

TexTiles: Exploring Patterns in Historical Discourse

Robert Roessler, Caiseen Kelly, Michael Behrisch, Johanna Beyer.

4th Workshop on Visualization for the Digital Humanities (at IEEE Vis 19), 2019.

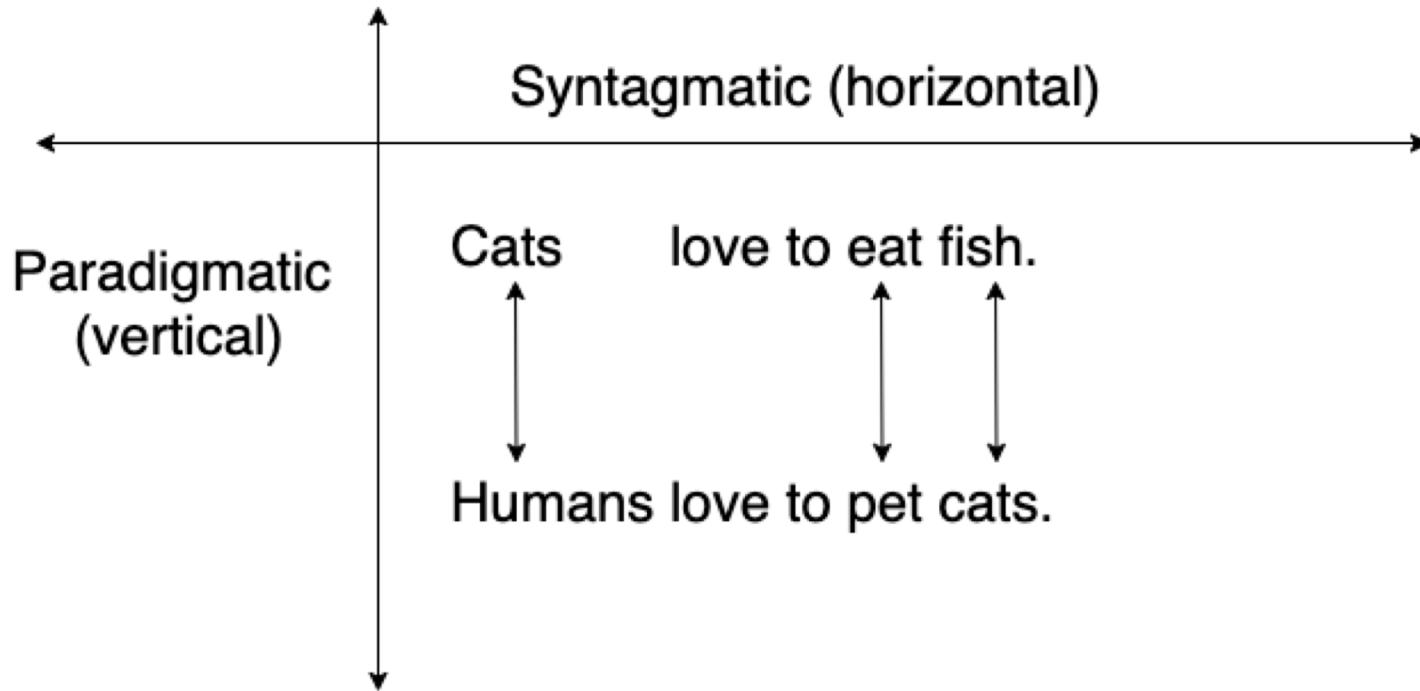
Background

- Discourse
 - A social boundary that defines what statements can be said about a topic
- Discourse analysis
 - A method in the humanities to analyze historical trends and knowledge configurations
- Corpus (pl. corpora)
 - a collection of written texts

User requirements

- Synchronic analysis
 - Explore a system of discourse at one point in time
- Diachronic analysis
 - Analyze discourse over time
- Syntagmatic analysis
 - Combination and connection of words surrounding a previously defined keyword
- Paradigmatic analysis
 - Interchangeability in certain word slots

User requirements



Abstract tasks

- Synchronic analysis
 - Co-occurrence of words
- Diachronic analysis
 - Relative popularity of the surrounding words

Word-Net view

Curate Your Corpus	Create TexTile Visualization for:	verdräng	Generate	Most Common Words in Surrounding		
		verdräng	Generate	verdräng (84)	aus (27)	vorstellung (26)
<input checked="" type="checkbox"/> Johann Friedrich Herbart <i>Psychologie als Wissenschaft (1824)</i>				verdräng (84)	aus (27)	vorstellung (26)
<input type="checkbox"/> Johann Friedrich Herbart <i>Lehrbuch zur Psychologie (1834)</i>				dem (25)	bewusstsein (18)	den (17)
<input type="checkbox"/> Ernst von Feuchtersleben <i>Zur Dietätik der Seele (1838)</i>				wunsch (16)	sich (14)	durch (14)
<input type="checkbox"/> Rudolf Hermann Lotze <i>Medizinische Psychologie oder Physiologie der Seele (1852)</i>				auf (12)	sie (12)	aber (12)
<input type="checkbox"/> Gustav Theodor Fechner <i>Elemente der Psychophysik (1860)</i>				dass (11)	zu (10)	empfindung (10)
<input checked="" type="checkbox"/> Wilhelm Wundt <i>Grundzüge der physiologischen Psychologie (1871)</i>						
<input type="checkbox"/> Franz Brentano <i>Psychologie vom empirischen Standpunkt (1874)</i>						
<input checked="" type="checkbox"/> Gustav Adolf Lindner <i>Lehrbuch der empirischen Psychologie als induktiver Wissenschaft (1883)</i>						
<input type="checkbox"/> Ernst Mach <i>Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen (1886)</i>						
<input checked="" type="checkbox"/> Sigmund Freud <i>Die Traumdeutung (1900)</i>						

