



Everyone is an NLP expert now

Natural Language Processing

Fall 2023

Programming Languages

C, C++, Java, Python, ...

- unambiguous
- fixed (relatively)
- designed
- learnable?
- known semantics

Natural Languages

French, English, Korean, Chinese, Tagalog, ...

- ambiguous
- evolving
- transmitted
- learnable
- unknown semantics

Why is NLP computationally hard?

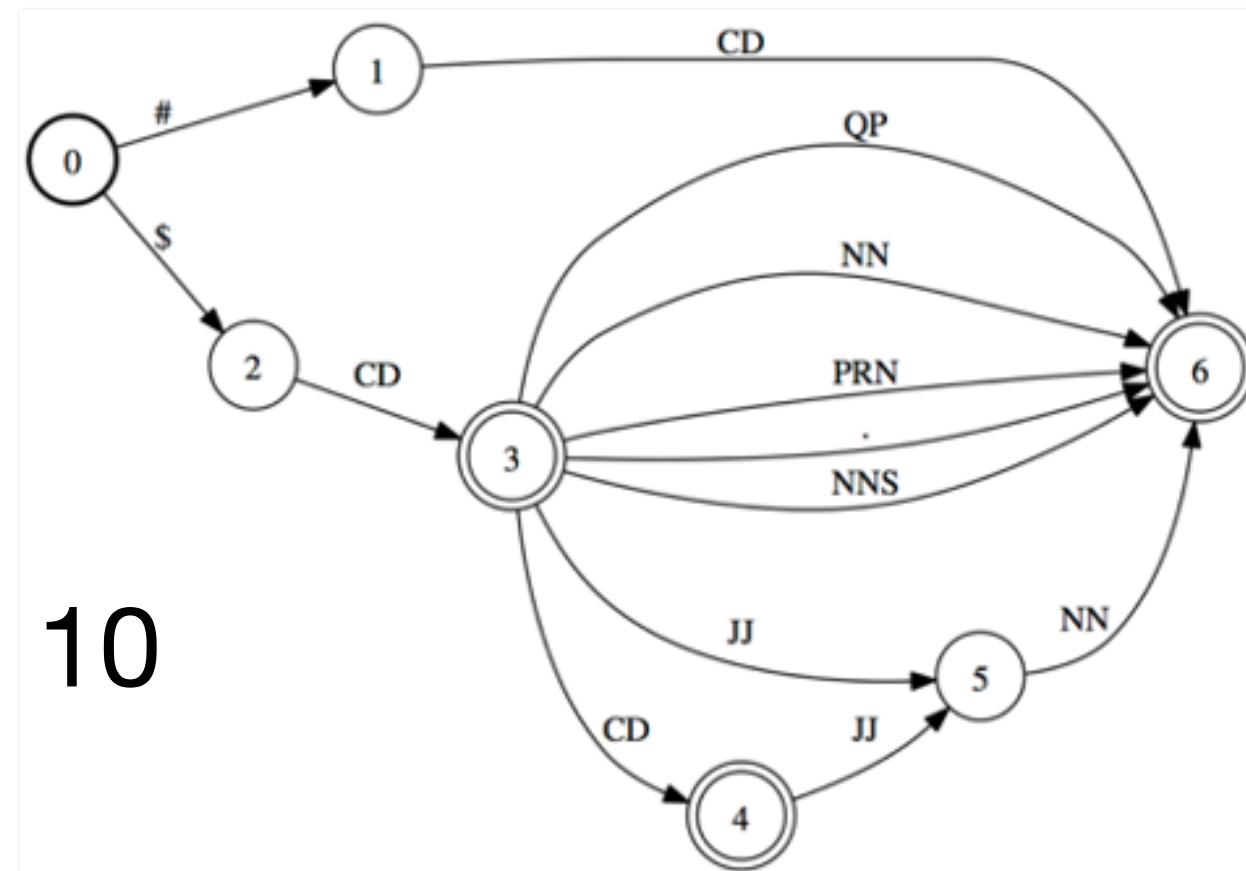
Language is Complex

\$ 3 monthly fee

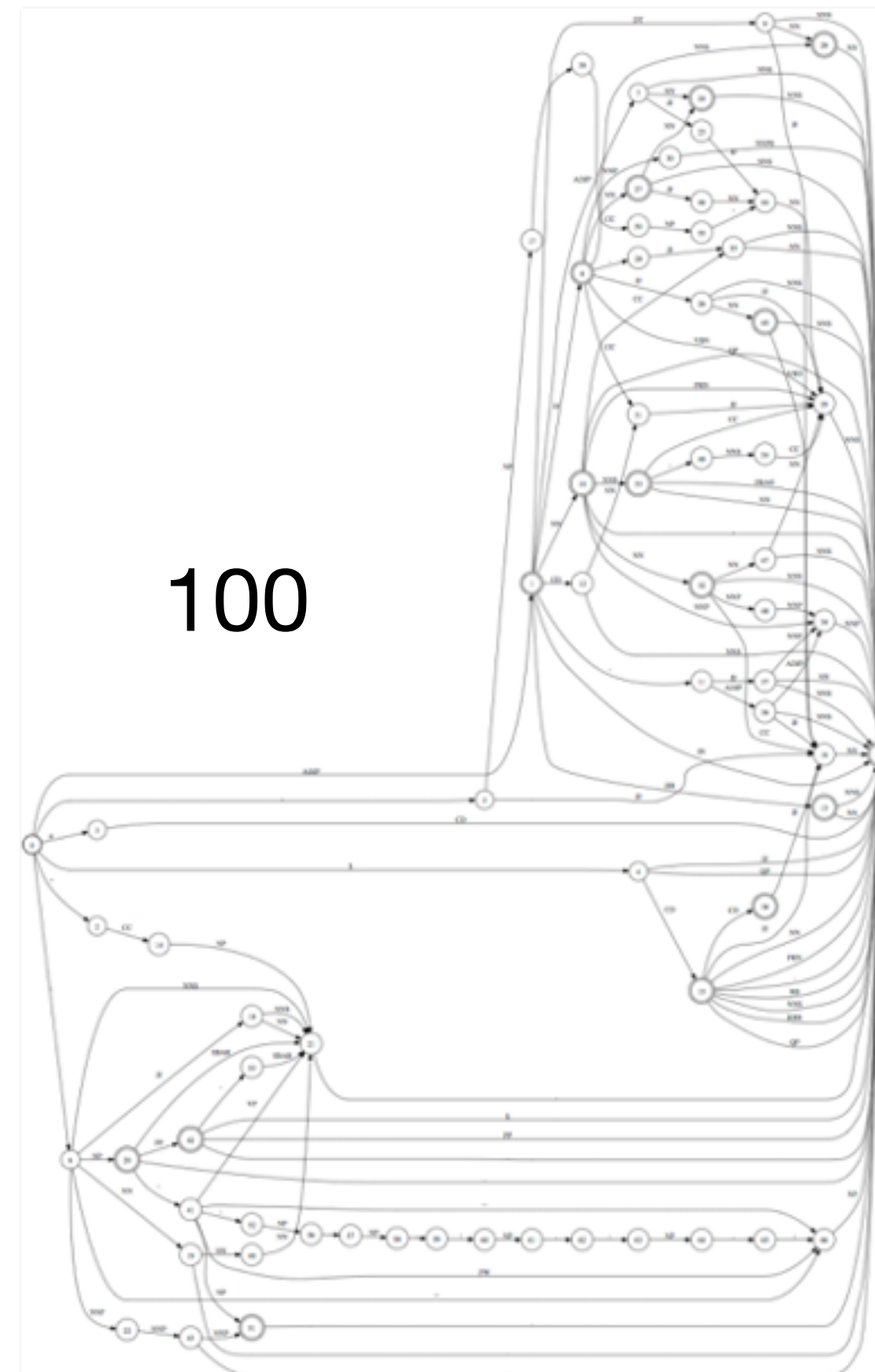
\$ 1 per-share pricing

✓ \$ number adjective noun

✗ \$ noun number adjective



10



100

1000



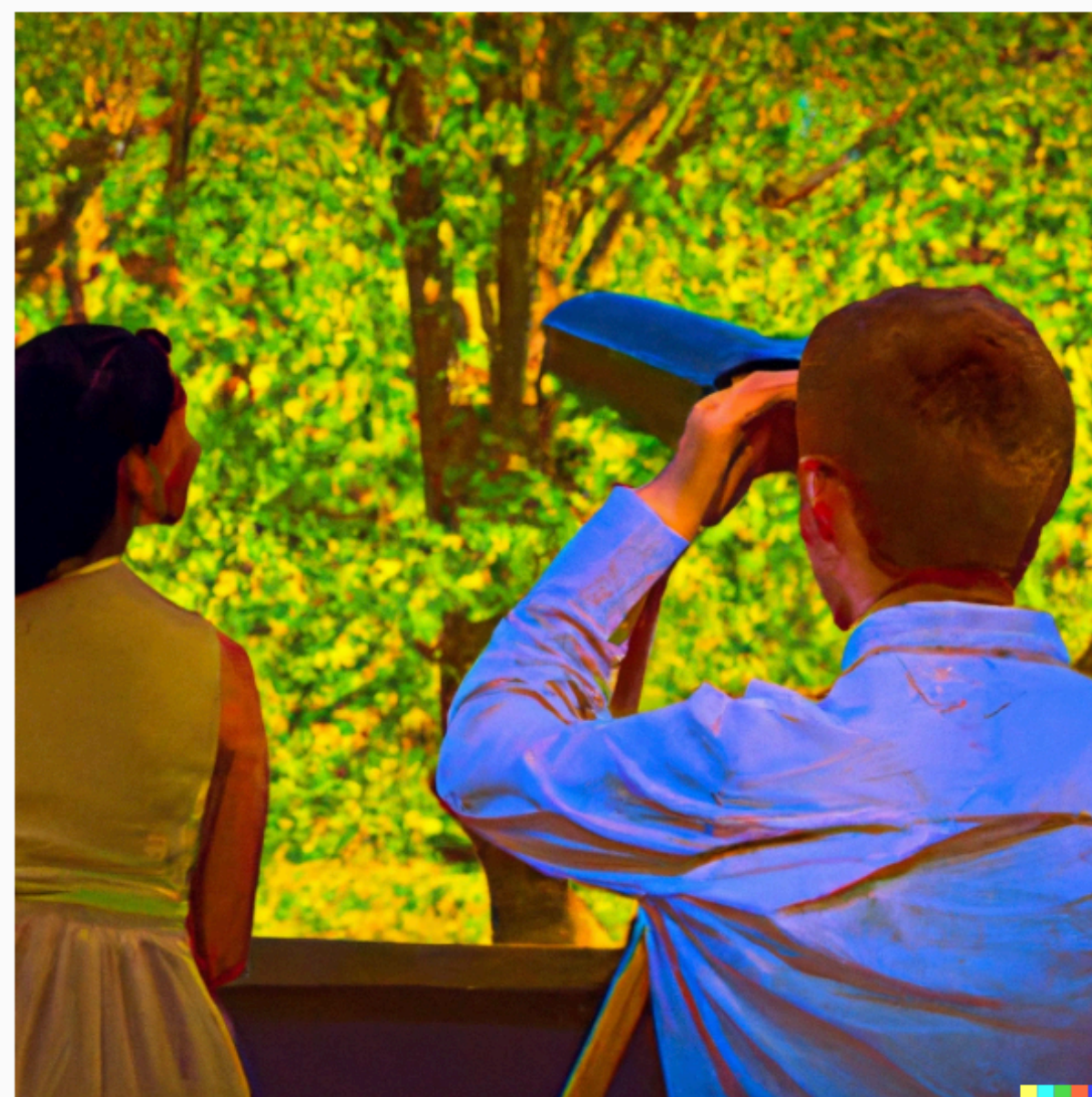
6000



Language is Complex

Edit the detailed description

oil painting of a man watching a woman in a park with a telescope



How many different meanings does this sentence have?

Language is Ambiguous

- Lung cancer in women mushrooms
 - Mushrooms is noun or a verb?
- Ban on nude dancing on governor's desk
 - Similar to “if-then-else” ambiguity
- Island Monks Fly in Satellite to Watch Pope Funeral
 - “get delivery” vs. “fly into space”
- British Left Waffles on Falkland Islands
 - Is it talking about the "British Left" political parties or "Left" as a verb?

Language is ambiguous

The image shows two screenshots of the Google Translate interface. The top screenshot shows the English input field with the text "The man saw the woman in the park with a telescope" and the Korean output field with the text "남자는 공원에서 망원경으로 여자를 보았다." The bottom screenshot shows the Korean input field with the text "남자는 공원에서 망원경으로 여자를 보았다." and the English output field with the text "The man looked at the woman through a telescope in the park." A large blue arrow points from the Korean text in the bottom screenshot to the English text in the top screenshot, highlighting the ambiguity of the Korean sentence.

English - Detected English Spanish French ▾

The man saw the woman in the park with a telescope ×

[Look up details](#)

50 / 5,000

Korean English Spanish ▾

남자는 공원에서 망원경으로 여자를 보았다. ☆

namjaneun gong-won-eseo mang-wongyeong-eulo yeojaleul boassda.

[Look up details](#)

Send feedback

Korean - Detected English Spanish French ▾

남자는 공원에서 망원경으로 여자를 보았다. ×

namjaneun gong-won-eseo mang-wongyeong-eulo yeojaleul boassda.

[Look up details](#)

23 / 5,000

Korean English Spanish ▾

The man looked at the woman through a telescope in the park. ☆

[Look up details](#)

Send feedback

Language is ambiguous

The image shows two screenshots of the Google Translate interface. The top screenshot shows the English input "The man saw the woman in the park with a telescope" being translated to Finnish as "Mies näki naisen puistossa kaukoputkella". The bottom screenshot shows the same Finnish sentence being translated back to English as "The man saw the woman in the park with binoculars". A large blue arrow points from the Finnish sentence in the bottom screenshot to the English translation "with binoculars".

