# Testing humour theory using word and sentence embeddings

Stephen Skalicky
Victoria University of Wellington
Aotearoa New Zealand

Salvatore Attardo
East Texas A&M University
USA

# Setting the stage

Computational humour studies focus on humour detection & generation. Incorporating humour theory into this work is important (Hempelmann, 2008)

- humour theory → improve computational approaches

Here we go **the other direction**, using computational methods to form additional tests of humour theory

- *predictions* of humour theory → tested computationally

# Incongruity

- The concept of *incongruity* is important for theoretical work on verbal humour
  - "the speaker says something unexpected, the soundness of which is thereupon recognized" (Aristotle, *Rhetoric*, Book III, Chapter 11, 1412b)
  - GTVH; humour as arising from a (semantic) incongruity (Attardo & Raskin, 1991)
  - humour understanding/comprehension as (cognitive) incongruity resolution (Forabosco, 1992, 2008)

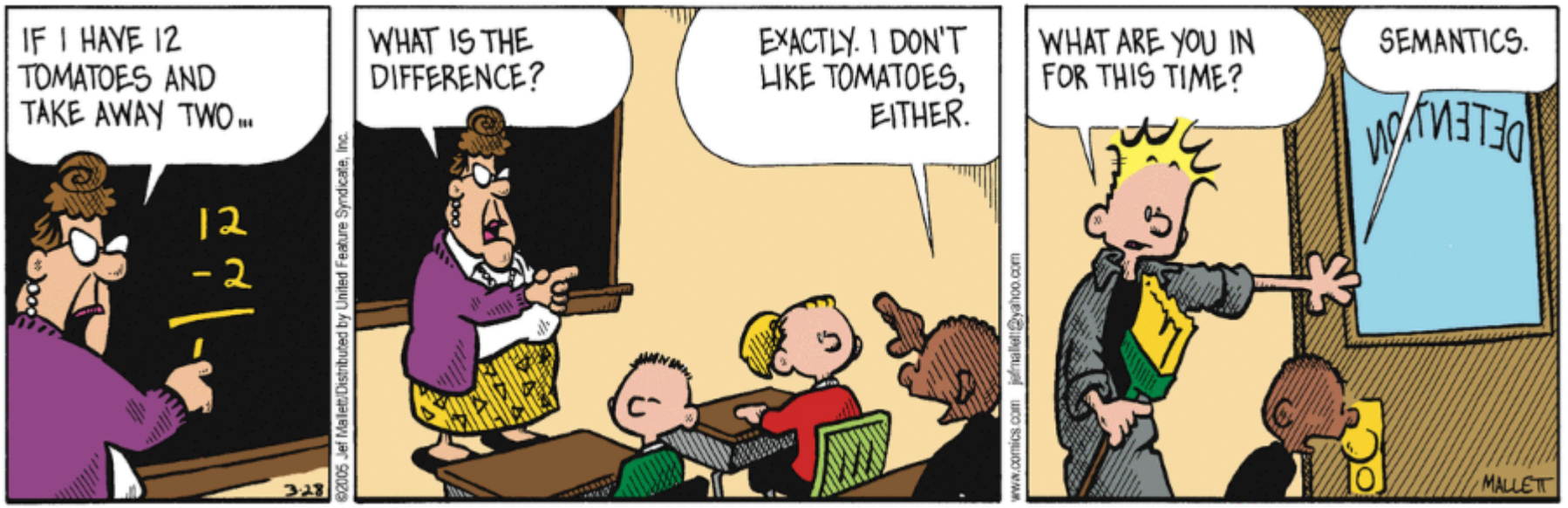
# Ritchie's "lowest common denominator"

> "…humour involves incongruity" (Ritchie, 2004)

But…variation in what we mean or how incongruity is defined in the context of humour (Ritchie, 2009)

> "All humour involves some degree of incongruity, but this incongruity is not random or arbitrary – it is systematically related to other aspects of the setting." (Ritchie, 2009, p. 299)

# Puns and semantic incongruity




https://www.gocomics.com/frazz/2005/03/28

The difference (math operation) ≈ The difference (who cares?)