word environment for "verdräng" in Sigmund Freud: *Die Traumdeutung* (1900):
der **traum** ist verkleidet erfüllung eines **unterdrückt** **verdräng** **wunsch** nun erübrigen noch **angst** **traum** der

The chart displays the popularity of various words over time, represented as a percentage of the total corpus. The x-axis shows years from 1840 to 1900, and the y-axis shows percentages from 0% to 100%. The words tracked are: verdräng (red), aus (teal), vorstellung (light green), dem (yellow), bewusstsein (orange), den (purple), wunsch (dark purple), sich (pink), durch (light blue), auf (blue), sie (light teal), aber (dark teal), dass (dark orange), zu (yellow-orange), empfindung (dark blue), and traum (yellow-green). The word 'verdräng' (red) shows a significant peak around 1880-1890, while 'aus' (teal) and 'vorstellung' (light green) show more gradual increases over the same period.

Word-Net view

Filter

visualize selected words only

visualize selected words +

visualize all words

Select Words

verdräng (84)	aus (27)	vorstellung (26)
dem (25)	bewusstsein (18)	den (17)
wunsch (16)	sich (14)	durch (14)
auf (12)	sie (12)	aber (12)
daß (11)	zu (10)	empfindung (10)
ein (10)	des (9)	diese (9)
zur (9)	eine (8)	im (8)
heb (8)	wir (8)	nach (7)

APPLY

The diagram illustrates semantic relationships between German words, centered around the word 'verdräng'. Nodes are represented by colored boxes, and edges by colored lines of varying thicknesses. The nodes are:

- verdräng (red)
- vorstellung (teal)
- heb (light green)
- bewusstsein (yellow)
- angst (yellow)
- wunsch (purple)
- traum (orange)
- vorbewusst (purple)
- unbewusst (green)
- empfindung (blue)

The connections are as follows:

- 'verdräng' is highly connected to 'bewusstsein' (thick red line), 'angst' (thick orange line), 'wunsch' (thick purple line), and 'traum' (thick orange line).
- 'bewusstsein' is connected to 'verdräng' (thick red line) and 'angst' (thin yellow line).
- 'angst' is connected to 'verdräng' (thick orange line) and 'wunsch' (thin purple line).
- 'wunsch' is connected to 'verdräng' (thick purple line) and 'traum' (thin orange line).
- 'traum' is connected to 'verdräng' (thick orange line) and 'wunsch' (thin purple line).
- 'vorbewusst' is connected to 'verdräng' (thin purple line) and 'unbewusst' (thin green line).
- 'unbewusst' is connected to 'vorbewusst' (thin green line) and 'verdräng' (thin purple line).
- 'vorstellung' is connected to 'heb' (thin teal line) and 'verdräng' (thin teal line).
- 'heb' is connected to 'vorstellung' (thin teal line) and 'verdräng' (thin teal line).
- 'empfindung' is connected to 'verdräng' (thin blue line) and 'bewusstsein' (thin teal line).
- 'bewusstsein' is connected to 'verdräng' (thin red line) and 'empfindung' (thin teal line).
- 'angst' is connected to 'verdräng' (thick orange line) and 'wunsch' (thin purple line).
- 'wunsch' is connected to 'verdräng' (thick purple line) and 'traum' (thin orange line).
- 'traum' is connected to 'verdräng' (thick orange line) and 'wunsch' (thin purple line).
- 'vorbewusst' is connected to 'verdräng' (thin purple line) and 'unbewusst' (thin green line).
- 'unbewusst' is connected to 'vorbewusst' (thin green line) and 'verdräng' (thin purple line).

Comparative Word-Matrix view



Analysis summary

- What: data
 - Text
- What: derived
 - Frequency of co-occurrence of words
- What: derived
 - Co-occurrence network
- How: encode
 - Time-series line charts
 - Force-directed placement
 - Matrix views of networks
- Scale
 - Thousands of pages of text, dozens of words

Critique

- Strengths
 - Two-dimensional layout augments keyword search
 - Matrix views of networks with color encoding
- Limitation
 - Paradigmatic analysis
 - Definition of co-occurrence
 - Line charts for too scarce and possibly categorical data

Thank you!