Natural language processing for subjectivity analysis in personal narratives

Gustave Cortal

Thesis director: Alain Finkel
Co-advisors: Patrick PAROUBEK and Lina YE

Laboratoire Méthodes Formelles

école normale supérieure paris–saclay

université PARIS-SACLAY

Introduction

Context

- Natural language processing for psychology is underexplored

Context

- Natural language processing for psychology is underexplored
- We build on an existing subfield: emotion analysis

Context

- Natural language processing for psychology is underexplored
- We build on an existing subfield: emotion analysis
- We study subjectivity (involving first-person perspective, meaning-making processes, and experiential content)

Context

- Natural language processing for psychology is underexplored
- We build on an existing subfield: emotion analysis
- We study subjectivity (involving first-person perspective, meaning-making processes, and experiential content)
- We focus on personal narratives (*e.g.*, dreams, emotional narratives)

Context

- Natural language processing for psychology is underexplored
- We build on an existing subfield: emotion analysis
- We study subjectivity (involving first-person perspective, meaning-making processes, and experiential content)
- We focus on personal narratives (*e.g.*, dreams, emotional narratives)

Context

- Natural language processing for psychology is underexplored
- We build on an existing subfield: emotion analysis
- We study subjectivity (involving first-person perspective, meaning-making processes, and experiential content)
- We focus on personal narratives (*e.g.*, dreams, emotional narratives)

We first address the *content* by classifying elements of personal narratives (*e.g.*, characters and emotions). Then, we study the *form* through the concept of style

Introduction

How to model subjective experience in personal narratives?

Introduction

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science

Introduction

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset

Introduction

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset
- Training of language models for emotion analysis

Introduction

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset
- Training of language models for emotion analysis
- Formalization of style in personal narratives

Introduction

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset
- Training of language models for emotion analysis
- Formalization of style in personal narratives
- Automatic thematic analysis in mental health narratives

Definition of objectives using cognitive science

G. Cortal and C. Bonard. Improving Language Models for Emotion Analysis: Insights from Cognitive Science. *CMCL, ACL 2024*.

Definition of objectives and scope using cognitive science

What are current limitations and interesting research directions?

Definition of objectives and scope using cognitive science

What are current limitations and interesting research directions?

We review psychological theories of emotion and emotion annotation schemes in NLP

Definition of objectives and scope using cognitive science

What are current limitations and interesting research directions?

We review psychological theories of emotion and emotion annotation schemes in NLP

- Different emotion theories lead to divergences in how to annotate them in the text

Definition of objectives and scope using cognitive science

What are current limitations and interesting research directions?

We review psychological theories of emotion and emotion annotation schemes in NLP

- Different emotion theories lead to divergences in how to annotate them in the text
- Some linguistic and cognitive science theories are not considered

Definition of objectives and scope using cognitive science

What are current limitations and interesting research directions?

We review psychological theories of emotion and emotion annotation schemes in NLP

- Different emotion theories lead to divergences in how to annotate them in the text
- Some linguistic and cognitive science theories are not considered
- There is no benchmark that evaluates the richness of the emotional phenomenon

Emotional analysis in text

What are the main ways to annotate emotion in text?

Psychological theories	In text, emotion is...	Example
Basic emotions theory	a **category**	"I love philosophy." → `joy`
Constructivist theories	a continuous value with an **affective** meaning	"His voice soothes me." → `valence` (4/5), `arousal` (1/5)
Appraisal theory	a continuous value with a **cognitive** meaning	"I received a surprise gift." → `sudden` (4/5), `control` (0/5)
	composed of **semantic roles**	"Louise (`experiencer`) was angry (`cue`) towards Paul (`target`), because he didn't inform her (`cause`)."

Integrated framework for emotion theories

How to integrate psychological theories of emotion?

Integrated framework for emotion theories

How to integrate psychological theories of emotion?

