Meaning in the N400: What can distributional language models tell us?

Carolin Dudschig

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Heisenberg-
Programm

Deutsche
Forschungsgemeinschaft
DFG

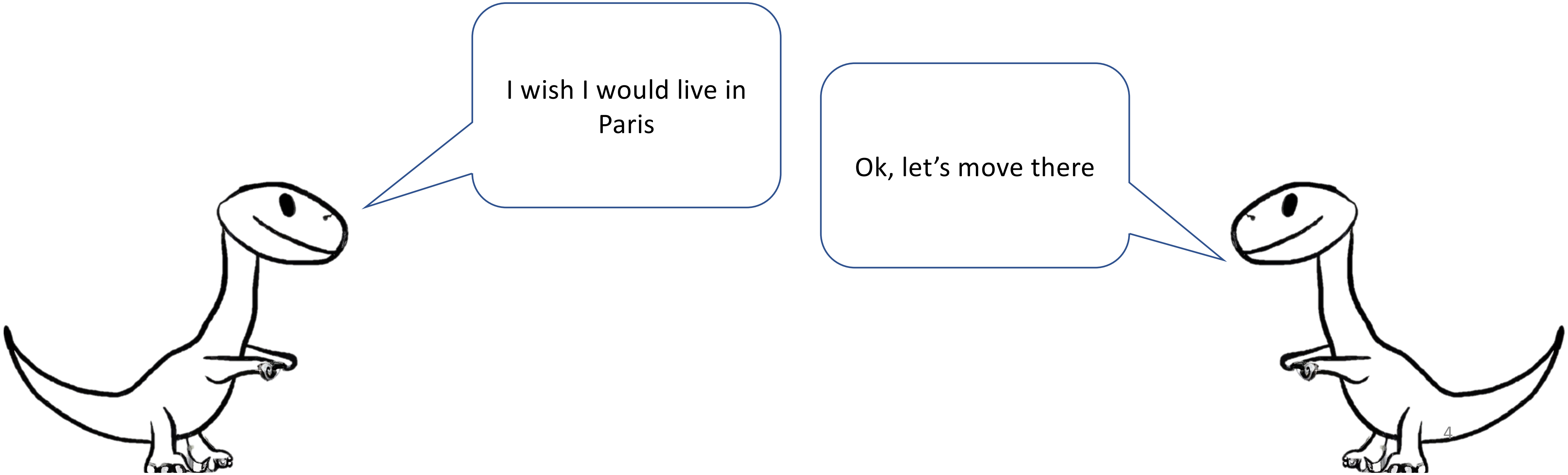
Overview

- Language
- Natural Language Processing (NLP) Techniques
  - LSA
  - Word2Vec
- Cognitive Reality?
  - Priming Studies
  - EEG Studies
- Conclusions

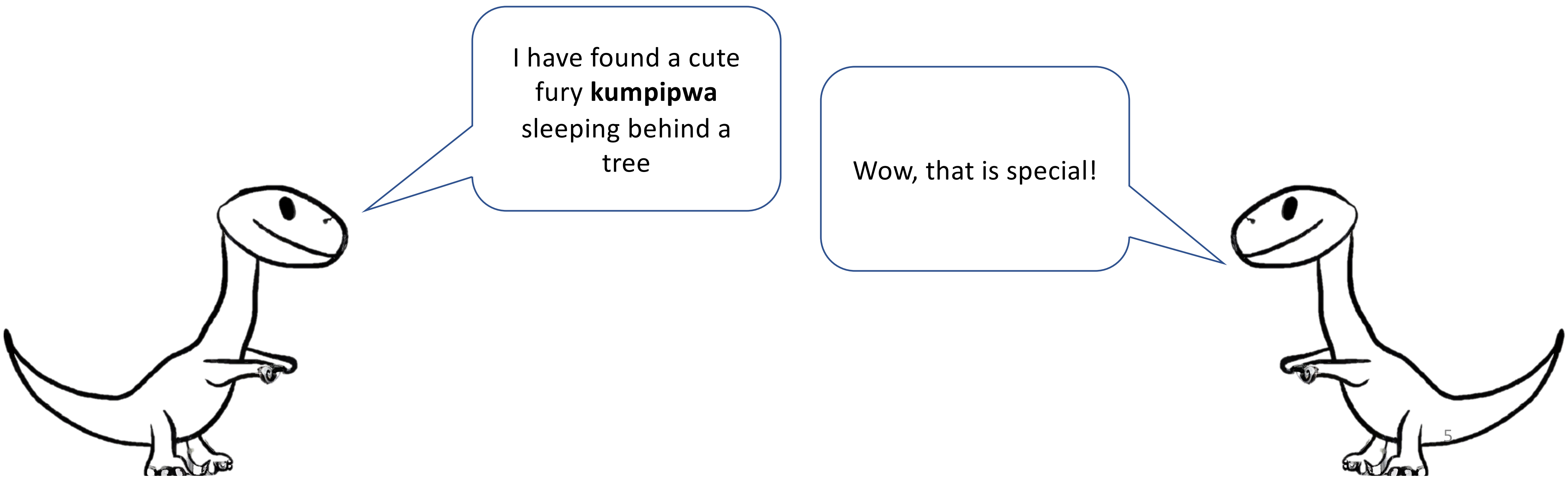
We all know:

Language is great, we can use it to communicate all kind of cool ideas …

Language is great, we can use it to communicate all kind of cool ideas …



Language is great, we can use it to communicate all kind of cool ideas …



Language is great, we can use it to communicate all kind of cool ideas …

I have found a cute fury **kumpipwa** sleeping behind a tree

Wow, that is special!

Language is great, we can use it to communicate all kind of cool ideas …

I have found a cute fury **kumpipwa** sleeping behind a tree

Wow, that is special!

Humans can infer meaning from linguistic context

“You shall know a word by the company it keeps”

J. R. Frith, 1957

Philological Society of London *A Synopsis of Linguistic Theory*

We become like the company we keep (Euripides, 406 B.C.)

Noscitur a sociis (She is known by her company) (Latin proverb)

"You shall know a word by the company it keeps"

J. R. Frith, 1957

Philological Society of London *A Synopsis of Linguistic Theory*

Maybe one of the most successful ideas in the 20th century in linguistics

What does this tell us about meaning?

"You shall know a word by the company it keeps"

J. R. Frith, 1957

Philological Society of London *A Synopsis of Linguistic Theory*

What does this tell us about meaning?

"You shall know a word by the company it keeps"

J. R. Frith, 1957

Philological Society of London *A Synopsis of Linguistic Theory*

If you know the context a word appears in or if you can predict the context a word appears in, then you can understand the word!

What does this tell us about meaning?

"You shall know a word by the company it keeps"

J. R. Frith, 1957

Philological Society of London *A Synopsis of Linguistic Theory*

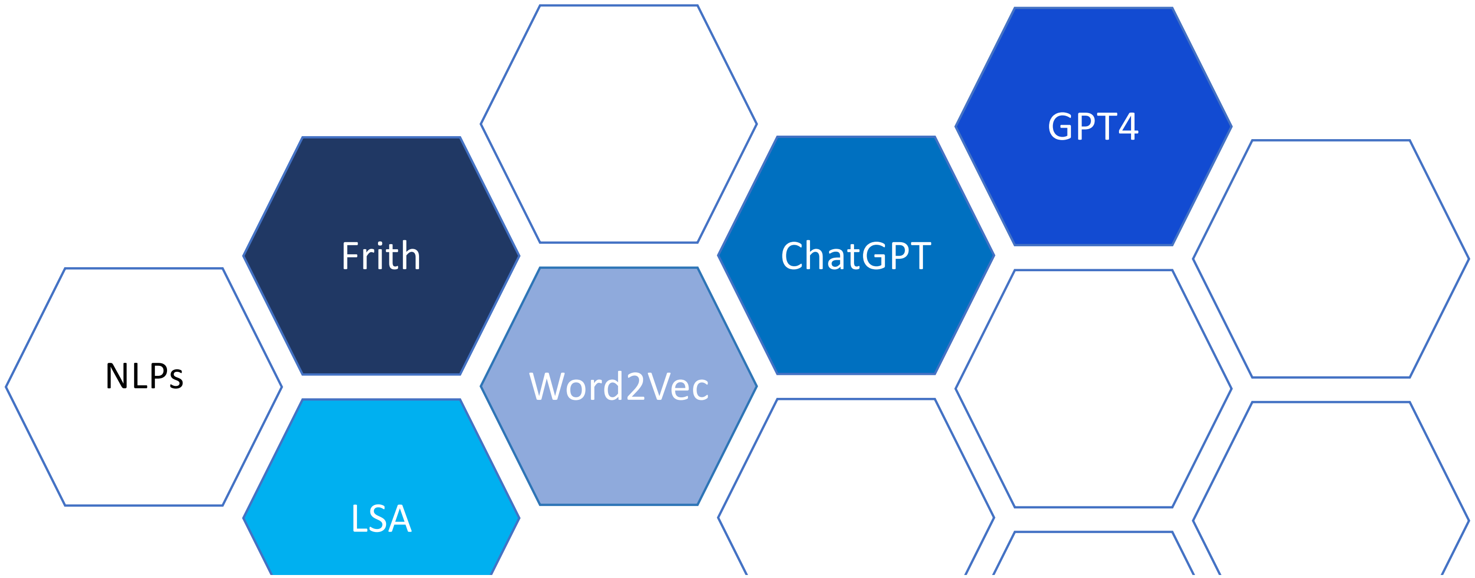
If you know the context a word appears in or if you can predict the context a word appears in, then you can understand the word!

How do we get machines to do that cool sort of thing?

What does this tell us about meaning?

"You shall know a word by the company it keeps"

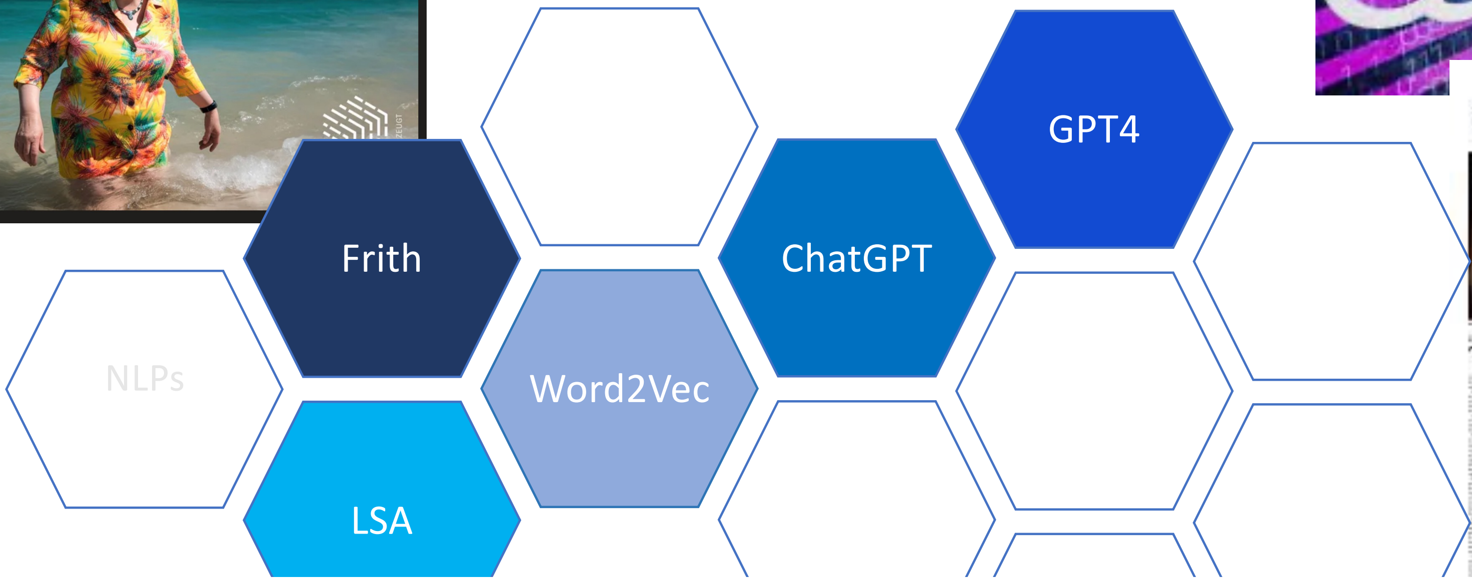
J. R. Frith, 1957



What does this tell us about meaning?

"You shall know a word by the company it keeps"

J. R. Frith, 1957



How do we get machines to do this?

- Can text be input in machines / deep learning -> NO!
- Can numbers be input in machines / deep learning -> YES!
- Encoding: Convert text into number!

“cats meat”

“dogs meat”

Unique Word	Encoding
cats	0
meat	1
dogs	2

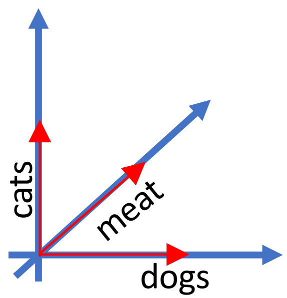
- Can text be input in machines / deep learning -> NO!
- Can numbers be input in machines / deep learning -> YES!
- Encoding: Convert text into number!

"cats meat"

"dogs meat"

	cats	meat	dogs
cats	1	0	0
meat	0	1	0
dogs	0	0	1

unique word	encoding
cats	(1,0,0)
meat	(0,1,0)
dogs	(0,0,1)



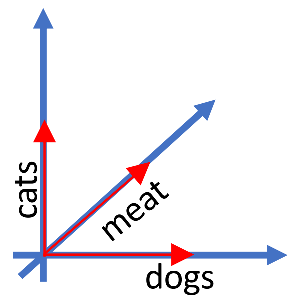
- Can text be input in machines / deep learning -> NO!
- Can numbers be input in machines / deep learning -> YES!
- Encoding: Convert text into number!

