GDR LIFT - ILFC Seminar

Interactions between formal and computational linguistics

1st June 2021

# What do neural models tell us about the nature of language?

*Juan Luis Gastaldi*

ETH Zürich (D-GESS, Turing Centre)

juan.luis.gastaldi@gess.ethz.ch

Turing Centre

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 839730

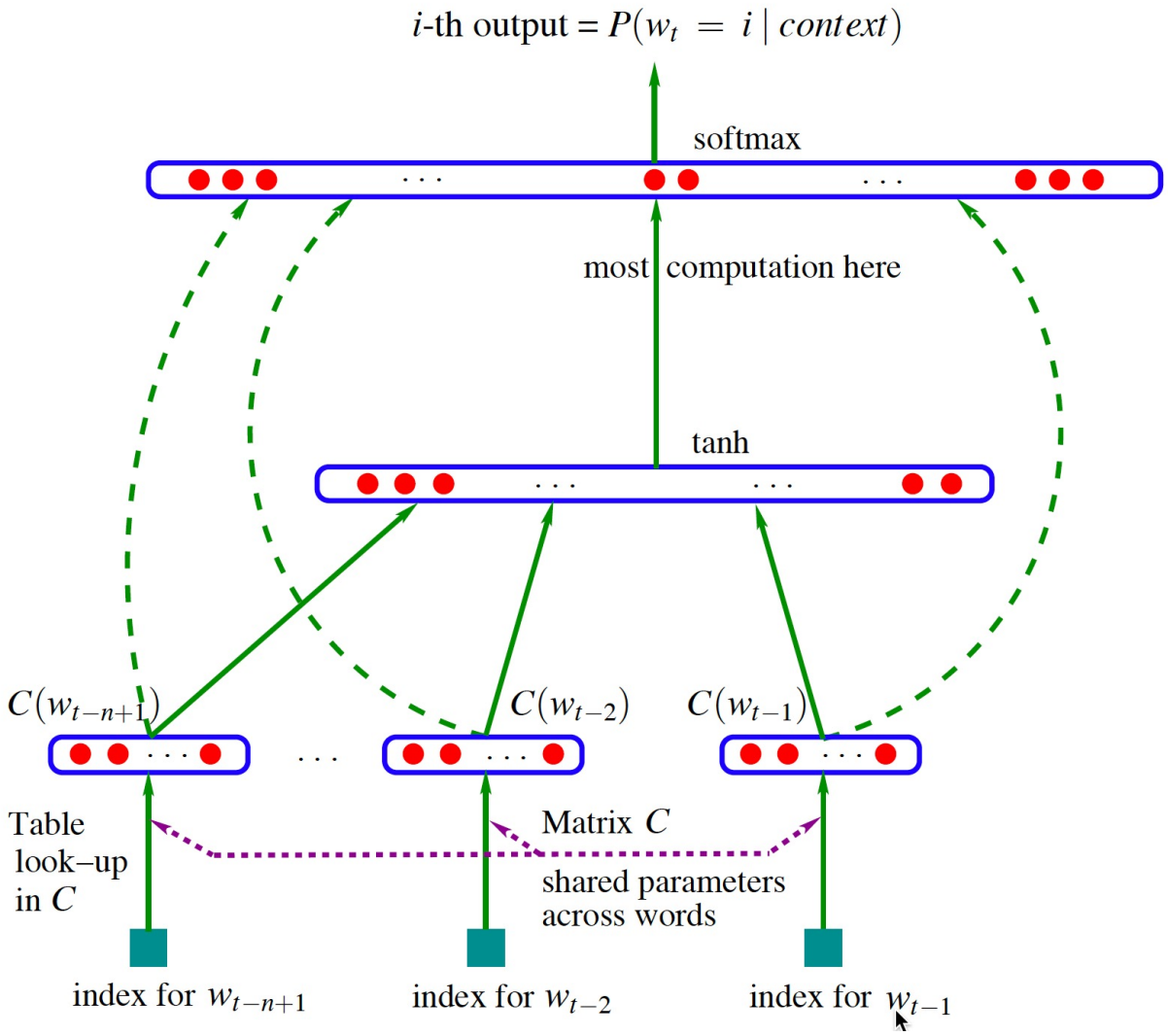
ETH zürich

There is no "philosophy" of language. There is only linguistics.