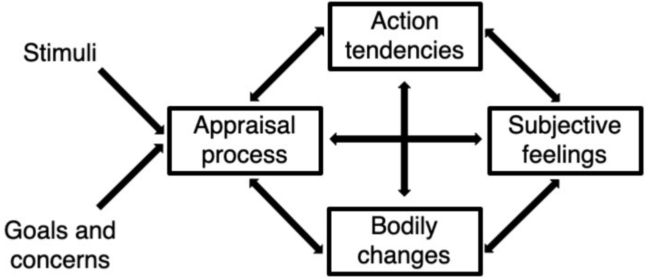


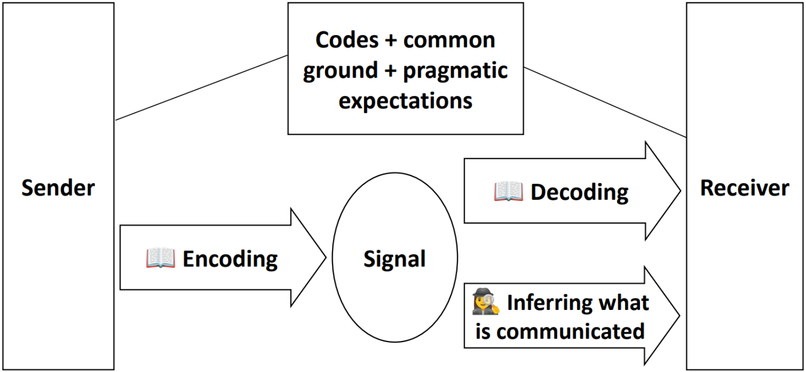
Figure: Emotional episodes are synchronized changes in four components (Scherer, 2022).

Detective analysis in pragmatics

What are the psychological mechanisms used to infer what is communicated?

Detective analysis in pragmatics

What are the psychological mechanisms used to infer what is communicated?



Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Raphaël Micheli categorizes a range of linguistic markers into three *emotion expression modes* (Micheli, 2013). The emotion can be:

Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Raphaël Micheli categorizes a range of linguistic markers into three *emotion expression modes* (Micheli, 2013). The emotion can be:

- *labeled* explicitly with an emotional term ("I am sad")

Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Raphaël Micheli categorizes a range of linguistic markers into three *emotion expression modes* (Micheli, 2013). The emotion can be:

- *labeled* explicitly with an emotional term ("I am sad")
- *shown* with utterance features such as interjections and punctuations ("Ah! That's great!")

Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Raphaël Micheli categorizes a range of linguistic markers into three *emotion expression modes* (Micheli, 2013). The emotion can be:

- *labeled* explicitly with an emotional term ("I am sad")
- *shown* with utterance features such as interjections and punctuations ("Ah! That's great!")
- *suggested* with the description of a situation which generally, in a given sociocultural context, leads to an emotion ("She gave me a gift")

Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Raphaël Micheli categorizes a range of linguistic markers into three *emotion expression modes* (Micheli, 2013). The emotion can be:

- *labeled* explicitly with an emotional term ("I am <u>sad</u>")
- *shown* with utterance features such as interjections and punctuations ("<u>Ah!</u> That's great<u>!</u>")
- *suggested* with the description of a situation which generally, in a given sociocultural context, leads to an emotion ("<u>She gave me a gift</u>")

Emotion expression modes

Which verbal signs are used to infer expressed emotions?

Raphaël Micheli categorizes a range of linguistic markers into three *emotion expression modes* (Micheli, 2013). The emotion can be:

- *labeled* explicitly with an emotional term ("I am sad")
- *shown* with utterance features such as interjections and punctuations ("Ah! That's great!")
- *suggested* with the description of a situation which generally, in a given sociocultural context, leads to an emotion ("She gave me a gift")

→ Different emotion expression modes are more or less difficult to interpret

Construction of an emotion dataset

Available at hf.co/datasets/gustavecortal/FrenchEmotionalNarratives

G. Cortal, A. Finkel, P. Paroubek, L. Ye. Emotion Recognition based on Psychological Components in Guided Narratives for Emotion Regulation. *SIGHUM, EACL 2023*.

French emotional narratives based on components

Goal: A more comprehensive understanding of emotional events

French emotional narratives based on components

Goal: A more comprehensive understanding of emotional events

Component	**Answer**
`Behavior`	I'm giving a lecture on a Friday morning at 8:30. A student goes out and comes back a few moments later with a coffee in his hand.
`Feeling`	My heart is beating fast, and I freeze, waiting to know how to act.
`Thinking`	I think this student is disrupting my class.
`Territory`	The student attacks my ability to be respected in class.