# Puns

> A pun is a textual occurrence in which a sequence of sounds must be interpreted with a formal reference to a second sequence of sounds, which may, but need not, be identical to the first sequence, for the full meaning of the text to be accessed. The perlocutionary goal or effect of the pun is to generate the perception of mirth or of the intention to do so. (Attardo, 2020, pp. 177–178)

# Puns

I call my horse mayo and sometimes **mayo neighs**

1. Pun: Sometimes Mayo (*Proper noun*) neighs (*verb*)
2. Target: Sometimes I call (*verb*) my horse mayonnaise (*Proper noun*)

The tomb of Karl Marx is just another **communist plot**

1. Pun: plot – a conspiracy, a scheme
2. Target: plot – a piece of land for a grave

# Current Study

For puns to work, both meanings of the Pun & Target should be viable, but also exist in a state of incongruity.

**Can we test this prediction as a function of cosine distance between vector representations of pun/target words?**

# Data

- Corpus of 1182 pun-target pairs (Hempelmann, 2003) from a larger set (Sobkowiak, 1991)
- Imperfect, heterophonic puns (i.e., not 100% sound overlap between pun-target)
- For example:

  a. *hens* & *hence*

  b. *comical* & *chemical*

  c. *chowder* & *showed her*

# Similarity comparisons - pre-trained vector spaces

**word2vec**

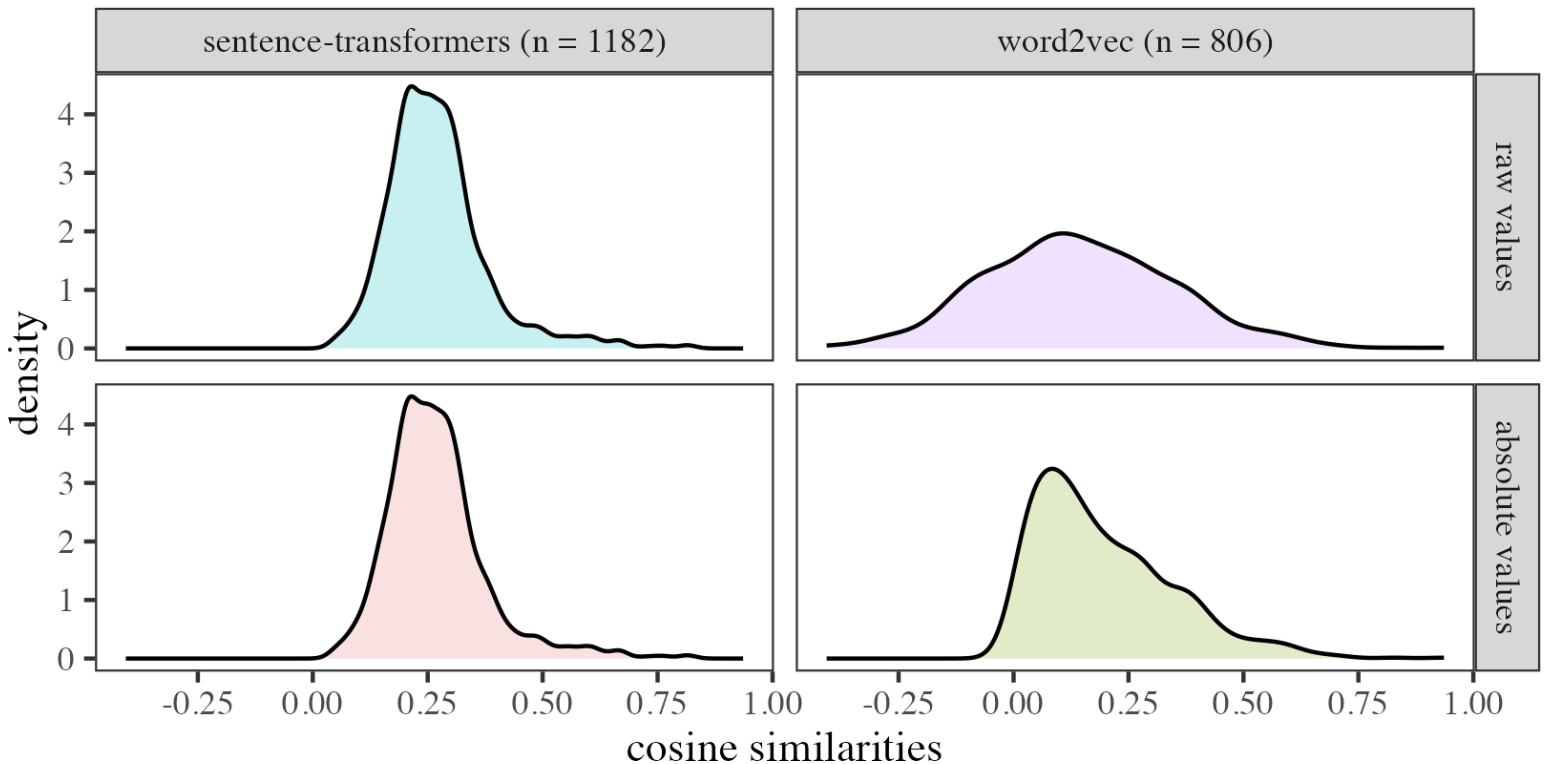
- word2vec-google-news-300
- 100 billion words
- 300 vectors
- for all single pun & target words