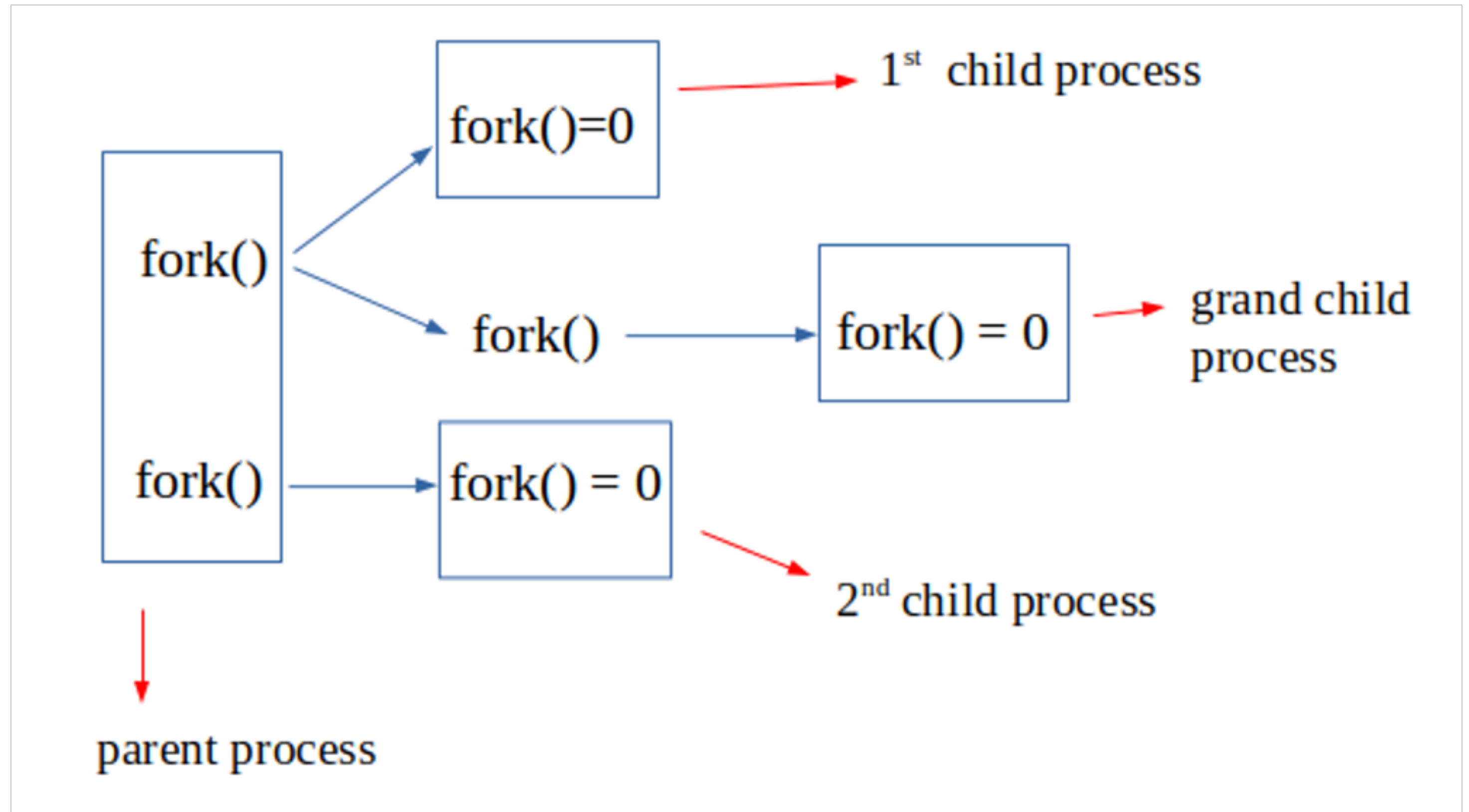
Top Screenshot:

- Language pair: English - Detected to Finnish
- Input: The man saw the woman in the park with a telescope
- Output: Mies näki naisen puistossa kaukoputkella
- Character count: 50 / 5,000

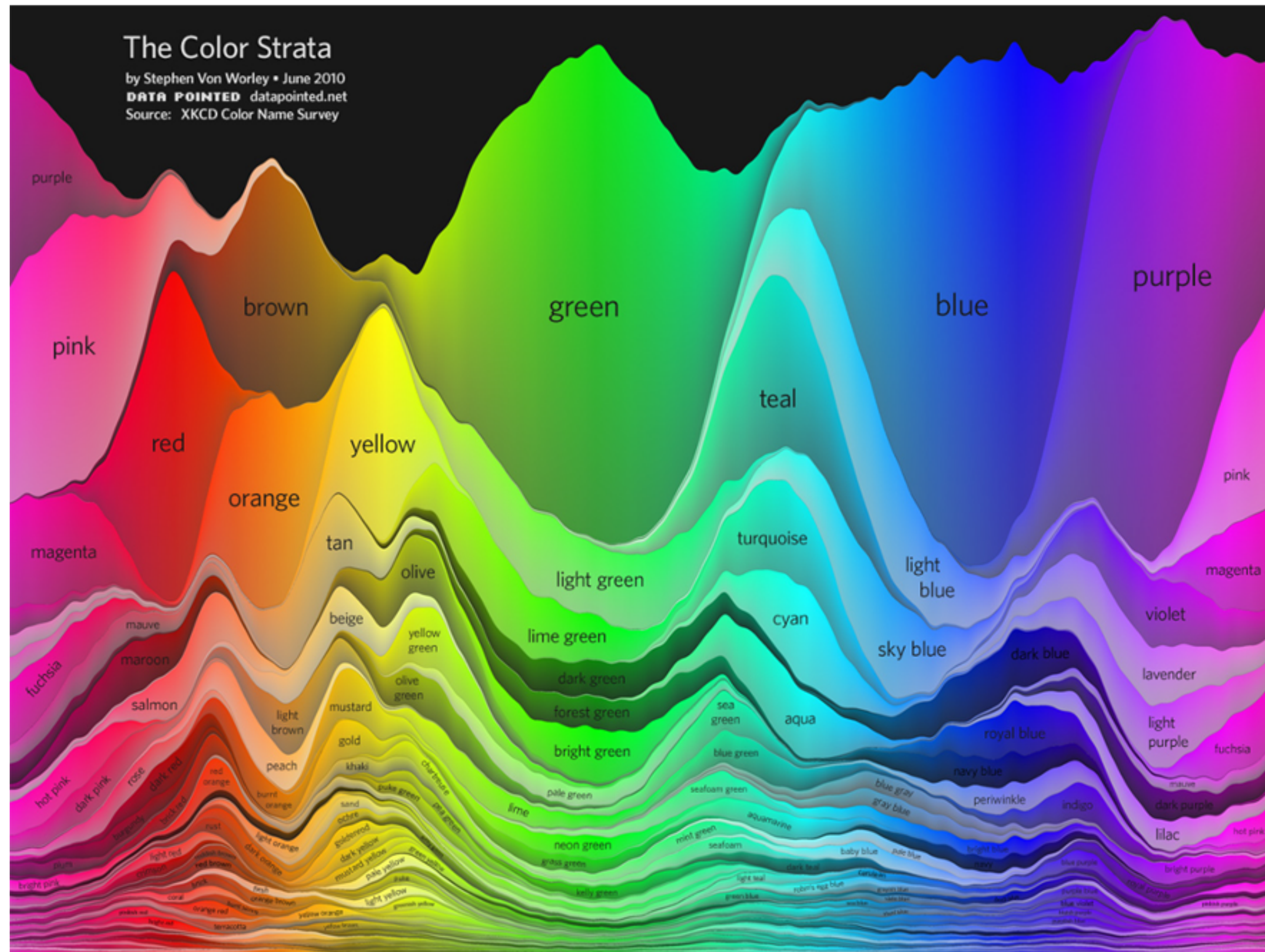
Bottom Screenshot:

- Language pair: Finnish - Detected to English
- Input: Mies näki naisen puistossa kaukoputkella
- Output: The man saw the woman in the park with binoculars
- Character count: 40 / 5,000
- Correction: Did you mean: Mies **näkee** naisen puistossa kaukoputkella

Language is ambiguous



Language is ambiguous



Language is ambiguous



Language is ambiguous

Chinese (Simplified) - Detected English Spanish French

English Spanish Arabic

锅

Guō

[Look up details](#)

2 / 5,000

pot

Noun 锅 壶 罐 煲 盆 锅子 [Look up details](#)

More translations Expand all

pan	Noun	锅 杆 平锅	▼
boiler	Noun	锅炉 锅 镬	▼
cauldron	Noun	釜 锅 镬 甑	▼
caldron	Noun	甑 釜 锅	▼
holloware	Noun	锅	▼

Language is ambiguous



Russian author Sasha Sokolov wrote *Between Dog and Wolf*, which had been deemed untranslatable from the original Russian since 1980. Alexander Boguslawski translated the work to English in December 2017.

<https://retrieverman.wordpress.com/2012/09/08/romeo-and-a-yellow-lab/>

<https://www.npr.org/2017/01/28/512199359/between-dog-and-wolf>

Language is temporal

- **Google's Computer Might Better Translation Tool**
 - New York Times March 8, 2010
- **Number of Lothian patients made ill by drinking rockets**
 - Edinburgh Evening News, March 4, 2010
- **Violinist linked to JAL crash blossoms**
 - <https://languagelog.idc.upenn.edu/nll/?p=1693>

Language is contextual

The town councillors refused to give the angry demonstrators a permit because they feared violence.

Who feared violence?

A) the town councillors

B) the angry demonstrators

Language is contextual

The town councillors refused to give the angry demonstrators a permit because they advocated violence.

Who advocated violence?

A) the town councillors

B) the angry demonstrators

Language is contextual

English - Detected English Spanish French

she took the easy way out

[Look up details](#)



25 / 5,000



Finnish English Korean

그녀는 쉬운 길을 택했다

geunyeoneun swiun gil-eul taeghaessda

[Look up details](#)



[Send feedback](#)

informal

Language is contextual

formal

English - Detected English Spanish French

her majesty, she took the easy way out

[Look up details](#)



38 / 5,000



Finnish English Korean

폐하, 그녀는 쉬운 길을 택했습니다

pyeha, geunyeoneun swiun gil-eul taeghaesseubnida

[Look up details](#)



[Send feedback](#)

Linguistics

The Scientific Study of Human Language Competence

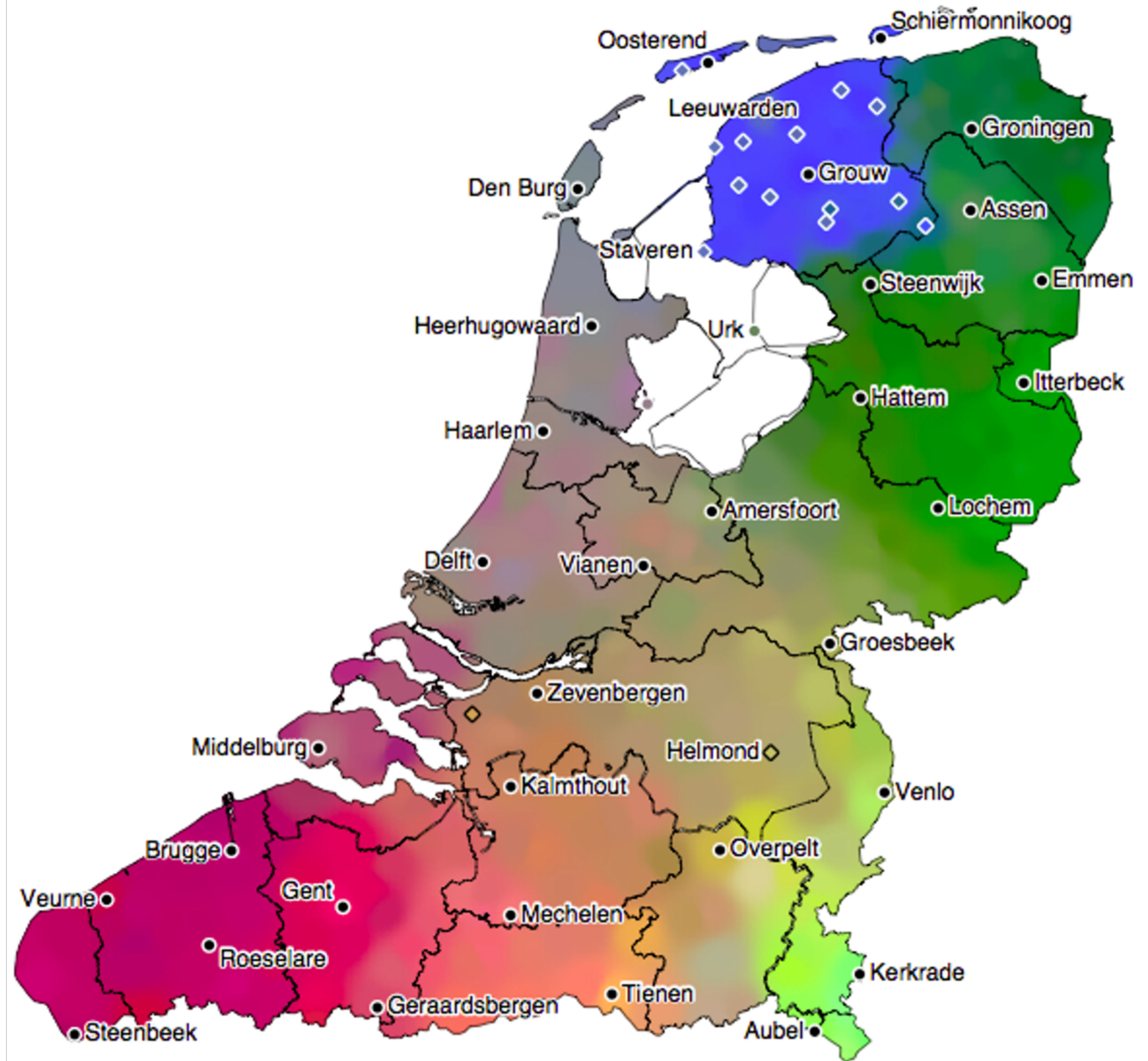
"Every time I fire a linguist, the performance of our speech recognition system goes up."

