

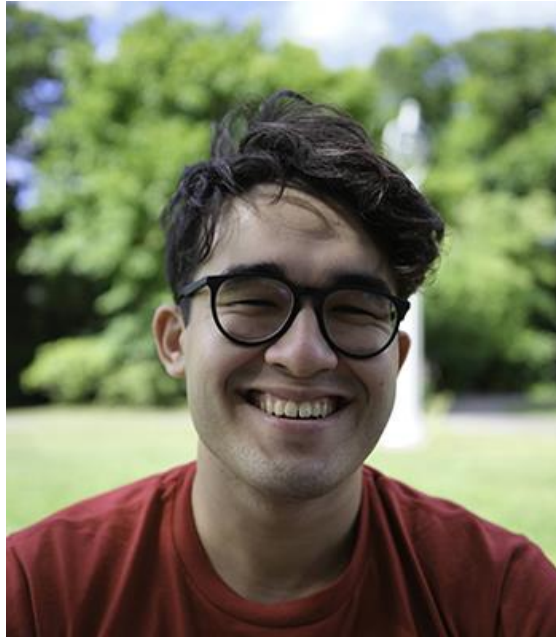
# **A first look at mind rhymes**

John R. Starr  
UC Berkeley Phorum  
February 25<sup>th</sup>, 2022

# The Team



Helena  
Aparicio



Me!  
(John R. Starr)



Marten van  
Schijndel

# Overview

1. Our cognitive model
2. Mind rhymes
3. Data
4. Experimental component
5. Computational component
6. Conclusion & future directions

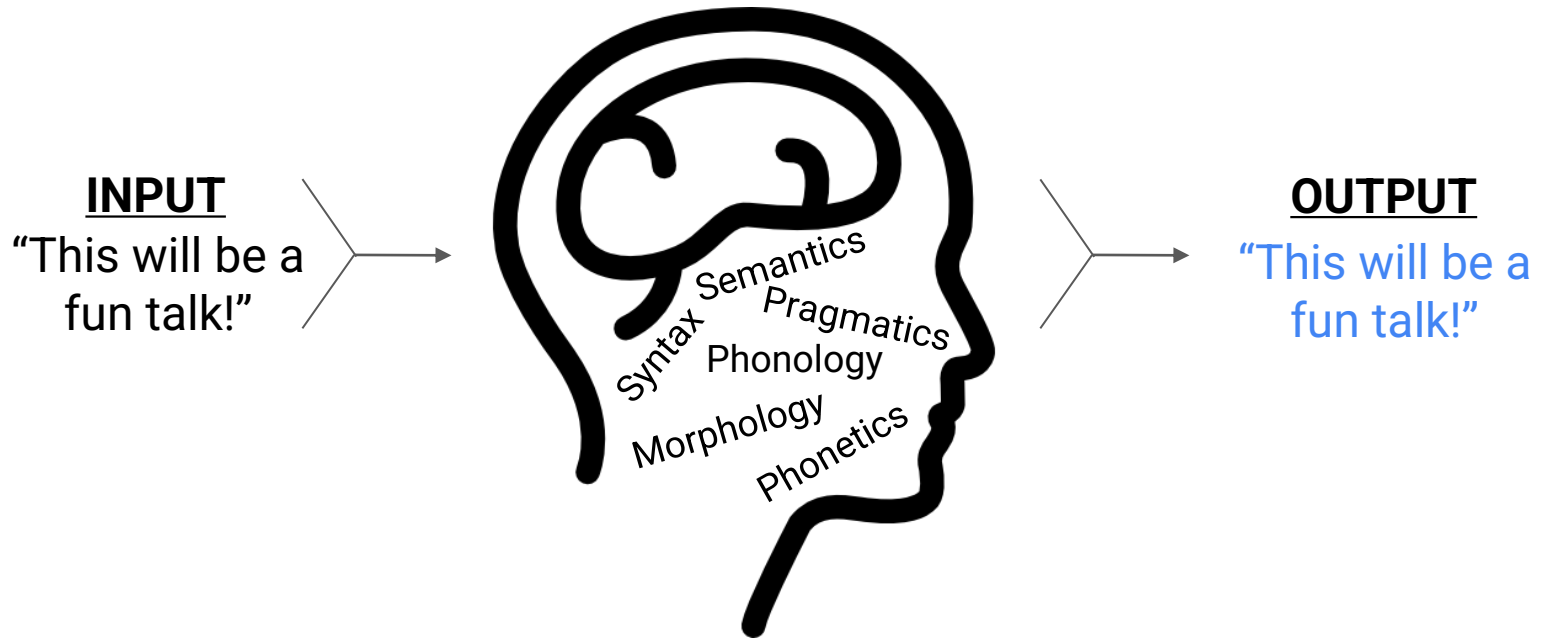
Brief disclaimer:  
taboo words referring to  
bodily fluids, derogatory or  
discriminatory language,  
and other NSFW topics  
may be discussed briefly

# Initial Takeaways

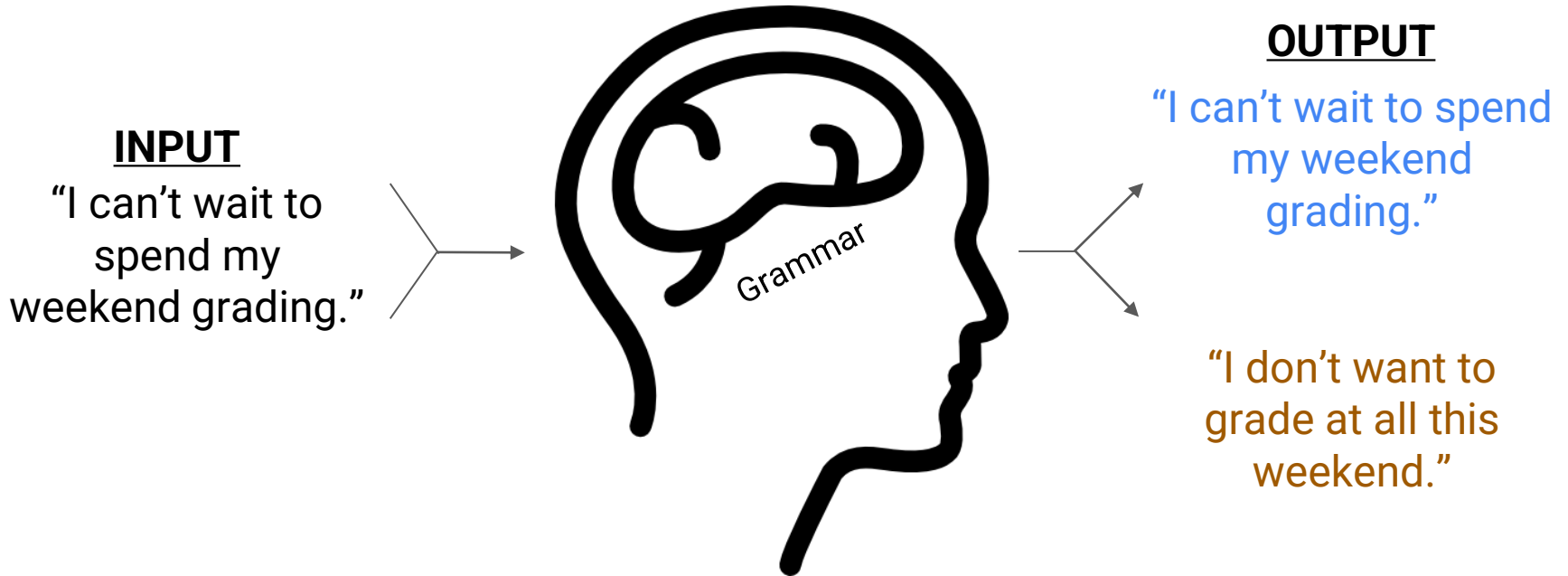
- Mind rhymes provide a clear look into linguistic phenomena that require dual-message processing and resolution
- Metrical structure improves rhyme prediction
- Global semantic cues drive the retrieval process when there is a phonological violation