Chosen emotion: `anger` (possible choices: `anger`, `fear`, `joy`, `sadness`)

More than 1,000 narratives were collected during emotion regulation sessions

Training language models for emotion analysis

G. Cortal, A. Finkel, P. Paroubek, L. Ye. Emotion Recognition based on Psychological Components in Guided Narratives for Emotion Regulation. *SIGHUM, EACL 2023*

G. Cortal. Sequence-to-Sequence Language Models for Character and Emotion Detection in Dream Narratives. *LREC-COLING 2024*

Discrete emotion detection based on components

	Logistic Regression			CamemBERT		
Component	Precision	Recall	F_1	Precision	Recall	F_1
All	71.2 (2.6)	69.1 (2.2)	67.8 (2.3)	**85.1**	**84.8**	**84.7**
Without `behavior`	77.4 (2.3)	75.8 (2.4)	74.5 (2.6)	80.3	79.8	79.7
Without `feeling`	64.3 (1.9)	61.5 (1.2)	61.3 (2.2)	81.6	79.8	79.9
Without `thinking`	70.9 (1.8)	69.1 (2.0)	68.3 (2.2)	79.6	78.5	78.7
Without `territory`	64.3 (4.1)	64.5 (2.4)	62.3 (2.8)	78.7	78.5	78.6
Only `behavior`	52.1 (3.5)	54.6 (2.9)	51.7 (2.9)	68.4	67.1	66.6
Only `feeling`	69.6 (1.5)	68.9 (2.1)	68.4 (2.0)	67.8	68.4	67.7
Only `thinking`	50.1 (3.4)	53.8 (2.3)	50.6 (2.7)	70.5	70.1	70.1
Only `territory`	68.2 (1.8)	66.8 (2.2)	66.6 (2.3)	71.4	68.4	68.9

Discrete emotion detection based on components

	Logistic Regression			CamemBERT		
Component	Precision	Recall	F_1	Precision	Recall	F_1
All	71.2 (2.6)	69.1 (2.2)	67.8 (2.3)	**85.1**	**84.8**	**84.7**
Without `behavior`	77.4 (2.3)	75.8 (2.4)	74.5 (2.6)	80.3	79.8	79.7
Without `feeling`	64.3 (1.9)	61.5 (1.2)	61.3 (2.2)	81.6	79.8	79.9
Without `thinking`	70.9 (1.8)	69.1 (2.0)	68.3 (2.2)	79.6	78.5	78.7
Without `territory`	64.3 (4.1)	64.5 (2.4)	62.3 (2.8)	78.7	78.5	78.6
Only `behavior`	52.1 (3.5)	54.6 (2.9)	51.7 (2.9)	68.4	67.1	66.6
Only `feeling`	69.6 (1.5)	68.9 (2.1)	68.4 (2.0)	67.8	68.4	67.7
Only `thinking`	50.1 (3.4)	53.8 (2.3)	50.6 (2.7)	70.5	70.1	70.1
Only `territory`	68.2 (1.8)	66.8 (2.2)	66.6 (2.3)	71.4	68.4	68.9

→ Each component improves prediction performance, the best results are achieved by jointly considering all components

Discrete emotion detection based on components

Component	Logistic Regression Precision	Logistic Regression Recall	Logistic Regression F_1	CamemBERT Precision	CamemBERT Recall	CamemBERT F_1
All	71.2 (2.6)	69.1 (2.2)	67.8 (2.3)	**85.1**	**84.8**	**84.7**
Without `behavior`	77.4 (2.3)	75.8 (2.4)	74.5 (2.6)	80.3	79.8	79.7
Without `feeling`	64.3 (1.9)	61.5 (1.2)	61.3 (2.2)	81.6	79.8	79.9
Without `thinking`	70.9 (1.8)	69.1 (2.0)	68.3 (2.2)	79.6	78.5	78.7
Without `territory`	64.3 (4.1)	64.5 (2.4)	62.3 (2.8)	78.7	78.5	78.6
Only `behavior`	52.1 (3.5)	54.6 (2.9)	51.7 (2.9)	68.4	67.1	66.6
Only `feeling`	69.6 (1.5)	68.9 (2.1)	68.4 (2.0)	67.8	68.4	67.7
Only `thinking`	50.1 (3.4)	53.8 (2.3)	50.6 (2.7)	70.5	70.1	70.1
Only `territory`	68.2 (1.8)	66.8 (2.2)	66.6 (2.3)	71.4	68.4	68.9