"cats meat"

"dogs meat"

	cats	meat	dogs
cats	1	0	0
meat	0	1	0
dogs	0	0	1

~~unique w d~~	~~encodin~~
~~cats~~	~~(1,0~~
~~meat~~	~~0)~~
~~dogs~~	~~(0,0,1~~



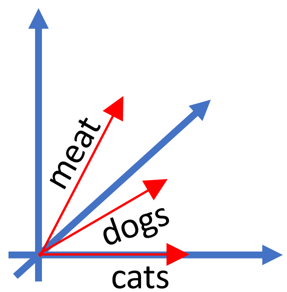
- Can text be input in machines / deep learning -> NO!
- Can numbers be input in machines / deep learning -> YES!
- Encoding: Convert text into number!

“cats meat”

“dogs meat”

	cats	meat	dogs
cats	1	0	0
meat	0	1	0
dogs	0	0	1

unique word	embedding
cats	(0.7,0.0,0.0)
meat	(0.2,0.9,0.3)
dogs	(0.3,0.0,0.2)



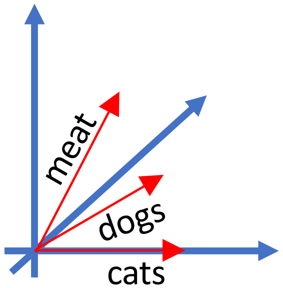
- Can text be input in machines / deep learning -> NO!
- Can numbers be input in machines / deep learning -> YES!
- Encoding: Convert text into number!

"cats meat"

"dogs meat"

	cats	meat	dogs
cats	1	0	0
meat	0	1	0
dogs	0	0	1

unique word	embedding
cats	(0.7,0.[illegible]
meat	(0.2,0.9,0.3)
dogs	(0.3,0.0,0.2)

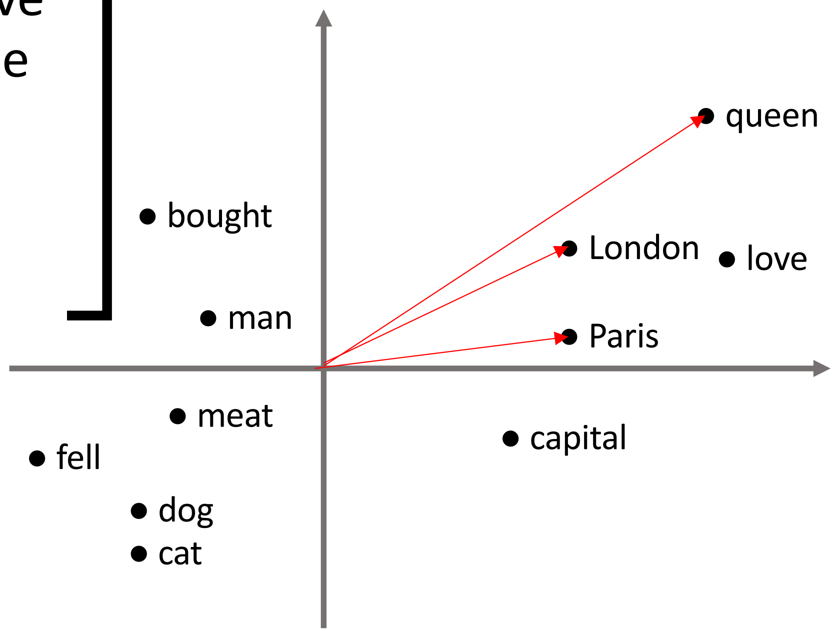


Text (Documents) -> Embeddings

- E-Books
- Movies (subtitles)
- Websites (newspapers, forums, blogs, Wiki, etc.)
- Apps

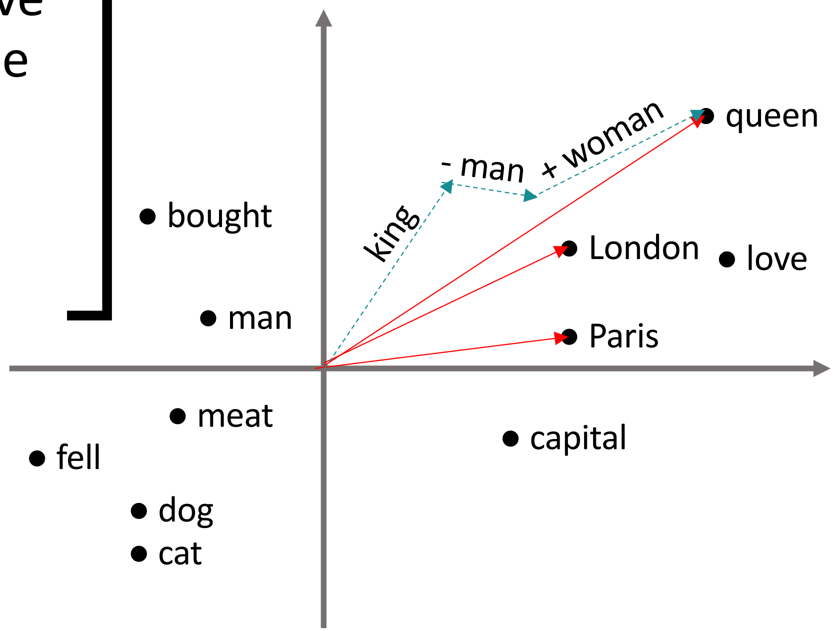
Text (Documents) -> Embeddings

["She bought a dog in the pet shop. The dog very much likes meat. Cats also like meat. She fell in love in Paris. Paris is the capital of France. London is the capital of England. The king lives in Buckingham palace. The Queen died in 2022 and Brits were queuing up to 24hrs to see her coffin."]



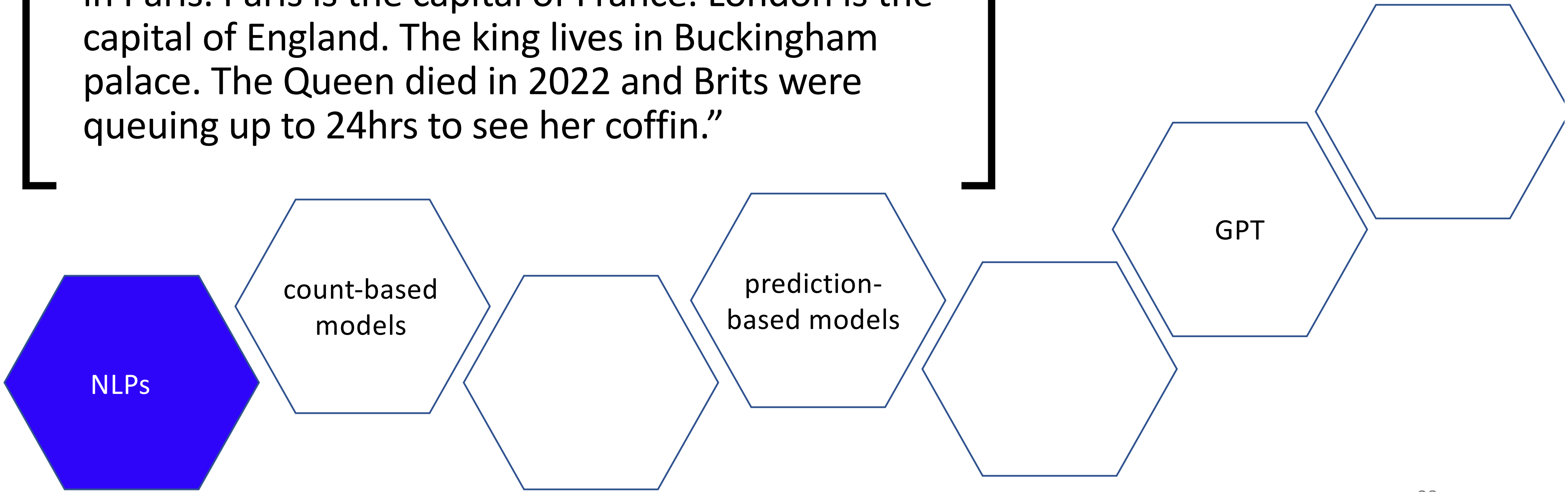
Text (Documents) -> Embeddings

[“She bought a dog in the pet shop. The dog very much likes meat. Cats also like meat. She fell in love in Paris. Paris is the capital of France. London is the capital of England. The king lives in Buckingham palace. The Queen died in 2022 and Brits were queuing up to 24hrs to see her coffin.”]



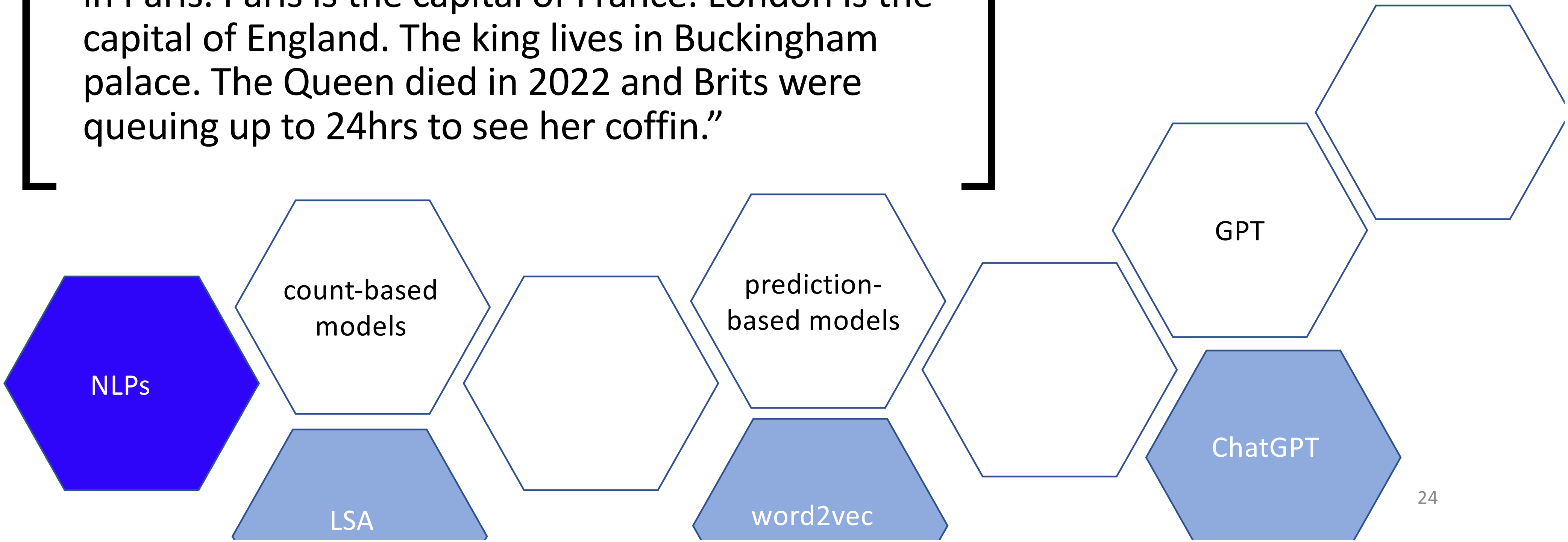
Text (Documents) -> Embeddings

[“She bought a dog in the pet shop. The dog very much likes meat. Cats also like meat. She fell in love in Paris. Paris is the capital of France. London is the capital of England. The king lives in Buckingham palace. The Queen died in 2022 and Brits were queuing up to 24hrs to see her coffin.”]



Text (Documents) -> Embeddings

[“She bought a dog in the pet shop. The dog very much likes meat. Cats also like meat. She fell in love in Paris. Paris is the capital of France. London is the capital of England. The king lives in Buckingham palace. The Queen died in 2022 and Brits were queuing up to 24hrs to see her coffin.”]



count-based models

LSA – Latent Semantic Analysis

Landauer & Dumais, 1997

- **Meaning of a word** = average of the meaning of all the passages in which it appear (Landauer, Foltz, Laham, 1998)
- If a word is an average of the meaning of the passages in which it appears, we might not need all the passages where the word appears -> dimensionality reduction

Fully automatic / statistical technique for extracting and inferring relations of expected contextual usage of words in passages or text -> Using matrix operations we can obtain these statistical approximations

count-based models

LSA – Latent Semantic Analysis

- She bought a dog in the pet shop.
- The dog very much likes meat.
- She fell in love in Paris.
- Paris is the capital of France.
- London is the capital of England.
- ....

	d1	d2	d3	d4	d5
bought	1	0	0	0	0
dog	1	1	0	0	0
pet	1	0	0	0	0
shop	1	0	0	0	0
very	0	1	0	0	0
much	0	1	0	0	0
likes	0	1	0	0	0
meat	0	1	0	0	0
fell	0	0	1	0	0
love	0	0	1	0	0
Paris	0	0	1	1	0
capital	0	0	0	1	1
France	0	0	0	1	0
London	0	0	0	0	1
England	0	0	0	0	1

sim(Paris, London) = 0

count-based models

LSA – Latent Semantic Analysis

- She bought a dog in the pet shop.
- The dog very much likes meat.
- She fell in love in Paris.
- Paris is the capital of France.
- London is the capital of England.
- ….

	d1	d2	d3	d4	d5
bought	1	0	0	0	0
dog	1	1	0	0	0
pet	1	0	0	0	0
shop	1	0	0	0	0
very	0	1	0	0	0
much	0	1	0	0	0
likes	0	1	0	0	0
meat	0	1	0	0	0
fell	0	0	1	0	0
love	0	0	1	0	0
Paris	0	0	**1**	**1**	0
capital	0	0	0	**1**	**1**
France	0	0	0	1	0
London	0	0	0	0	**1**
England	0	0	0	0	1

sim(Paris, London) = 0

count-based models

LSA – Latent Semantic Analysis

Singular Value Decomposition (SVD)

	d1	d2	d3	d4	d5
bought	1	0	0	0	0
dog	1	1	0	0	0
pet	1	0	0	0	0
shop	1	0	0	0	0
very	0	1	0	0	0
much	0	1	0	0	0
likes	0	1	0	0	0
meat	0	1	0	0	0
fell	0	0	1	0	0
love	0	0	1	0	0
Paris	0	0	1	1	0
capital	0	0	0	1	1
France	0	0	0	1	0

$$A_{m,n} = U_{m,m}\Sigma_{m,n}(V^T)_{n,n}$$

$A_{m,n}$ original word x document matrix with m words and n docs

$U_{m,m}$ describes the original row entities as vectors of derived orthogonal factor values