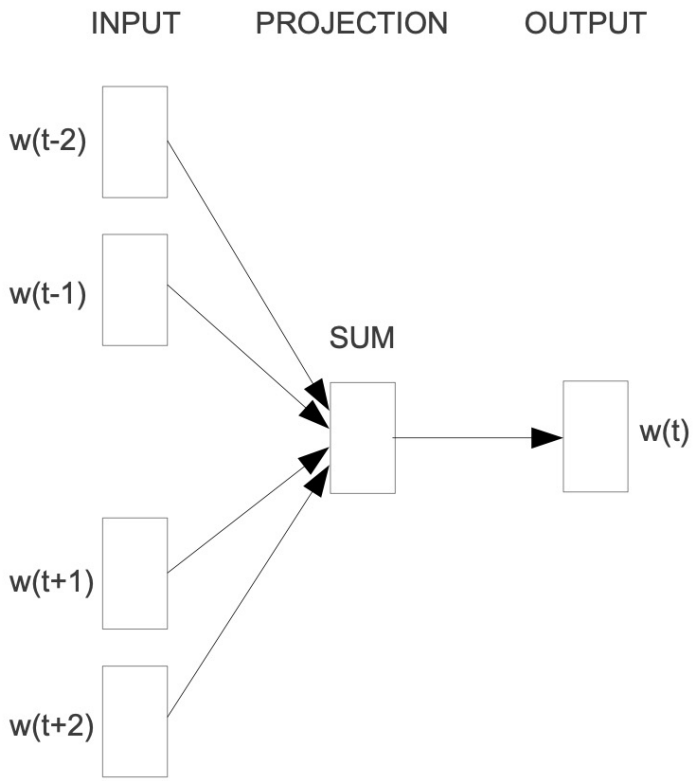
Louis Hjelmslev

*Principes de Grammaire Générale*, 1928

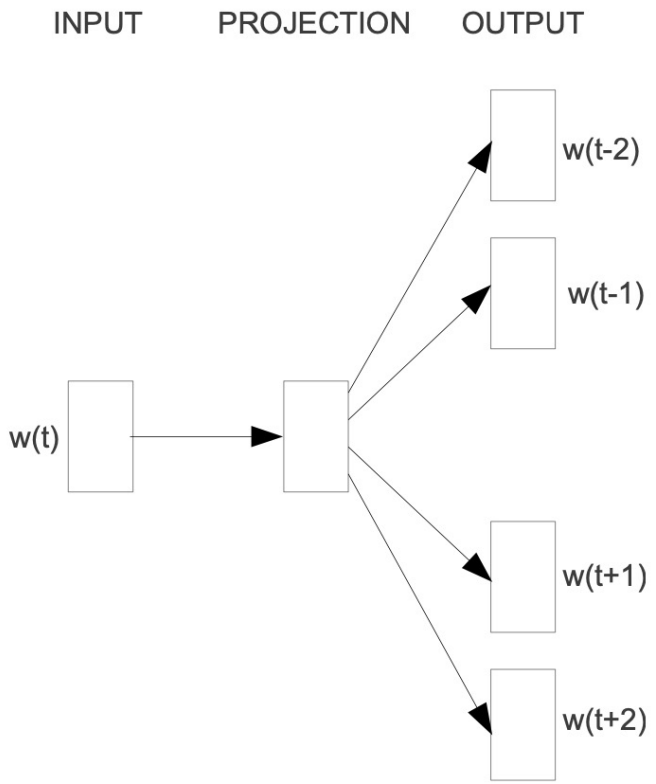
# DNN and Natural Language




Source: Bengio et al., 2003



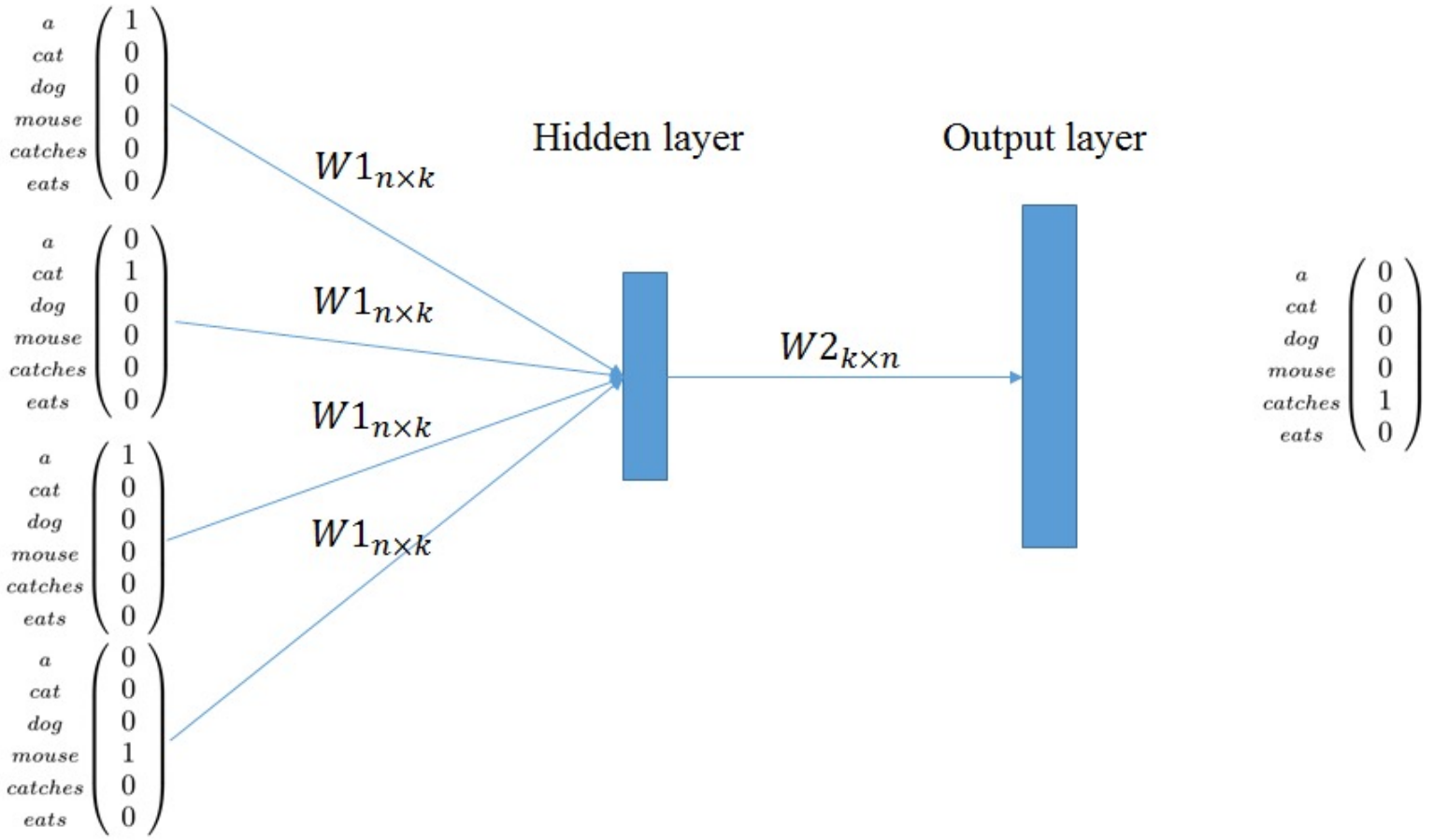

**CBOW**




**Skip-gram**

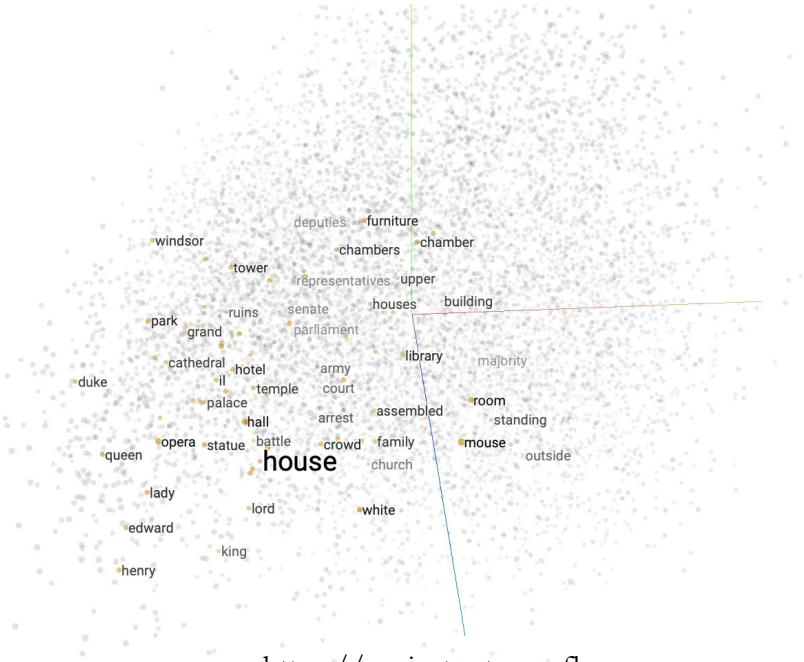
Mikolov et al., 2013

*a cat catches a mouse*

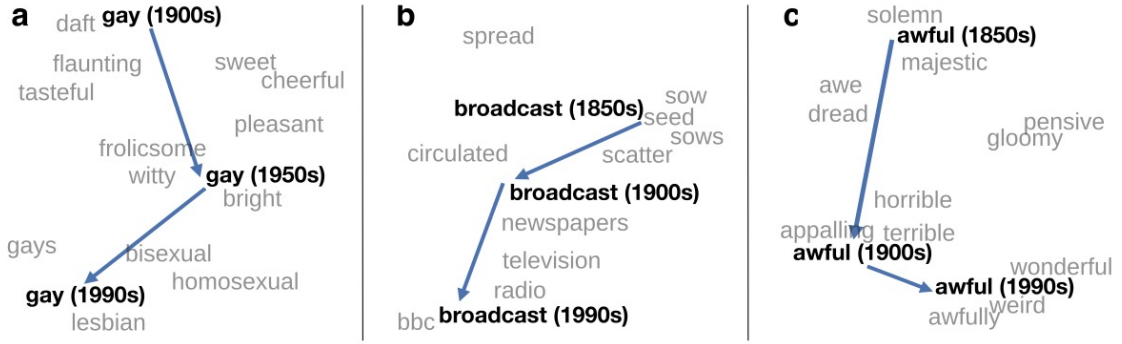



Source: Ferrone et al., 2017

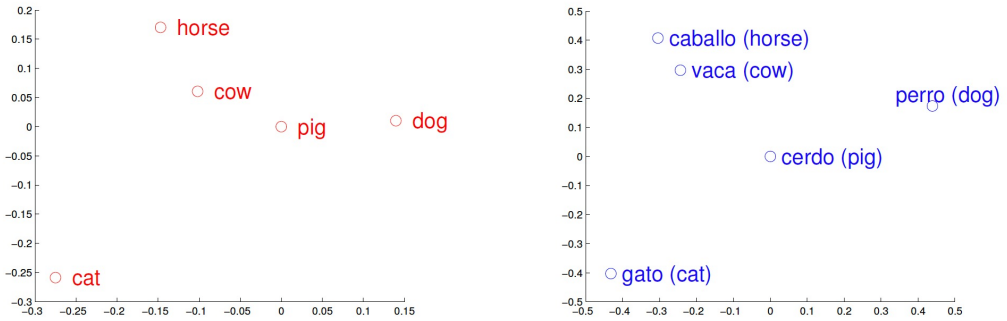
# Word Embeddings



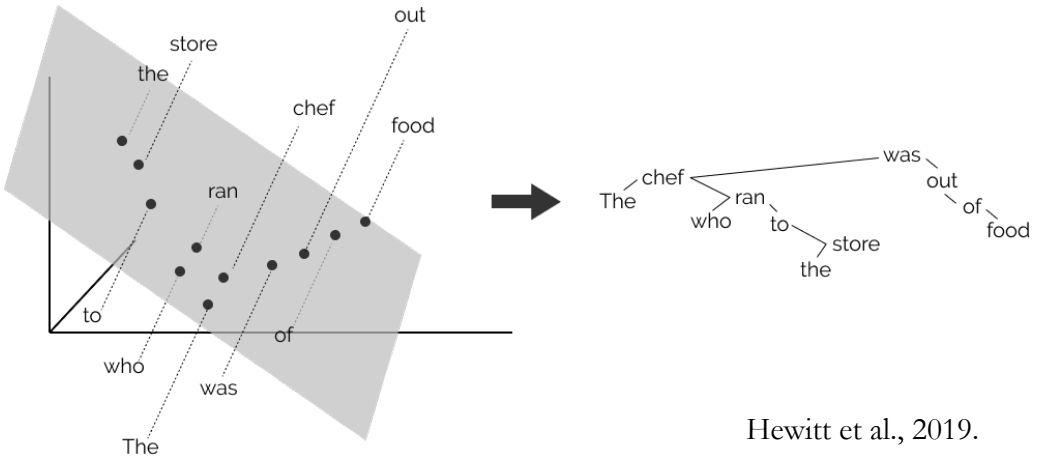

https://projector.tensorflow.org




Hamilton et al., 2016.




Mikolov et al., 2013.

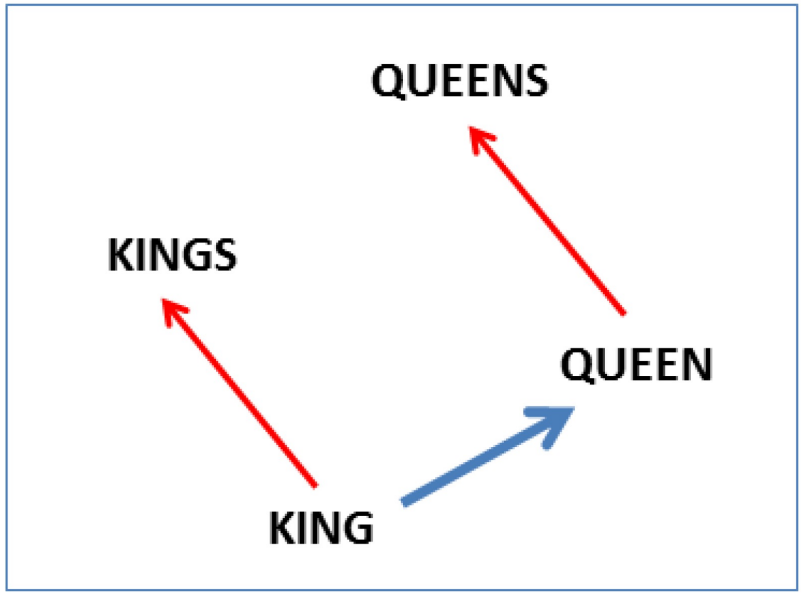



Hewitt et al., 2019.

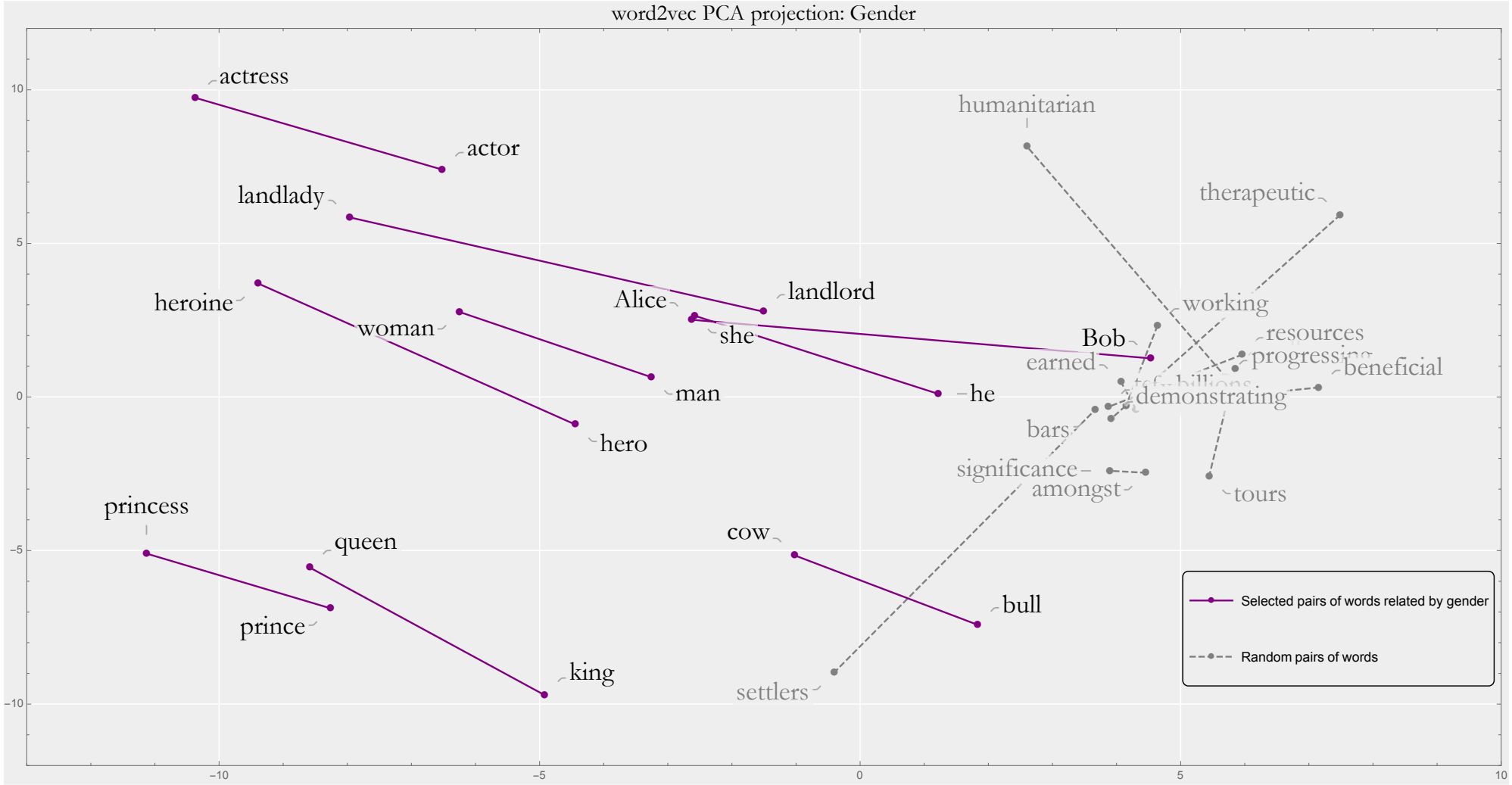
$$v_{house} - v_{city} + v_{countryside} \approx v_{farmhouse}$$