→ Each component improves prediction performance, the best results are achieved by jointly considering all components

→ Some components benefit from contextual understanding and world knowledge (*e.g.*, `behavior` and `thinking`)

Quantitative analysis of dream narratives

Need other datasets with narrative structure, emotional content, and available for research

Quantitative analysis of dream narratives

Need other datasets with narrative structure, emotional content, and available for research

Quantitative dream analysis most relies on dream databases (*e.g.*, DreamBank composed of 27,000 dreams) and annotation schemes (*e.g.*, Hall and Van de Castle system)

Quantitative analysis of dream narratives

Need other datasets with narrative structure, emotional content, and available for research

Quantitative dream analysis most relies on dream databases (*e.g.*, DreamBank composed of 27,000 dreams) and annotation schemes (*e.g.*, Hall and Van de Castle system)

The annotation process is complex and costly

Quantitative analysis of dream narratives

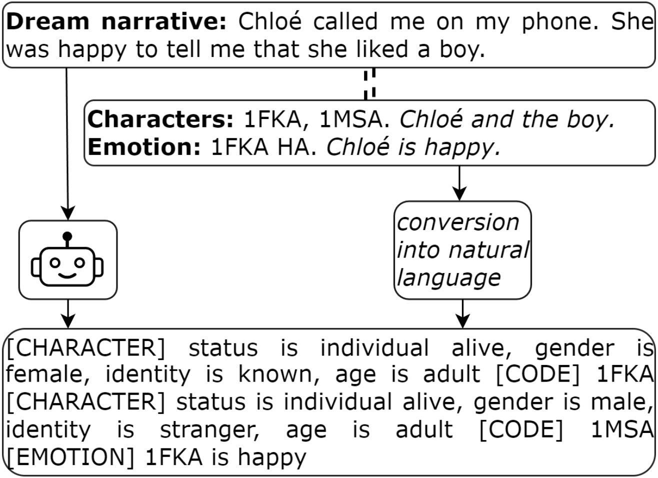
Need other datasets with narrative structure, emotional content, and available for research

Quantitative dream analysis most relies on dream databases (*e.g.*, DreamBank composed of 27,000 dreams) and annotation schemes (*e.g.*, Hall and Van de Castle system)

The annotation process is complex and costly

How to automate the annotation process?

Character and emotion detection in dream narratives



Results

Baseline is LaMini-Flan-T5 finetuned on 1823 dream narratives

Model	Status	Gender	Identity	Age	Character	Emotion
Baseline	82.87	78.02	76.17	86.21	64.74	75.13
$\texttt{No}_{\texttt{semantics}}$	71.37	56.54*	61.0	90.51	41.79*	75.79
$\texttt{No}_{\texttt{names}}$	80.66*	74.32**	74.2	83.95*	60.93**	73.04*
$\texttt{Size}_{\texttt{small}}$	78.35**	72.13**	70.25**	81.66**	56.79**	70.15**
$\texttt{Size}_{\texttt{large}}$	84.51*	80.3**	78.63**	87.29	67.63**	74.71
$\texttt{First}_{\texttt{group}}$	82.33	77.71	74.86	85.61	63.71	71.94
$\texttt{First}_{\texttt{individual}}$	80.59**	76.14	74.22*	83.87**	62.67	67.32
$\texttt{First}_{\texttt{emotion}}$	83.92	78.74	77.06	87.63	64.97	72.03
$\texttt{Conversion}_{\texttt{comma}}$	84.02**	79.84**	77.67**	87.08*	66.69**	73.68
$\texttt{Conversion}_{\texttt{marker}}$	82.39	78.45	76.53	86.09	65.44	74.36
Cross-validation	86.28	81.9	79.51	89.52	68.64	76.18

Table: F_1-scores for character and emotion detection. Significant differences from baseline: ** ($p < 0.01$), * ($p < 0.05$).