$V_{n,n}$ describes the original column entities in the same way

$\Sigma_{m,n}$ Is a diagonal matrix containing scaling values

Singular Value Decomposition (SVD)

$$U = \begin{pmatrix} -0.11 & 0.29 & -0.41 & -0.11 & -0.34 & -0.52 & 0.06 \\ -0.07 & 0.14 & -0.55 & 0.28 & 0.5 & 0.07 & 0.01 \\ 0.04 & -0.16 & -0.59 & -0.11 & -0.25 & 0.3 & -0.06 \\ 0.06 & -0.34 & 0.1 & 0.33 & 0.38 & -0. & 0. \\ -0.17 & 0.36 & 0.33 & -0.16 & -0.21 & 0.17 & -0.03 \\ 0.11 & -0.43 & 0.07 & 0.08 & -0.17 & -0.28 & 0.02 \\ 0.11 & -0.43 & 0.07 & 0.08 & -0.17 & -0.28 & 0.02 \\ -0.14 & 0.33 & 0.19 & 0.11 & 0.27 & -0.03 & 0.02 \\ 0.27 & -0.18 & -0.03 & -0.54 & 0.08 & 0.47 & 0.04 \\ 0.49 & 0.23 & 0.02 & 0.59 & -0.39 & 0.29 & -0.25 \\ 0.62 & 0.22 & 0. & -0.07 & 0.11 & -0.16 & 0.68 \\ 0.45 & 0.14 & -0.01 & -0.3 & 0.28 & -0.34 & -0.68 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2.54 & 0 & 0 & 0 & 0 \\ 0 & 2.35 & 0 & 0 & 0 \\ 0 & 0 & 1.64 & 0 & 0 \\ 0 & 0 & 0 & 1.50 & 0 \\ 0 & 0 & 0 & 0 & 1.31 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.61 & -0.46 & -0.54 & -0.28 & -0. & -0.01 & -0.02 \\ 0.17 & -0.13 & -0.23 & 0.11 & 0.19 & 0.44 & 0.62 \\ -0.5 & 0.21 & 0.57 & -0.51 & 0.1 & 0.19 & 0.25 \\ -0.03 & 0.04 & 0.27 & 0.15 & 0.02 & 0.02 & 0.01 \\ -0.21 & 0.38 & -0.21 & 0.33 & 0.39 & 0.35 & 0.15 \\ -0.26 & 0.72 & -0.37 & 0.03 & -0.3 & -0.21 & 0. \\ 0.43 & 0.24 & -0.26 & -0.67 & 0.34 & 0.15 & -0.25 \\ -0.05 & -0.01 & 0.02 & 0.06 & -0.45 & 0.76 & -0.45 \\ -0.24 & -0.02 & 0.08 & 0.26 & 0.62 & -0.02 & -0.52 \end{pmatrix}$$

Singular Value Decomposition (SVD)

$$U = \begin{pmatrix} -0.11 & 0.29 & -0.41 & -0.11 & -0.34 & -0.52 & 0.06 \\ -0.07 & 0.14 & -0.55 & 0.28 & 0.5 & 0.07 & 0.01 \\ 0.04 & -0.16 & -0.59 & -0.11 & -0.25 & 0.3 & -0.06 \\ 0.06 & -0.34 & 0.1 & 0.33 & 0.38 & -0. & 0. \\ -0.17 & 0.36 & 0.33 & -0.16 & -0.21 & 0.17 & -0.03 \\ 0.11 & -0.43 & 0.07 & 0.08 & -0.17 & -0.28 & 0.02 \\ 0.11 & -0.43 & 0.07 & 0.08 & -0.17 & -0.28 & 0.02 \\ -0.14 & 0.33 & 0.19 & 0.11 & 0.27 & -0.03 & 0.02 \\ 0.27 & -0.18 & -0.03 & -0.54 & 0.08 & 0.47 & 0.04 \\ 0.49 & 0.23 & 0.02 & 0.59 & -0.39 & 0.29 & -0.25 \\ 0.62 & 0.22 & 0. & -0.07 & 0.11 & -0.16 & 0.68 \\ 0.45 & 0.14 & -0.01 & -0.3 & 0.28 & -0.34 & -0.68 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2.54 & 0 & 0 & 0 & 0 \\ 0 & 2.35 & 0 & 0 & 0 \\ 0 & 0 & 1.64 & 0 & 0 \\ 0 & 0 & 0 & 1.50 & 0 \\ 0 & 0 & 0 & 0 & 1.31 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.61 & -0.46 & -0.54 & -0.28 & -0. & -0.01 & -0.02 \\ 0.17 & -0.13 & -0.23 & 0.11 & 0.19 & 0.44 & 0.62 \\ -0.5 & 0.21 & 0.57 & -0.51 & 0.1 & 0.19 & 0.25 \\ -0.03 & 0.04 & 0.27 & 0.15 & 0.02 & 0.02 & 0.01 \\ -0.21 & 0.38 & -0.21 & 0.33 & 0.39 & 0.35 & 0.15 \\ -0.26 & 0.72 & -0.37 & 0.03 & -0.3 & -0.21 & 0. \\ 0.43 & 0.24 & -0.26 & -0.67 & 0.34 & 0.15 & -0.25 \\ -0.05 & -0.01 & 0.02 & 0.06 & -0.45 & 0.76 & -0.45 \\ -0.24 & -0.02 & 0.08 & 0.26 & 0.62 & -0.02 & -0.52 \end{pmatrix}$$

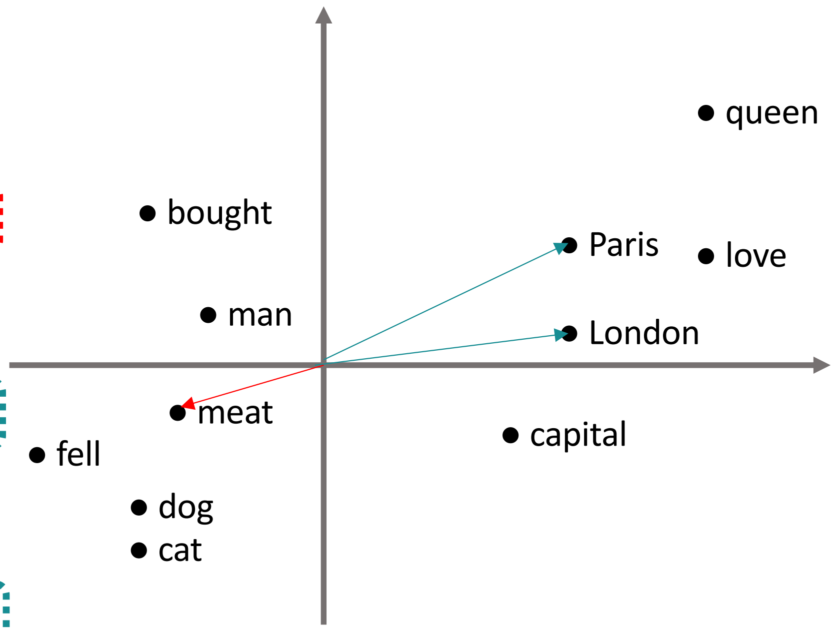
```
up,sp,vp=u[:,0:2],np.diag(s[0:2]),vh[:,0:2]
```

Truncated SVD

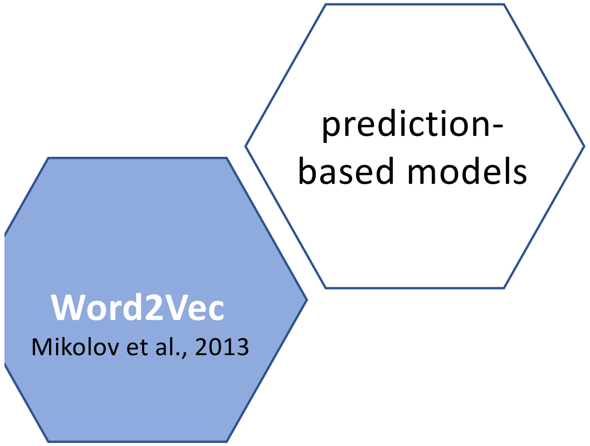
Latent / hidden variables that are hidden in the data

-> For LSA these represent topics

	topic1	topic2
bought	0.665	0.555
dog	0.234	0.543
pet	0.233	0.294
shop	0.198	0.214
very	0.189	-0.333
much	0.222	0.222
likes	0.264	0.331
meat	-0.179	-0.222
fell	0.205	0.122
love	0.321	-0.667
Paris	0.263	0.832
capital	0.169	0.333
France	0.665	0.555
London	0.234	0.587
England	0.273	0.294



Sim (France, London) = 0.89
Sim (London, meat) = -0.27


prediction-
based models
Word2Vec
Mikolov et al., 2013

prediction-based models

Word2Vec

Paris capital of France

Queen beautiful woman

training data

prediction-based models

Word2Vec

Paris capital ~~of~~ France

Queen beautiful woman

training data

window size 2 -> skip-gram

word	neighbour
Paris	capital
Paris	France
capital	Paris
capital	France
France	capital
France	Paris
queen	beautiful
queen	woman
beautiful	queen
beautiful	women
woman	queen
woman	beautiful

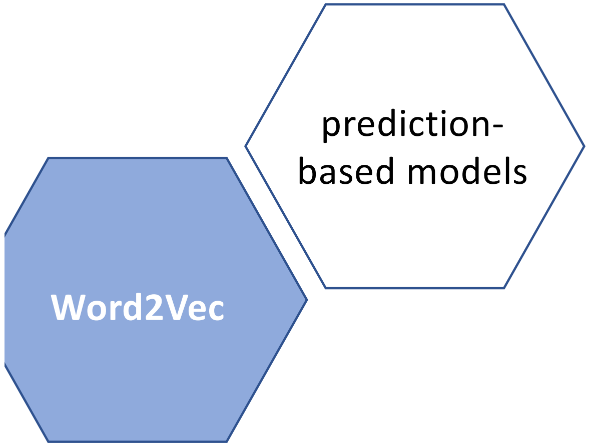
window size 1 -> skip-gram

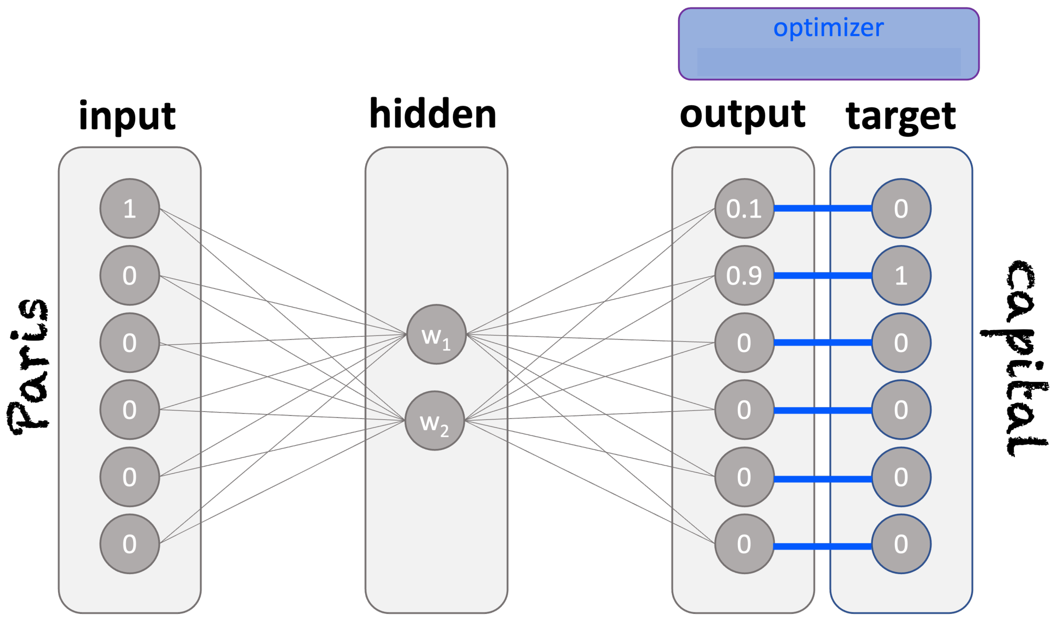
word	neighbour
Paris	capital
capital	Paris
capital	France
France	capital
queen	beautiful
beautiful	queen
beautiful	woman

