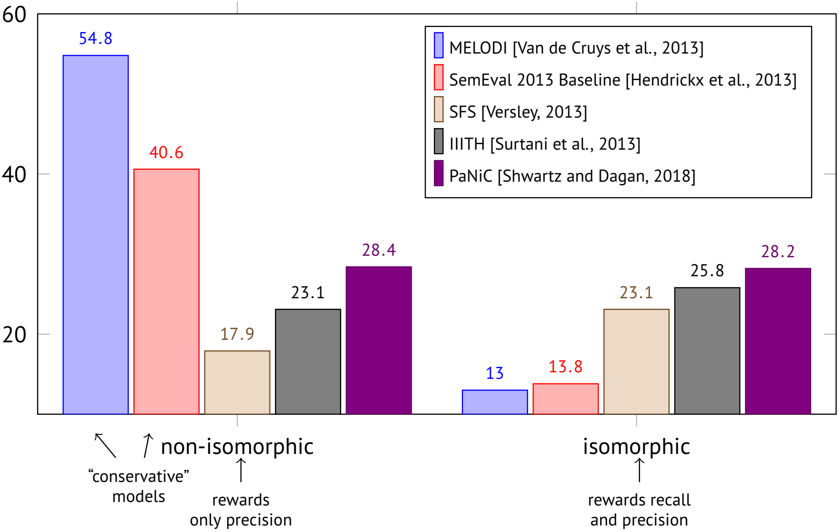

60
40
20
54.8
40.6
17.9
23.1
28.4
13
13.8
23.1
25.8
28.2
MELODI [Van de Cruys et al., 2013]
SemEval 2013 Baseline [Hendrickx et al., 2013]
SFS [Versley, 2013]
IIITH [Surtani et al., 2013]
PaNiC [Shwartz and Dagan, 2018]
non-isomorphic
isomorphic
rewards recall
and precision

Results



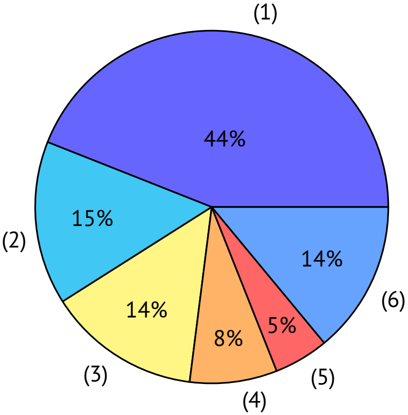
60
40
20
54.8
40.6
17.9
23.1
28.4
13
13.8
23.1
25.8
28.2
MELODI [Van de Cruys et al., 2013]
SemEval 2013 Baseline [Hendrickx et al., 2013]
SFS [Versley, 2013]
IIITH [Surtani et al., 2013]
PaNiC [Shwartz and Dagan, 2018]
non-isomorphic
isomorphic
rewards
only precision
rewards recall
and precision

Results



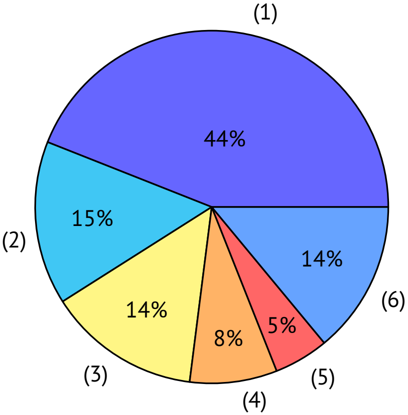
60
40
20
54.8
40.6
17.9
23.1
28.4
13
13.8
23.1
25.8
28.2
MELODI [Van de Cruys et al., 2013]
SemEval 2013 Baseline [Hendrickx et al., 2013]
SFS [Versley, 2013]
IIITH [Surtani et al., 2013]
PaNiC [Shwartz and Dagan, 2018]
non-isomorphic
isomorphic
"conservative" models
rewards only precision
rewards recall and precision

Error Analysis
False Positive



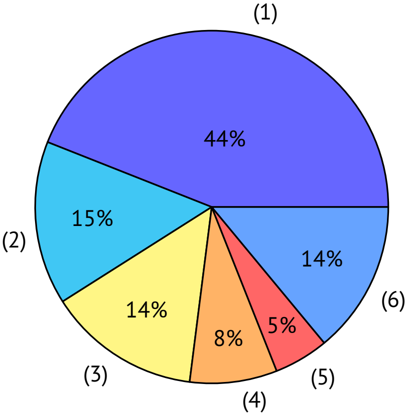
1. Valid, missing from gold-standard
("discussion by group")

