

Quantitative Drama Analytics

Heidelberg Computational Humanities Summer School July 15, 2019

https://quadrama.github.io





Part 1: Theory and overview

Introduction (Marcus)

- QuaDramA: People
- QuaDramA: Research

Operationalisation: A key challenge in Computational Literary Studies (Benjamin)

Use cases in quantitative drama analysis

- Utterances (Janis)
- Semantics of character speech (Benjamin)
- Network analysis (Nils)

Part 2: Practice and lab session

Introduction (Nils)

- R Basics
- The R package DramaAnalysis

Lab session (Benjamin/Janis/Marcus/Nils) Work by yourselves on the data, using a corpus of English or German plays

Wrap-Up (Nils) Observations, insights, follow-ups, questions



quadrama.github.io/blog/2019/03/08/quadrama-tutorial.en



Introduction - QuaDramA: People

Dr. Nils Reiter



Janis Pagel





Dr. Marcus Willand



Benjamin Krautter





Introduction - QuaDramA: People

Students (Theses):

Sonja Eberhardt

Hebbel's fathers and children from a psychoanalytical perspective [Working title]

Christiane Schneider

"You whistle / how the beak has grown for you" -The Scaramutza character in the German-speaking drama

Dominik Wabersich

Especially Typical. For the understanding of the Faustfigure around 1800

Nathalie Wiedmer

Automatic extraction and quantitative analysis of relations of dramatic characters

Nina Wittmann *Character types in Expressionist drama* [Working title]

Introduction - QuaDramA: People

Students (Assistants):

Annotation / (Meta) Data

Sonja Eberhard, Nathalie Schürmann, Martin Glasbrenner, Ivan Vidakovic, Laura Friedrichsohn (Anja Braun, Dominik Wabersich, Martin Kuhn, Annika Haag, Anja Schmelzle, Tanja Preuß, Nina Stark, Alexander Frank)

Technical Support

Tim Strohmayer

Overall goal:

Test hypotheses about dramatic characters on large corpora using NLP methods

Overall goal:

Test hypotheses about dramatic characters on large corpora using NLP methods

Hypothesis: A proposed explanation for a phenomenon.

Overall goal:

Test hypotheses about dramatic characters on large corpora using NLP methods

Hypothesis: A proposed explanation for a phenomenon.

Scientific hypothesis: Requires that one can test it.

Overall goal:

Test hypotheses about dramatic characters on large corpora using NLP methods

Hypothesis: A proposed explanation for a phenomenon.

Scientific hypothesis: Requires that one can test it.

Hypothesis in literary studies?

Overall goal:

Test hypotheses about dramatic characters on large corpora using NLP methods

Hypothesis: A proposed explanation for a phenomenon.

Scientific hypothesis: Requires that one can test it. Hypothesis in literary studies? The play is brilliant because god made its author brilliant vs. The play is brilliant because its characters speak in a very unique way

Overall goal:

Test hypotheses about dramatic characters on large corpora using NLP methods

aesthetics

Hypothesis: A proposed explanation for a phenomenon.

Scientific hypothesis:

Requires that one can test it.

Hypothesis in literary studies?

The play is brilliant because god made its author brilliant

VS.

structuralism The play is brilliant because its characters speak in a very unique way

428

Research objects:

Dramatic Texts

= Plays≠ Prose≠ Poetry

EIGENSCHAFTEN FIKT LIT \rightarrow DLS vs xy

Operationalisierung als Grundherausforderung

MARY J. WINDLE.

violent death; your voice has condemned him-your voice can yet save him. If it be your country's weal that you desire, that object has been already sufficiently answered by the example of his trial; or, if it is to further the cause of the Lord of Hosts that you place yourself at the head of Britain in his place, be assured that he who would assert his power by surrounding himself with a pomp like this, is no delegate of One who commissioned Moses to lead his people through the wilderness, a sharer in the common lot, and a houseless wanderer like themselves. Bethink you, therefore, what must be the doom of him, who-for the sake of ambition and pride -in order that he might for the brief space of his life enjoy luxury and power-under the borrowed name, too, of that God who views the act with horror and detestation-stains his hands with parricidal blood. Yes, General Cromwell, for thy own soul's, if not for mercy's sake, I entreat thee, in whom alone lies the power, to cause Charles Stuart's sentence to be remitted."