Results

Baseline is LaMini-Flan-T5 finetuned on 1823 dream narratives

Model	Status	Gender	Identity	Age	Character	Emotion
Baseline	82.87	78.02	76.17	86.21	64.74	75.13
$\texttt{No}_{\texttt{semantics}}$	71.37	56.54*	61.0	90.51	41.79*	75.79
$\texttt{No}_{\texttt{names}}$	80.66*	74.32**	74.2	83.95*	60.93**	73.04*
$\texttt{Size}_{\texttt{small}}$	78.35**	72.13**	70.25**	81.66**	56.79**	70.15**
$\texttt{Size}_{\texttt{large}}$	84.51*	80.3**	78.63**	87.29	67.63**	74.71
$\texttt{First}_{\texttt{group}}$	82.33	77.71	74.86	85.61	63.71	71.94
$\texttt{First}_{\texttt{individual}}$	80.59**	76.14	74.22*	83.87**	62.67	67.32
$\texttt{First}_{\texttt{emotion}}$	83.92	78.74	77.06	87.63	64.97	72.03
$\texttt{Conversion}_{\texttt{comma}}$	84.02**	79.84**	77.67**	87.08*	66.69**	73.68
$\texttt{Conversion}_{\texttt{marker}}$	82.39	78.45	76.53	86.09	65.44	74.36
Cross-validation	86.28	81.9	79.51	89.52	68.64	76.18

Table: F_1-scores for character and emotion detection. Significant differences from baseline: ** ($p < 0.01$), * ($p < 0.05$).

→ Language models can effectively address character and emotion detection in dream narratives

Results

`StableBeluga`$_i$ is a 7B model with in-context learning using *i* examples

Model	**Status**	**Gender**	**Identity**	**Age**	**Character**	**Emotion**
Baseline	82.87	78.02	76.17	86.21	64.74	75.13
StableBeluga$_1$	43.95**	39.76**	31.25**	56.16**	15.65**	-
StableBeluga$_3$	52.44**	46.49**	38.46**	63.88**	21.06**	-
StableBeluga$_5$	55.89**	46.29**	42.61**	63.73**	24.86**	-

Table: F_1-scores for character and emotion detection. Significant differences from baseline: ** ($p < 0.01$), * ($p < 0.05$).

Results

`StableBeluga`$_i$ is a 7B model with in-context learning using i examples

Model	Status	Gender	Identity	Age	Character	Emotion
`Baseline`	82.87	78.02	76.17	86.21	64.74	75.13
`StableBeluga`$_1$	43.95**	39.76**	31.25**	56.16**	15.65**	-
`StableBeluga`$_3$	52.44**	46.49**	38.46**	63.88**	21.06**	-
`StableBeluga`$_5$	55.89**	46.29**	42.61**	63.73**	24.86**	-

Table: F_1-scores for character and emotion detection. Significant differences from `baseline`: ** ($p < 0.01$), * ($p < 0.05$).

→ Compared to `StableBeluga`, our supervised models perform better while having 28 times fewer parameters (248M vs 7B)

Case study on the war veteran

Group	Category	% Vet	% Total	Δ
Identity	known*	24.9	51.6	-26.7
	prominent	1.9	2.5	-0.6
	occupational*	22.4	8.0	14.4
	ethnic*	4.1	0.9	3.1
	unknown*	46.8	37.0	9.8
Gender	male*	56.2	43.0	13.1
	female*	24.1	33.1	-9.0
	joint	10.9	12.2	-1.3
	undefined	7.9	8.7	-0.9

Table: Identity and gender proportions for the veteran (n=566 narratives) versus other dreamers. Δ shows the difference in percentage points. * indicates significant difference ($p < 0.05$).

Case study on the war veteran

Group	Category	% Vet	% Total	Δ
Identity	known*	24.9	51.6	-26.7
	prominent	1.9	2.5	-0.6
	occupational*	22.4	8.0	14.4
	ethnic*	4.1	0.9	3.1
	unknown*	46.8	37.0	9.8
Gender	male*	56.2	43.0	13.1
	female*	24.1	33.1	-9.0
	joint	10.9	12.2	-1.3
	undefined	7.9	8.7	-0.9

Table: Identity and gender proportions for the veteran (n=566 narratives) versus other dreamers. Δ shows the difference in percentage points. * indicates significant difference ($p < 0.05$).

→ The war veteran dreams more about *occupational*, *ethnic*, and *unknown* identities compared to other dreamers.

Formalization of style in personal narratives

G. Cortal and A. Finkel. Formalizing Style in Personal Narratives. *EMNLP 2025*.

Introduction

How is subjective experience communicated in narratives?

Introduction

How is subjective experience communicated in narratives?