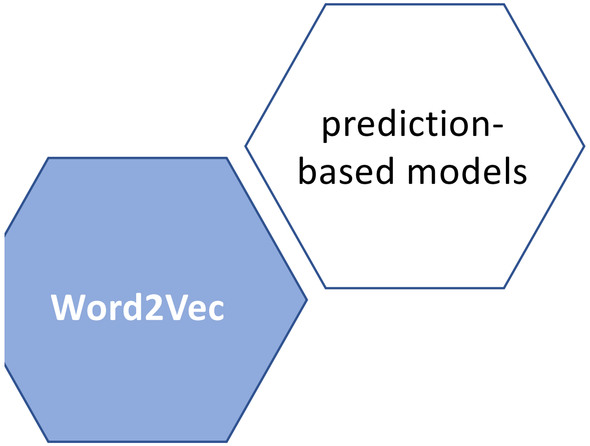
prediction-based models

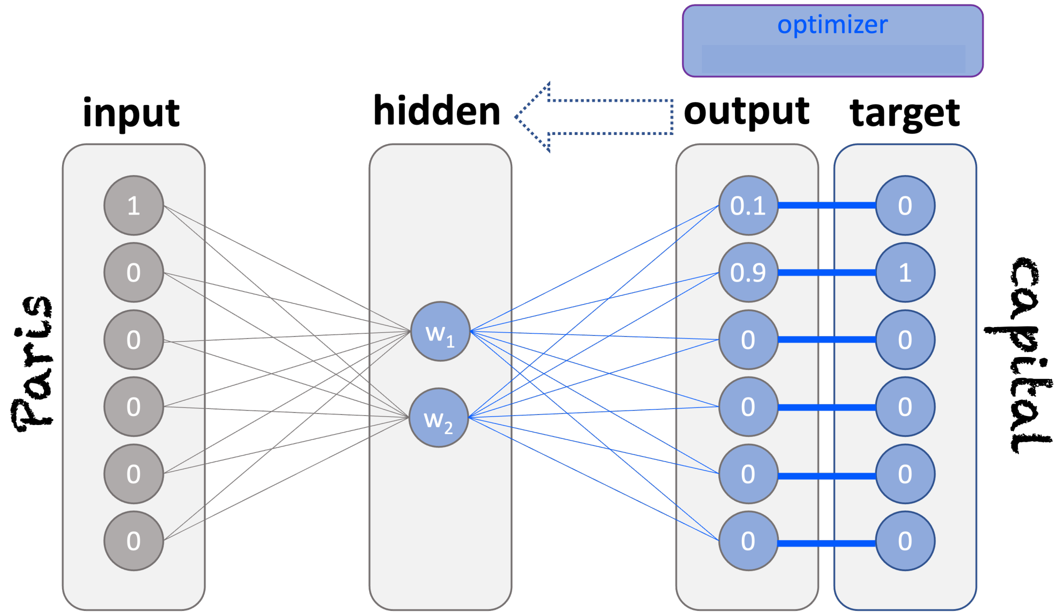
Word2Vec

word	input	neighbour	target
Paris	[1,0,0,0,0,0]	capital	[0,1,0,0,0,0]
Paris	[1,0,0,0,0,0]	France	[0,0,1,0,0,0]
capital	[0,1,0,0,0,0]	Paris	[1,0,0,0,0,0]
capital	[0,1,0,0,0,0]	France	[0,0,1,0,0,0]
France	[0,0,1,0,0,0]	Paris	[1,0,0,0,0,0]
France	[0,0,1,0,0,0]	capital	[0,1,0,0,0,0]
queen	[0,0,0,1,0,0]	beautiful	[0,0,0,0,1,0]
queen	[0,0,0,1,0,0]	woman	[0,0,0,0,0,1]
beautiful	[0,0,0,0,1,0]	queen	[0,0,0,1,0,0]
beautiful	[0,0,0,0,1,0]	women	[0,0,0,0,0,1]
woman	[0,0,0,0,0,1]	queen	[0,0,0,1,0,0]
woman	[0,0,0,0,0,1]	beautiful	[0,0,0,0,1,0]


prediction-based models
Word2Vec

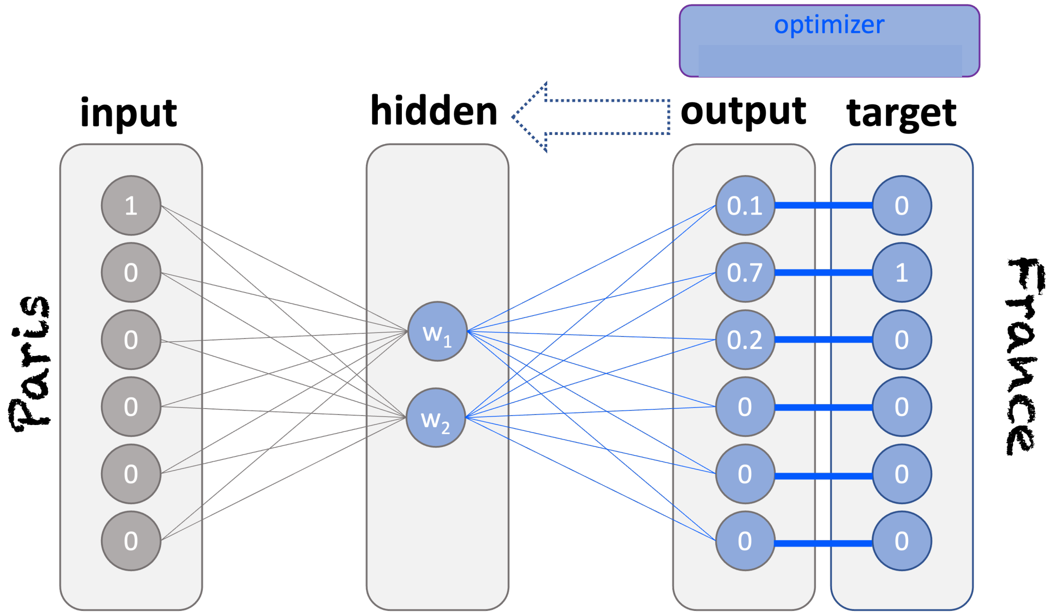

optimizer
input
hidden
output
target
Paris
capital
1
0
0
0
0
0
w_1
w_2
0.1
0.9
0
0
0
0
0
1
0
0
0
0

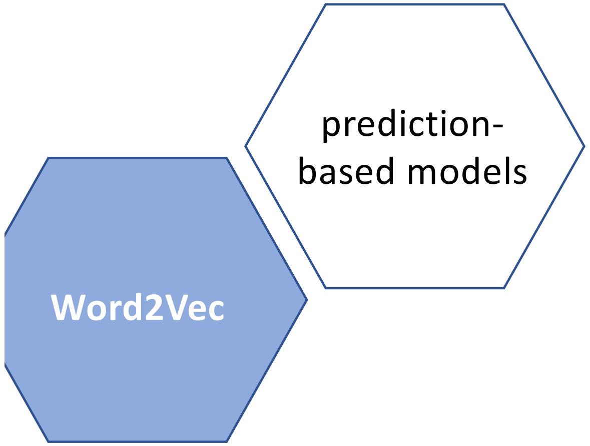

prediction-
based models
Word2Vec

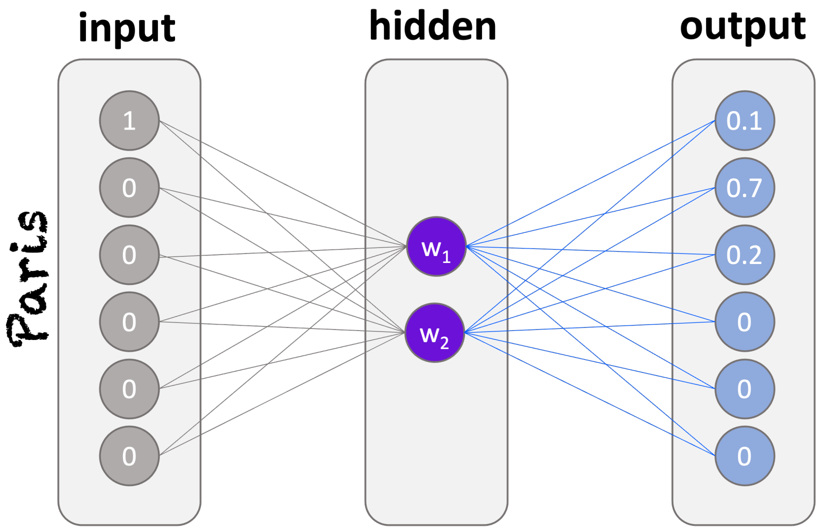

optimizer
input
hidden
output
target
Paris
1
0
0
0
0
0
w1
w2
0.1
0.9
0
0
0
0
0
1
0
0
0
0
capital

prediction-based models

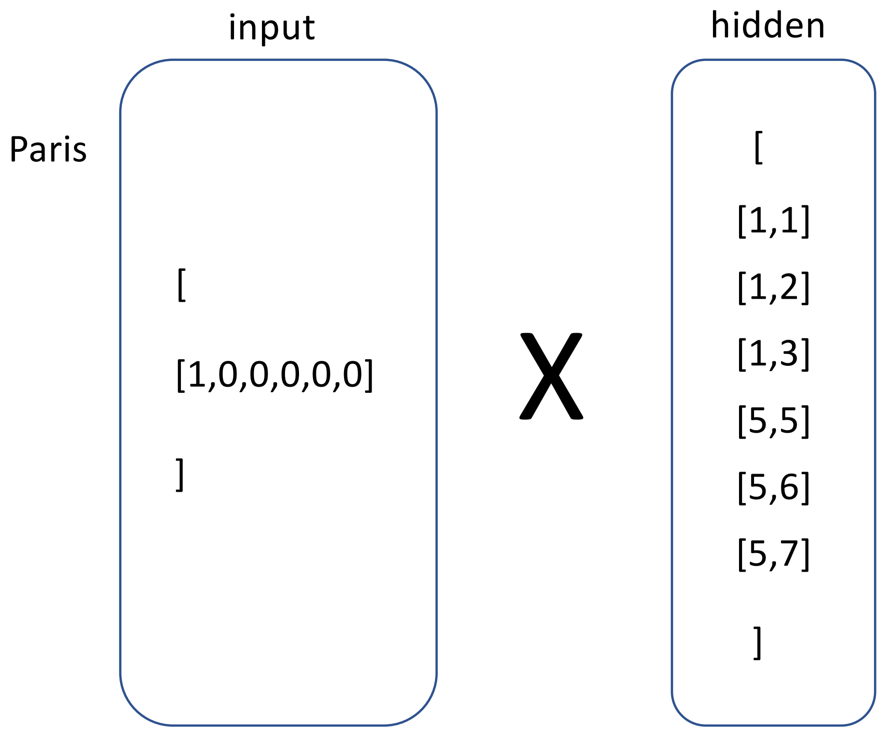
Word2Vec



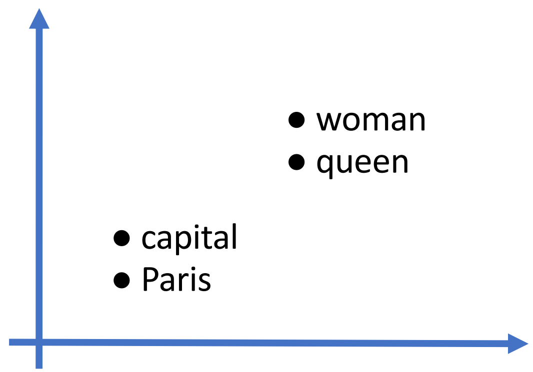




word	embedding
Paris	[1,1]
capital	[1,2]
France	[1,3]
queen	[5,5]
beautiful	[5,5]
woman	[5,6]

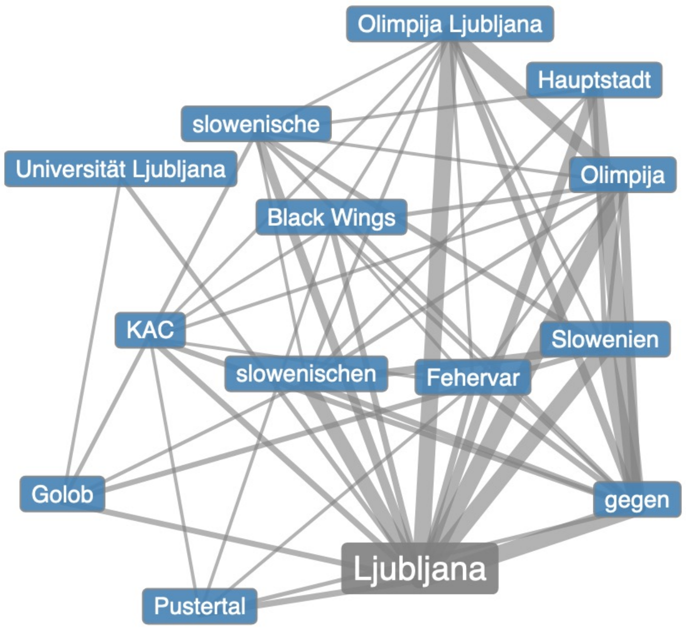


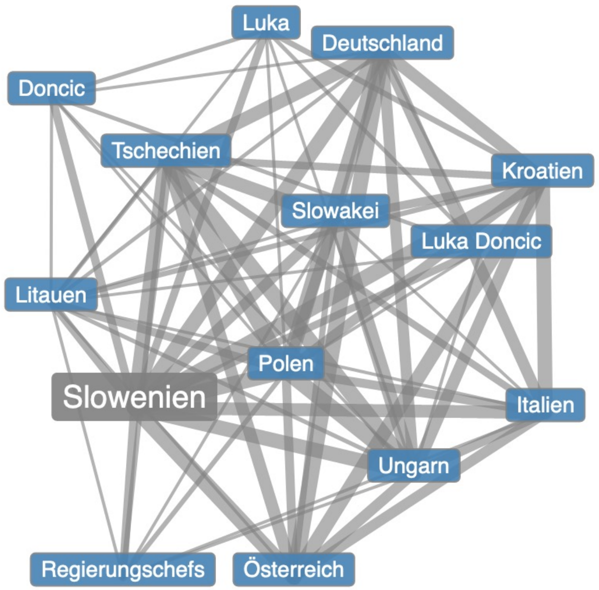
Unique word	encoding	word2vec embedding
Paris	[1,0,0,0,0,0]	[1,1]
capital	[0,1,0,0,0,0]	[1,3]
queen	[0,0,0,1,0,0]	[5,5]
woman	[0,0,0,0,0,1]	[5,7]

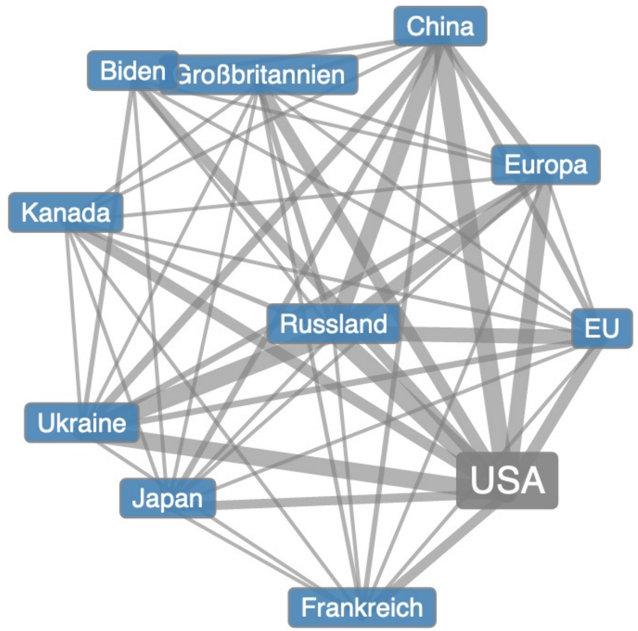


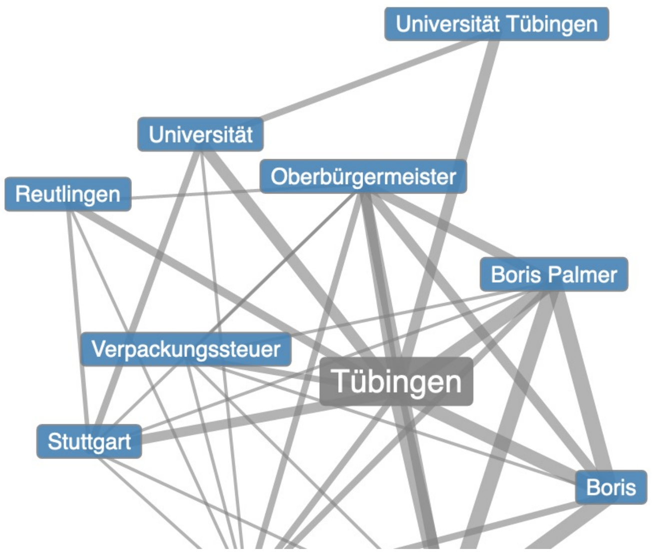
So this was all a bit abstract …

- How does the result of such models look in the best case?