After a few moments' hesitation, during which Alice looked in his face with the deepest anxiety, and awaited his answer, he said, "Go to, young woman, who presumest to interfere between a judge raised up for the redemption of England, and a traitor king, whom the Lord hath permitted to be condemned to the axe. As my soul liveth, and as He liveth, who will one day make me a ruler in Israel, thou hast more than the vanity of thy sex, in hoping by thy foolish speech to more me to lift up my hand against the decree of the Almighty. Truly..."

"Nay, General Cromwell," said Alice, interrupting him, as soon as she perceived he was about to enter into one of his lengthy and pointless harangues, "nay, you evade the matter both with me and with the conscience whose workings I have for the last few moments beheld in the disorder of your frame. Have its pleadings—for to them I look and not to any eloquence of mine own—been of no avail ? Will it please you to do aught for the king ?"

"Young lady," replied Cromwell, bursting into tears, which he was occasionally wont to do, "a man like me, who is called to perform great acts in Israel, had need to be immovable to feelings of human charities. Think you not it is painful to our mortal sym-

Act 1 Scene 1

Verona's Streets

Two C in the Balthe	Capulet servants, Gregory and Sampson, are talking street. Two Montague servants, Abraham and asar, come by.
Gregory:	Here comes two of the house of the Montagues.
Sampson:	Quarrel; I will back thee.
Gregory:	How? Turn thy back and run?
Sampson:	Fear me not.
Gregory:	I will frown as they pass by and let them take it as they list.
Sampson:	Nay as they dare. I will bite my thumb at them which is disgrace to them if they bear it.
Samp servar	son aims this insulting gesture at the Montague nts who cannot ignore it.
Abraham:	Do you bite your thumb at us, sir?
Sampson:	I do bite my thumb, sir.
Abraham:	Do you bite your thumb at us, sir?
Gregory:	Do you quarrel, sir?
Abraham:	Quarrel, sir? No, sir.
Sampson:	But if you do, sir, I am for you. I serve as good a man as you. Draw if you be men.
A figh opposi	t breaks out. Benvolio and Tybalt enter from ite directions.
Benvolio:	Part fools! Put up your swords; you know not what you do.
Tybali	t approaches Benvolio, who speaks to him
	I do but keep the peace: put up thy sword.

Research objects:

- > 465 german dramatic texts (soon some more)
 Textgrid / DLINA (Frank Fischer, Peer Trilcke et al.)
- ~ 1740 1930
- canonical / non-canonical
- Tragedy / Comedy
 - subgenres



structural drama research: non-digital past CONFIGURATION MATRICES

	1	2	3	4	5	6	7	8	9	10
Dirne	1	0	0	0	0	0	0	0	0	1
Soldat	1	1	0	0	0	0	0	0	0	0
Stubenmädchen	0	1	1	0	0	0	0	0	0	0
Junger Herr	0	0	1	1	0	0	0	0	0	0
Junge Frau	0	0	0	1	1	0	0	0	0	0
Ehemann	0	0	0	0	1	1	0	0	0	0
Süßes Mädel	0	0	0	0	0	1	1	0	0	0
Dichter	0	0	0	0	0	0	1	1	0	0
Schauspielerin	0	0	0	0	0	0	0	1	1	0
Graf	0	0	0	0	0	0	0	0	1	1

Pfister: Das Drama (2001[1977]), S.