**sentence-transformers**

- all-MiniLM-L6-v2 (HuggingFace)
- 1 billion related sentence pairs
- 384 vectors
- for all pun & target words

Pairwise comparisons of semantic distance as cosine distance between pun & target words

# Results: pun vs. targets




Note: x-axis values below 0 only possible for raw word2vec distribution

| sentence-transformers | word2vec |
|---|---|
| $M$ = 0.279 (0.109) | $M$ = 0.143 (0.203) |
| | $\lvert M \rvert$ = 0.198 (0.150) |

# Results: pun vs. targets

# WordNet 3.0

- An ontology of **synsets** - different senses and their related words (called *lemmas*) for thousands of English words (Fellbaum, 1998)

all the synsets for the word **humour**

1. temper
2. wit
3. liquid body substance
4. humour (experiencing humour)
5. humour (being humorous)
6. humour (sense of humour)
7. humour (humorous mood)

all lemmas for **2. wit**

- *wit*
- *humour*
- *witticism*
- *wittiness*

# WordNet Baseline Method

- Synset lemmas should be semantically congruent with their seed words
- Calculate cosine distance between pun or target for all WordNet synset lemmas
- excluding repetitions of pun/target word in lemmas

**invisible**

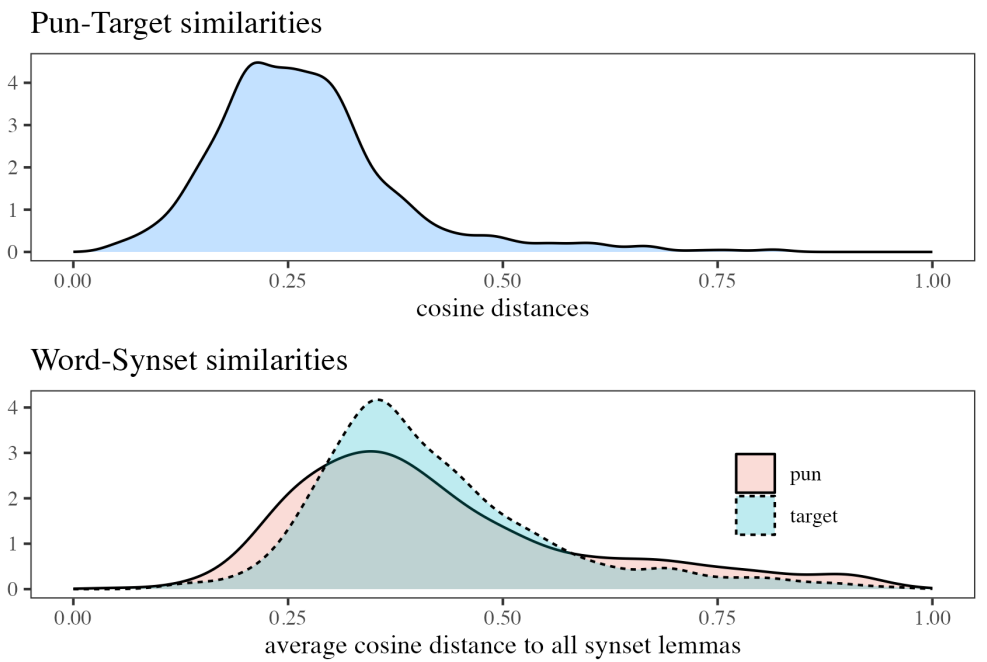
| synset | lemmas |
|---|---|
| 1. invisible (hard to see) | *invisible*, unseeable |
| 2. invisible (not prominent) | inconspicuous, *invisible* |