$$v_{king} - v_{man} + v_{woman} \approx v_{queen}$$

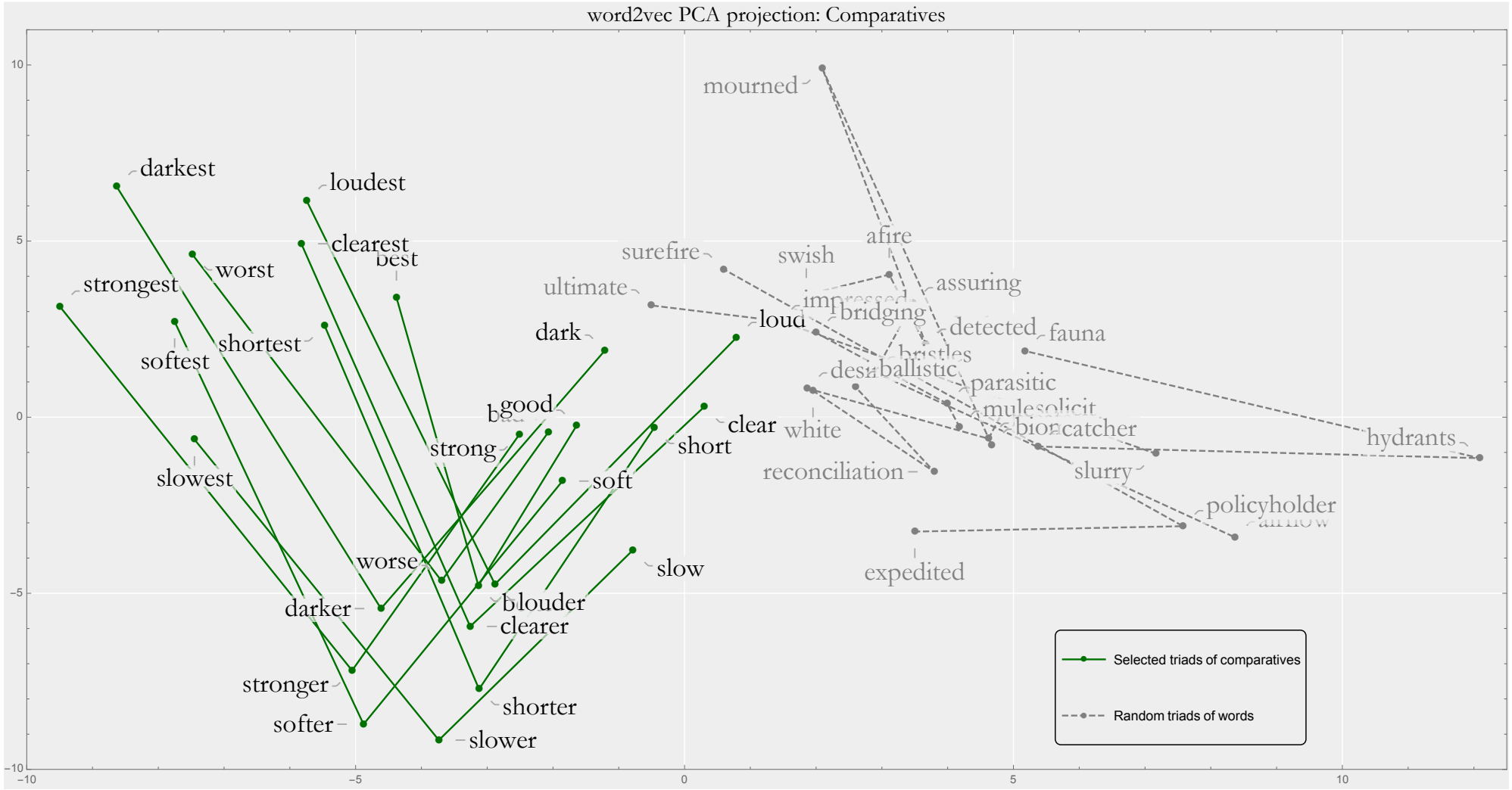
$$v_{king} - v_{queen} \approx v_{man} - v_{woman}$$



# Word Embeddings



# Word Embeddings



| Type of relationship | Word Pair 1 | | Word Pair 2 | |
|---|---|---|---|---|
| Common capital city | Athens | Greece | Oslo | Norway |
| All capital cities | Astana | Kazakhstan | Harare | Zimbabwe |
| Currency | Angola | kwanza | Iran | rial |
| City-in-state | Chicago | Illinois | Stockton | California |
| Man-Woman | brother | sister | grandson | granddaughter |
| Adjective to adverb | apparent | apparently | rapid | rapidly |
| Opposite | possibly | impossibly | ethical | unethical |
| Comparative | great | greater | tough | tougher |
| Superlative | easy | easiest | lucky | luckiest |
| Present Participle | think | thinking | read | reading |
| Nationality adjective | Switzerland | Swiss | Cambodia | Cambodian |
| Past tense | walking | walked | swimming | swam |
| Plural nouns | mouse | mice | dollar | dollars |
| Plural verbs | work | works | speak | speaks |

Source: Mikolov et al., 2013.

# Theoretical Consequences of Word Embeddings

1. The automatic reconstruction of the underlying organization of language does not require more human intervention than the one implied in the most ordinary use of language as recorded in a practically raw linguistic corpus.

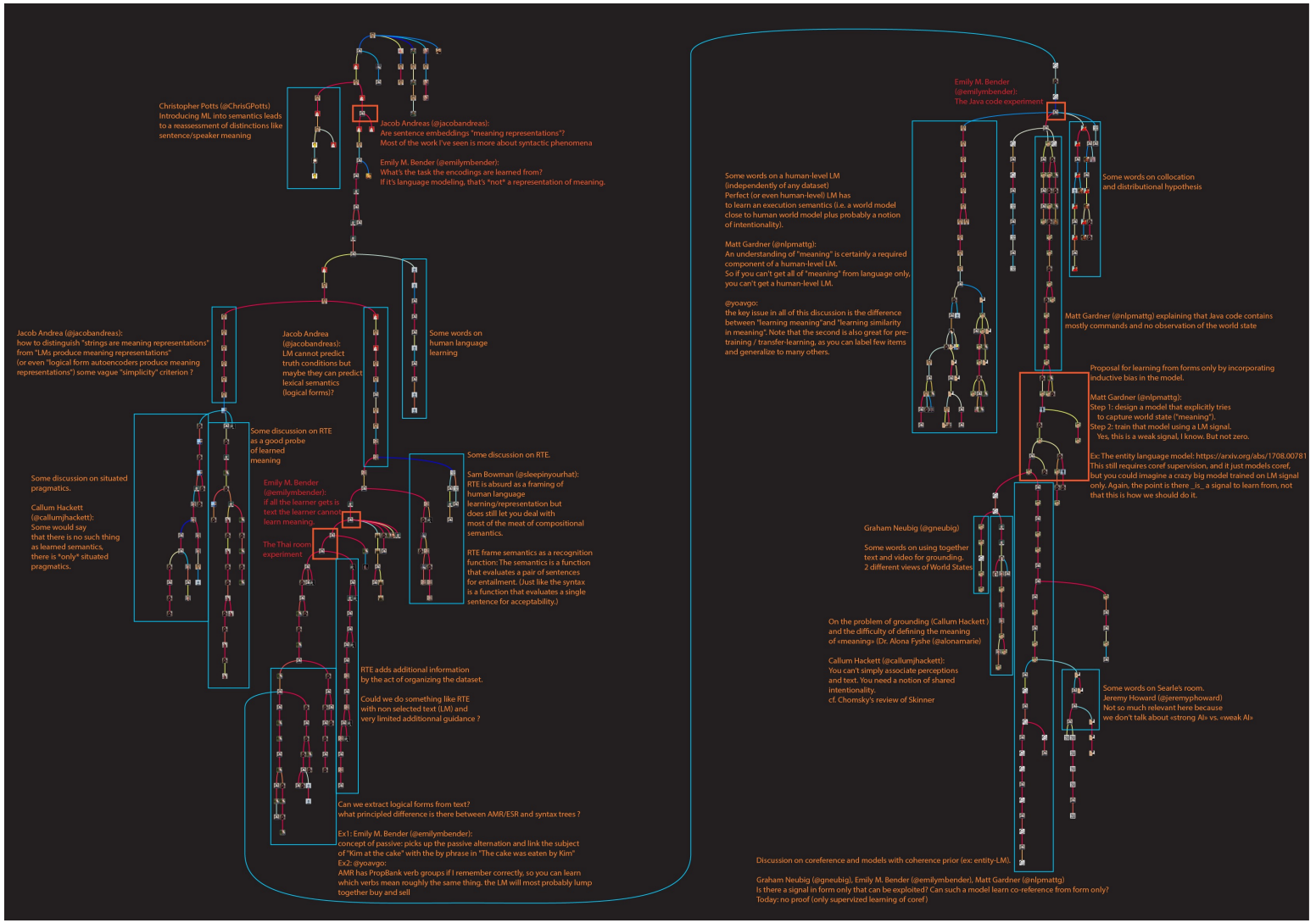
2. In that reconstruction, both semantic and syntactic contents of words are determined at once and as the result of the same procedure.

3. Word vector representations are not simply disposed in similarity neighbourhoods, but that the vector space itself is also structured following precise directions at the crossroads of which syntactic and semantic contents are established.