# 1. Our Cognitive Model

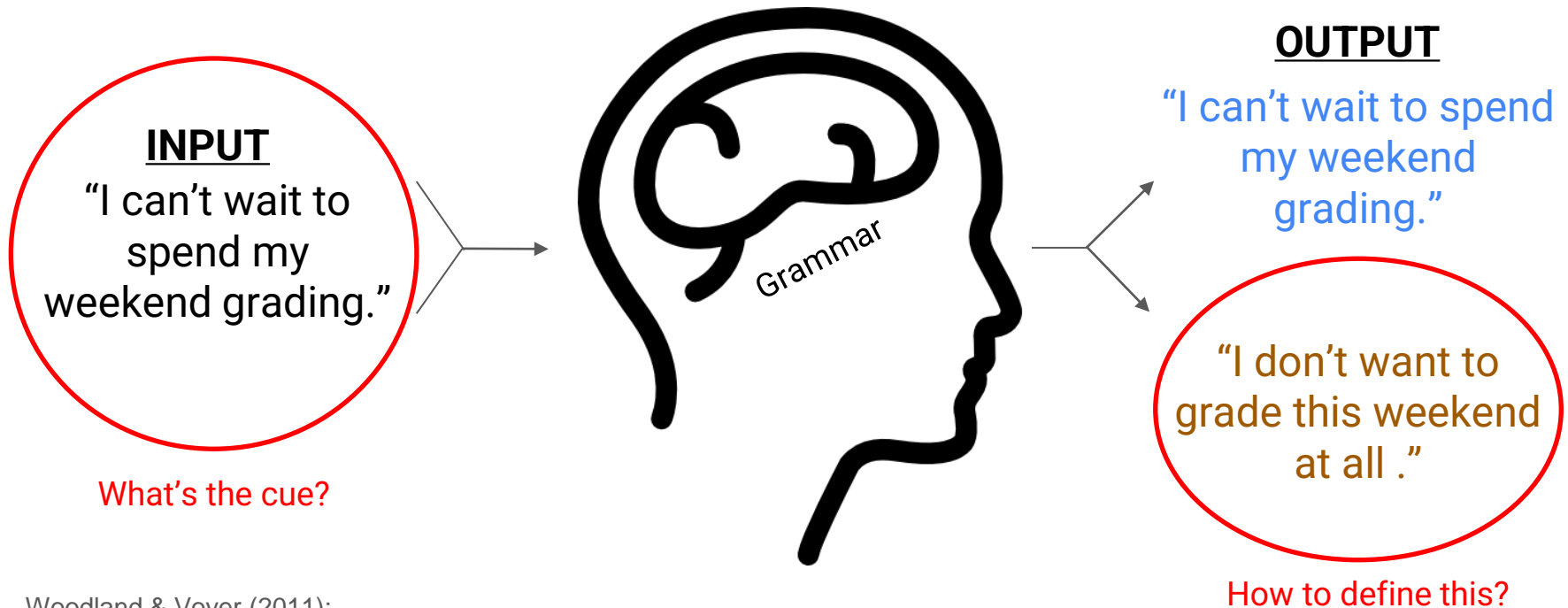
# Dissecting Simple Linguistic Signals



# Dissecting Complex Linguistic Signals:



# Dissecting Complex Linguistic Signals



Woodland & Voyer (2011);  
Nakassis & Snedeker (2002)

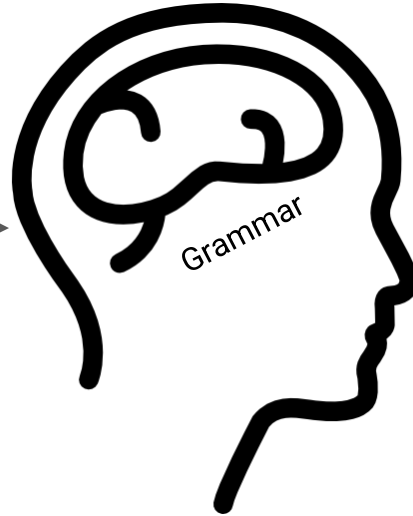


## 2. Mind Rhymes

# Enter the Mind Rhyme!

## INPUT

Good food's rare as Halley's comet  
This stuff makes me want to \_\_\_\_\_.



## OUTPUT

“leave”

“vomit”

Blue: Cue  
Brown: Implicit Message

# A Simple Definition of Mind Rhyme

“A form of rhyme subversion where the rhyming intended target (IT) is substituted by an unrhyming, overt word (OT), often for humorous effect.”

# Basic Examples of Mind Rhymes

Call me old and make me cry /  
Laughing's like a piece of \_\_\_\_\_ . (cake / pie)

Winter's a good time to stay in and cuddle /  
But put me in summer and I'll be a \_\_\_\_\_ . (happy snowman / puddle)

Now they're going to bed /  
And my stomach is sick /  
'cause it's all in my head /  
But she's touching his \_\_\_\_\_ . (chest / dick)

# Nuances of Mind Rhymes

Work all night, I'm always tired /  
I hope my boss doesn't get me \_\_\_\_\_.

(laid off = fired)

My uncle thinks I'm barmy /  
because I don't pack my bag and join the \_\_\_\_\_.


(navy ~ army)

He's limber-slouched against a post /  
and tells a friend what matters \_\_\_\_\_.

(least ≠ most)

# Other Features of Mind Rhymes – Meter, Taboo

I have a sad story to tell you /  
It may hurt your feelings a bit /  
Last night I walked into my bathroom /  
And stepped in a big pile of \_\_\_\_\_.

ternary feet  
  
(shaving cream / shit)

People say that summer's never cool /  
but when I'm hot I jump inside the \_\_\_\_\_.