- Fred Jelinek, IBM Speech Recognition group circa 1980s

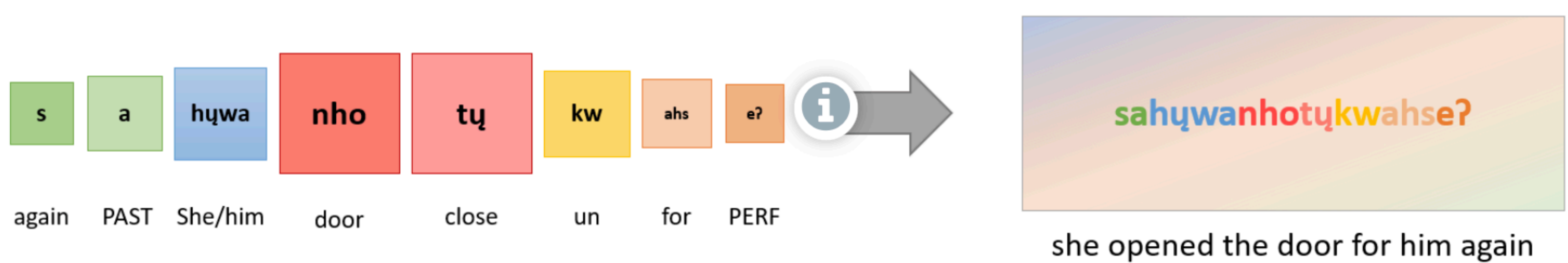
https://en.wikiquote.org/wiki/Fred_Jelinek

Sounds

Pronunciation dialect map of the Netherlands based on phonetic edit-distance (W. Heeringa Phd thesis, 2004)



Words and morphemes



Mohawk or Kanien'kéha

Image credit: Dinesh Ramoo

<https://opentextbc.ca/psyclanguage/chapter/morphology-of-different-languages/>

Words and morphemes

- There is often ambiguity in how morphemes get combined
- Take the word in English: *delighted*
- Can you find two different ways to create this word using morphemes?

**BREAKTHROUGH:
A COMPUTER THAT
UNDERSTANDS YOU
LIKE YOUR MOTHER.**

Having to learn letter-perfect software languages can be frustrating to the average person trying to tap the power of a computer.

But practical thinkers at our McDonnell Douglas Computer Systems Company have created the first computer that accepts you as you are—human.

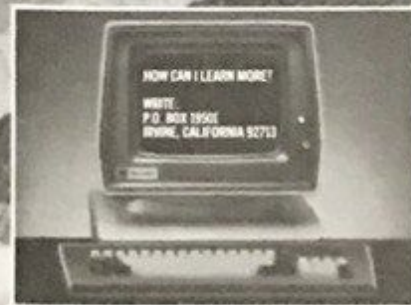
They emulated the two halves of the human brain with two-level software: One level with a dictionary of facts and a second level to interpret them. The resulting Natural Language processor understands everyday conversational English. So it knows what you mean, no matter how you express yourself. It also learns your idiosyncrasies, forgives your errors, and tells you how to find what you're looking for.

Now, virtually anyone who can read and write can use a computer.

We're creating breakthroughs not only in artificial intelligence but also in health care, space manufacturing and aircraft.

We're McDonnell Douglas.

**MCDONNELL
DOUGLAS**



© 1985 McDonnell Douglas Corporation

Syntax

(A computer that understands)
(you like your mother)

(A computer that understands you)
(like your mother)

1985 McDonnell Douglas Print Ad

<https://www.ebay.com/itm/134242735420>

Syntax

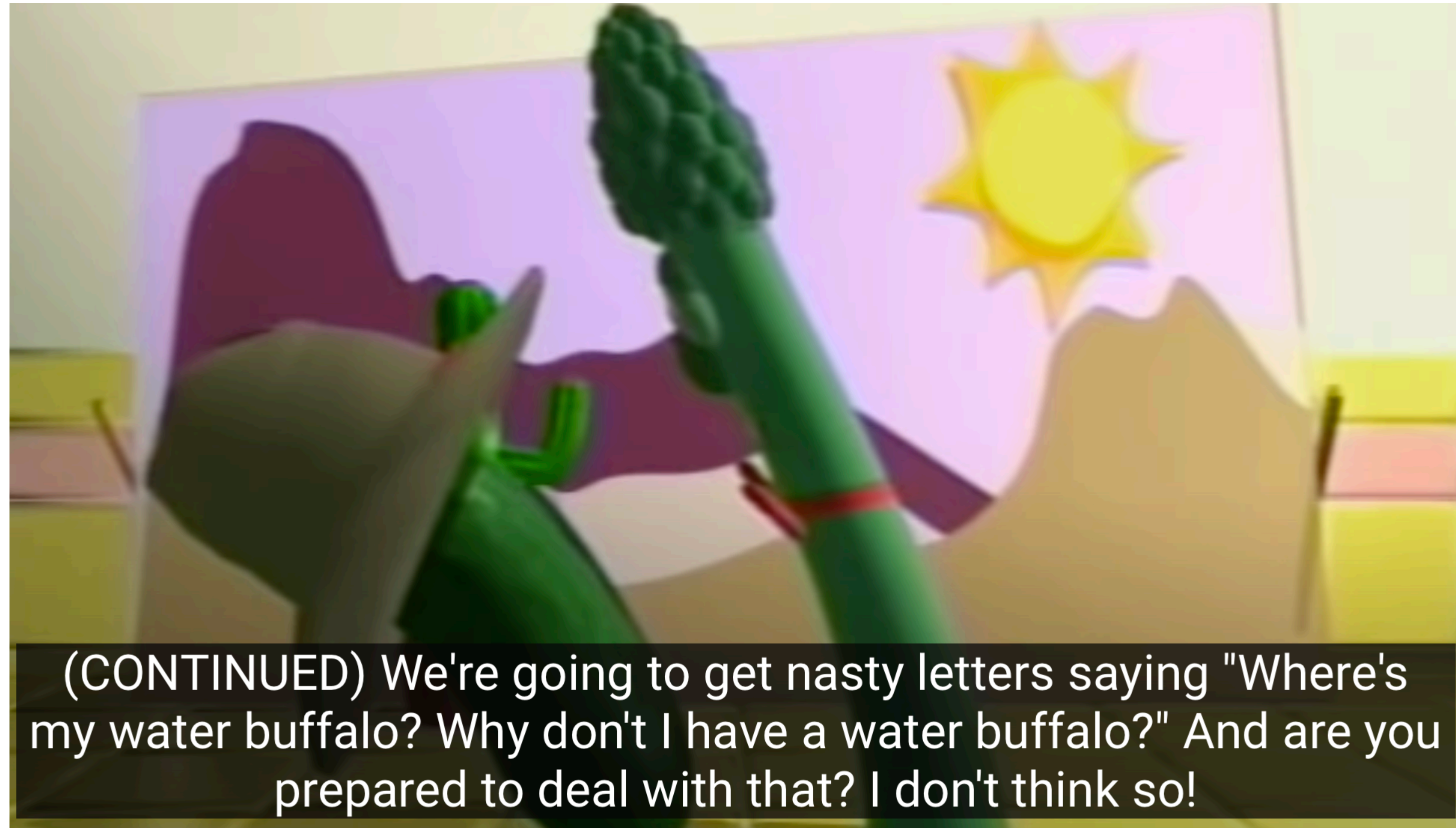
Phenomenon	N	Acceptable Example	Unacceptable Example
ANAPHOR AGR.	2	<i>Many girls insulted <u>themselves</u>.</i>	<i>Many girls insulted <u>herself</u>.</i>
ARG. STRUCTURE	9	<i>Rose wasn't <u>disturbing</u> Mark.</i>	<i>Rose wasn't <u>boasting</u> Mark.</i>
BINDING	7	<i>Carlos said that Lori helped <u>him</u>.</i>	<i>Carlos said that Lori helped <u>himself</u>.</i>
CONTROL/RAISING	5	<i>There was <u>bound</u> to be a fish escaping.</i>	<i>There was <u>unable</u> to be a fish escaping.</i>
DET.-NOUN AGR.	8	<i>Rachelle had bought that <u>chair</u>.</i>	<i>Rachelle had bought that <u>chairs</u>.</i>
ELLIPSIS	2	<i>Anne's doctor cleans one <u>important</u> book and Stacey cleans a few.</i>	<i>Anne's doctor cleans one book and Stacey cleans a few <u>important</u>.</i>
FILLER-GAP	7	<i>Brett knew <u>what</u> many waiters find.</i>	<i>Brett knew <u>that</u> many waiters find.</i>
IRREGULAR FORMS	2	<i>Aaron <u>broke</u> the unicycle.</i>	<i>Aaron <u>broken</u> the unicycle.</i>
ISLAND EFFECTS	8	<i>Which <u>bikes</u> is John fixing?</i>	<i>Which is John fixing <u>bikes</u>?</i>
NPI LICENSING	7	<i>The truck has <u>clearly</u> tipped over.</i>	<i>The truck has <u>ever</u> tipped over.</i>
QUANTIFIERS	4	<i>No boy knew <u>fewer than</u> six guys.</i>	<i>No boy knew <u>at most</u> six guys.</i>
SUBJECT-VERB AGR.	6	<i>These casseroles <u>disgust</u> Kayla.</i>	<i>These casseroles <u>disgusts</u> Kayla.</i>