Error Analysis
False Positive



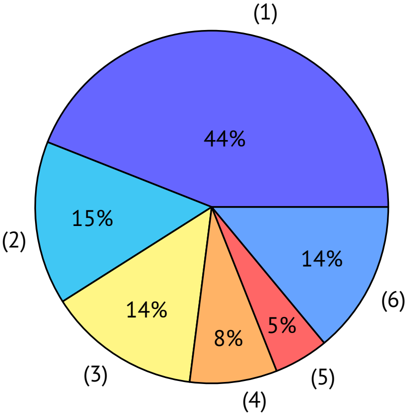
1. Valid, missing from gold-standard
   ("discussion by group")
2. Too specific
   ("life *of women in* community")

Error Analysis
False Positive



1. Valid, missing from gold-standard
   ("discussion by group")
2. Too specific
   ("life *of women in* community")
3. Incorrect prepositions
   E.g., n-grams don't respect syntactic structure: "rinse away the oil from baby 's head" ⇒ "oil from baby"

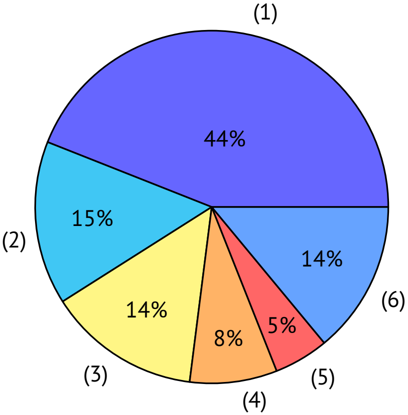
Error Analysis
False Positive



1. Valid, missing from gold-standard
   ("discussion by group")
2. Too specific
   ("life *of women in* community")
3. Incorrect prepositions
   E.g., n-grams don't respect syntactic structure: "rinse away the oil from baby 's head" ⇒ "oil from baby"
4. Syntactic errors

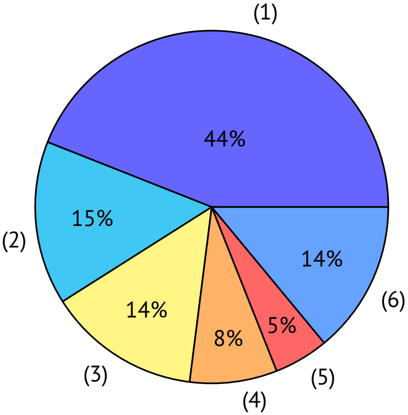
Error Analysis

False Positive



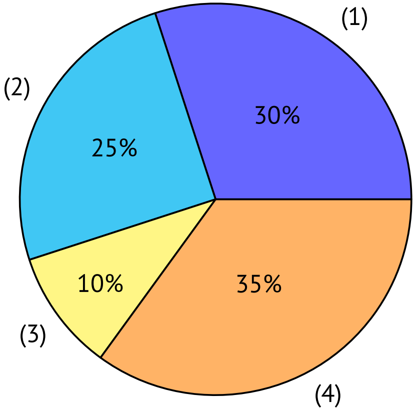
1. Valid, missing from gold-standard
   (“discussion by group”)
2. Too specific
   (“life *of women in* community”)
3. Incorrect prepositions
   E.g., n-grams don’t respect syntactic structure: “rinse away the oil from baby ’s head” ⇒ “oil from baby”
4. Syntactic errors
5. Borderline grammatical
   (“force of coalition forces”)

Error Analysis
False Positive



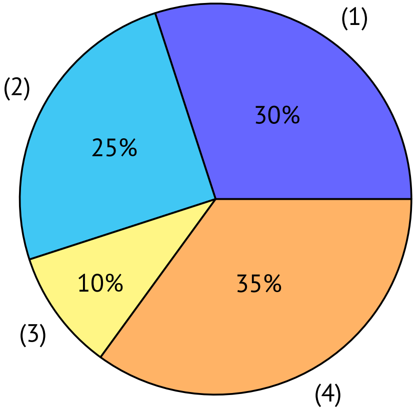
1. Valid, missing from gold-standard
   ("discussion by group")
2. Too specific
   ("life *of women in* community")
3. Incorrect prepositions
   E.g., n-grams don't respect syntactic structure: "rinse away the oil from baby 's head" ⇒ "oil from baby"
4. Syntactic errors
5. Borderline grammatical
   ("force of coalition forces")
6. Other errors