239 Schnitzlers Round Dance

structural drama research: digital present **NETWORK-GRAPHS**

	1	2	3	4	5	6	7	8	9	10
Dirne	1	0	0	0	0	0	0	0	0	1
Soldat	1	1	0	0	0	0	0	0	0	0
Stubenmädchen	0	1	1	0	0	0	0	0	0	0
Junger Herr	0	0	1	1	0	0	0	0	0	0
Junge Frau	0	0	0	1	1	0	0	0	0	0
Ehemann	0	0	0	0	1	1	0	0	0	0
Süßes Mädel	0	0	0	0	0	1	1	0	0	0
Dichter	0	0	0	0	0	0	1	1	0	0
Schauspielerin	0	0	0	0	0	0	0	1	1	0
Graf	0	0	0	0	0	0	0	0	1	1



https://dracor.org/ger/schnitzler-reigen

structural drama research: digital present

	1	2	3	4	5	6	7	8	9	10
Dirne	1	0	0	0	0	0	0	0	0	1
Soldat	1	1	0	0	0	0	0	0	0	0
Stubenmädchen	0	1	1	0	0	0	0	0	0	0
Junger Herr	0	0	1	1	0	0	0	0	0	0
Junge Frau	0	0	0	1	1	0	0	0	0	0
Ehemann	0	0	0	0	1	1	0	0	0	0
Süßes Mädel	0	0	0	0	0	1	1	0	0	0
Dichter	0	0	0	0	0	0	1	1	0	0
Schauspielerin	0	0	0	0	0	0	0	1	1	0
Graf	0	0	0	0	0	0	0	0	1	1

Speech-based analysis



Focus: Dramatic characters

3 Dimensions

- Character types
 - By gender, action, social class, stock character
 - How are fathers marked as being 'tender'?
- Relations between characters
 - What are shared topics/emotions?
 - What do characters say about each other?
- Character type development
 - How do topics change?
 - How are relations changing?



Operationalization: A key challenge in Computational Literary Studies



Question: How to measure a theoretical concept?

- P. W. Bridgman: *Logic of Modern Physics* (1927) Example:

"The concept of length is therefore fixed when the operations by which length is fixed are fixed: that is, the concept of length involves as much and nothing more than the set of operations by which length is determined."

- Franco Moretti: "Operationalizing": or, the function of measurement in modern literary theory (2013)

"[...] In our case: from the concepts of literary theory, through some form of quantification, to literary texts."

Question: How to measure a theoretical concept in literary studies?

- An adequate operationalization is important for the acceptance of the results in literary studies
- How to: **direct** operationalization and/or **approximate** operationalization

Some potential problems:

- Focus of the theoretical frame: author, text, reader?
- Different approaches: a concept as an analytic tool for text analysis or text interpretation to sharpen the concept
- Concepts range from being rather formalistic (e.g., narratology) to (deliberately) vague
- Oftentimes, literary concepts cannot be measured directly

Direct operationalization:

. . .

- Size of the stage personnel (how many characters?)
- Number of acts and/or scenes
- When are characters introduced?
- How do characters appear and disappear on stage? (*liaison de scène*)

Most of the times, these questions are relevant for the comparison within a large(r) text corpus, e.g., for studies in literary history

Approximate Operationalization:

- Is it possible to operationalize vague concepts? How can we measure concepts that were not established with quantification in mind?
- Possible concepts:
 - Who is the protagonist of a literary text?
 - What are the themes/topics characters talk about?
 - Is it possible to distinguish different character types? (lover, father, villain,...)
 - Is it possible to determine the genre of a given text?

- ...

- Poetological concepts: e.g., the concept of **natural language**, a character development (e.g., from Katilina to Brutus) or a specific aesthetic effect for the reader/spectator

Approximate Operationalization - natural language:

- Approach: split the concept in individual aspects that are easier to measure; combine the individual aspects to get an approximation
- criteria/Instrumental variables (Graham Sack: Simulating Plot, 2011)
 - sentence length
 - utterance length
 - use of punctuation (e.g., dashes)
 - use of vocabulary (type/token ratio)

- ...