# Recent Evolution of Neural NLP Models

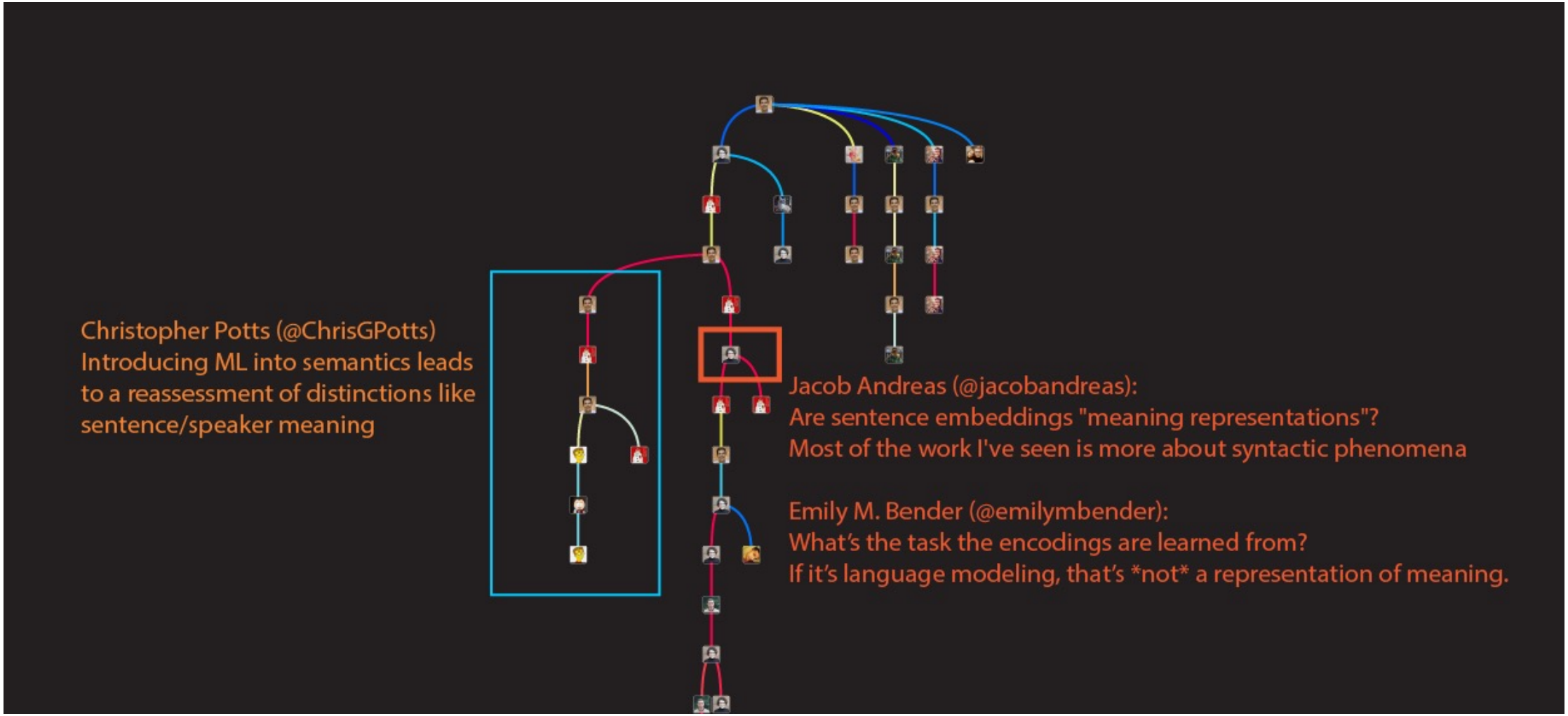
- Word Embeddings (word2vec, GloVe)
- Recurrent Neural Nets (LSTM)
- Encoding-Decoding (seq2seq)
- Transformers (attention, GPT-3)

# The Semantics Mega-Thread

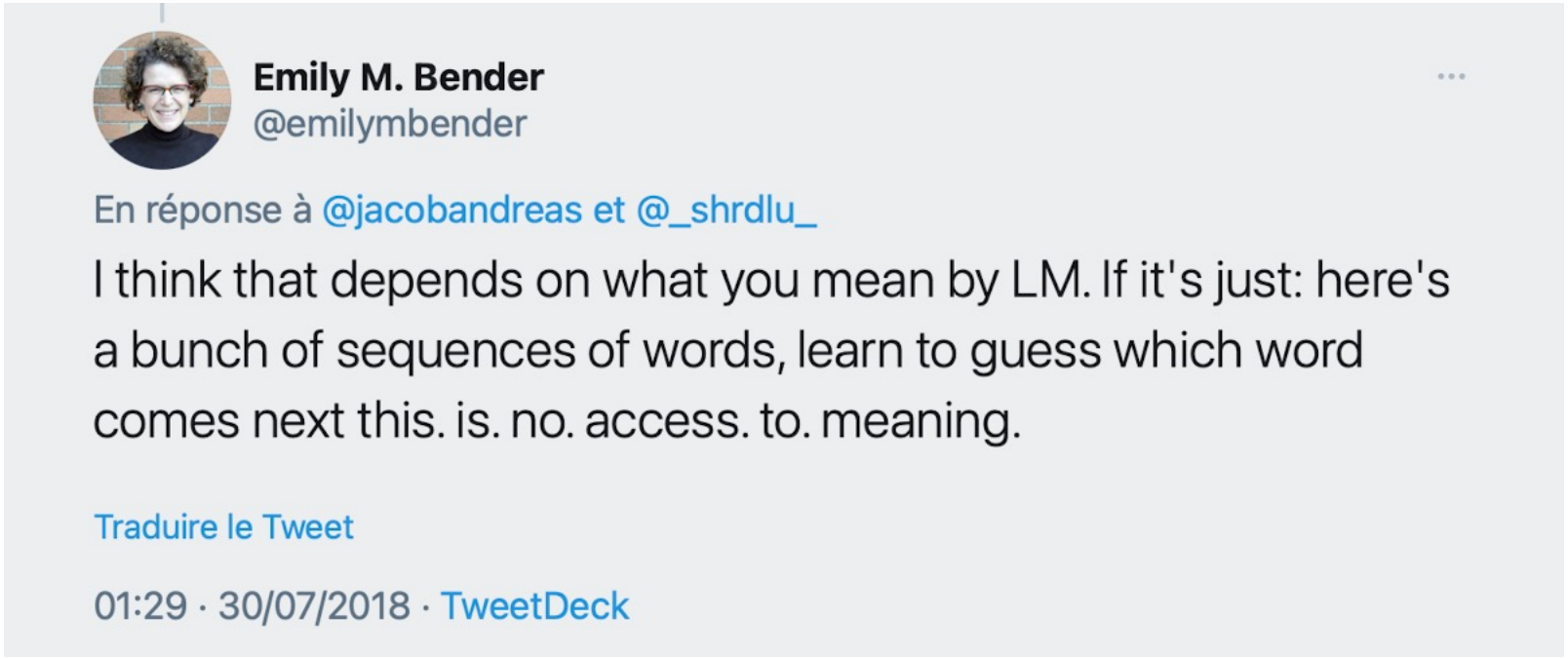



https://medium.com/huggingface/learning-meaning-in-natural-language-processing-the-semantics-mega-thread-9c0332dfe28e

# The Semantics Mega-Thread



# The Semantics Mega-Thread




Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. https://doi.org/10.18653/v1/2020.acl-main.463

# The Distributional Hypothesis

- "You shall know a word by the company it keeps!" (Firth, 1957)
- "Words which are similar in meaning occur in similar contexts" (Rubenstein & Goodenough 1965)
- "Words with similar meanings will occur with similar neighbors if enough text material is available" (Schütze & Pedersen 1995)
- "A representation that captures much of how words are used in natural context will capture much of what we mean by meaning" (Landauer & Dumais 1997)
- "Words that occur in the same contexts tend to have similar meanings" (Pantel 2005)
- "The degree of semantic similarity between two linguistic expressions A and B is a function of the similarity of the linguistic contexts in which A and B can appear" (Lenci, 2010)

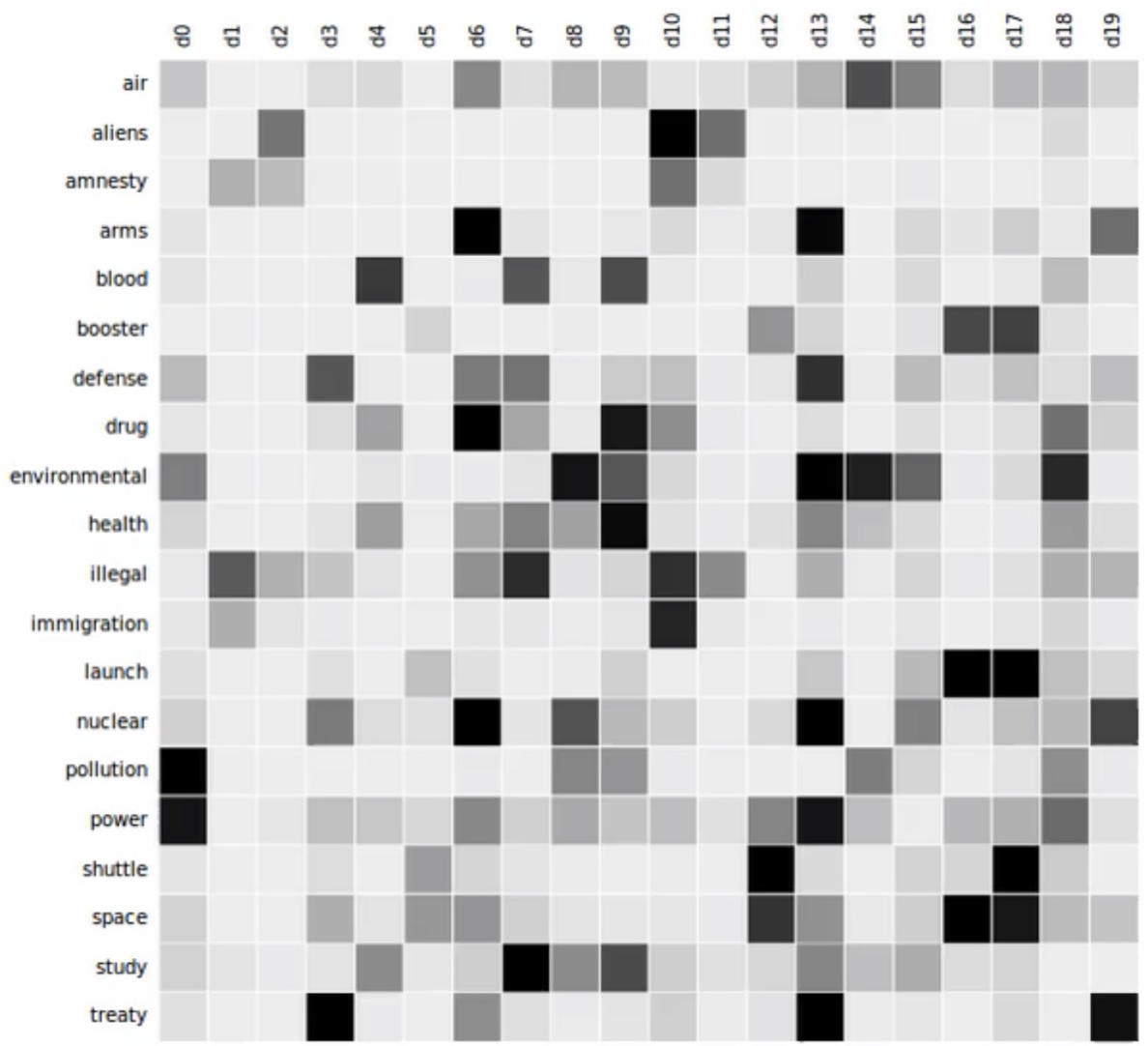
# The Distributional Hypothesis

- Theory of (linguistic) meaning as “usage” (Wittgenstein)
  - “the meaning of a word is defined by the circumstances of its use” (Manning & Schütze, 1999)

- Two versions of the DH:
  - Weak: Correlation between context and word meaning (Spence & Owens, 1990)
  - Strong: Causality attributed to contextual distributions (Miller & Charles, 1991)