Olimpija Ljubljana
Hauptstadt
slowenische
Universität Ljubljana
Olimpija
Black Wings
KAC
Slowenien
slowenischen
Fehervar
Golob
gegen
Ljubljana
Pustertal


Luka
Deutschland
Doncic
Tschechien
Kroatien
Slowakei
Luka Doncic
Litauen
Polen
Slowenien
Italien
Ungarn
Regierungschefs
Österreich


China
Biden
Großbritannien
Europa
Kanada
Russland
EU
Ukraine
USA
Japan
Frankreich


Universität Tübingen
Universität
Oberbürgermeister
Reutlingen
Boris Palmer
Verpackungssteuer
Tübingen
Stuttgart
Boris

Cognitive Reality?

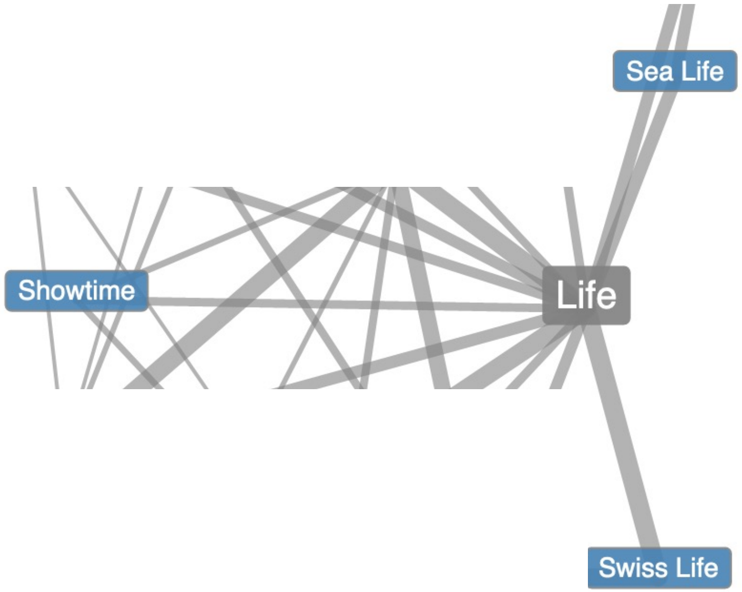
- Stochastic Parrot (Emily Bender, 2021)
  - Maybe we are all stochastic parrots?

- No human in their entire lifetime will have that amount of language experience like LLMs
  - So just because LLMs do it that way -> does not mean that humans learn / process language that way (?)

- Do children learn language like LLMs?

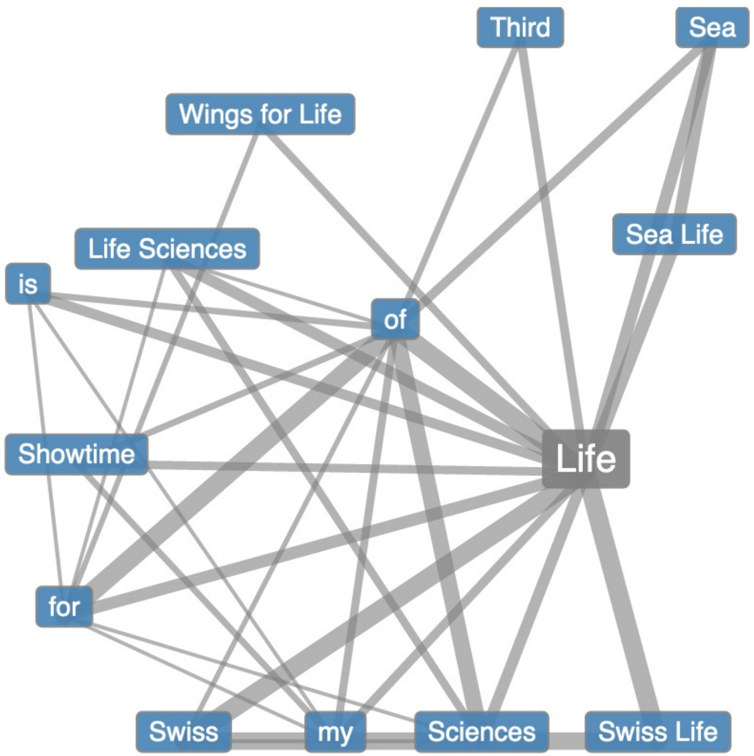
Cognitive Reality?

Cognitive Reality?



SwissLife

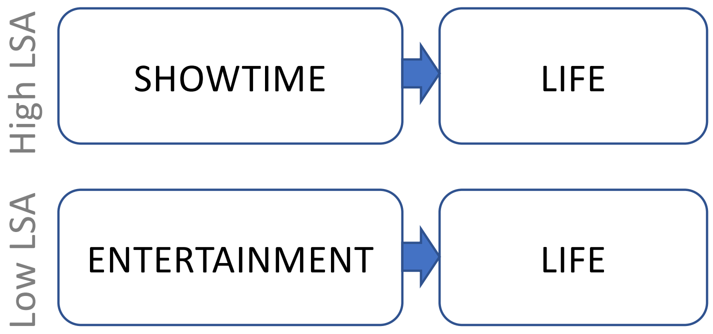
Cognitive Reality?

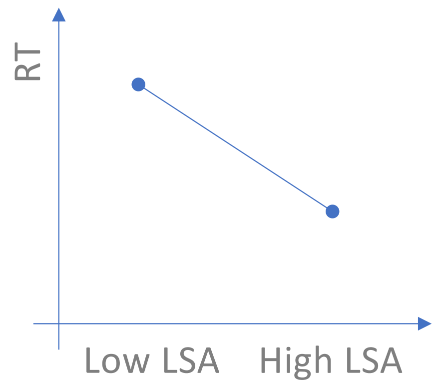


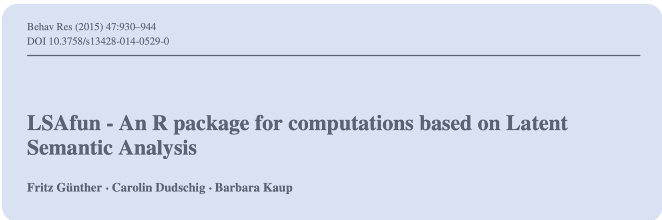
-> Many priming studies in this field
-> Chose words that are not intuitively related / unrelated

Cognitive Reality: How is this tested?

- RT Priming studies

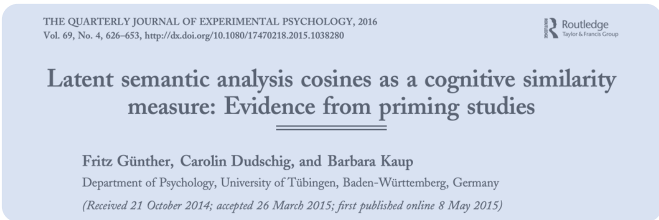





Behav Res (2015) 47:930–944
DOI 10.3758/s13428-014-0529-0

LSAfun - An R package for computations based on Latent Semantic Analysis

Fritz Günther · Carolin Dudschig · Barbara Kaup

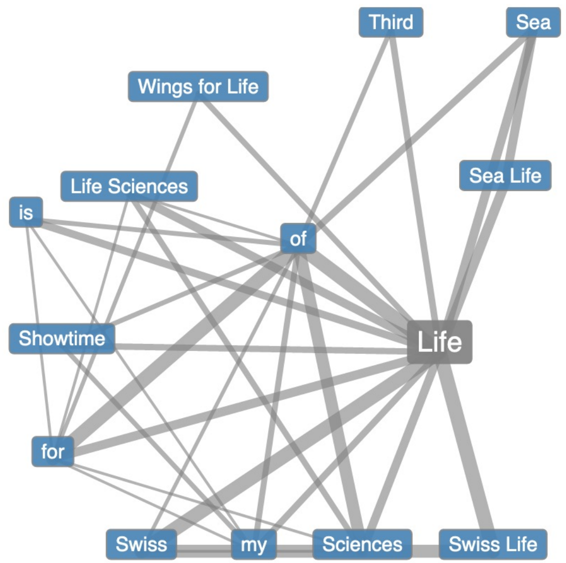

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2016
Vol. 69, No. 4, 626–653, http://dx.doi.org/10.1080/17470218.2015.1038280
Routledge

Latent semantic analysis cosines as a cognitive similarity measure: Evidence from priming studies

Fritz Günther, Carolin Dudschig, and Barbara Kaup
Department of Psychology, University of Tübingen, Baden-Württemberg, Germany

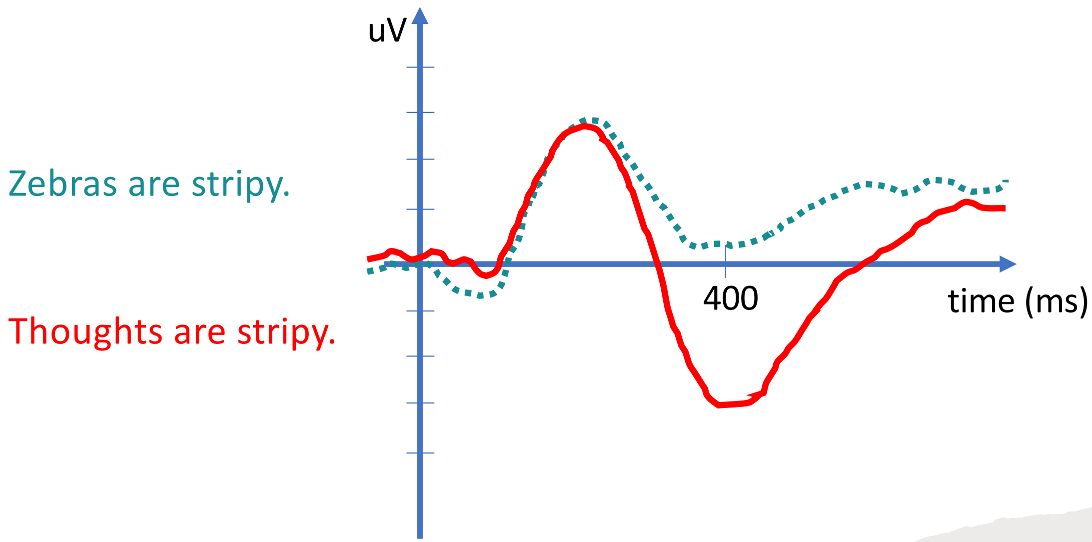
(Received 21 October 2014; accepted 26 March 2015; first published online 8 May 2015)

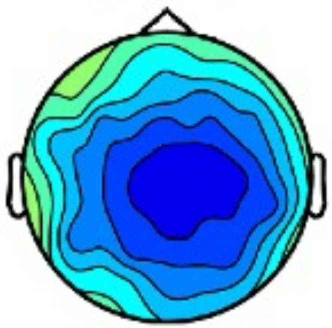
- N400 Studies



N400

- Negative going centro-parietal component. Peaking around 400ms post-stimulus onset
- Part of the brain's response to **meaningful** stimuli (e.g., words, pictures, environmental sounds)





What is the N400?

- Integration View: Sentence integration results / Compositional

Zebras are really stripy.
Zebras are <u>not</u> stripy.

Same N400 amplitude

- Lexical View: Long term memory associations

- Prediction-based View: Best prediction in current context

What is the N400?

- Integration View: Sentence integration results / Compositional

Zebras are really stripy.
Zebras are <u>not</u> stripy.