Operationalization through annotation

- Goal: try to sharpen the meaning/sense of a concept through manual annotation
- Approach: textual work on individual cases
- Needs annotation guidelines that are continuously updated: Guideline as applicable operationalization of the concept
- Application can result in a training corpus, e.g., for machine learning (second operationalization)

Use Cases in Quantitative Drama Analysis

Use Cases in Quantitative Drama Analysis **Detecting Protagonists**



Protagonists

Need to make literary view operationalizable

Protagonists

Need to make literary view operationalizable

Need for clear guidelines for annotation

Protagonists

Need to make literary view operationalizable

Need for clear guidelines for annotation

Our definition:

- 1. Figure causes or solves the central dramatic conflict
- 2. Either actively or passively

Features

Come up with good features/characteristics that capture what it means/entails to be a protagonist

Utterances





Emilia Galotti

Group discussion

What could be reasonable features to characterize and classify protagonists in plays?

- Number of uttered tokens
-

 - ...

- Features that we have experimented with
 - Uttered tokens

- Features that we have experimented with
 - Uttered tokens
 - Stage presence (active and passive)

- Features that we have experimented with
 - Uttered tokens
 - Stage presence (active and passive)
 - Centrality measures
 - (Weighted) Degree
 - Eigenvalue centrality
 - Betweenness
 - Closeness

- Features that we have experimented with
 - Uttered tokens
 - Stage presence (active and passive)
 - Centrality measures
 - (Weighted) Degree
 - Eigenvalue centrality
 - Betweenness
 - Closeness
 - Topic modeling

- Features that we have experimented with
 - Uttered tokens
 - Stage presence (active and passive)
 - Centrality measures
 - (Weighted) Degree
 - Eigenvalue centrality
 - Betweenness
 - Closeness
 - Topic modeling
 - Appears in last act?

- Features that we have experimented with
 - Uttered tokens
 - Stage presence (active and passive)
 - Centrality measures
 - (Weighted) Degree
 - Eigenvalue centrality
 - Betweenness
 - Closeness
 - Topic modeling
 - Appears in last act?
 - Genre/Epoch

Classification

Next step:

Train model to decide: protagonist or not-protagonist (based on features)

Feature importance



Results on 34 plays with 171 protagonists

Shapley analysis



Emilia Galotti

Use Cases in Quantitative Drama Analysis Word field analysis

Kleist's plays

/ dikto't. ctatorial tator. 2 over orially adv. (Latin TATOR diction /'dık $\int(\partial n) n/n$. ciation in speaking dictio from dico dictdictionary /'dikfana book listing (usu. al explaining the words giving corresponding language. 2 reference terms p

Word field analysis

How can we measure prevalent themes in the character speech of German

plays (1750-1830) that are interpretable?

- Dictionary based approach, created by a domain expert; currently 5 dictionaries:
 - Love (91 words)
 - Familiy (73 words)
 - Reason (108 words)
 - War (112 words)
 - Religion (57 words)
 - others are wip (politics, economy, ...)

Word field analysis

 adore, adored, adoring, sincere, sincerity, covetousness, desire, eagerness, mourn, relationship, burning, ceremony, copulation, copulate, honest, feeling, sentiment, angel, delight, feeling, spark, beloved, lover, spouse, wife, luck, marriage, heart, wedding, homage, cuddle, kiss, passion, love, gracious, graciousness, caress, lip, lust, rose, beauty, soul, sensual, sensuality, marriage, marry, tender, tenderness, ceremony, lovest, [...]

Word field analysis - single character



Word field analysis - Kleist's Familie Schroffenstein













Year

Use Cases in Quantitative Drama Analysis Relations and Character Presence



Configuration

- Represents the character stage presence
 - Right: Romeo & Juliet, by scene
- Extension in two directions:
 - Copresence networks
 - Passive presence

character 🌐 🚊	1 ‡	2 [‡]	3 [‡]	4 [‡]	5 [‡]	6 0
Benvolio	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE
Tybalt	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE
Capulet	TRUE	TRUE	FALSE	FALSE	TRUE	FALSE
Lady Capulet	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE
Montague	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
Escalus	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
Romeo	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE
Paris	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
Nurse	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
Juliet	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
Mercutio	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE
Friar Lawrence	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Configuration \rightarrow Copresence Network