(creek / pool)

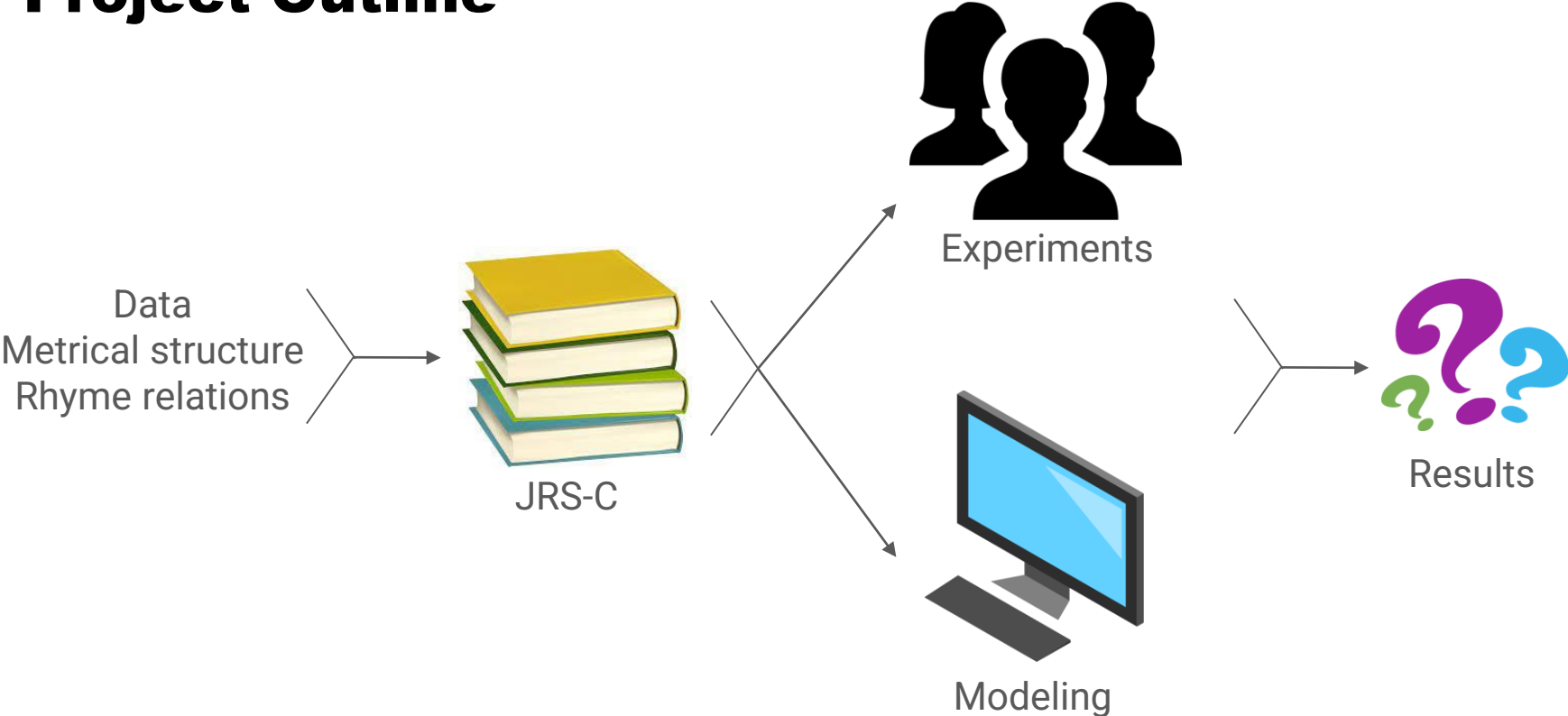
binary feet  


Schmidt-Kassow & Kotz (2009); Rothermich et al. (2011); Pitt & Samuel (1990); Quene & Port (2005); Niruala et al. (2020); Jay et al. (2008)

# Our Questions (Broadly)

1. What linguistic signals cue mind rhyme retrieval?
2. How do quantitative measures of information processing align/differ with empirical studies?

# Project Outline





# 3. Data

# John's Rhyme Subversion Corpus (JRS-C)

- Mind rhymes gathered from:
  - Humor websites
  - Scraping lyrics from comedy musicians
  - My own brain
  - The graciousness of others
- Data annotated for:
  - Taboo
  - Metrical structure
  - Relationship between targets
  - Word frequency (COCA)
- 210 rhyming pairs (420 total targets) altogether have been processed

# Chicago Rhyming Poetry Corpus (CRPC)

- Corpus of several hundred poems from the 15<sup>th</sup>-20<sup>th</sup> centuries
- All poems annotated for rhyme scheme

18 RHYME a a b c b c d e f d f e

19

20 Exile of immortality, strongly wise,

21 Strain through the dark with undesirous eyes

22 To what may lie beyond it. Sets your star,

23 O heart, for ever! Yet, behind the night,

24 Waits for the great unborn, somewhere afar,

25 Some white tremendous daybreak. And the light,

# 4. Experimental Component

# Experimental Questions

- Are mind rhymes truly a legitimate and predictable phenomenon?
- Do people prefer phonological or semantic continuations when facing a violation of the linguistic signal?

# Cloze Task

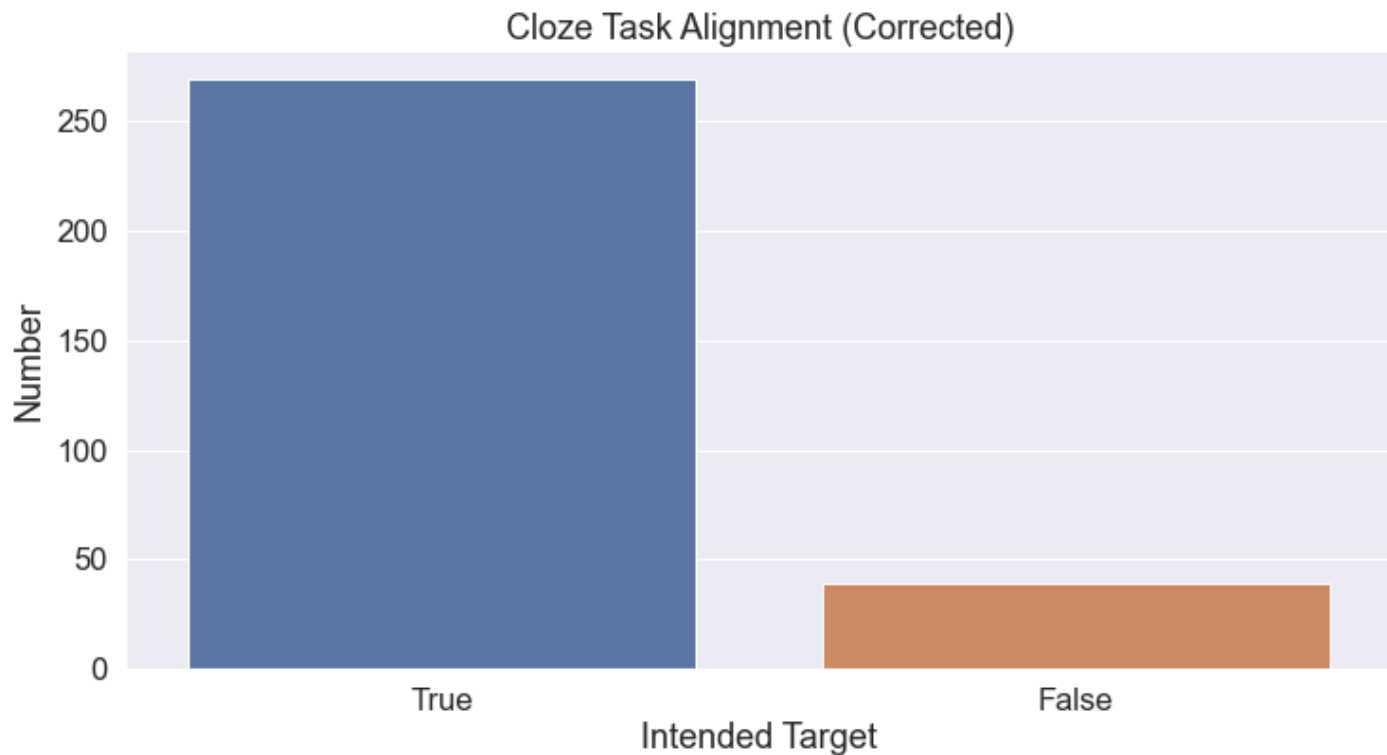
The poems I write  
are a real delight!  
So please be polite  
when the rhyme is not \_\_\_\_\_.