Same N400 amplitude

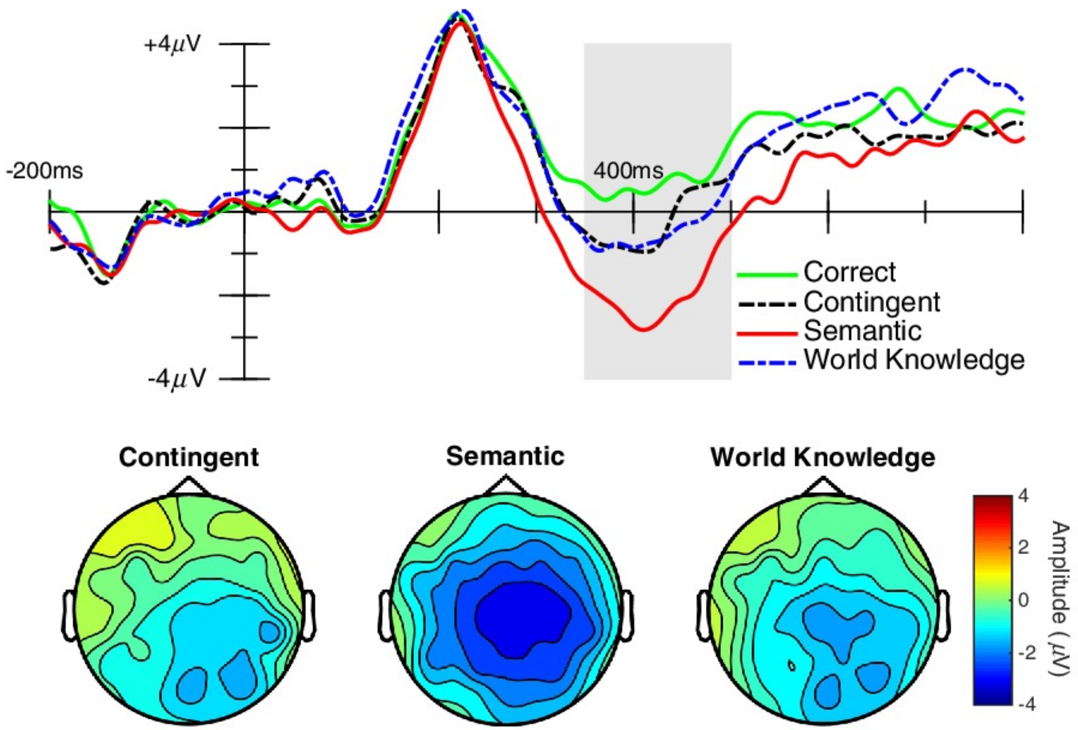
- Lexical View: Long term memory associations

LSA

- Prediction-based View: Best prediction in current context

word2vec

Methods



Zebras are stripy.

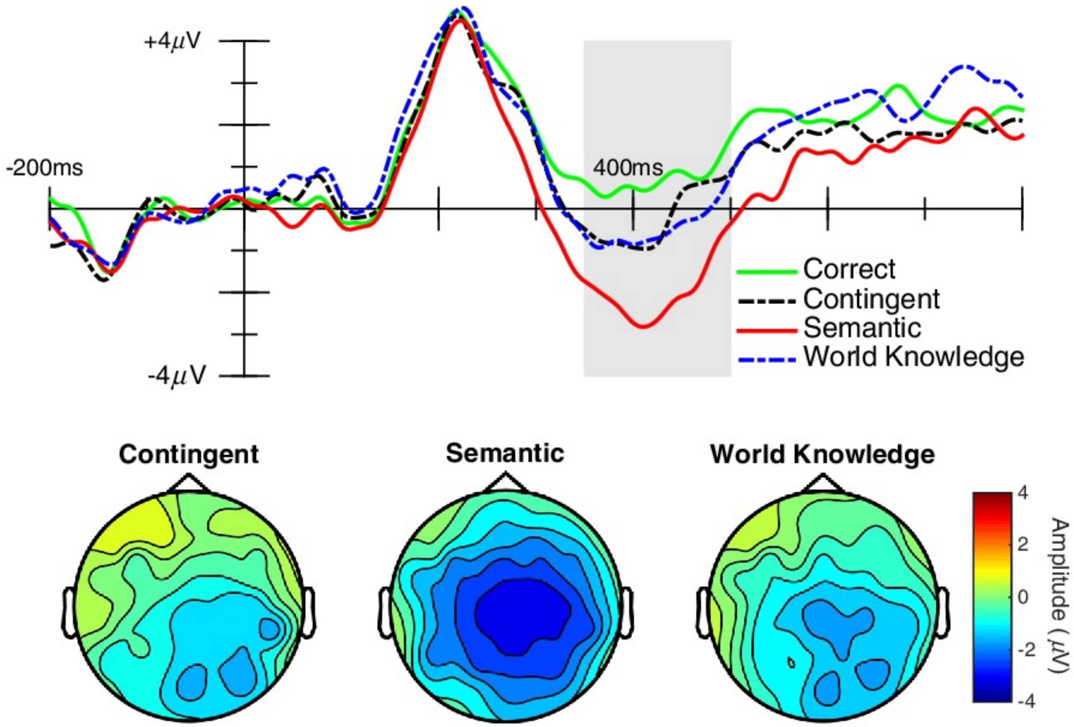
Thoughts are stripy.

Ladybirds are stripy.

Trousers are stripy.

Hagoort et al. (2004) *Science*

Methods: Data




Biological Psychology 169 (2022) 108282

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Language and non-linguistic cognition: Shared mechanisms and principles reflected in the N400

Carolin Dudschig

University of Tübingen, Schleichstr. 4, 72076 Tübingen, Germany


Brain and Cognition 103 (2016) 38–49

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Is there a difference between stripy journeys and stripy ladybirds? The N400 response to semantic and world-knowledge violations during sentence processing

Carolin Dudschig*, Claudia Maienborn, Barbara Kaup

University of Tübingen, Germany


Neuroscience Letters 630 (2016) 141–146

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research article

These lemons are sour: Investigating the influence of demonstrative determiners on the N400 complex

Carolin Dudschig*, Claudia Maienborn, Barbara Kaup

University of Tübingen, Germany


LANGUAGE, COGNITION AND NEUROSCIENCE
https://doi.org/10.1080/23273798.2018.1535127

Routledge
Taylor & Francis Group

REGULAR ARTICLE

Negation and the N400: investigating temporal aspects of negation integration using semantic and world-knowledge violations

Carolin Dudschig, Ian Grant Mackenzie, Claudia Maienborn, Barbara Kaup and Hartmut Leuthold

Fachbereich Psychologie, University of Tübingen, Tübingen, Germany

Methods: EEG

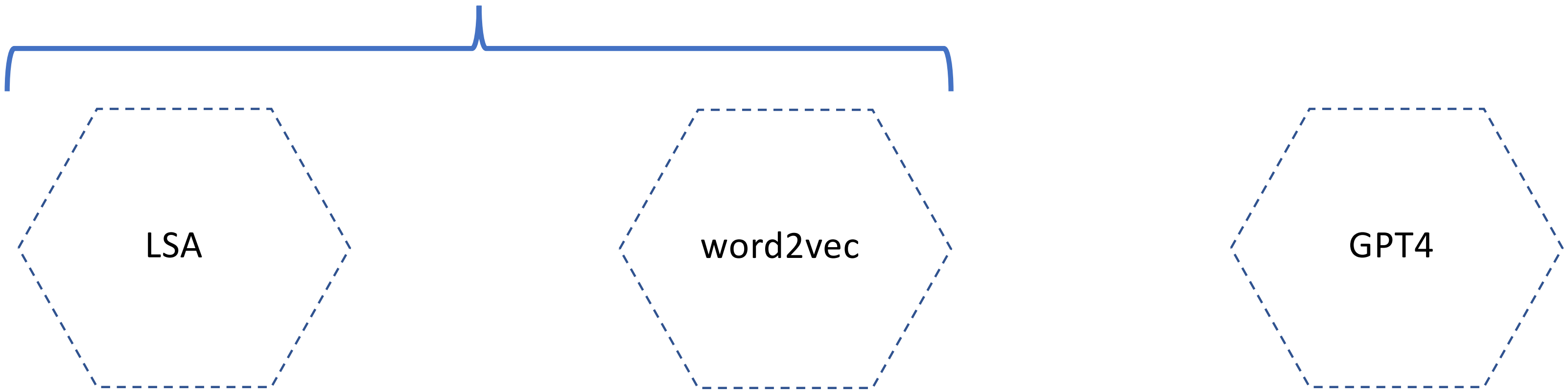
- **107.023** single EEG trials

- Same pre-processing pipeline:

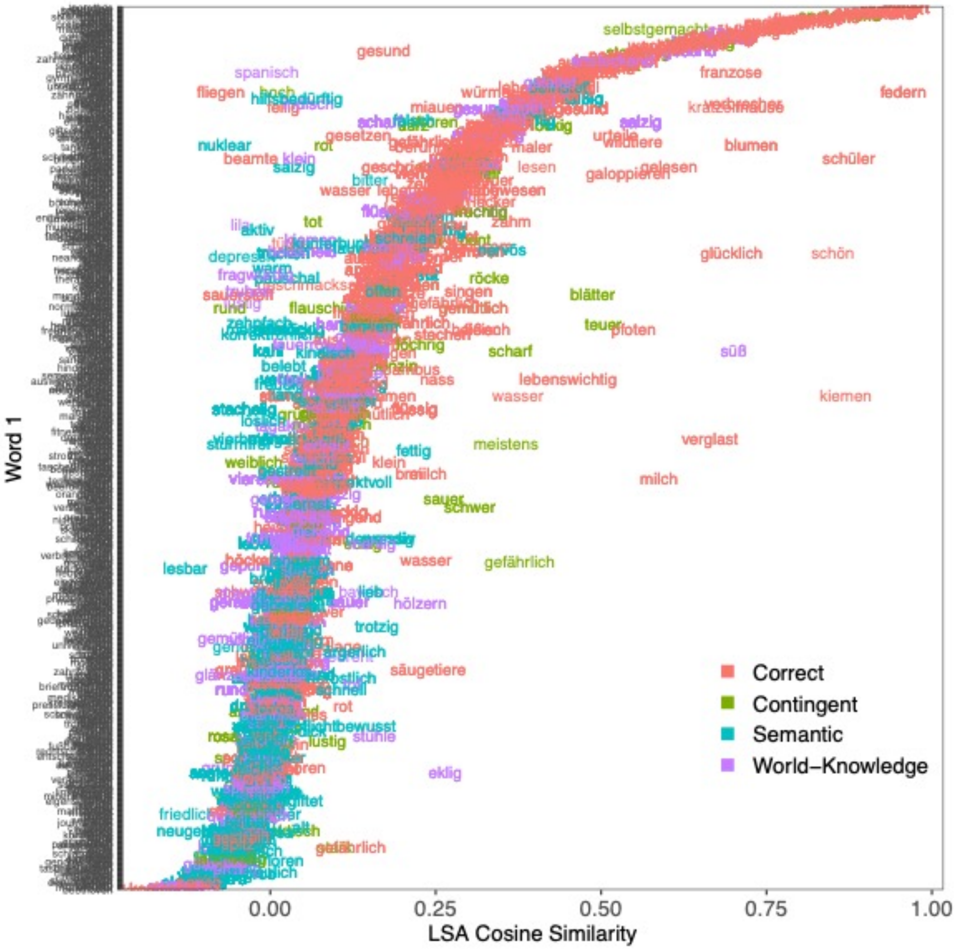
  - 64 standard electrode locations + M1/M2 (mastoids) + IO1/IO2/F9/F10 vEOG/hEOG electrodes combinations
  - All datasets downloaded to lowest common sample rate of 256 Hz
  - Ocular artifacts removed via ICA
  - 0.1 Hz high-pass filter to continuous data
  - 30 Hz low-pass filter to epoched data
  - Epoch -500 to 1500 ms around critical word (second word)
  - Baseline to 200ms interval before the onset of critical word
  - N400 analysis is mean amplitude in time-window 300 to 500 ms (mean: C1, Cz, C2, CP1, Cpz, CP2, P1, Pz and P2)

Methods: NLPs

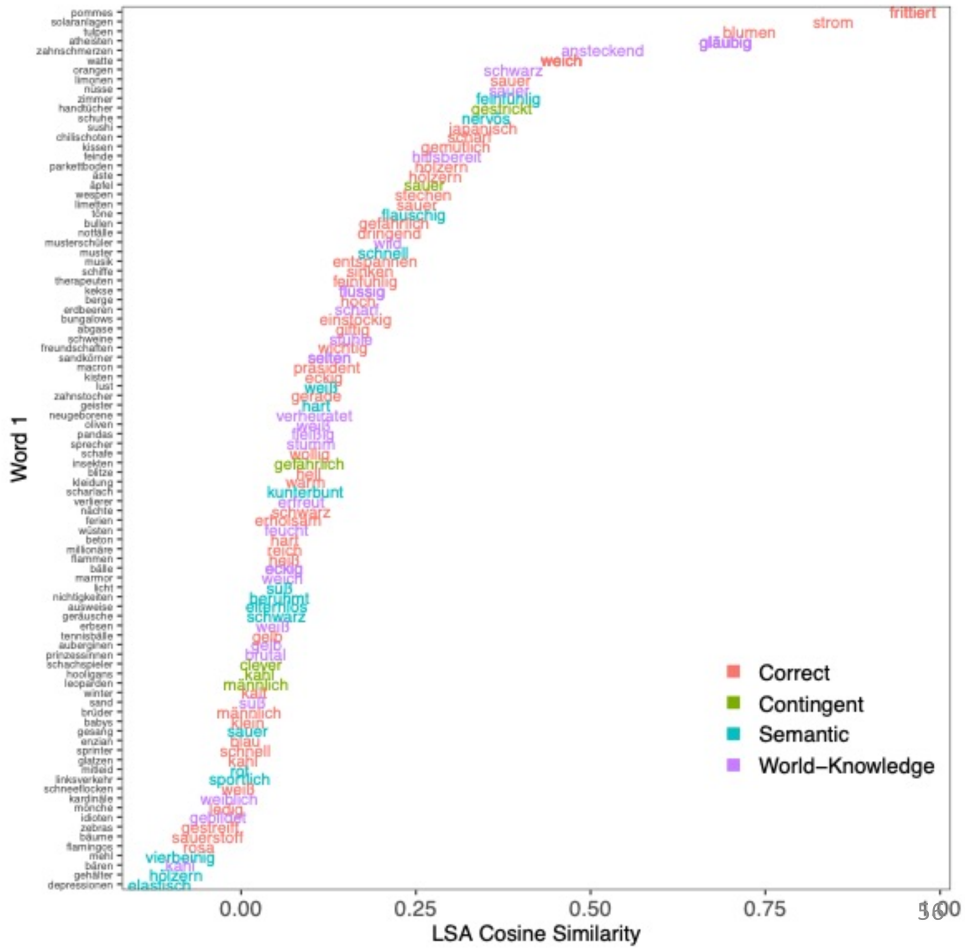
- German Wikipedia: dewiki-latest-pages-articles.xml.bz2 (June 2023, ~ 7GB)
- gensim 4.3.1 (https://pypi.org/project/gensim/)
  - removed words that appear in fewer than 10 documents
  - removed words that appear in more than 10% of documents
  - keep 500.000 most frequent words