Table 1: Minimal pairs from each of the twelve linguistic phenomenon categories covered by BLiMP. Differences are underlined. *N* is the number of 1,000-example minimal pair paradigms within each broad category.

Semantics



Semantics



Not every viewer has a water buffalo (but some do)

Every single viewer does not have a water buffalo

Discourse

Plan	Premise	Premise: A new law grad returns home to start her career, but struggles with the broken justice system.
	Setting	Setting: The story is set in a small town in the United States.
	Characters	<ol style="list-style-type: none">Character Portrait: Liza Turner is a 22-year-old woman.Character Portrait: Peyton Turner is Liza's older sister.
	Outline	<p>Outline the main plot points of the story.</p> <ol style="list-style-type: none">Liza Turner graduates from law school.She moves back to her hometown to start her career.She struggles with the reality of the broken justice system.

Discourse

Plan	Premise	Premise: A new law grad returns home to start her career, but struggles with the broken justice system.
	Setting	Setting: The story is set in a small town in the United States.
	Characters	1. Character Portrait: Liza Turner is a 22-year-old woman. 2. Character Portrait: Peyton Turner is Liza's older sister.
	Outline	Outline the main plot points of the story. 1. Liza Turner graduates from law school. 2. She moves back to her hometown to start her career. 3. She struggles with the reality of the broken justice system.