Fill in the blank:

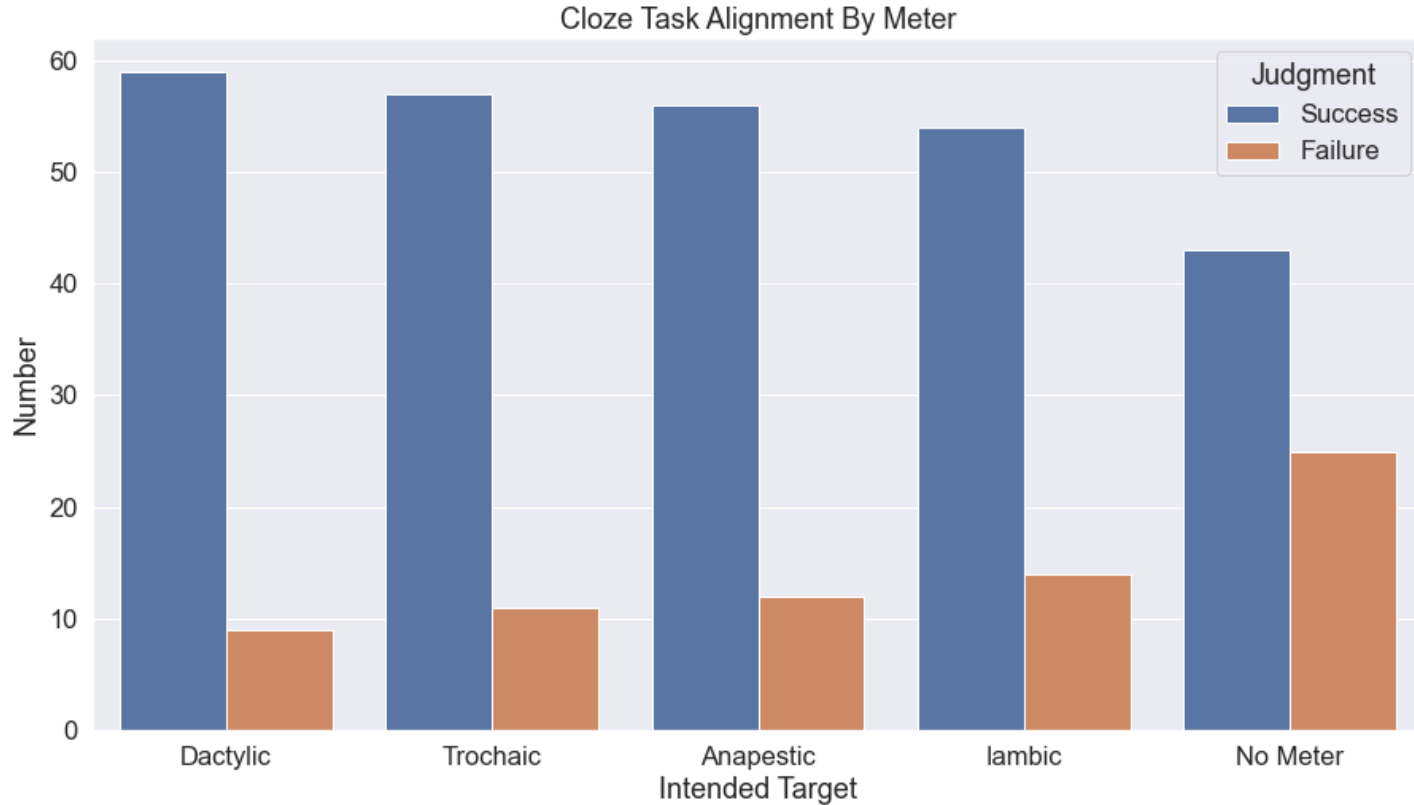
Details: 19 participants, 16 mind rhyme examples per participant

Predictions: people will choose the intended target (IT) over all potential targets

# Cloze Task (Results pt. 1)




# Cloze Task (Results pt. 2)






# The Covered Box Task and Mind Rhymes

  
*progress*

My favorite color's green:  
of that I'm sure is true.  
My son prefers the color red  
but he despises \_\_\_\_\_.






**orange**


F J

# Experimental Items

  
*progress*

My favorite color's green:  
of that I'm sure is true.  
My son prefers the color red  
but he despises \_\_\_\_\_.





**orange**

F J

1. Intended Target (IT):
2. Overt Target (OT):
3. IT rhyme competitor (IT-RC):
4. Semantically distant word from IT:

“blue”  
“orange”  
“stew”  
“couches”

# The Covered Box Paradigm and MRs



*progress*

My favorite color's green:  
of that I'm sure is true.  
My son prefers the color red  
but he despises \_\_\_\_\_.