Character	1	2	3	4	5
Romeo	\checkmark	\checkmark	X	\checkmark	X
Juliet	X	\checkmark	\checkmark	\checkmark	\checkmark
Paris	X	X	\checkmark	X	X

1: Configuration

Configuration \rightarrow Copresence Network

Character	1	2	3	4	5
Romeo	\checkmark	\checkmark	X	\checkmark	X
Juliet	X	\checkmark	\checkmark	\checkmark	\checkmark
Paris	X	X	\checkmark	X	X

1: Configuration

	Romeo	Juliet	Paris
Romeo	3	2	0
Juliet	2	4	1
Paris	0	1	1

2: Adjacency matrix

Configuration → Copresence Network

Character	1	2	3	4	5
Romeo	\checkmark	\checkmark	X	\checkmark	X
Juliet	X	\checkmark	\checkmark	\checkmark	\checkmark
Paris	X	X	\checkmark	X	X

1: Configuration



	Romeo	Juliet	Paris
Romeo	3	2	0
Juliet	2	4	1
Paris	0	1	1

2: Adjacency matrix

Copresence Network

Network metrics

• Avg. degree, centrality, densely connected clusters, ...



Trilcke et al. (2015)

Copresence Network

Network metrics

- Avg. degree, centrality, densely connected clusters, ...
- Drama-historic developments



Trilcke et al. (2015)

Copresence Network

Network metrics

- Avg. degree, centrality, densely connected clusters, ...
- Drama-historic developments
- Character relations
- Visualisation



Trilcke et al. (2015)

Configuration and Presence

- Configuration represents the character stage presence
- Active presence: Character on stage
 - Misleading

Act I, Scene 1.

Prince. Complaints; nothing but complaints! Petitions; nothing but petitions! [...] Emilia? (*opening a petition, and looking at the signature.*) An Emilia? Yes - but an Emilia Bruneschi - not Galotti. Not Emilia Galotti. What does she want, this Emilia Bruneschi? (*Reads*) She asks much--too

Pirro -Orsina Odoardo Galotti Marinelli Emilia Der Prinz Conti Claudia Appiani Angelo

much. But her name is

Configuration and Presence

- Configuration represents the character stage presence
- Active presence: Character on stage
 - Misleading
- Passive presence: Character not on stage, but talked about
- Coreference resolution
 - Proper names
 - Pronouns
 - Nominal phrases

We are still working on that!

Act I, Scene 1.

Prince. Complaints; nothing but complaints! Petitions; nothing but petitions! [...] Emilia? (*opening a petition, and looking at the signature.*) An Emilia? Yes - but an Emilia Bruneschi - not Galotti. Not Emilia Galotti. What does she want, this Emilia Bruneschi? (*Reads*) She asks much--too much. But her name is Emilia. It is granted (*signs the paper, and rings*).

Active and Passive Presence

Character	Active scenes	Passive scenes	Presence
Appiani	5	25	-0.465
Claudia	13	12	0.023
Der Prinz	14	26	-0.279
Emilia	7	34	-0.627
Marinelli	19	23	-0.093
Odoardo	12	20	-0.186
Orsina	6	16	-0.232



Presence of Non-Characters

Coreference annotation also gives us a grasp on objects and abstract entities



References

Peer Trilcke, Frank Fischer, Mathias Göbel and Dario Kampkaspar. **Comedy vs. Tragedy: Network values by genre**. 2015. URL <u>https://dlina.github.io/Network-Values-by-Genre/</u>.

Franco Moretti. **"Operationalizing": or the Function of Measurement in Modern Literary Theory**. In: Literay Lab Pamphlet 6 (2013).

P. W. Bridgman. The logic of modern physics. Macmillan, NY, 1927.

Marcus Willand and Nils Reiter. **Geschlecht und Gattung. Digitale Analysen von Kleists 'Familie Schroffenstein'**. In: Kleist Jahrbuch (2017), pp. 177-195.

Graham Alexander Sack. **Simulating Plot: Towards a Generative Model of Narrative Structure**. Papers from the AAAI Fall Symposium (2011).

Final Discussion