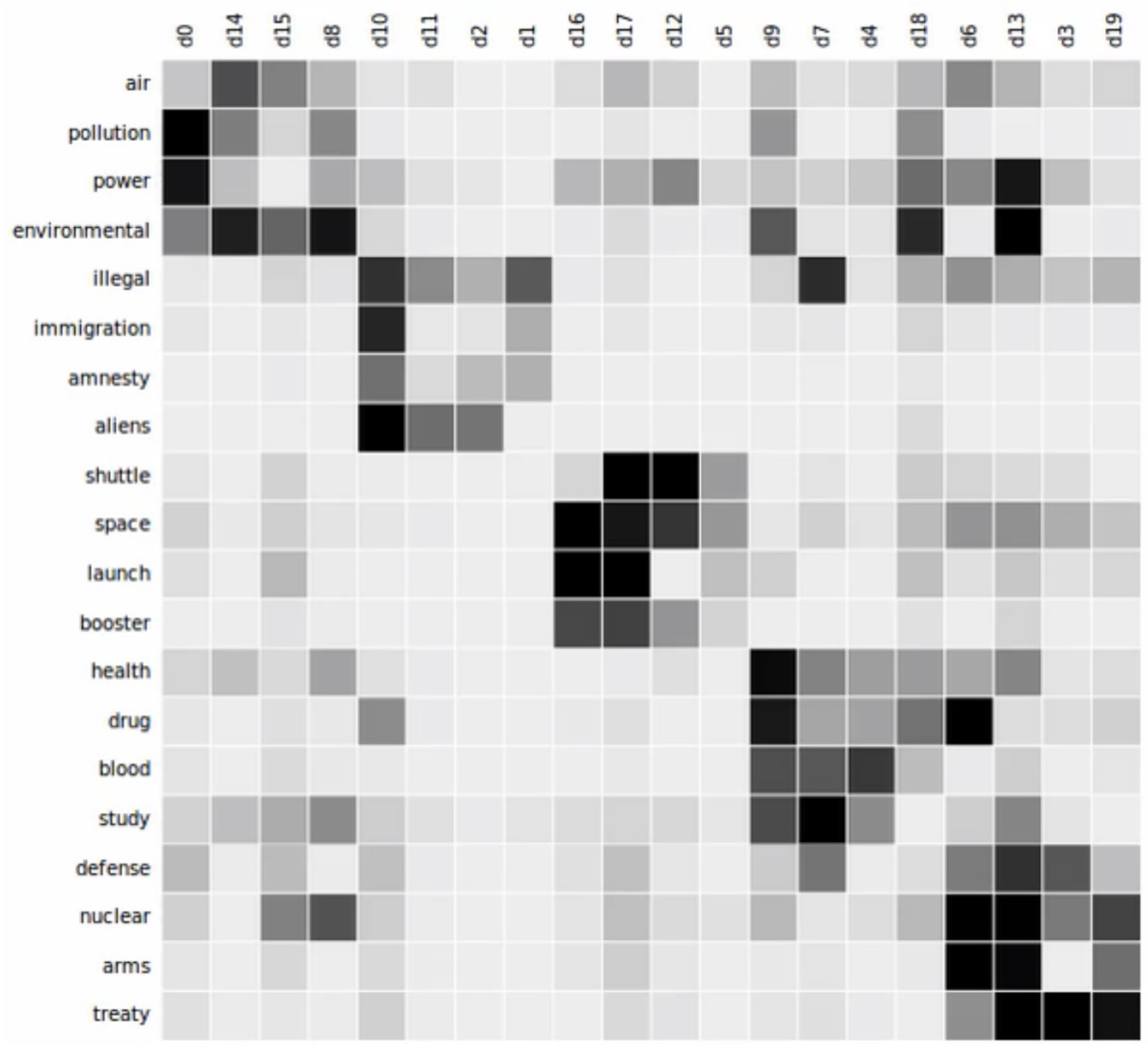
- Context: the domain or scope within which entities of the same nature can be presented together (“co-occur”), in such a way that they can be associated by a cognitive agent.

# Word Embeddings as Matrix Factorization




Source: topicmodels.west.uni-koblenz.de

# Word Embeddings as Matrix Factorization




topicmodels.west.uni-koblenz.de

# Word Embeddings as Matrix Factorization

| | ... | w | x | y | z | ... |
|---|---|---|---|---|---|---|
| ... | ... | 0 | 0 | 0 | 0 | ... |
| a | ... | 0 | 1 | 1 | 0 | ... |
| b | ... | 0 | 0 | 1 | 1 | ... |
| c | ... | 1 | 0 | 0 | 1 | ... |
| ... | ... | 0 | 0 | 0 | 0 | ... |

| Context | Term |
|---|---|
| (w,x) | (a,c) |
| (w,y) | (a,b,c) |
| (w,z) | (b,**c**) |
| (x,y) | (**a**,b) |
| (y,z) | (a,**b**,c) |

| Term | Context |
|---|---|
| (a,c) | (w,x,y,z) |
| (a,b,c) | (w,x,**y**,**z**) |
| (b,c) | (w,y,**z**) |
| (a,b) | (x,**y**,z) |
| (a,b,c) | (w,x,**y**,**z**) |

# Word Embeddings as Matrix Factorization

| | |
|---|---|
| a = your<br>c = my | w = apartment<br>x = house<br>y = chair<br>z = stool |

| | ... | w | x | y | z | ... |
|---|---|---|---|---|---|---|
| ... | ... | 0 | 0 | 0 | 0 | ... |
| a | ... | 0 | 1 | 1 | 0 | ... |
| b | ... | 0 | 0 | 1 | 1 | ... |
| c | ... | 1 | 0 | 0 | 1 | ... |
| ... | ... | 0 | 0 | 0 | 0 | ... |

your : house

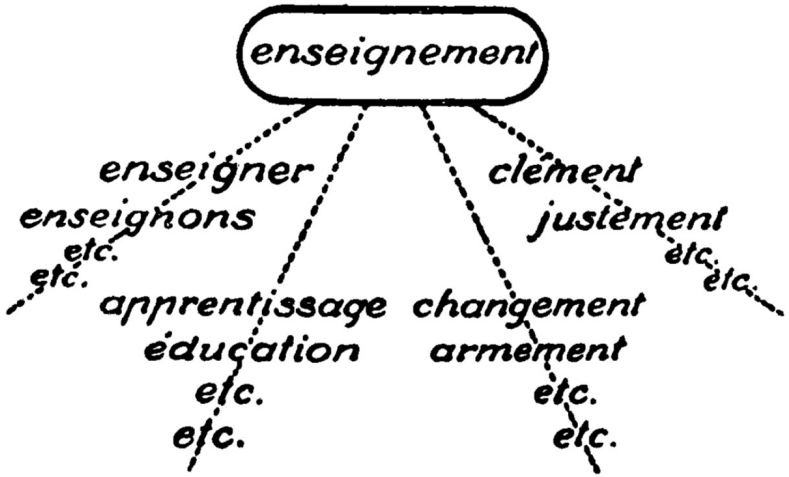
my : apartment

"The day she came to your house in Paris"

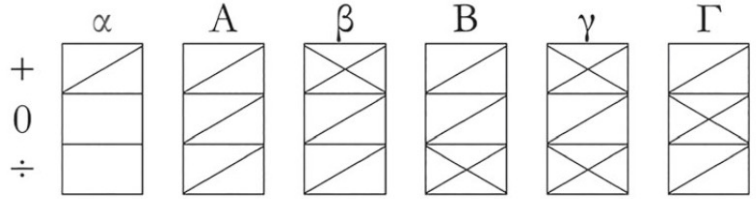
"She bought a house and a bungalow"

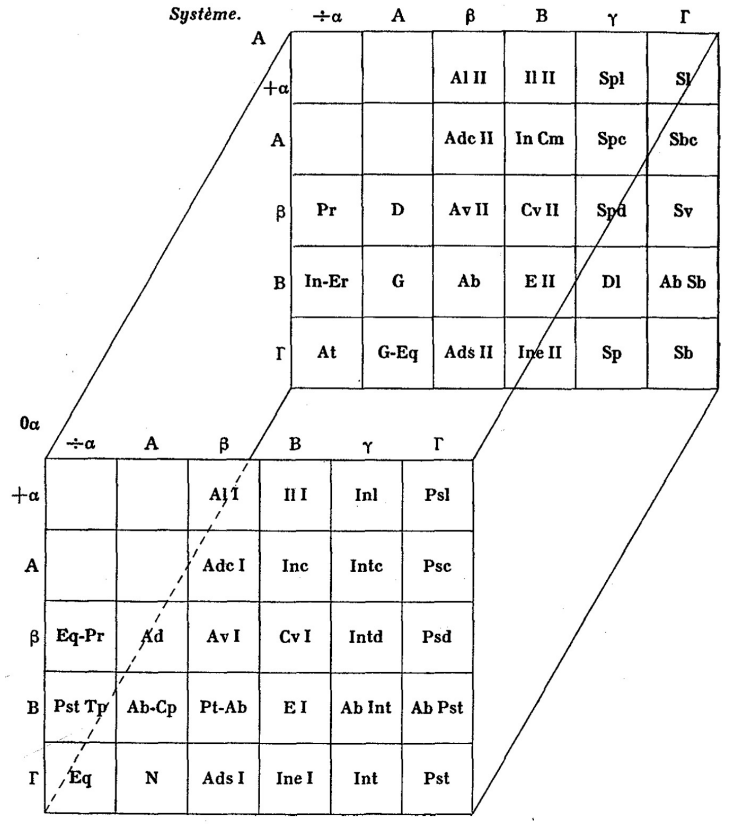
"Vous vous trompez"

- Saussure, Hjelmslev, Harris…




(Saussure, 1916)




(Hjelmslev, 1935)

| Segments | Environments | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | #— r | #— r | #— l | e i —C | æ—C | a o u —C | s— e i | s—æ | s— a o u | … | t — | C³— |
| t | ✓ | | | | | | | | | | | |
| t | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| K | | | | | | ✓ | | | ✓ | | | |
| k | | ✓ | ✓ | | ✓ | | | ✓ | | | | |
| ĸ | | | | ✓ | | | ✓ | | | | | |
| G | | | | | | ✓ | | | | | | |
| g | | ✓ | ✓ | | ✓ | | | | | | | |
| ɢ | | | | ✓ | | | | | | | | |
| r | | | | ✓ | ✓ | ✓ | | | | | | ✓ |
| ɾ | | | | | | | | | | | ✓ | |
| | | | | | | | | | | | | |

(Harris, 1951)

| SEGMENTS | ENVIRONMENTS | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | # — r̥ | # — r | # — l | e i — C | æ — C | a o u — C | s — e i | s — æ | s — a o u | … | t̥ — | C³ — |
| t̥ | √ | | | | | | | | | | | |
| t | | √ | | √ | √ | √ | √ | √ | √ | | | |
| K | | | | | | √ | | | √ | | | |
| k | | √ | √ | | √ | | | √ | | | | |
| ĸ | | | | √ | | | √ | | | | | |
| G | | | | | | √ | | | | | | |
| g | | √ | √ | | √ | | | | | | | |
| ɢ | | | | √ | | | | | | | | |
| r | | | | √ | √ | √ | | | | | | √ |
| r̥ | | | | | | | | | | | √ | |
| | | | | | | | | | | | | |

(Harris, 1951)