Error Analysis
False Negative



1. Long paraphrase ($n > 5$)

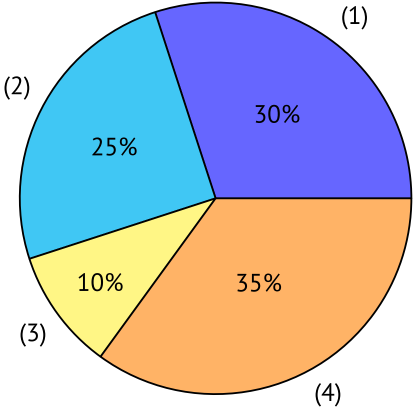
Error Analysis
False Negative



1. Long paraphrase ($n > 5$)
2. Determiners
   ("mutation of **a** gene")

Error Analysis

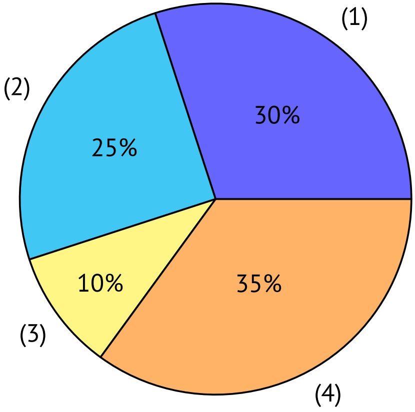
False Negative



1. Long paraphrase ($n > 5$)
2. Determiners
   ("mutation of **a** gene")
3. Inflected constituents
   ("holding of share**s**")

Error Analysis

False Negative



1. Long paraphrase ($n > 5$)
2. Determiners
   (“mutation of **a** gene”)
3. Inflected constituents
   (“holding of share**s**”)
4. Other errors

Evaluation 2: Classification

Model

- Represent a noun compound using its paraphrase vectors

Model

- Represent a noun compound using its paraphrase vectors
- Predict the *k* most likely paraphrases and average their vectors
  - weighted by confidence scores

Model

- Represent a noun compound using its paraphrase vectors
- Predict the *k* most likely paraphrases and average their vectors
  - weighted by confidence scores
- Train a classifier to predict the relation

Model

- Represent a noun compound using its paraphrase vectors
- Predict the *k* most likely paraphrases and average their vectors
  - weighted by confidence scores
- Train a classifier to predict the relation
- With (**integrated**) or without (**paraphrase**) concatenating the constituent word embeddings

Model

- Represent a noun compound using its paraphrase vectors
- Predict the *k* most likely paraphrases and average their vectors
  - weighted by confidence scores
- Train a classifier to predict the relation
- With (**integrated**) or without (**paraphrase**) concatenating the constituent word embeddings
- Baseline: just the constituent embeddings (**distributional**)

Lexical Memorization

- Supervised distributional methods suffer from *"lexical memorization"* [Dima, 2016]

Lexical Memorization

- Supervised distributional methods suffer from *"lexical memorization"* [Dima, 2016]
- [Levy et al., 2015]:
  overfitting to the most common relation of a specific word
  - Training: *(olive oil,* source*)*, *(coconut oil,* source*)*, *(soybean oil,* source*)*, ...
  - Model: *x oil* is in the source relation, regardless of *x*

Lexical Memorization

- Supervised distributional methods suffer from *"lexical memorization"* [Dima, 2016]
- [Levy et al., 2015]:
  overfitting to the most common relation of a specific word
  - Training: *(olive oil,* source*)*, *(coconut oil,* source*)*, *(soybean oil,* source*)*, ...
  - Model: *x oil* is in the source relation, regardless of *x*