Results

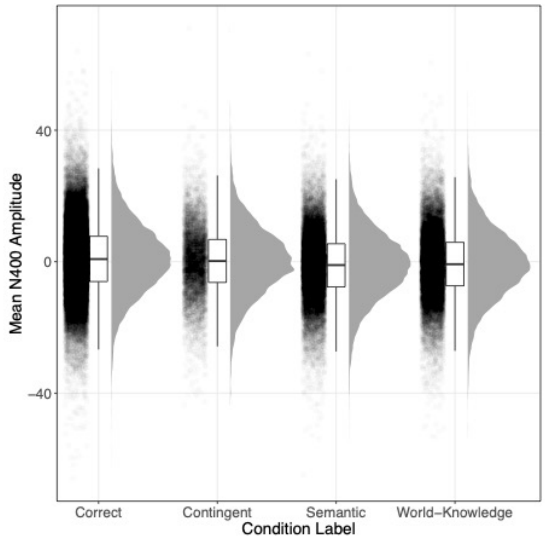


Word 1
0.00
0.25
0.50
0.75
1.00
LSA Cosine Similarity
Correct
Contingent
Semantic
World−Knowledge

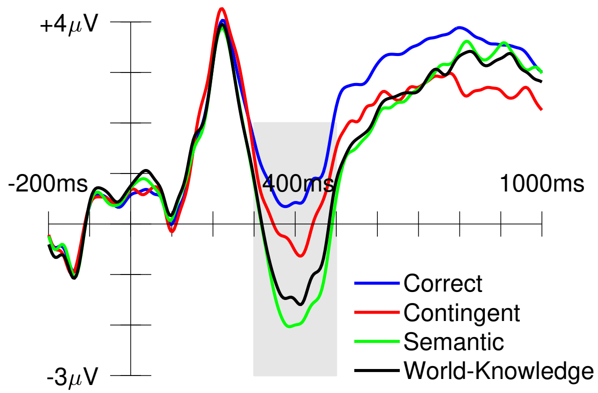


Word 1
0.00
0.25
0.50
0.75
1.00
LSA Cosine Similarity
Correct
Contingent
Semantic
World−Knowledge

Results

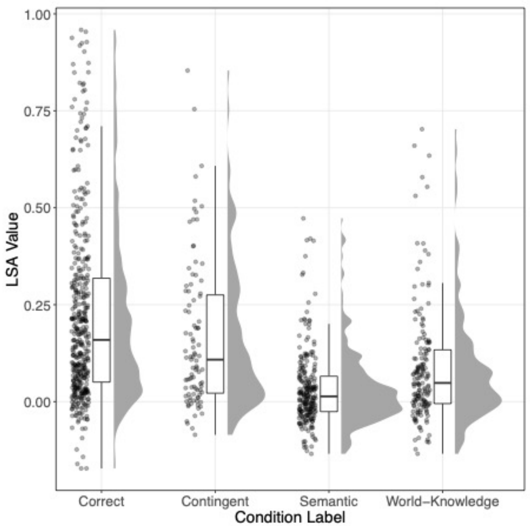


Mean N400 Amplitude
40
0
-40
Correct
Contingent
Semantic
World–Knowledge
Condition Label

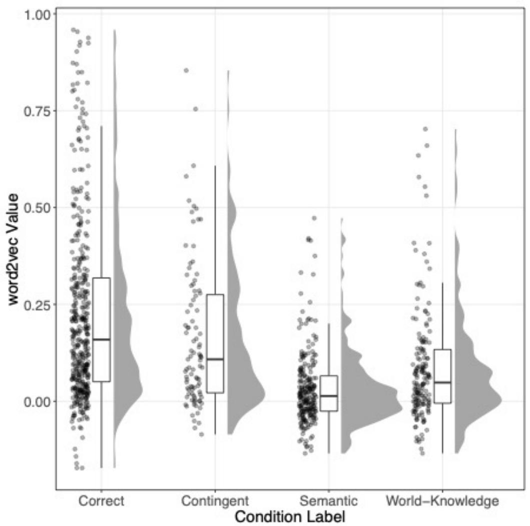


+4μV
-200ms
400ms
1000ms
-3μV
Correct
Contingent
Semantic
World-Knowledge

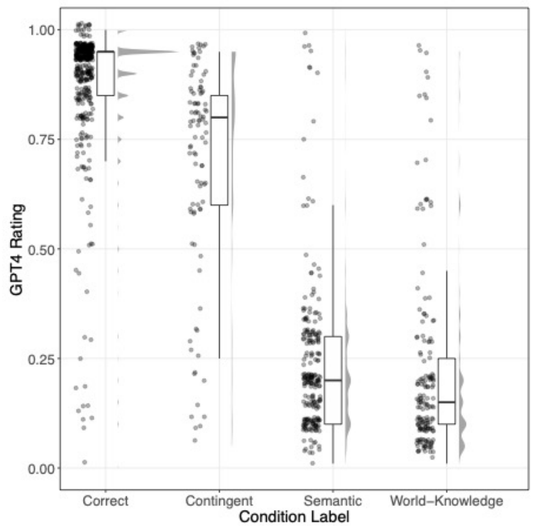
Results



1.00
0.75
0.50
0.25
0.00
LSA Value
Correct
Contingent
Semantic
World−Knowledge
Condition Label

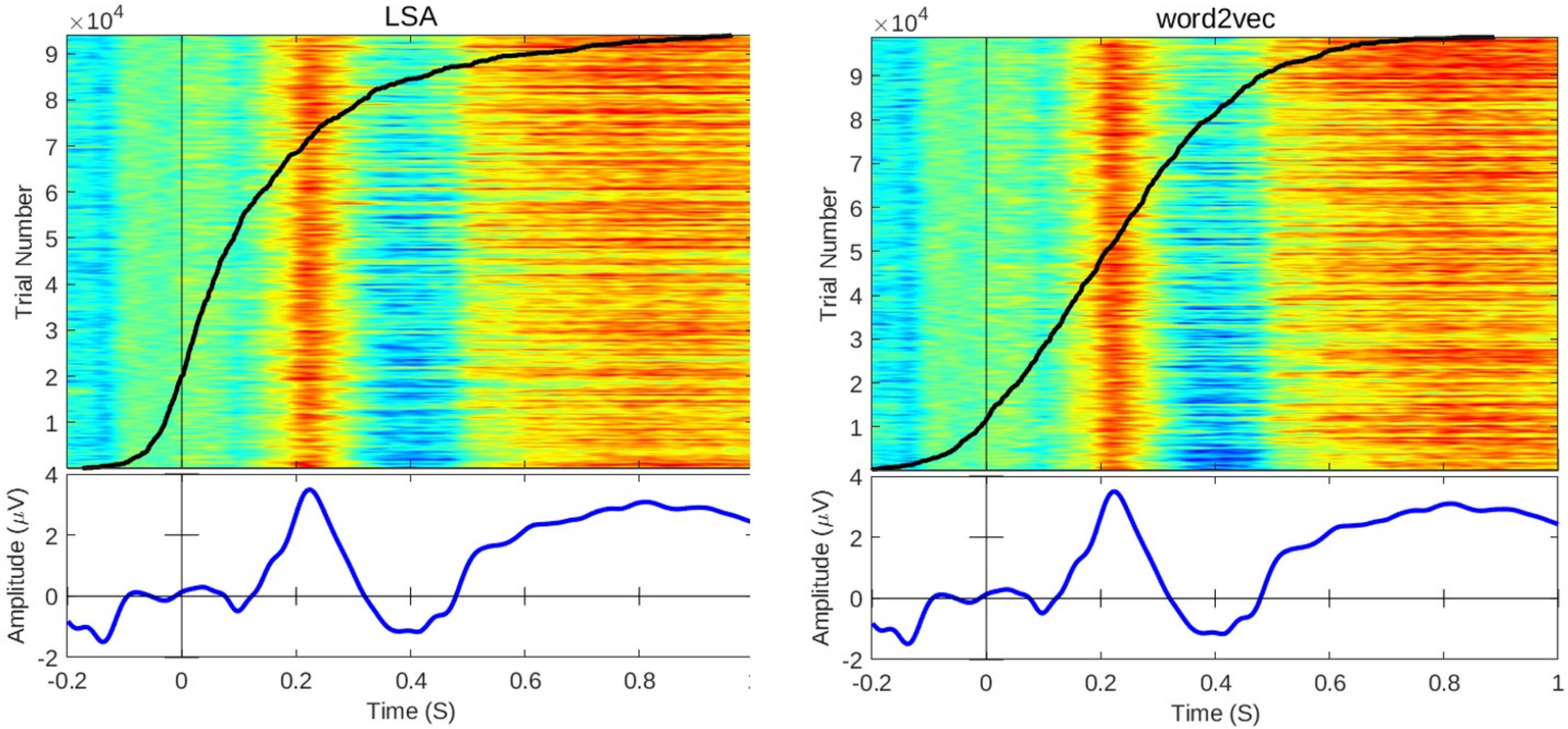


1.00
0.75
0.50
0.25
0.00
word2vec Value
Correct
Contingent
Semantic
World−Knowledge
Condition Label



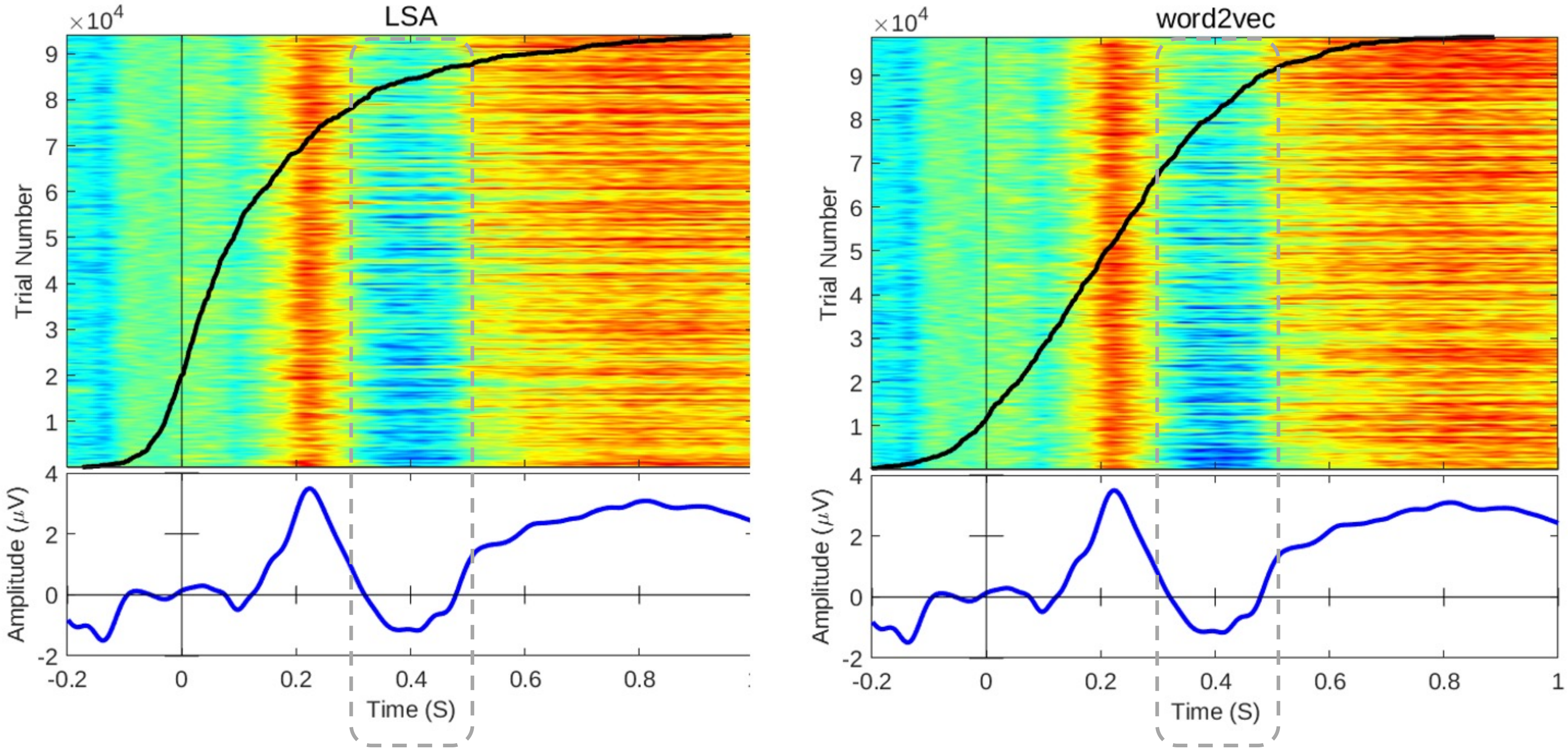
1.00
0.75
0.50
0.25
0.00
GPT4 Rating
Correct
Contingent
Semantic
World−Knowledge
Condition Label