# Structuralist Hypothesis

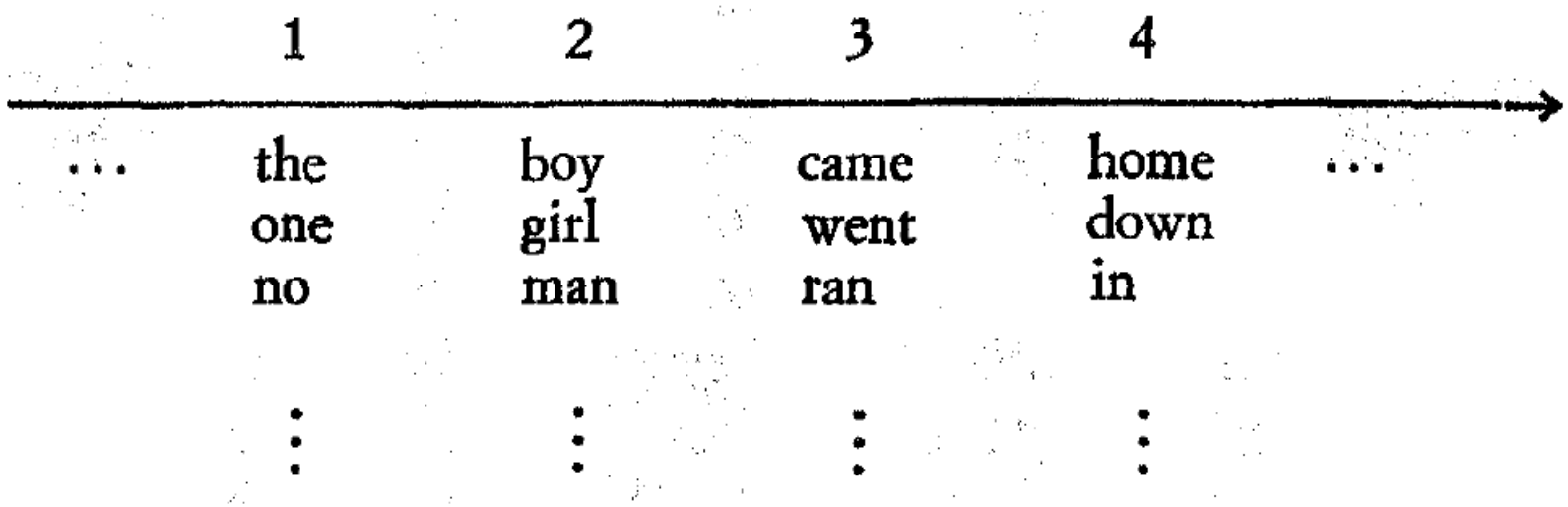
- F. de Saussure
- R. Jakobson
- L. Hjelmslev
- L. Bloomfield
- Z. Harris

- Linguistic content (including essential aspects of linguistic meaning) is the effect of a virtual structure of classes and dependencies at multiple levels, underlying (and derivable from) the mass of things said or written in a given language.
- The task of linguistic analysis is not just that of identifying loose similarities between words out of distributional properties of a corpus, but rather this other one of explicitly drawing from that corpus the system of strict dependencies between implicit linguistic categories.

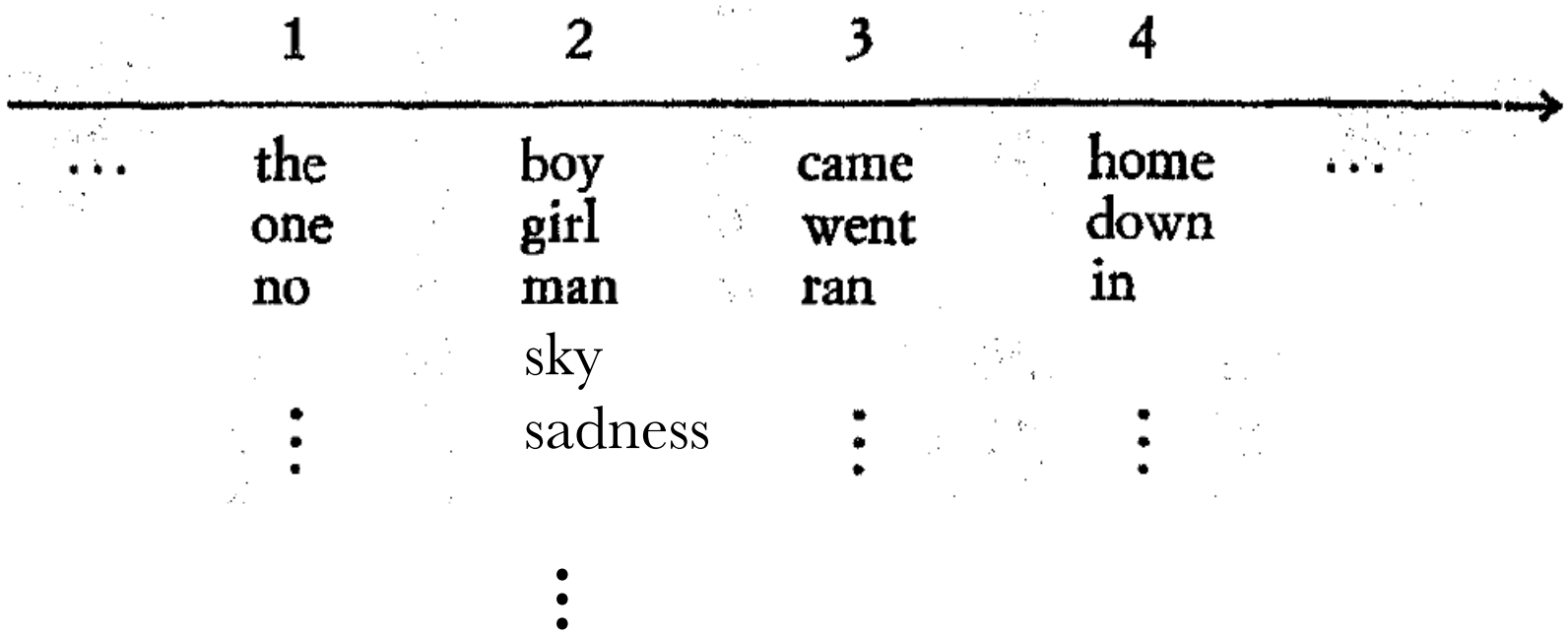
*A priori* it would seem to be a generally valid thesis that for every *process* there is a corresponding *system*, by which the process can be analyzed and described by means of a limited number of premises. It must be assumed that any process, can be analyzed into a limited number of elements recurring in various combinations. Then, on the basis of this analysis, it should be possible to order these elements into classes according to their possibilities of combination. And it should be further possible to set up a general and exhaustive calculus of the possible combinations.

Hjelmslev (1943)