We use narratives to express our representations of reality and make sense of the world (Bruner, 1990)

Introduction

How is subjective experience communicated in narratives?

We use narratives to express our representations of reality and make sense of the world (Bruner, 1990)

In everyday usage, style refers to a distinctive manner of expression

Introduction

How is subjective experience communicated in narratives?

We use narratives to express our representations of reality and make sense of the world (Bruner, 1990)

In everyday usage, style refers to a distinctive manner of expression

We use style as a proxy to study how subjective experience is linguistically communicated

Introduction

How is subjective experience communicated in narratives?

We use narratives to express our representations of reality and make sense of the world (Bruner, 1990)

In everyday usage, style refers to a distinctive manner of expression

We use style as a proxy to study how subjective experience is linguistically communicated

We narrow the general definition of style: *a distinctive manner of communicating subjective experience in narratives*

Contributions

How to give an operational definition of style?

Contributions

How to give an operational definition of style?

Hypothesis: An individual uses some redundant choices of features that characterize its style

Contributions

How to give an operational definition of style?

Hypothesis: An individual uses some redundant choices of features that characterize its style

1. A sequence-based framework defining style as patterns in sequences of linguistic choices

Contributions

How to give an operational definition of style?

Hypothesis: An individual uses some redundant choices of features that characterize its style

1. A sequence-based framework defining style as patterns in sequences of linguistic choices
2. A methodology for identifying patterns using sequence analysis

Contributions

How to give an operational definition of style?

Hypothesis: An individual uses some redundant choices of features that characterize its style

1. A sequence-based framework defining style as patterns in sequences of linguistic choices
2. A methodology for identifying patterns using sequence analysis
3. A case study on dream narratives

Categorizing linguistic features

What linguistic features encode subjective experience?

Categorizing linguistic features

What linguistic features encode subjective experience?

According to systemic functional linguistics, language represents experience through *processes*, *participants* and *circumstances* (Halliday et al., 2014)

Categorizing linguistic features

What linguistic features encode subjective experience?

According to systemic functional linguistics, language represents experience through *processes*, *participants* and *circumstances* (Halliday et al., 2014)

Processes	**Examples**
`Action`: actions and events in the physical world.	[He]$_{\text{Actor}}$ [**takes**]$_{\text{Action}}$ [the valuable]$_{\text{Affected}}$ [Members of my cult]$_{\text{Actor}}$ [**have made**]$_{\text{Action}}$ [1500 euros]$_{\text{Result}}$ [I]$_{\text{Actor}}$ [**give**]$_{\text{Action}}$ [her]$_{\text{Recipient}}$ [a chance]$_{\text{Range}}$
`Mental`: internal experiences such as thoughts, perceptions, and feelings.	[We]$_{\text{Senser}}$ [**believe**]$_{\text{Mental}}$ [women are the leaders of change]$_{\text{Phenomenon}}$ [The moon]$_{\text{Senser}}$ [**sees**]$_{\text{Mental}}$ [the earth]$_{\text{Phenomenon}}$ [He]$_{\text{Senser}}$ [**disliked**]$_{\text{Mental}}$ [Gilbert's writing]$_{\text{Phenomenon}}$
`Verbal`: acts of communication.	[David]$_{\text{Sayer}}$ [**said**]$_{\text{Verbal}}$ ["the corrupt, criminals and money launderers"]$_{\text{Verbiage}}$
`State`: states of being, having, or existence.	There [**was**]$_{\text{Existential}}$ [a swimming pool]$_{\text{Existent}}$ [John]$_{\text{Carrier}}$ [**is**]$_{\text{State}}$ [an interesting teacher]$_{\text{Attribute}}$ [Hadrian's Wall]$_{\text{Possessor}}$ [**has**]$_{\text{State}}$ [something for everyone]$_{\text{Possessed}}$

Pipeline for our sequence-based framework

Clause	**Process (symbol)**	**Participants**
I wake in a dark room	`Action` (**a**)	`Actor`
I feel a cold wind	`Mental` (**m**)	`Senser`, `Phenomenon`
I tell myself to move	`Verbal` (**v**)	`Sayer`, `Recipient`