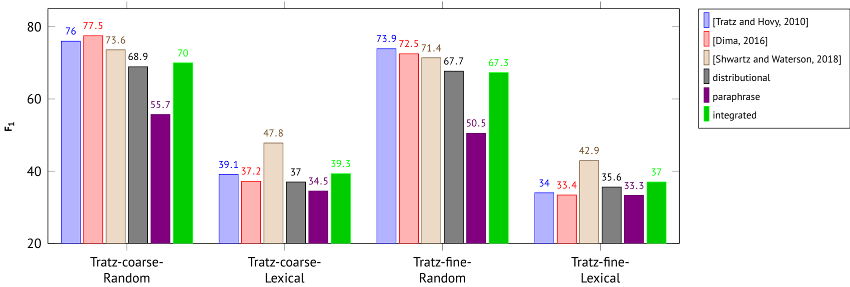
Dataset

- Dataset: [Tratz, 2011]
  - 19,158 instances
  - 37 relations (fine-grained) / 12 relations (coarse-grained)

Dataset

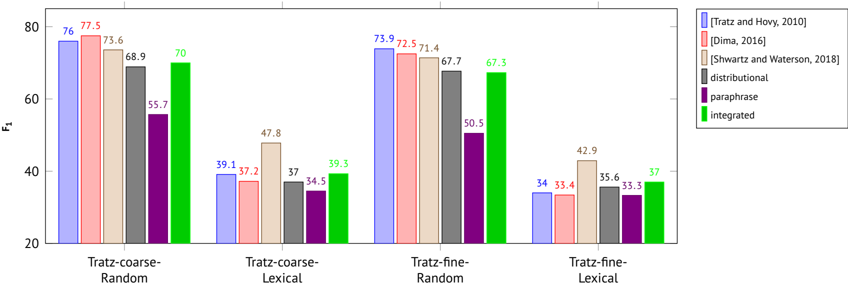
- Dataset: [Tratz, 2011]
  - 19,158 instances
  - 37 relations (fine-grained) / 12 relations (coarse-grained)
- Dataset splits:
  - Random 75:20:5 (like previous work)
  - Lexical [Levy et al., 2015]

Results



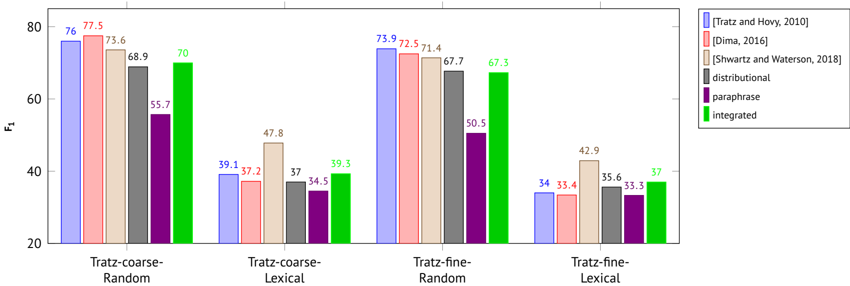
80
60
40
20
F1
76
77.5
73.6
68.9
55.7
70
39.1
37.2
47.8
37
34.5
39.3
73.9
72.5
71.4
67.7
50.5
67.3
34
33.4
42.9
35.6
33.3
37
Tratz-coarse-Random
Tratz-coarse-Lexical
Tratz-fine-Random
Tratz-fine-Lexical
[Tratz and Hovy, 2010]
[Dima, 2016]
[Shwartz and Waterson, 2018]
distributional
paraphrase
integrated

Results



Paraphrases are beneficial
for lexical splits:
can't default to
the common relation
of each constituent

Results



Paraphrases are beneficial
for lexical splits:
can't default to
the common relation
of each constituent

Always performs worse than a paraphrase-based model
trained especially for classification [Shwartz and Waterson, 2018]

Analysis - When did paraphrases help?

Example Noun-compounds	Gold	Distributional	Example Paraphrases
printing plant	purpose	objective	[w_2] engaged in [w_1]
marketing/development expert	topical	objective	[w_2] in [w_1], [w_2] knowledge of [w_1]
alcohol damage	causal	other	[w_2] caused by [w_1], [w_2] by [w_1]
rubber band, rice cake	containment	purpose	[w_2] of [w_1], [w_2] made of [w_1], [w_2] is made of [w_1]
laboratory animal	location/part-whole	attribute	[w_2] in [w_1], [w_2] used in [w_1]

- Distributional often defaults to a common label in the dataset

Analysis - When did paraphrases help?