# Paradigm Derivation

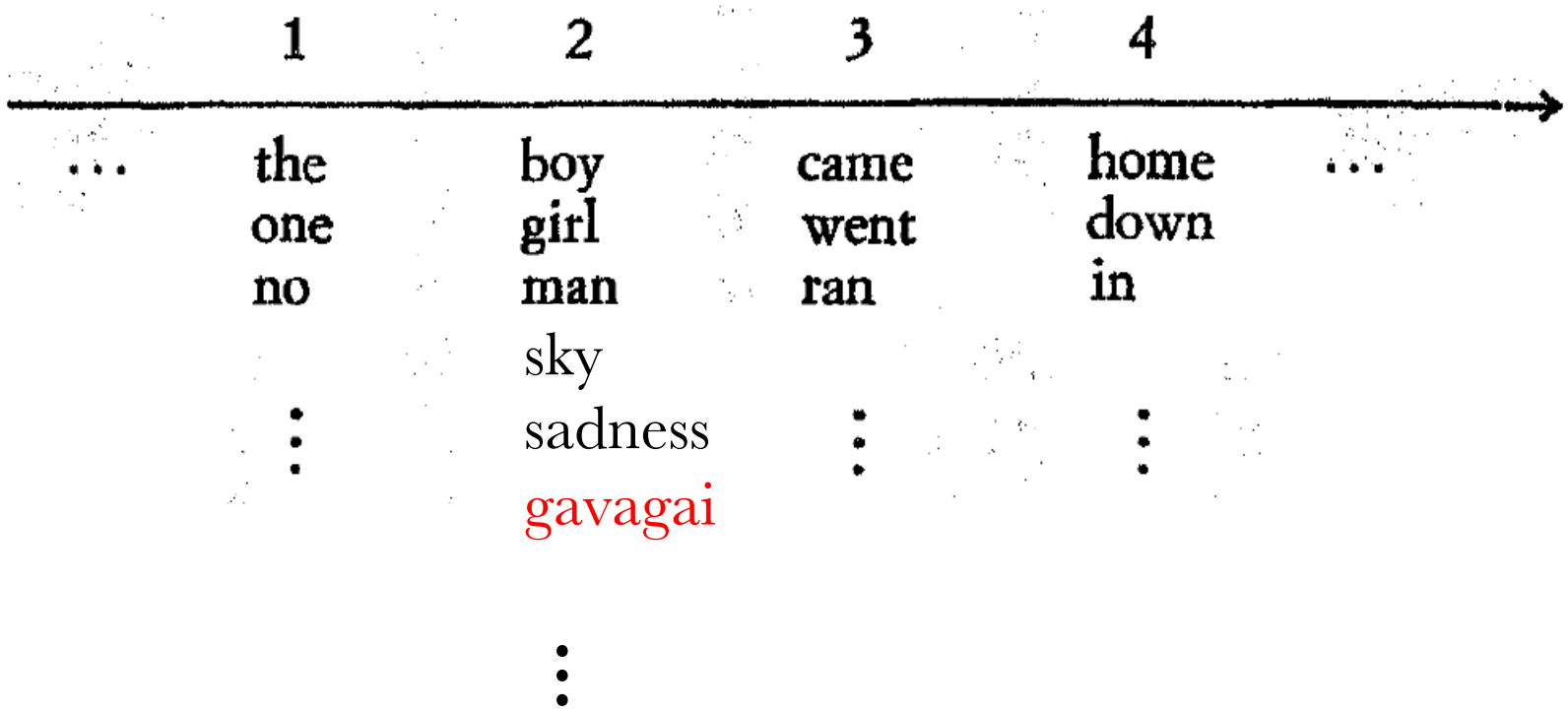


# Paradigm Derivation

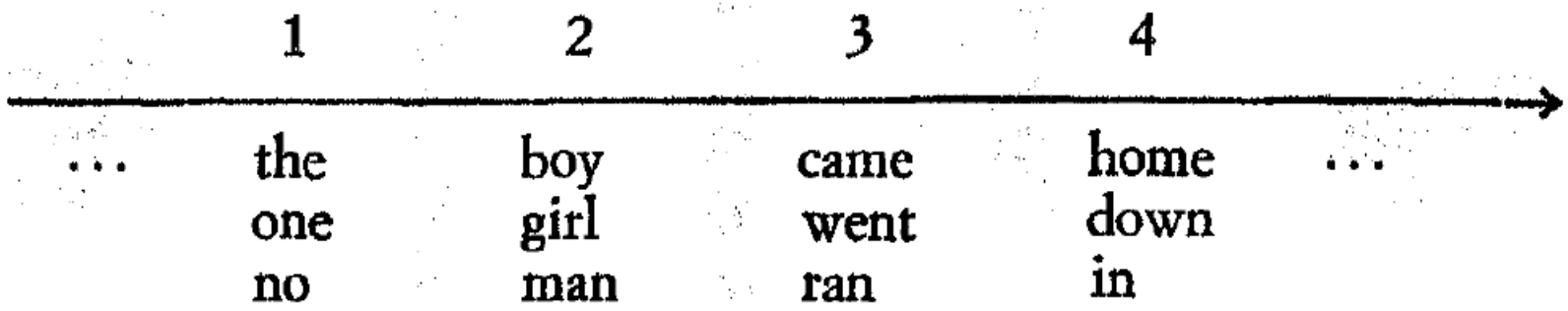


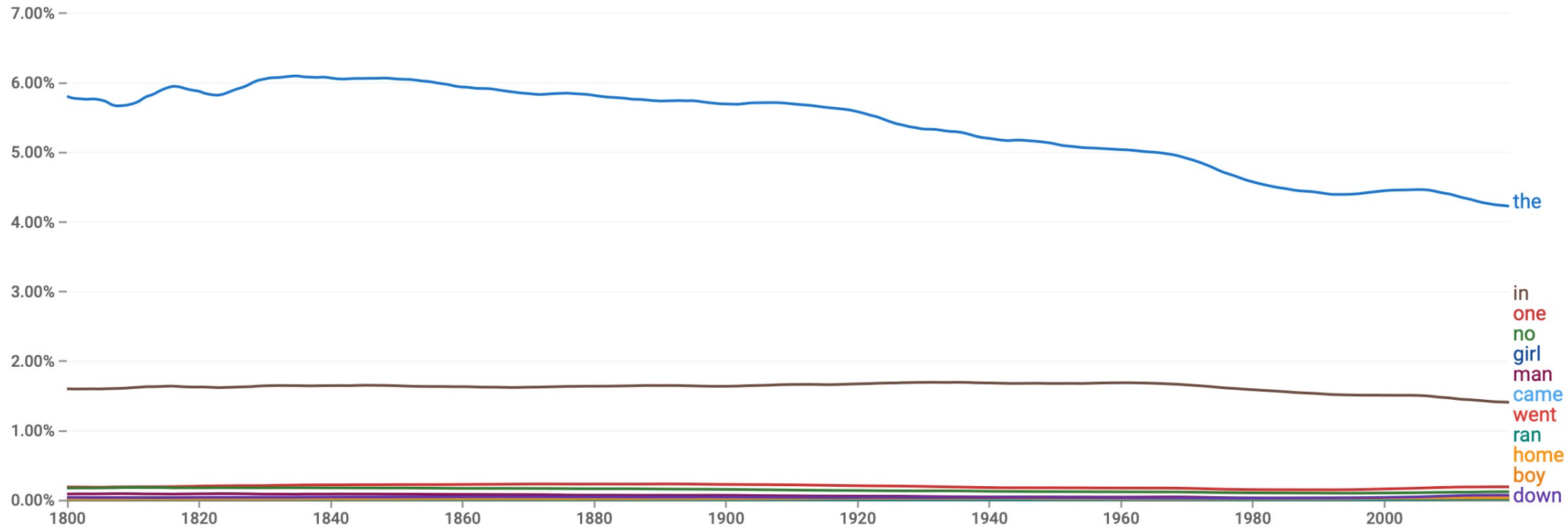

- Colorless green ideas sleep furiously ✓
- Furiously sleep ideas green colorless ✗