Sequence: *amv* | **Substrings:** {am, mv}

Pipeline for our sequence-based framework

Clause	Process (symbol)	Participants
I wake in a dark room	`Action` (**a**)	`Actor`
I feel a cold wind	`Mental` (**m**)	`Senser`, `Phenomenon`
I tell myself to move	`Verbal` (**v**)	`Sayer`, `Recipient`

Sequence: *amv* | **Substrings:** {am, mv}

1. We first segment *"I wake in a dark room. I feel a cold wind. I tell myself to move."* into clauses

Pipeline for our sequence-based framework

Clause	Process (symbol)	Participants
I wake in a dark room	`Action` (**a**)	`Actor`
I feel a cold wind	`Mental` (**m**)	`Senser`, `Phenomenon`
I tell myself to move	`Verbal` (**v**)	`Sayer`, `Recipient`

Sequence: *amv* | **Substrings:** {am, mv}

1. We first segment *"I wake in a dark room. I feel a cold wind. I tell myself to move."* into clauses
2. Identify features (*e.g.*, processes and participants) for each clause using in-context learning with large language models

Pipeline for our sequence-based framework

Clause	Process (symbol)	Participants
I wake in a dark room	`Action` (**a**)	`Actor`
I feel a cold wind	`Mental` (**m**)	`Senser`, `Phenomenon`
I tell myself to move	`Verbal` (**v**)	`Sayer`, `Recipient`

Sequence: *amv* | **Substrings:** {am, mv}

1. We first segment *"I wake in a dark room. I feel a cold wind. I tell myself to move."* into clauses
2. Identify features (*e.g.*, processes and participants) for each clause using in-context learning with large language models
3. Each narrative is mapped to a symbolic sequence using an alphabet based on identified features

Pipeline for our sequence-based framework

Clause	**Process (symbol)**	**Participants**
I wake in a dark room	`Action` (**a**)	`Actor`
I feel a cold wind	`Mental` (**m**)	`Senser`, `Phenomenon`
I tell myself to move	`Verbal` (**v**)	`Sayer`, `Recipient`

Sequence: *amv* | **Substrings:** {am, mv}

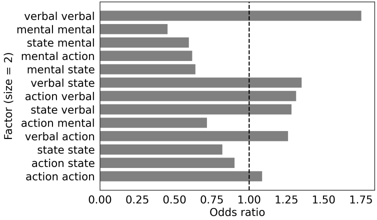
1. We first segment *"I wake in a dark room. I feel a cold wind. I tell myself to move."* into clauses
2. Identify features (*e.g.*, processes and participants) for each clause using in-context learning with large language models
3. Each narrative is mapped to a symbolic sequence using an alphabet based on identified features
4. We perform sequence analysis to identify patterns such as frequent substrings and representative sequences

Results on the war veteran

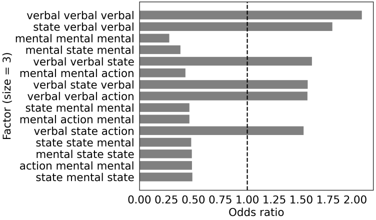
We compare the proportion of sequences containing a given substring

Results on the war veteran

We compare the proportion of sequences containing a given substring



(a) Size 2.

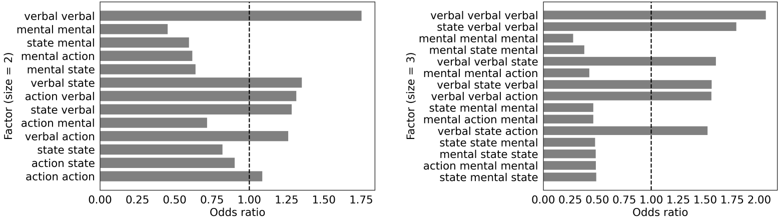


(b) Size 3.

Figure: Top substring odds ratio between the veteran and the norm

Results on the war veteran

We compare the proportion of sequences containing a given substring



(a) Size 2. (b) Size 3.