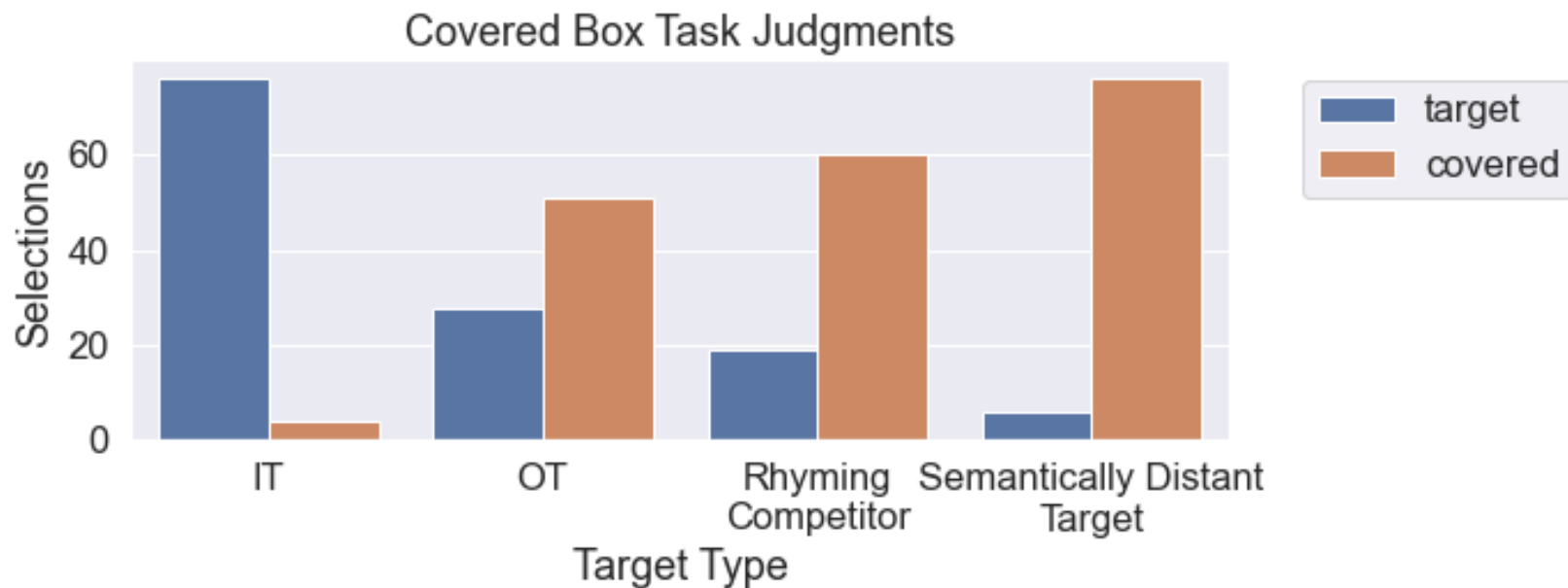


orange

F J

- Details: 24 participants, 10 mind rhyme examples per participant
- Predictions:
  - IT should be accepted every time
  - OT should be accepted sometimes, rejected in others
  - IT rhyming competitors should be rejected nearly every time
  - IT distant competitors should be rejected every time

# The Covered Box Paradigm (Results)



# Experimental Conclusions

- Mind rhyme is a salient phenomenon (to native speakers of English, at least)
- Despite the obvious rhyme scheme, people prefer the OT over the rhyming competitors
- What is special about the OT?

# 5. Computational Component

# Computational Questions We are Examining

1. Can models distinguish between mind rhymes and true rhyming data?
2. Can models distinguish between OT and IT targets within mind rhymes?
3. Is mind rhyme retrieval motivated by a cue outside of the syntactic or phonological domains, as suggested by our experimental work?

# How Can We Quantify Information?

- If we have a large amount of data and a tough computer, we can get probability distributions for every word within the data!
- In this way, we can determine which words are more or less surprising given a context:



- Or: **surprisal(x) =  $-\log_2 p(x)$**



# Language Models

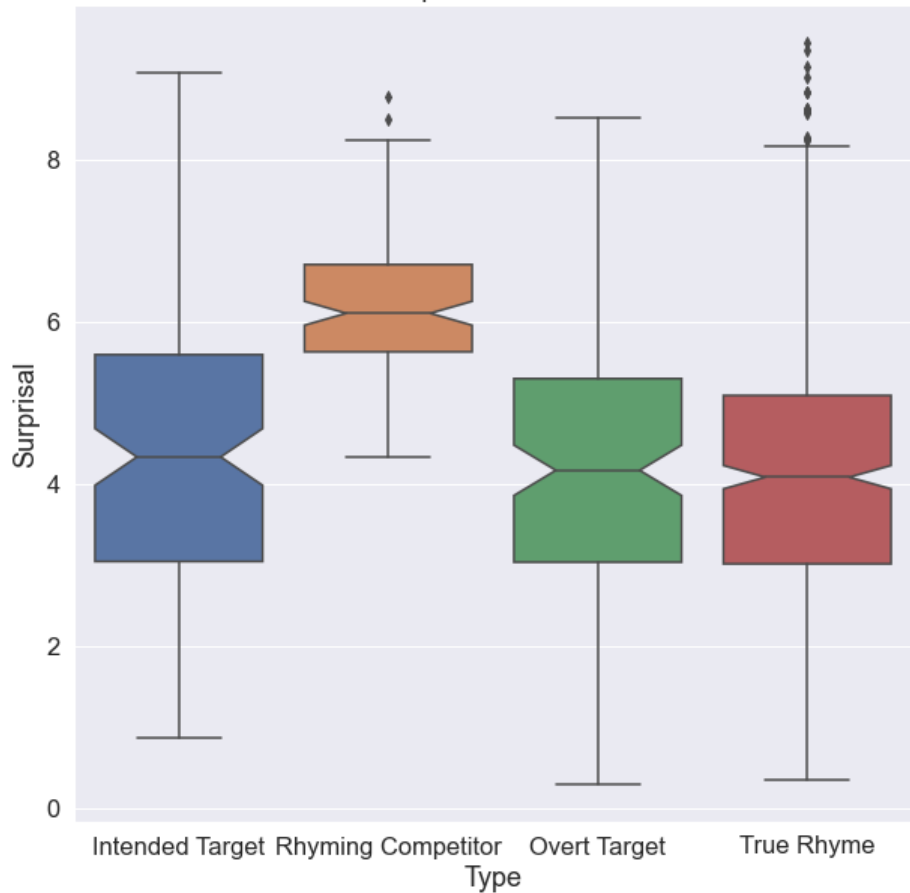
## 1. KenLM

- N-gram language model
- Trained on 6 million sentences from Project Gutenberg
- Captures local-coherence surprisal