# Paradigm Derivation

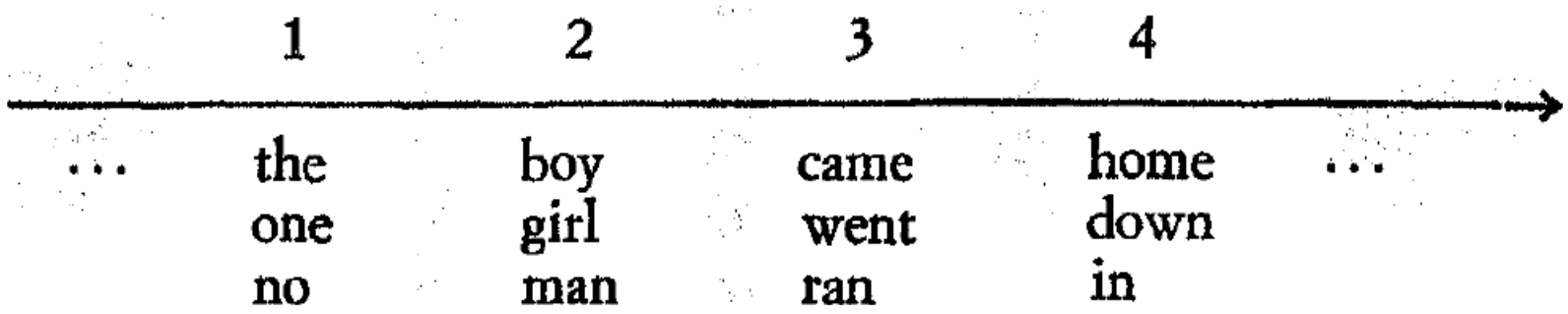


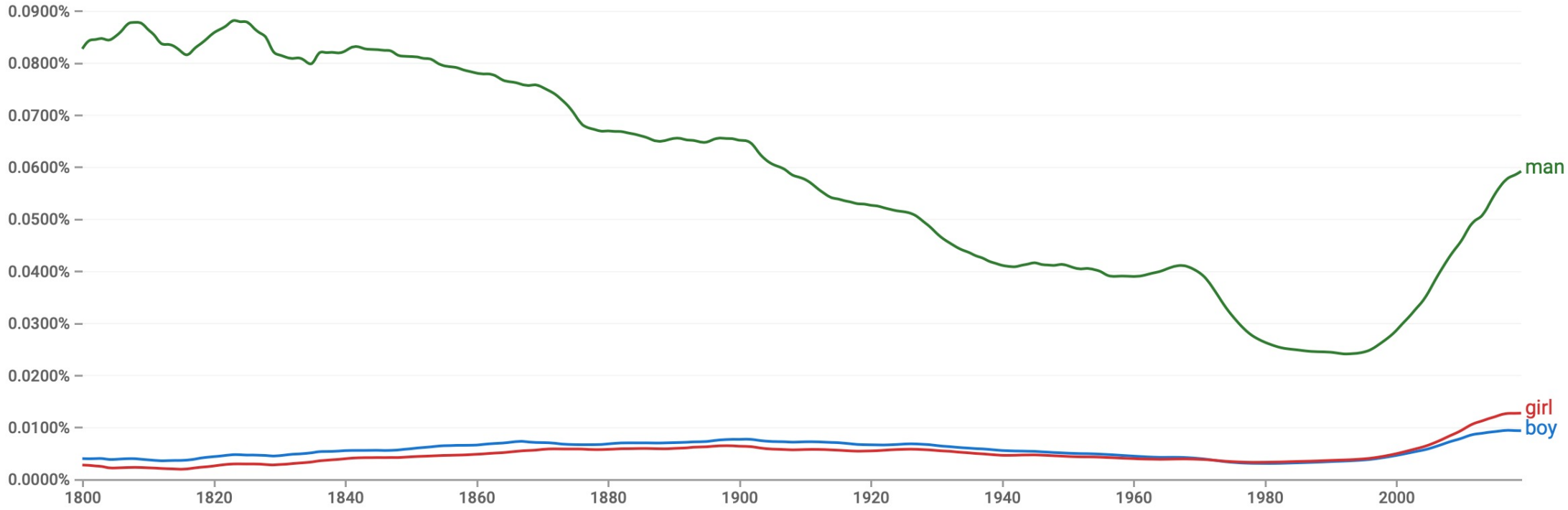
# Paradigm Derivation



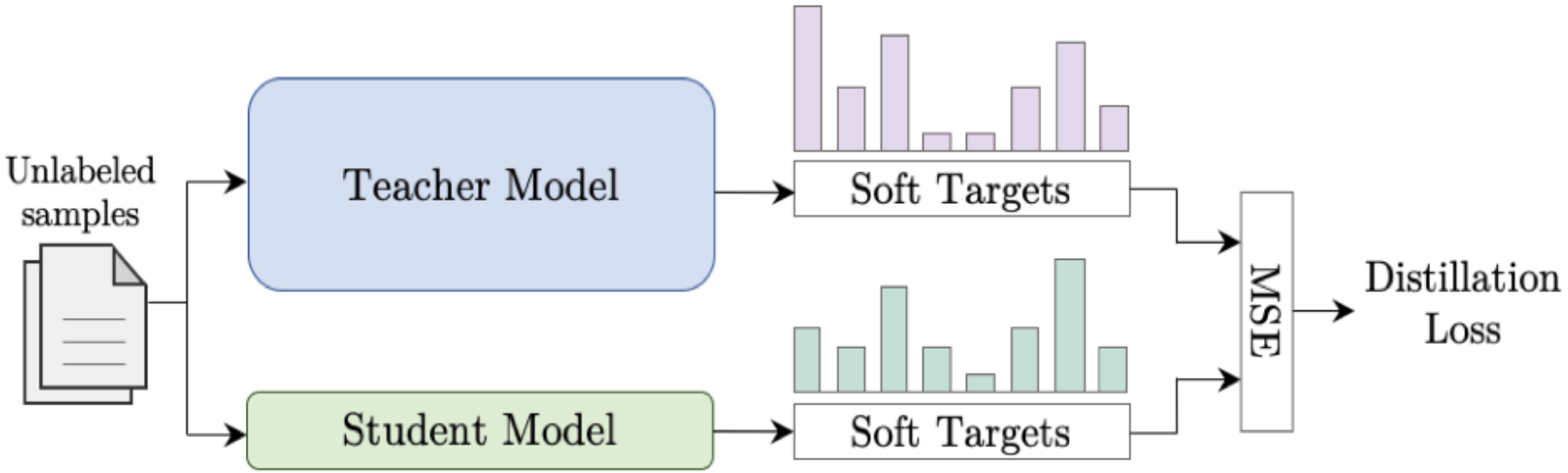



# Paradigm Derivation



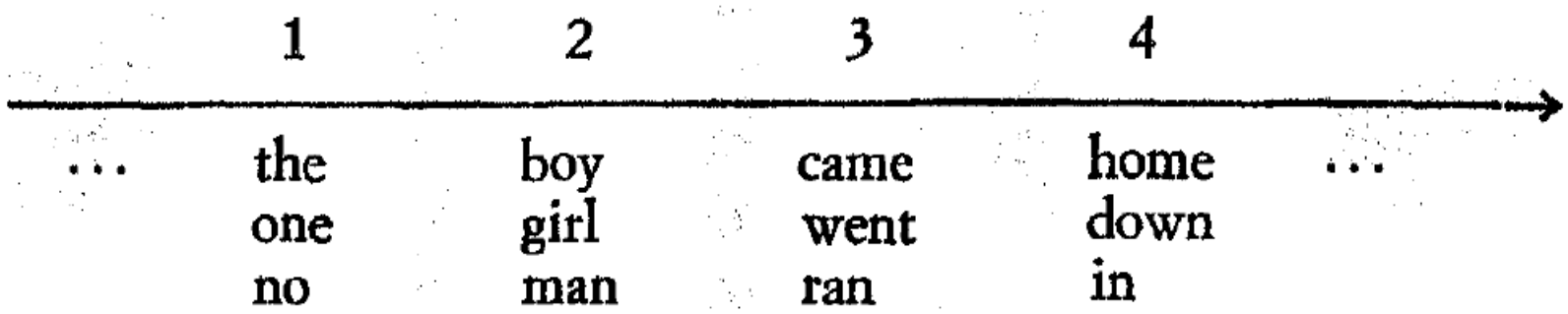



# Distillation

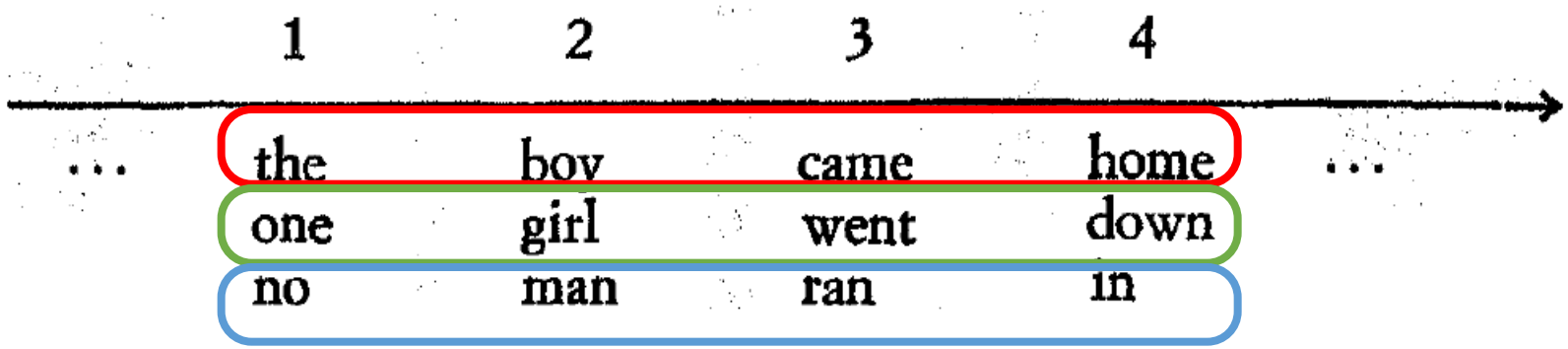



Source: Wasserblat, M., Pereg, O., & Izsak, P. (2020)

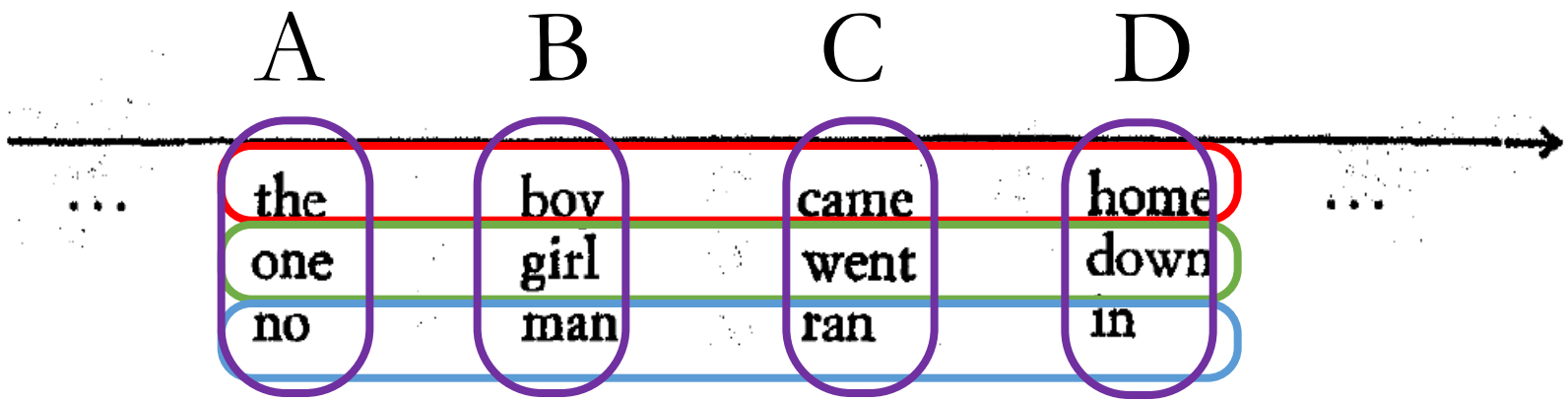
# Paradigm Derivation



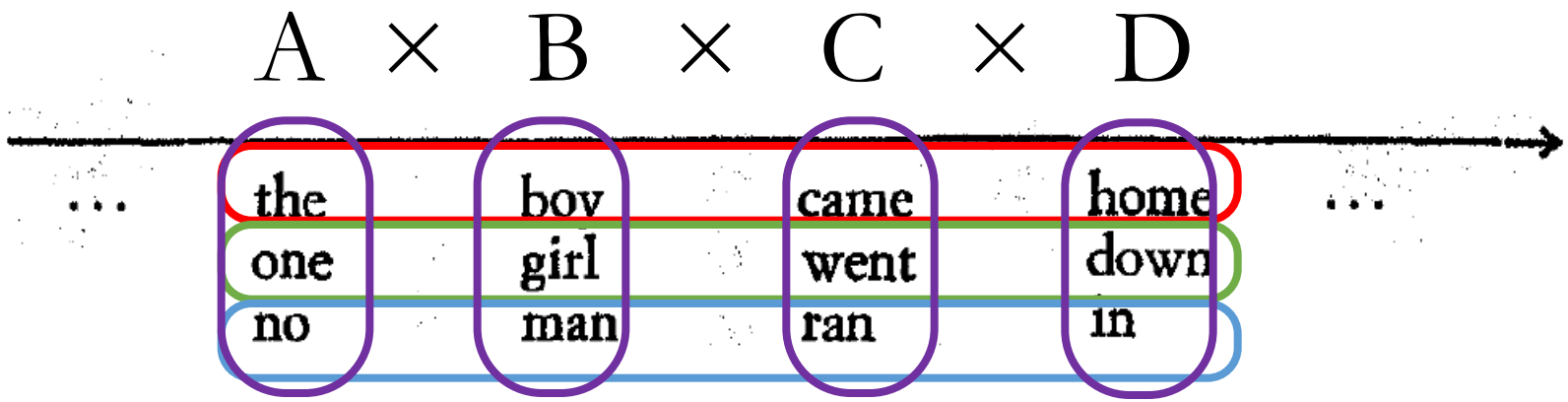
# Paradigm Derivation



# Paradigm Derivation



# Paradigm Derivation




the girl ran home

# Non-Neural Distributional Model

| Procedure Step | Neural | Non-Neural |
|---|---|---|
| Segmentation | Sub-word Tokenization<br>Character level DNNs | BPE |
| Classification | Word Embedding<br>Distillation | Matrix models |
| Dependencies | Attention | Biorthogonal Typing<br>Linear Logic |

# Some Conclusions

- Generic model
- Non-necessarily neural
- Producing explicit representations
- Supporting logical relations
- Capturing significant aspects of meaning

# Reference Papers for this Talk

J. L. Gastaldi. **“Why Can Computers Understand Natural Language?”**.
In: *Philosophy & Technology* 34.1 (2021), pp. 149–214. DOI: 10.1007/s13347-020-00393-9
Link

J. L. Gastaldi and Luc Pellissier: **“The Calculus of Language: Explicit Representation of Emergent Linguistic Structure Through Type-Theoretical Paradigms”**
In: *Interdisciplinary Science Reviews* (2021), DOI: 10.1080/03080188.2021.1890484

# References


- Baroni, Marco, Georgiana Dinu, and Germán Kruszewski (2014). “Don’t count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors”. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Baltimore, Maryland: Association for Computational Linguistics, pp. 238–247.
- Bender, Emily M. and Alexander Koller (July 2020). “Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data”. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online: Association for Computational Linguistics, pp. 5185–5198.
- Bengio, Yoshua et al. (Mar. 2003). “A Neural Probabilistic Language Model”. In: J. Mach. Learn. Res. 3, pp. 1137–1155. ISSN: 1532-4435.
- Gastaldi, Juan Luis (May 2020). “Why Can Computers Understand Natural Language?: The Structuralist Image of Language Behind Word Embeddings”. In: Philosophy & Technology.
- Gastaldi, Juan Luis and Luc Pellissier (2021). “The Calculus of Language: Explicit Representation of Emergent Linguistic Structure Through Type-Theoretical Paradigms”. In: Interdisciplinary Science Reviews.

# References


- Hamilton, William L., Jure Leskovec, and Dan Jurafsky (2016). “Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change”. In: CoRR abs/1605.09096.
- Harris, Zellig (1960). Structural linguistics. Chicago: University of Chicago Press. ISBN: 0226317714 0226217714.
- Hewitt, John and Christopher D. Manning (June 2019). “A Structural Probe for Finding Syntax in Word Representations”. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis, Minnesota: Association for Computational Linguistics, pp. 4129–4138.
- Hjelmslev, Louis (1935). La catégorie des cas. Munchen: Wilhelm Fink Verlag.
- — (1953). Prolegomena to a Theory of Language. Baltimore: Wawerly Press.
- — (1975). Résumé of a Theory of Language. Travaux du Cercle linguistique de Copenhague 16. Copenhagen: Nordisk Sprog-og Kulturforlag. ISBN: 0-299-07040-9.
- Lenci, Alessandro (2008). “Distributional semantics in linguistic and cognitive research”. In: From context to meaning: distributional models of the lexicon in linguistics and cognitive science, Italian Journal of Linguistics 1.20. Ed. by Alessandro Lenci, pp. 1–31.

# References


- Levy, Omer and Yoav Goldberg (2014a). “Linguistic Regularities in Sparse and Explicit Word Representations”. In: Proceedings of the Eighteenth Conference on Computational Natural Language Learning, CoNLL 2014, Baltimore, Maryland, USA, June 26-27, 2014, pp. 171–180.
- — (2014b). “Neural Word Embedding As Implicit Matrix Factorization”. In: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. NIPS’14. Montreal, Canada: MIT Press, pp. 2177–2185.
- Manning, Christopher D. and Hinrich Schütze (1999). Foundations of Statistical Natural Language Processing. Cambridge, MA, USA: MIT Press. ISBN: 0-262-13360-1.
- Mikolov, Tomáš, Kai Chen, et al. (2013). “Efficient Estimation of Word Representations in Vector Space”. In: CoRR abs/1301.3781.
- Mikolov, Tomáš, Quoc V. Le, and Ilya Sutskever (2013). “Exploiting Similarities among Languages for Machine Translation”. In: CoRR abs/1309.4168.
- Mikolov, Tomáš, Ilya Sutskever, et al. (2013). “Distributed Representations of Words and Phrases and their Compositionality”. In: CoRR abs/1310.4546.

# References


- Mikolov, Tomáš, Wen-tau Yih, and Geoffrey Zweig (2013). "Linguistic Regularities in Continuous Space Word Representations". In: Proceedings of the 2013 Conference of the North American Chapter of the ACL: Human Language Technologies. Atlanta, Georgia: ACL, pp. 746–751.
- Sahlgren, Magnus (2008). "The Distributional Hypothesis". In: Special issue of the Italian Journal of Linguistics 1.20. Ed. by Alessandro Lenci, pp. 33–53.
- Saussure, Ferdinand de (1959). Course in General Linguistics. Translated by Wade Baskin. New York: McGraw-Hill.

GDR LIFT - ILFC Seminar

Interactions between formal and computational linguistics

1st June 2021

# What do neural models tell us about the nature of language?

*Juan Luis Gastaldi*

ETH Zürich (D-GESS, Turing Centre)

juan.luis.gastaldi@gess.ethz.ch

Turing Centre

This project has received funding from the
*European Union's Horizon 2020 research and innovation programme*
under grant agreement No 839730

ETH zürich