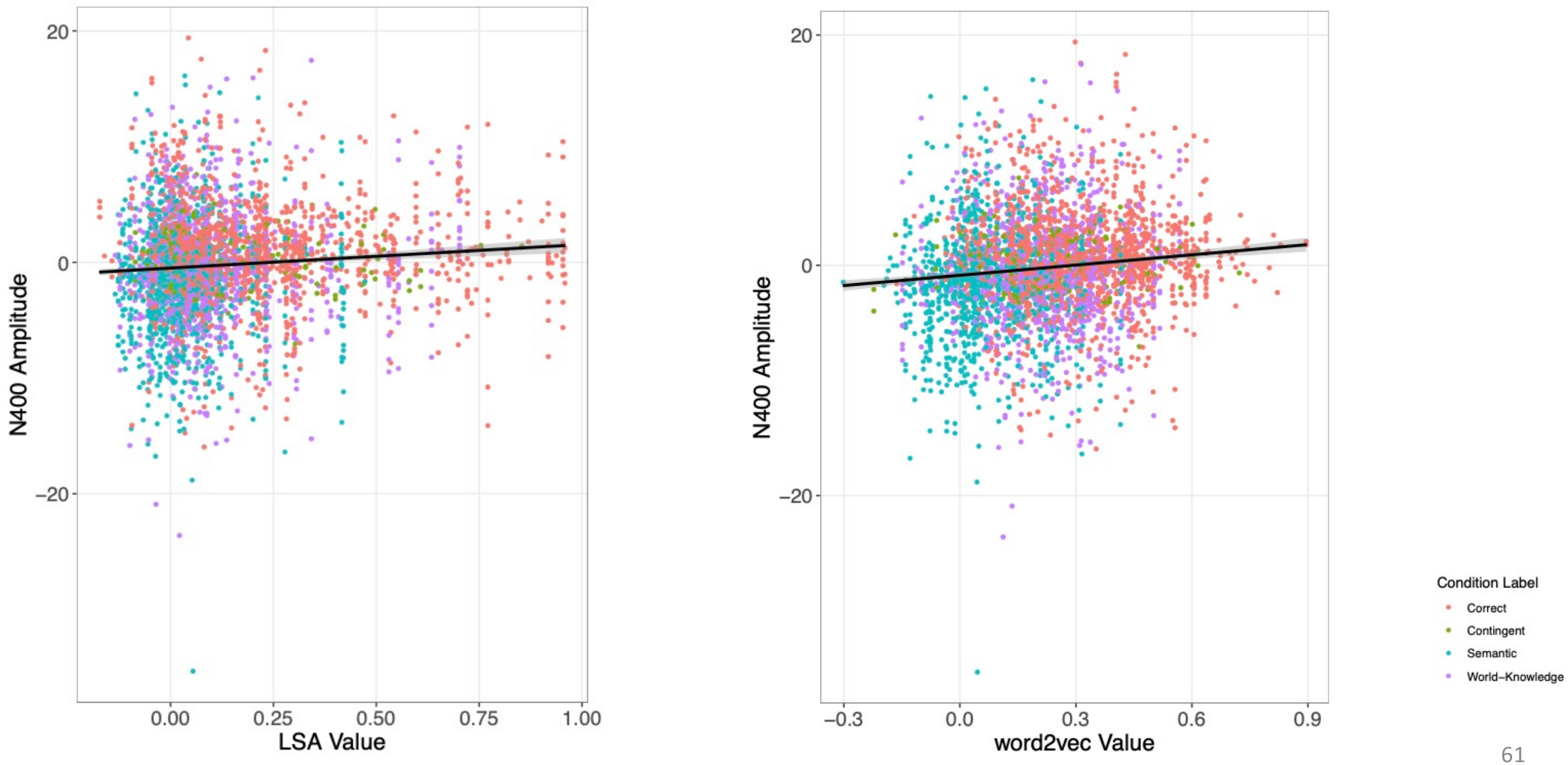
Results

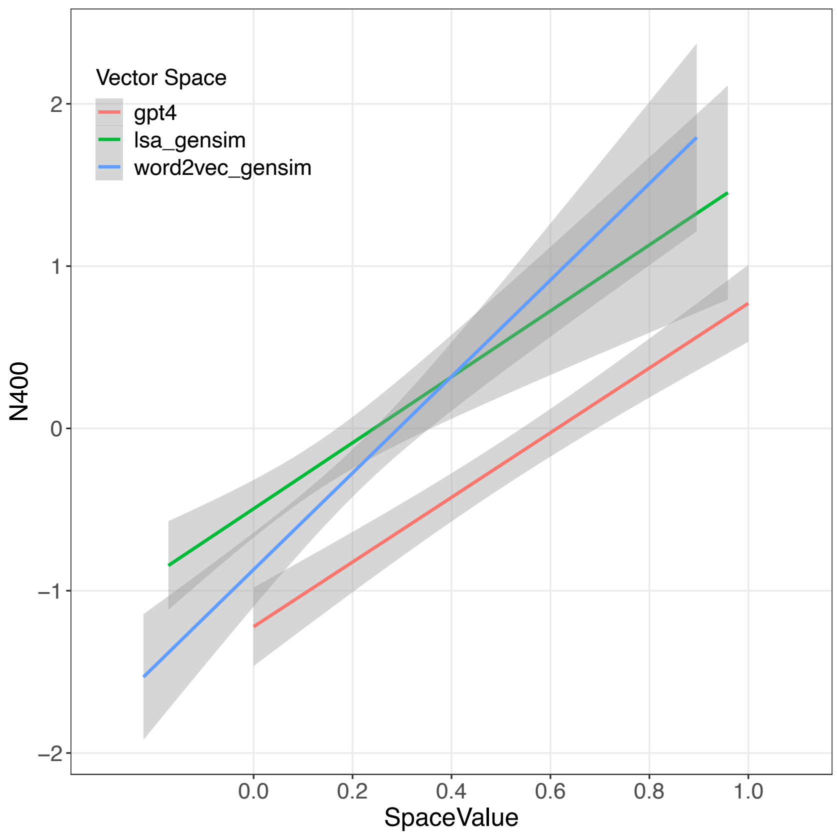


LSA
word2vec
×10^4
Trial Number
Amplitude (μV)
Time (S)
-0.2
0
0.2
0.4
0.6
0.8
1

Results


LSA
×10^4
Trial Number
Amplitude (μV)
Time (S)
word2vec
×10^4
Trial Number
Amplitude (μV)
Time (S)


20
0
−20
N400 Amplitude
0.00
0.25
0.50
0.75
1.00
LSA Value
20
0
−20
N400 Amplitude
−0.3
0.0
0.3
0.6
0.9
word2vec Value
Condition Label
Correct
Contingent
Semantic
World−Knowledge




```
Model: N400 ~ lsa_gensim + word2vec_gensim + GPT4 + (1 | Subject) +
Model:     (1 | Item)
Data: .
Df full model: 7
           Effect df      Chisq p.value
1      lsa_gensim  1       1.97    .160
2 word2vec_gensim  1     4.41 *    .036
3            GPT4  1 135.00 ***   <.001
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘+’ 0.1 ‘ ’ 1
```

Discussion

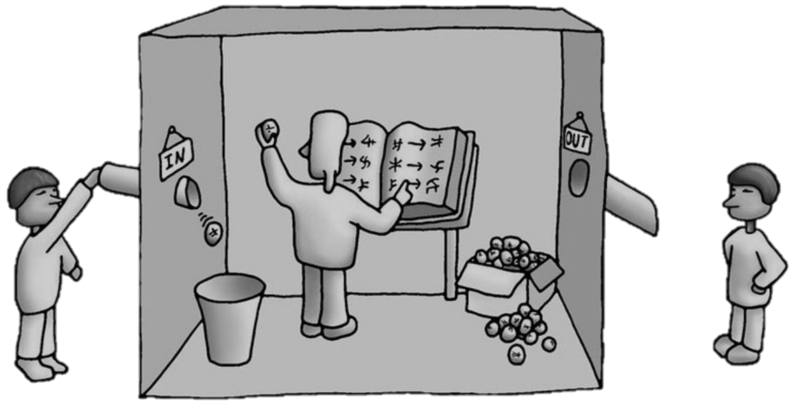
- All NLP techniques investigated here explain certain amount of data
- GPT4 > Word2Vec > LSA
- What does this suggest? Prediction account of N400?

Future

- Random sampling of items across certain LSA, word2vec & GPT4 values
- Single Trial analysis -> Extremely noisy data in EEG -> Improvement?

Future – Is this *all* about meaning?

- Likely: No!
- More likely an interplay of different factors
- Chinese Room problem

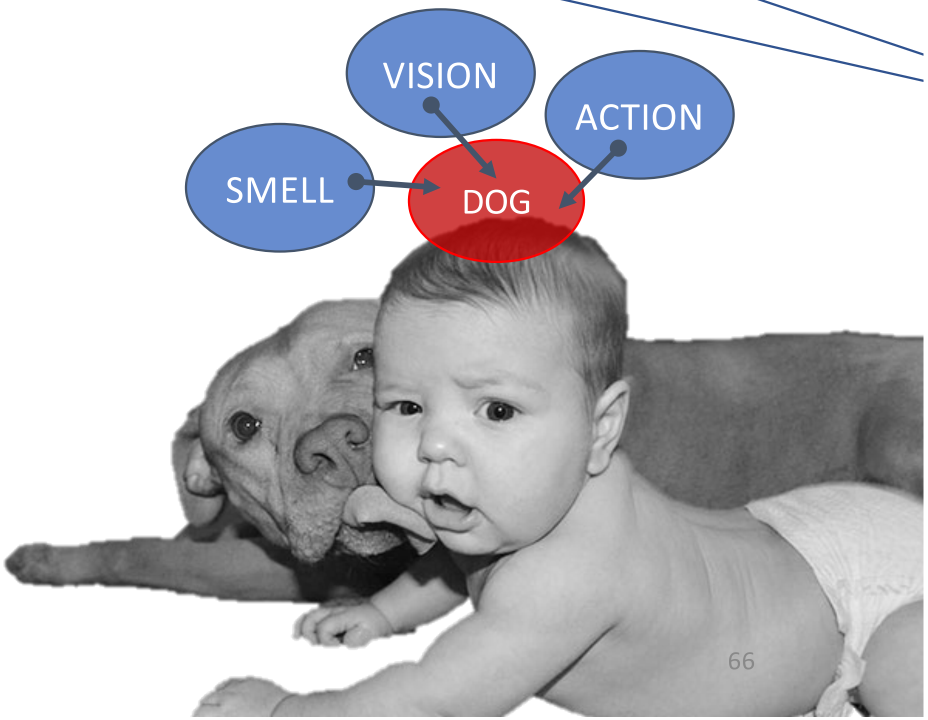


- LLMs are trained on the end products of our culture
- We don't learn language in such an abstract manner

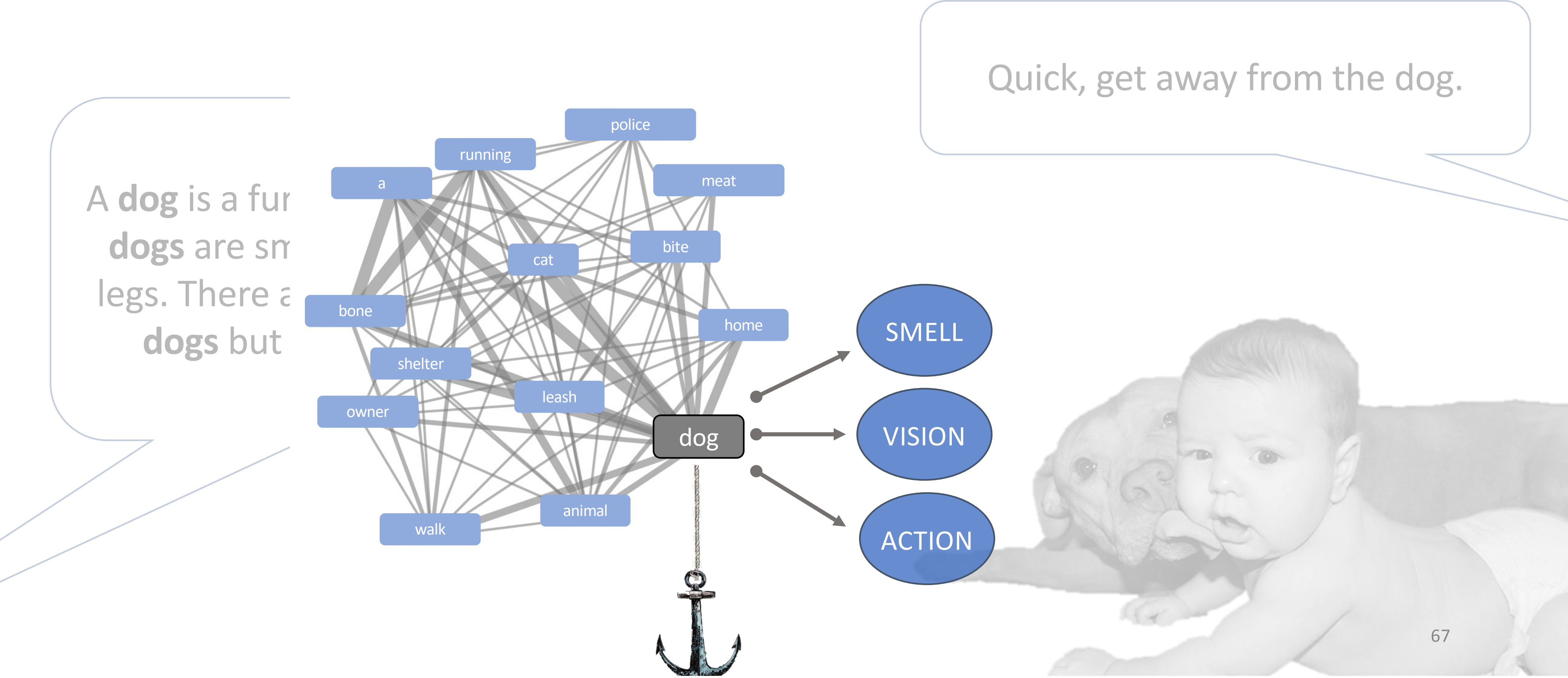
Future – Is this *all* about meaning from a cognitive perspective?

Quick, get away from the dog.

A **dog** is a fury animal. Sometimes **dogs** are smelly, they have four legs. There are different types of **dogs** but they all eat meat. They often look cute and run fast.



Future – Is this *all* about meaning from a cognitive perspective?



Thank you for your attention

- Thanks Ian Mackenzie & Fritz Günther