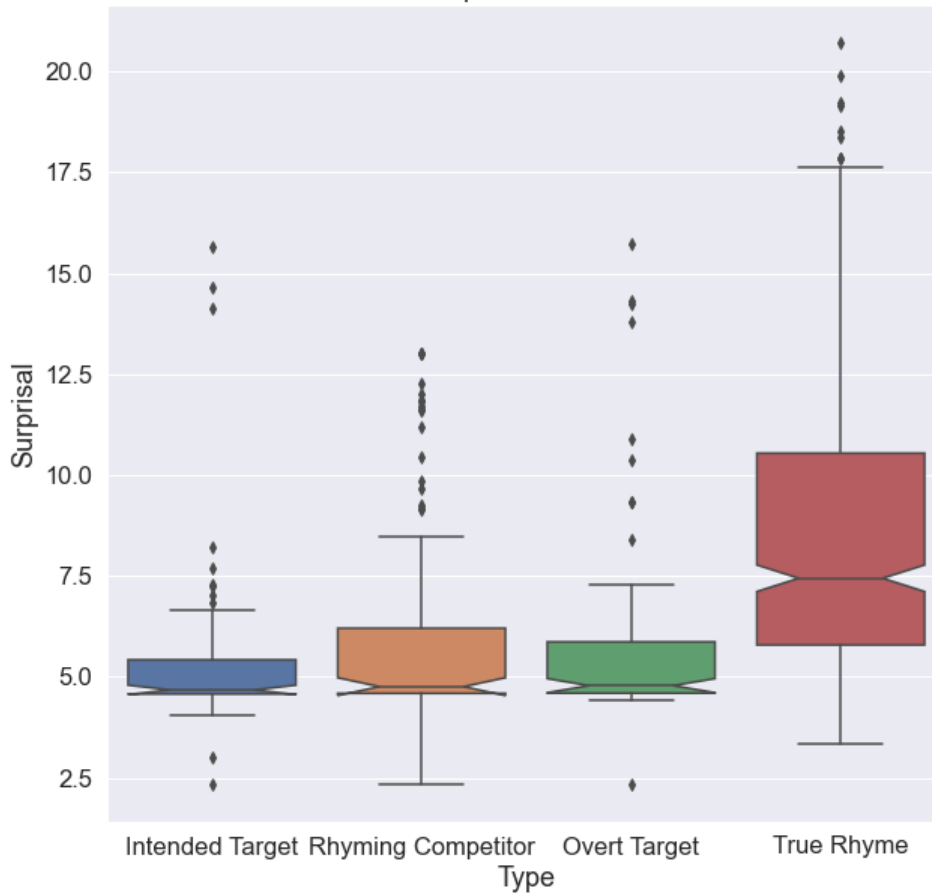
## 2. Roark (2001) Parser

- Top-down syntactic parser
- Trained on sections 2-21 of the Penn WSJ Treebank
- Captures syntactic surprisal

KLM Surprisal Across Models



Roark Surprisal Across Models



# Interim Conclusions

1. Language models can decipher differences between real rhymes, fake rhymes, and mind rhymes
2. Local and syntactic surprisal does not motivate the special status of the OT, suggesting this cue must be driven by another aspect of the linguistic signal.

# Measuring Semantic Similarity

- Take the following sentence:

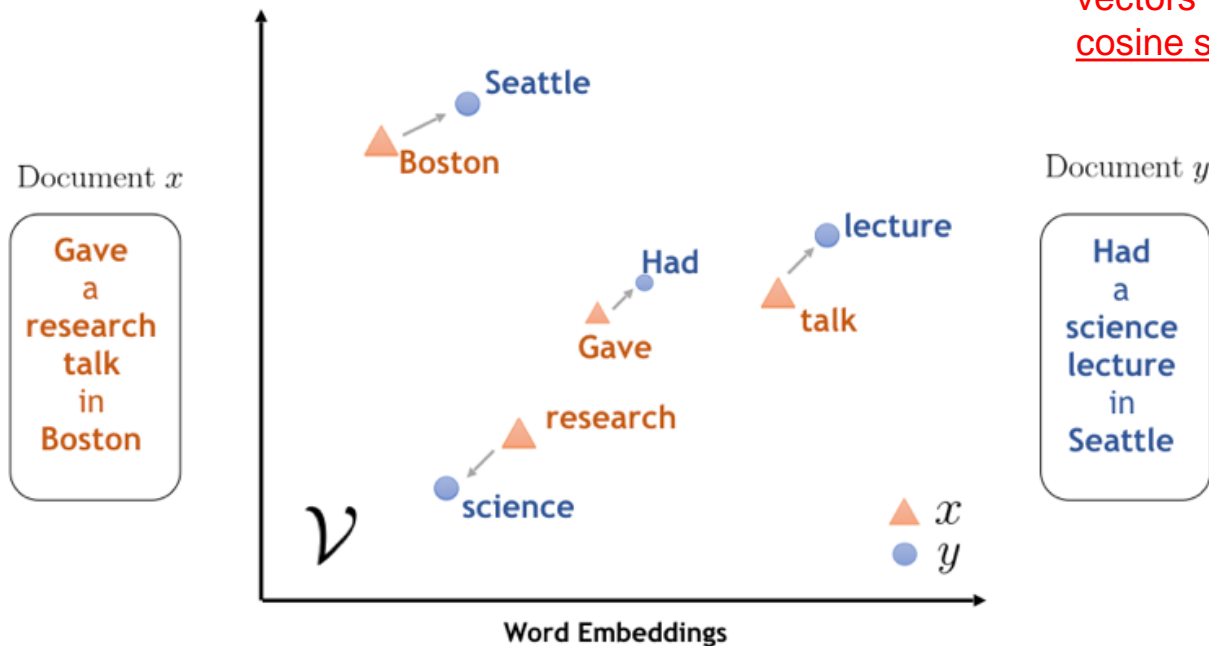
*I have always wanted a pet dog.*

- What other words can take the position of the underlined word?