Example Noun-compounds	Gold	Distributional	Example Paraphrases
printing plant	purpose	objective	[w_2] engaged in [w_1]
marketing/development expert	topical	objective	[w_2] in [w_1], [w_2] knowledge of [w_1]
alcohol damage	causal	other	[w_2] caused by [w_1], [w_2] by [w_1]
rubber band, rice cake	containment	purpose	[w_2] of [w_1], [w_2] made of [w_1], [w_2] is made of [w_1]
laboratory animal	location/part-whole	attribute	[w_2] in [w_1], [w_2] used in [w_1]

- Distributional often defaults to a common label in the dataset
- Paraphrases can indicate a different relation

Recap

- A model for generating paraphrases for given noun-compounds

Recap

- A model for generating paraphrases for given noun-compounds
- Better generalization abilities:
  - Generalize for unseen noun-compounds
  - Embed semantically-similar paraphrases in proximity

Recap

- A model for generating paraphrases for given noun-compounds
- Better generalization abilities:
  - Generalize for unseen noun-compounds
  - Embed semantically-similar paraphrases in proximity
- Evaluation on noun compound classification and paraphrasing

Recap

- A model for generating paraphrases for given noun-compounds
- Better generalization abilities:
  - Generalize for unseen noun-compounds
  - Embed semantically-similar paraphrases in proximity
- Evaluation on noun compound classification and paraphrasing
  - Improved performance in challenging evaluation settings

Recap

- A model for generating paraphrases for given noun-compounds
- Better generalization abilities:
  - Generalize for unseen noun-compounds
  - Embed semantically-similar paraphrases in proximity
- Evaluation on noun compound classification and paraphrasing
  - Improved performance in challenging evaluation settings

Thank you!

References

[Dima, 2016] Dima, C. (2016). On the compositionality and semantic interpretation of english noun compounds. In *Proceedings of the 1st Workshop on Representation Learning for NLP*.

[Hendrickx et al., 2013] Hendrickx, I., Kozareva, Z., Nakov, P., Ó Séaghdha, D., Szpakowicz, S., and Veale, T. (2013). Semeval-2013 task 4: Free paraphrases of noun compounds. In *SemEval*, pages 138–143.

[Levy et al., 2015] Levy, O., Remus, S., Biemann, C., and Dagan, I. (2015). Do supervised distributional methods really learn lexical inference relations? In *NAACL*, pages 970–976.

[Nakov, 2013] Nakov, P. (2013). On the interpretation of noun compounds: Syntax, semantics, and entailment. *Natural Language Engineering*, 19(03):291–330.

[Nakov and Hearst, 2006] Nakov, P. and Hearst, M. (2006). Using verbs to characterize noun-noun relations. In *International Conference on Artificial Intelligence: Methodology, Systems, and Applications*, pages 233–244. Springer.

[Pantel et al., 2007] Pantel, P., Bhagat, R., Coppola, B., Chklovski, T., and Hovy, E. (2007). ISP: Learning inferential selectional preferences. In *ACL*, pages 564–571.

[Shwartz and Dagan, 2018] Shwartz, V. and Dagan, I. (2018). Paraphrase to explicate: Revealing implicit noun-compound relations. In *ACL*, Melbourne, Australia.

[Shwartz and Waterson, 2018] Shwartz, V. and Waterson, C. (2018). Olive oil is made of olives, baby oil is made for babies: Interpreting noun compounds using paraphrases in a neural model. In *NAACL*.

[Surtani et al., 2013] Surtani, N., Batra, A., Ghosh, U., and Paul, S. (2013). Iiit-h: A corpus-driven co-occurrence based probabilistic model for noun compound paraphrasing. In *SemEval*, pages 153–157.

[Tratz, 2011] Tratz, S. (2011). *Semantically-enriched parsing for natural language understanding*. University of Southern California.

[Tratz and Hovy, 2010] Tratz, S. and Hovy, E. (2010). A taxonomy, dataset, and classifier for automatic noun compound interpretation. In *ACL*, pages 678–687.

[Van de Cruys et al., 2013] Van de Cruys, T., Afantenos, S., and Muller, P. (2013). Melodi: A supervised distributional approach for free paraphrasing of noun compounds. In *SemEval*, pages 144–147.

[Versley, 2013] Versley, Y. (2013). Sfs-tue: Compound paraphrasing with a language model and discriminative reranking. In *SemEval*, pages 148–152.