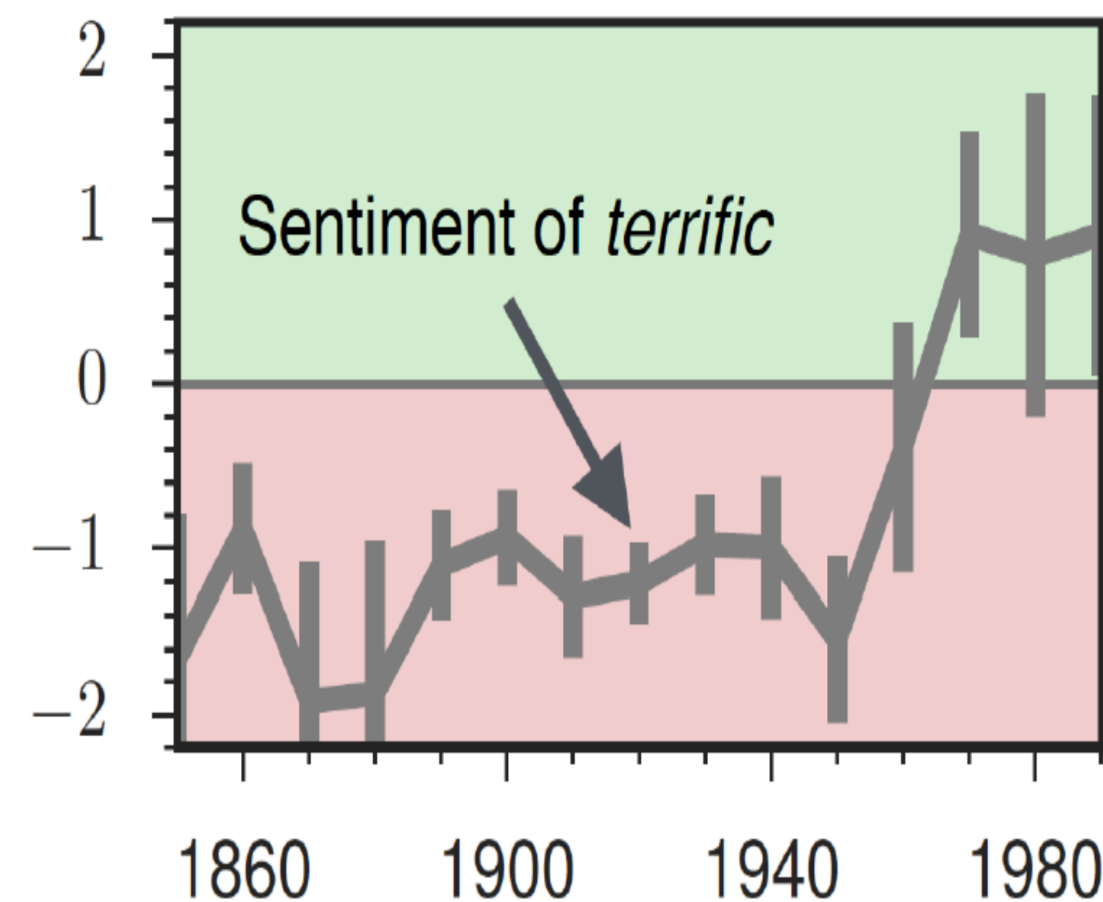
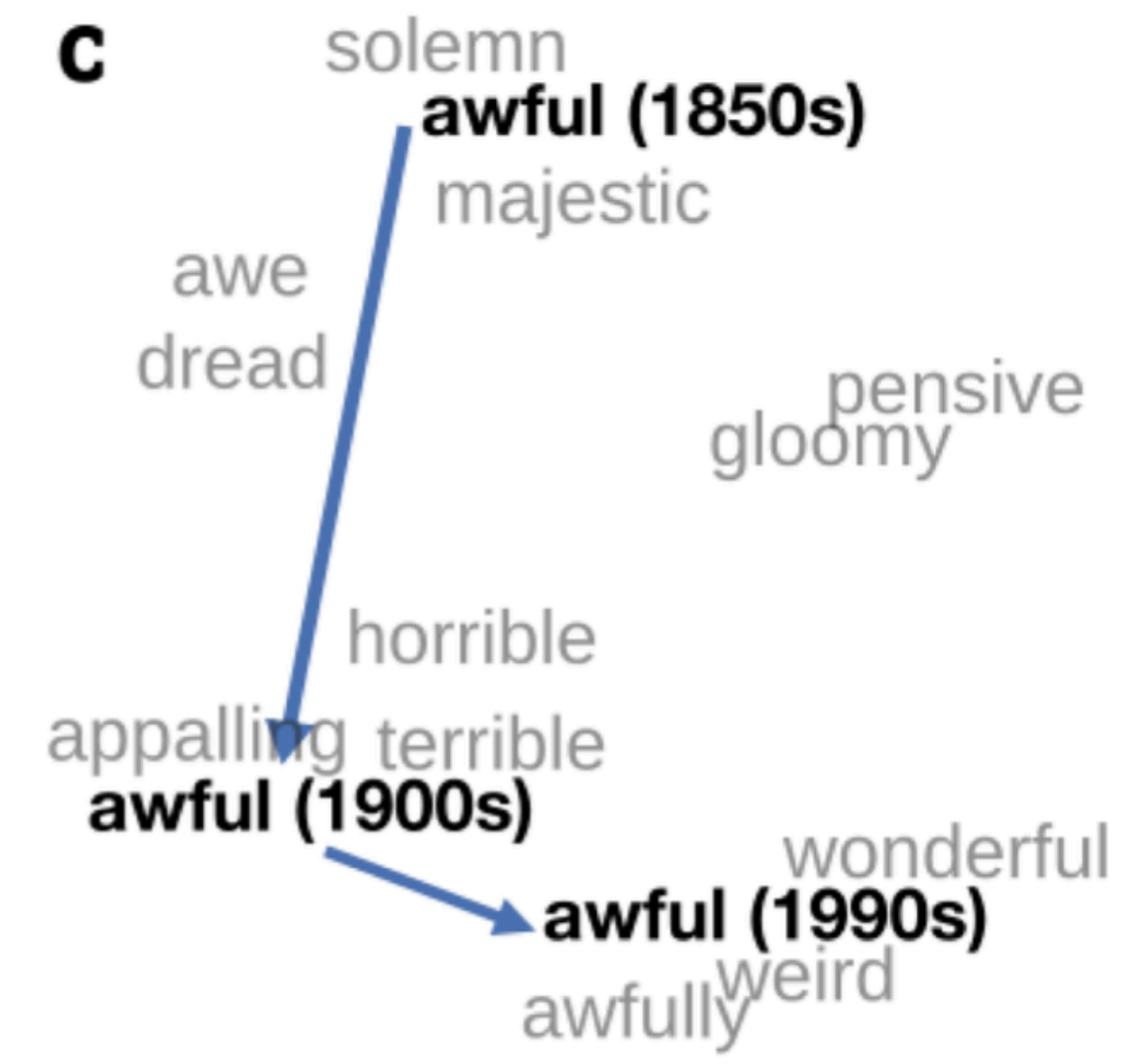
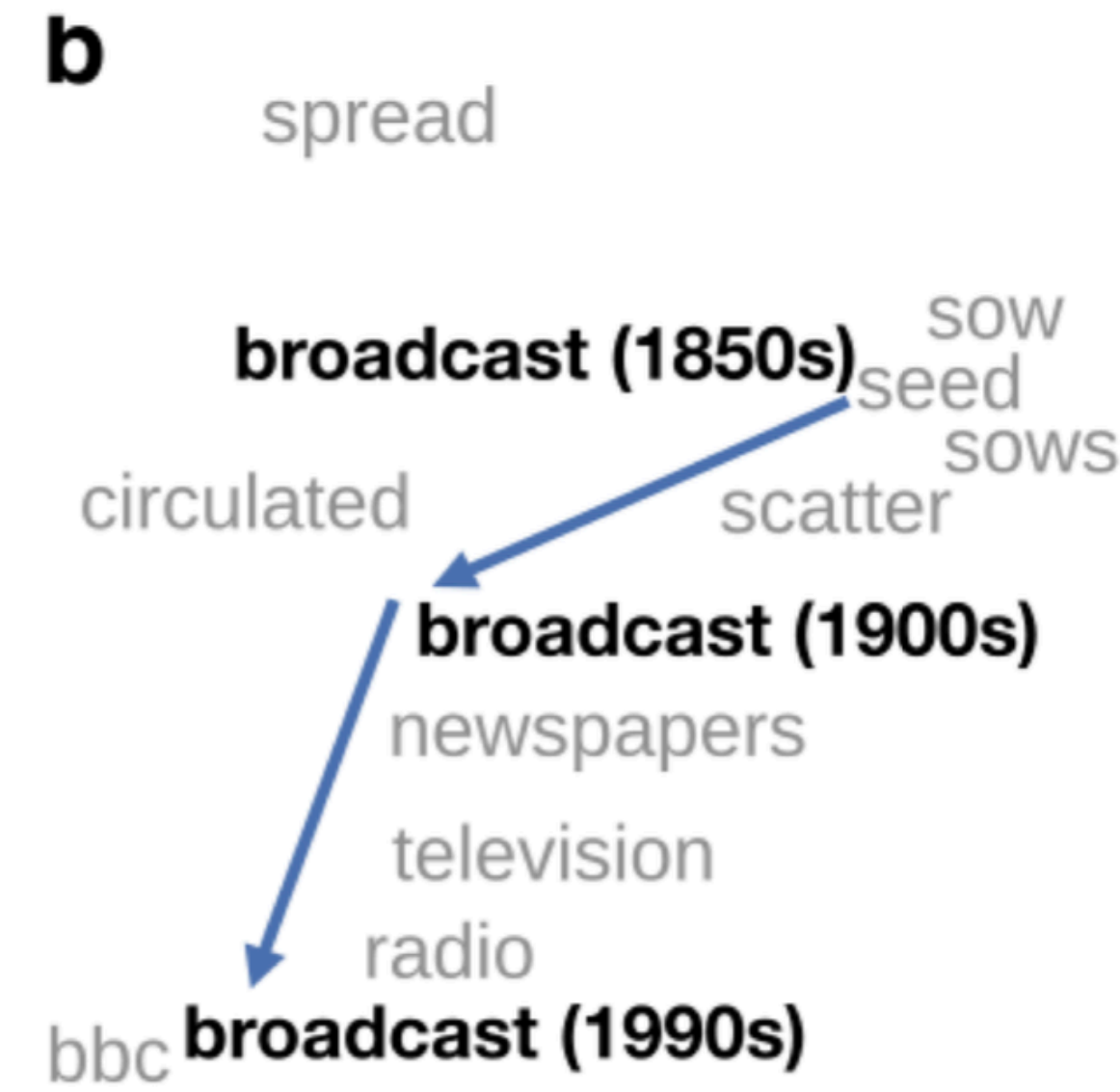
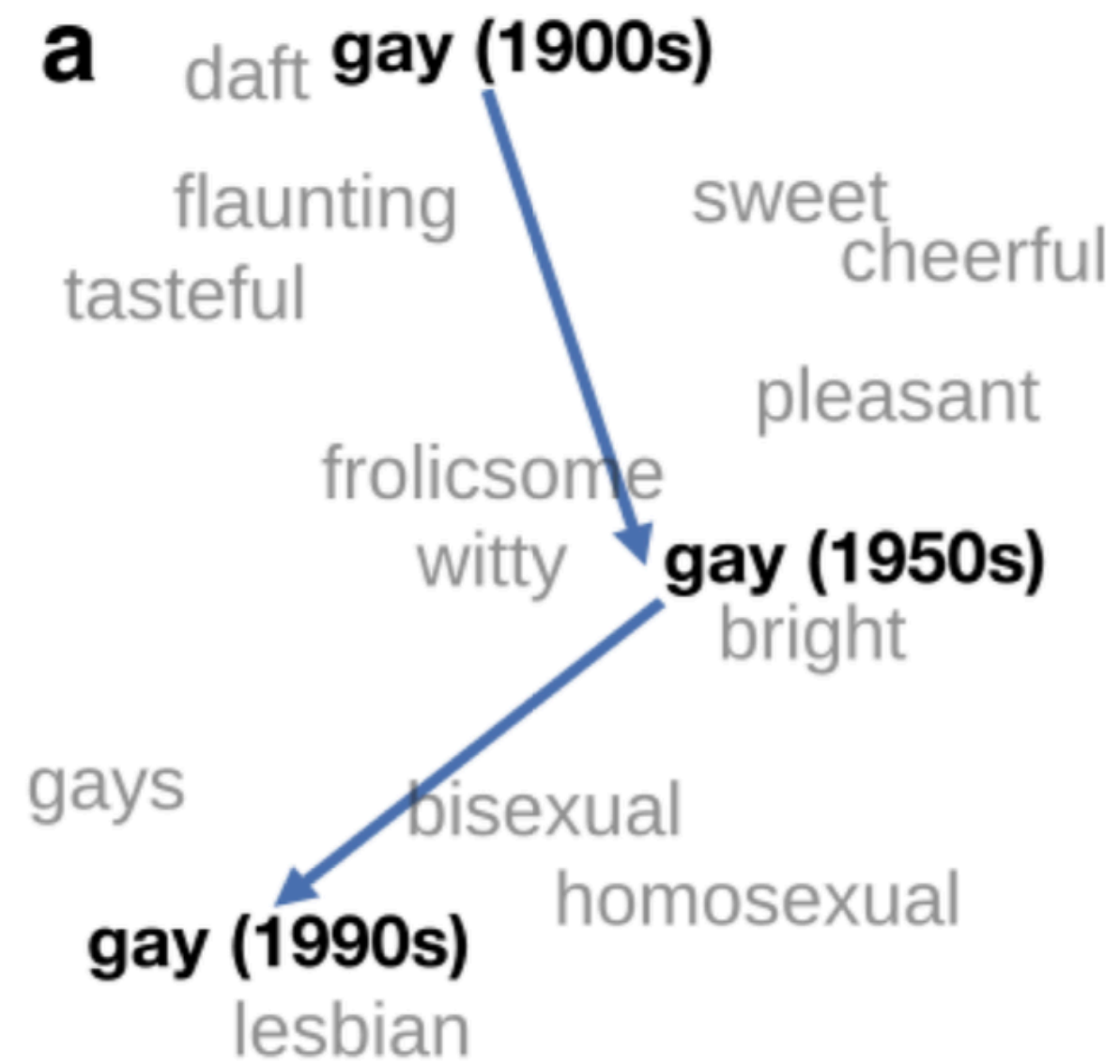
Draft	Relevant Context	Relevant context: Liza Turner is a 22-year-old woman. Peyton Turner is Liza's older sister.
	Previous Sections' Outlines	Previous story summary: Liza Turner graduates from law school.
	Recent Story Summary	Immediately before current passage: Liza Turner returns home to her small town, feeling both familiar and unsafe. She is unsure if she wants to live there, but hesitant to leave.
	Upcoming Section Outline	In the upcoming passage, She moves back to her hometown to start her career.
	Auto-regressive Context	Full text below: She locked up the car and carried her things into the house through the back door in case Peyton was home still.