*I have always wanted a pet cat.*

*? I have always wanted a pet rooster.*

# Word Embeddings



Measure the  
difference  
between two  
vectors with  
cosine similarity

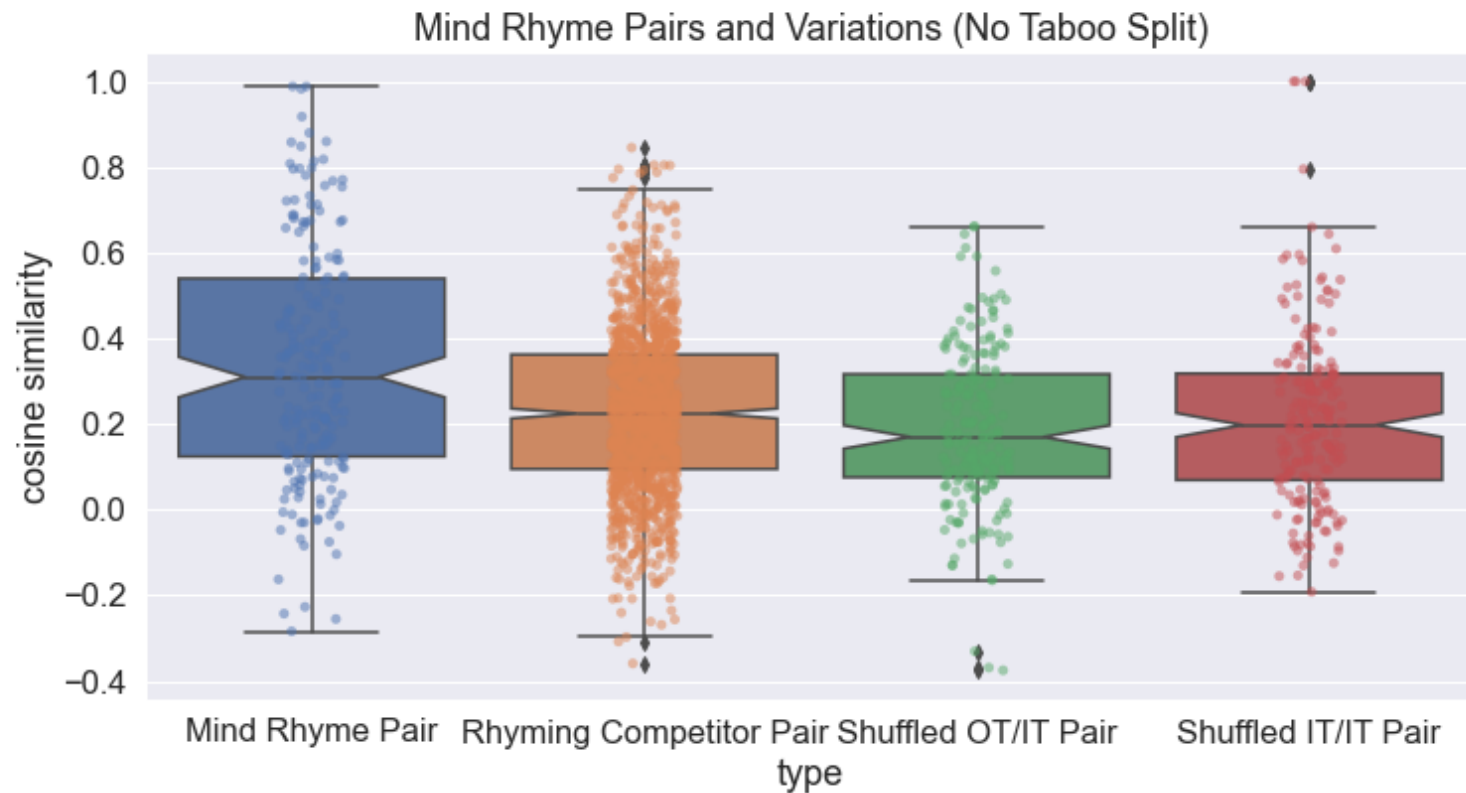
Pennington et al. (2014)

# Cosine Similarity Analysis

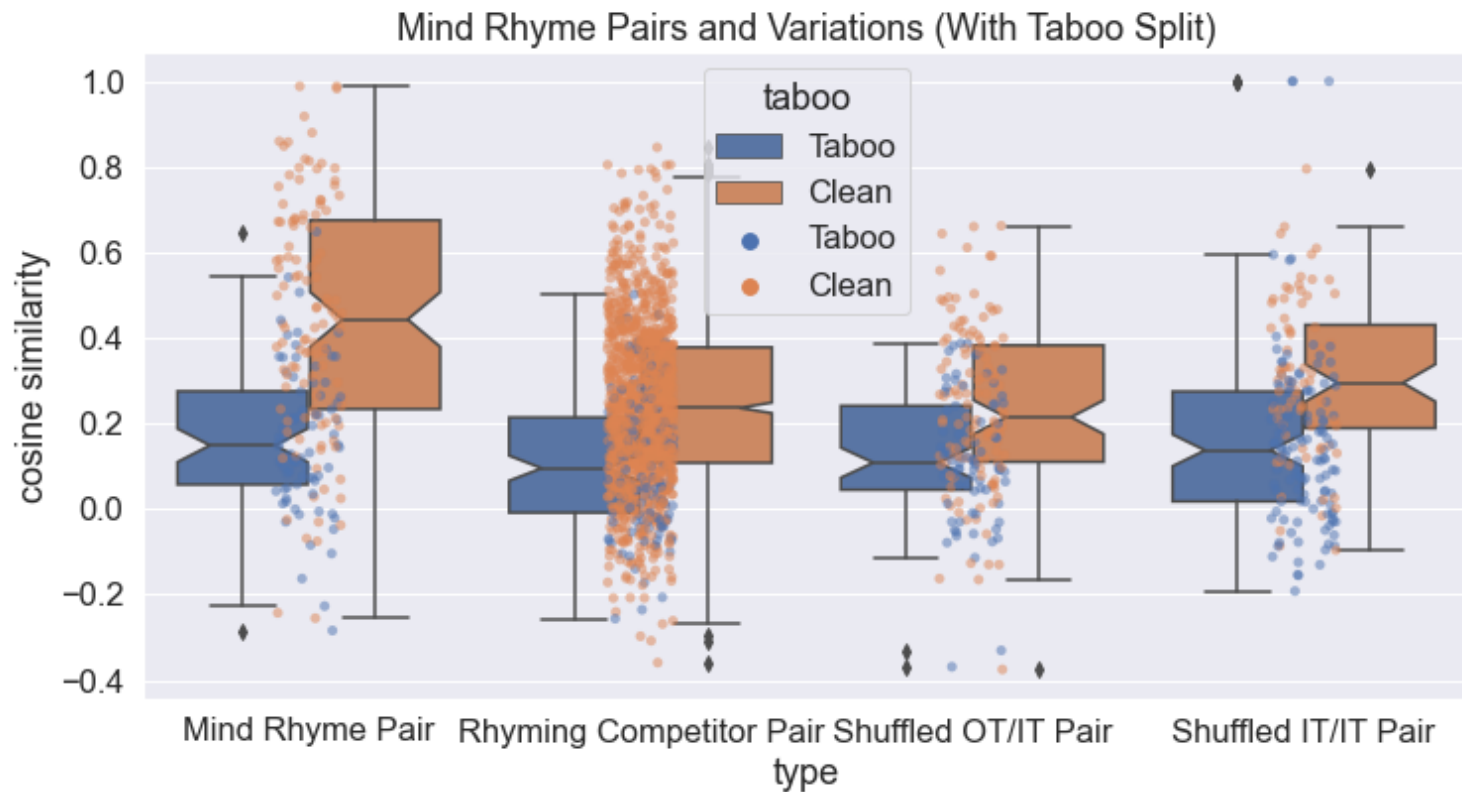
Looking at our data, we take the cosine similarity of these pairs:

1. OT/IT
  - The mind rhyme pair!
2. OT/IT-rhyme-competitor
  - Covers potential relationship with other rhyming targets
3. Shuffled-OT/IT
  - Covers if there is a general pairing across OTs to a random target
4. IT/IT pairs
  - Covers the idea that perhaps all ITs are similar to one another

# Cosine Similarity (Results pt. 1)



# Cosine Similarity (Results pt. 2)





# Summary of Computational Component

Regarding our previous questions:

1. Models can use surprisal to distinguish between a real rhyme and a mind rhyme target (OT or IT), but they cannot use surprisal to distinguish between the OT and IT.
2. The global semantic relationship between OT and IT is likely the cue that aids retrieval

# 6. Conclusion & Future Directions

# Conclusions

- Mind rhymes permit a concrete pathway into investigation of phenomena such as sarcasm, irony, and humor
- Mind rhymes suggest that the necessary cue for retrieving an implicit message must be the global semantic relationship between the intended target and the overt target
  - Holds even when phonological signal is predictable

# Future Directions

- Can we alter people's predictions by giving them an incorrect OT?