**visible**

| synset | lemmas |
|---|---|
| 1. visible (capable of being seen) | *visible*, seeable |
| 2. visible (obvious) | *visible* |

# WordNet Baseline Result

Average WN similarity (sentence-transformers): 0.422 (0.156)




**average similarities to WN lemmas significantly higher**

| measure | mean difference | 95%CI | *t* | *p* |
|---|---|---|---|---|
| pun-WN | 0.158 | 0.144, 0.171 | 23.325 | < .001 |
| target-WN baseline | 0.139 | 0.128, 0.151 | 23.704 | < .001 |

# Discussion

Our results show support for theoretical claims of incongruity theory

Specifically, semantic incongruity for puns

- both meanings are possible
- but exist in a state of incongruity

For puns, words must be somewhat related to be appropriate in same sentence context

# Limitations & Future Directions

- Data set is somewhat old
- Puns lack sentence context
  - would be useful for context-aware embeddings
- Incongruity is dependent on humour type (Ritchie, 2009)
  - Same method may require adaptation to other humour forms
- Variation in embedding models; different models, different vectors
  - Compare degree of difference between pun-target & baseline using different embedding models

# Conclusion

- Much excitement in computational generation & detection of humour (and related constructs)
  - Evidence that incorporation of humour theory is good for these approaches
- Our study tests potential for using modern embeddings to further empirically test humour theory
- Results tend to support the hypothesis, but much more work to be done

# Thank You

- Further ideas, questions, and collaborations are welcome!

Contact:

Stephen Skalicky
stephen.skalicky@vuw.ac.nz

Salvatore Attardo
salvatore.attardo@tamuc.edu

# References


Attardo, S. (2020). *The linguistics of humor: An introduction*. Oxford University Press.

Attardo, S., & Raskin, V. (1991). Script theory revis(it)ed: Joke similarity and joke representation model. *Humor - International Journal of Humor Research*, *4*(3-4), 293–348.

Fellbaum, C. (1998). *WordNet: An electronic lexical database*. MIT Press.

Forabosco, G. (1992). Cognitive aspects of the humor process: The concept of incongruity. *Humor*, *5*(1/2), 45–68.

Forabosco, G. (2008). Is the concept of incongruity still a useful construct for the advancement of humor research? *Lodz Papers in Pragmatics*, *4*(1), 45–62. https://doi.org/10.2478/v10016-008-0003-5

Hempelmann, C. (2003). *Paronomasic puns: Target recoverability towards automatic generation* [PhD thesis]. Purdue University.

Hempelmann, C. (2008). Computational humor: Beyond the pun? In V. Raskin (Ed.), *The primer of humor research* (pp. 333–360).

Ritchie, G. (2004). *The linguistic analysis of jokes*. Routledge.

Ritchie, G. (2009). Variants of Incongruity Resolution. *Journal of Literary Theory*, *3*(2), 313–332. https://doi.org/10.1515/JLT.2009.017

Sobkowiak, W. (1991). *Metaphonology of English Paronomasic Puns*. P. Lang.