Rewrite	Draft Continuation 1	All the lights were off and there was no sign of Peyton. She shrugged and decided to go out and spend the rest of her evening at one of New York City's many bars.	Coherence + Relevance -1.7 ✘
	Draft Continuation 2	She knew Peyton was probably working late at his restaurant so he wouldn't come home early to see her, but she wouldn't put it past him to do it anyway.	Coherence + Relevance 2.0 ✔

Edit	Selected Continuation	She knew Peyton was probably working late at his restaurant so he wouldn't come home early to see her, but she wouldn't put it past him to do it anyway.
	Inferred Facts	<u>Peyton Turner</u> Peyton Turner is male. Peyton works at a restaurant.
	Attribute Dictionary	<u>Peyton Turner</u> Younger sister: Liza Turner Gender: female male Workplace: restaurant
	Editing Instruction	Edit so that Peyton Turner is female.
	Final Edited Continuation	She knew Peyton was probably working late at her restaurant so she wouldn't come home early to see her, but she wouldn't put it past her to do it anyway.

Sociolinguistics

- Dialects
- Bias
- Toxicity
- Sentiment



Negative words change faster than positive words

Sociolinguistics



The Unbeatable Squirrel Girl (2015) #4 (Marvel comics)

Reasoning and Inference

That professor dresses even worse than Anoop

What can you infer from this?

Reasoning and Inference

Text	Judgments	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction C C C C C	The man is sleeping
An older and younger man smiling.	neutral N N E N N	Two men are smiling and laughing at the cats playing on the floor.
A black race car starts up in front of a crowd of people.	contradiction C C C C C	A man is driving down a lonely road.
A soccer game with multiple males playing.	entailment E E E E E	Some men are playing a sport.
A smiling costumed woman is holding an umbrella.	neutral N N E C N	A happy woman in a fairy costume holds an umbrella.

- Material for some of these slides comes from ideas borrowed from Angel Chang, Danqi Chen, Karthik Narasimhan and others on the Slack group for teaching NLP.