Christmas time is full of cheer.

(worry / fear)

All the children are without \_\_\_\_\_.

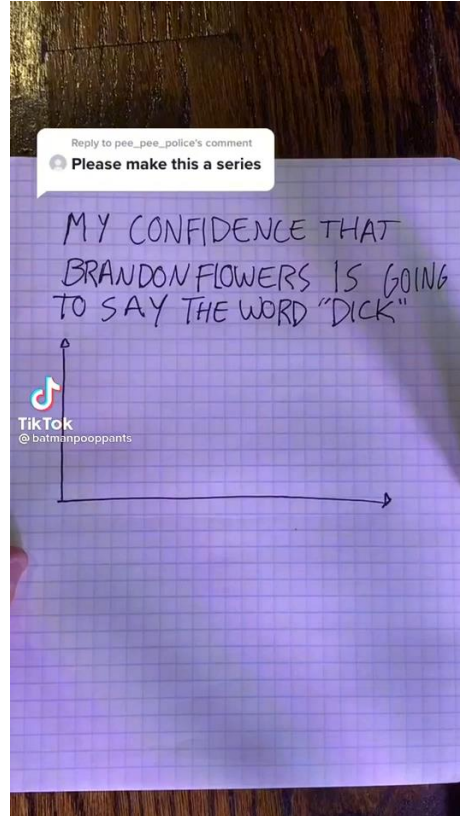
(alcohol / beer)

- Why do taboo targets behave so differently?
- Puns?


# References

1. Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In Second meeting of the north american chapter of the association for computational linguistics.
2. Heafield, K. (2011, July). KenLM: Faster and smaller language model queries. In *Proceedings of the sixth workshop on statistical machine translation* (pp. 187-197).
3. Holdefer, C. (2009). "Shaving Cream" and other mind rhymes. *The Antioch Review*, 67(1), 158–163.
4. Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126- 1177.
5. Nakassis, C., & Snedeker, J. (2002). Context as relational cues in children's recognition of irony. *Proceedings of the Annual Boston University Conference on Language Development*, 26(2), 429–440.
6. Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532-1543.
7. Pitt, M. A., & Samuel, A. G. (1990). Attentional allocation during speech perception: How fine is the focus? *Journal of Memory and Language*, 29(5), 611–632.  
[https://doi.org/10.1016/0749-596X\(90\)90055-5](https://doi.org/10.1016/0749-596X(90)90055-5)
8. Quene, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica*, 62(1), 1-13.
9. Read, K., Macauley, M., & Furay, E. (2014). The Seuss boost: Rhyme helps children retain words from shared storybook reading. *First Language*, 34(4), 354–371.  
<https://doi.org/10.1177/0142723714544410>
10. Roark, B. (2001). Probabilistic top-down parsing and language modeling. *Computational linguistics*, 27(2), 249-276.
11. Roark, B., Bachrach, A., Cardenas, C., & Pallier, C. (2009, August). Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. In *Proceedings of the 2009 conference on empirical methods in natural language processing* (pp. 324-333).
12. Rothermich, K., Schmidt-Kassow, M., & Kotz, S. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, 50, 232–244.
13. Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, 21(9), 1693–1708. 19
14. Schwarz, F., Bill, C., & Romoli, J. (2016). Reluctant acceptance of the literal truth: Eye tracking in the covered box paradigm. In *Proceedings of Sinn und Bedeutung*, 61-78.
15. Woodland, J., & Voyer, D. (2011). Context and intonation in the perception of sarcasm. *Metaphor and Symbol*, 26, 227–239.

# A Summary by my Co-author, @batmanpooppants



# Norming Tasks


  
*progress*

**red**  
**insane**

How similar do these words sound on a scale of 1-5?

*very different*      *very similar*

*1 = very different , 5 = very similar*

  
*progress*

**predicament**  
**six**

What is the meaning relationship of the two words on a scale from 1-5?

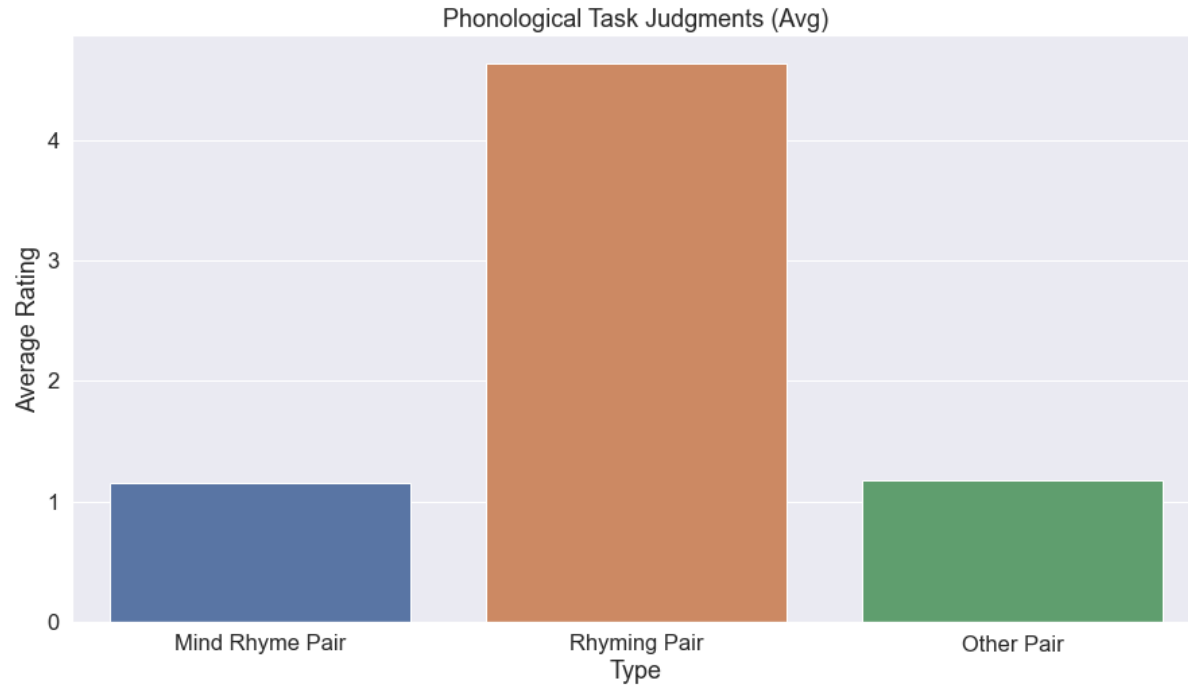
*no relationship*      *strong relationship*

*1 = no meaning relationship , 5 = strong meaning relationship*

- Predictions:

- MR pairs should be “very different” for sound, but show degrees of semantic relatedness
- Intended rhyming targets should be “very similar” to other rhyming words, but show no semantic relatedness
- Fillers should be unrelated in both meaning and sound

# Relatedness Task (Phonological Results)





# Relatedness Task (Semantic Results)