Figure: Top substring odds ratio between the veteran and the norm

We show a preference for the war veteran to remain in a verbal process, as indicated by substrings such as *verbal.verbal* and *verbal.verbal.verbal* with high odds ratios (respectively 2.00 and 1.75)

Results on the war veteran

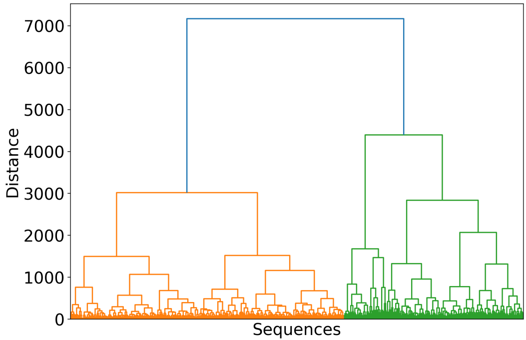


Figure: Dendrogram with Ward linkage and cosine similarity

Results on the war veteran

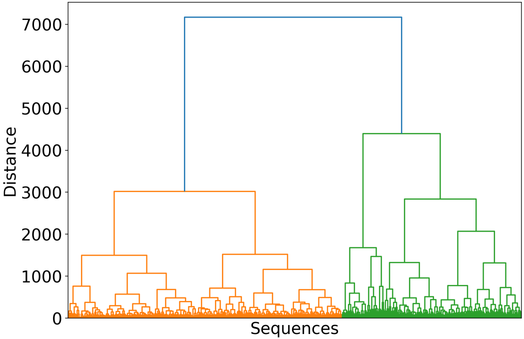


Figure: Dendrogram with Ward linkage and cosine similarity

Representative sequences: *savamasasaaamaaasavvvaaaaaaavssaaaaa* and *sssssavaavssvsavvvvsmasasaasasaamaamvmsss* with $a = action, m = mental, s = state, v = verbal$

Perspectives

- Authorship profiling

Perspectives

- Authorship profiling
- Style-conditioned narrative generation

Perspectives

- Authorship profiling
- Style-conditioned narrative generation
- Applying methods from complexity science and formal language theory

Conclusion

How to model subjective experience in personal narratives?

Conclusion

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science

Conclusion

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset

Conclusion

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset
- Training of language models for emotion analysis

Conclusion

How to model subjective experience in personal narratives?

- Definition of objectives and scope using cognitive science
- Construction of an emotion dataset
- Training of language models for emotion analysis
- Formalization of style in personal narratives

Appendix

Selected open-source projects

My research models are publicly hosted on Hugging Face and were trained using the Jean Zay supercomputer.

Selected open-source projects

My research models are publicly hosted on Hugging Face and were trained using the Jean Zay supercomputer.

A repo for lightweight preference optimization using LoRA and ORPO.

Selected open-source projects

My research models are publicly hosted on Hugging Face and were trained using the Jean Zay supercomputer.

A repo for lightweight preference optimization using LoRA and ORPO.

Piaget, a model fine-tuned for psychological reasoning, and Beck, a model aligned with psychotherapeutic preferences.

Selected open-source projects

My research models are publicly hosted on Hugging Face and were trained using the Jean Zay supercomputer.

A repo for lightweight preference optimization using LoRA and ORPO.

Piaget, a model fine-tuned for psychological reasoning, and Beck, a model aligned with psychotherapeutic preferences.

PsychologicalReasoning-15k, open psychological and philosophical reasoning traces.

Selected open-source projects

My research models are publicly hosted on Hugging Face and were trained using the Jean Zay supercomputer.

A repo for lightweight preference optimization using LoRA and ORPO.

Piaget, a model fine-tuned for psychological reasoning, and Beck, a model aligned with psychotherapeutic preferences.

PsychologicalReasoning-15k, open psychological and philosophical reasoning traces.

Oneirogen, a model for dream generation, and Dream-T5, a model for emotion and character prediction in dream narratives.

Selected research papers

Constant Bonard and Gustave Cortal (2024). "Improving Language Models for Emotion Analysis: Insights from Cognitive Science". In: *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*. Ed. by Tatsuki Kuribayashi et al. Bangkok, Thailand: Association for Computational Linguistics, pp. 264–277. DOI: 10.18653/v1/2024.cmcl-1.23

Gustave Cortal, Alain Finkel, et al. (2023). "Emotion Recognition Based on Psychological Components in Guided Narratives for Emotion Regulation". In: *Proceedings of the 7th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Ed. by Stefania Degaetano-Ortlieb et al. Dubrovnik, Croatia: Association for Computational Linguistics, pp. 72–81. DOI: 10.18653/v1/2023.latechclfl-1.8

Gustave Cortal (2024). "Sequence-to-Sequence Language Models for Character and Emotion Detection in Dream Narratives". In: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. Ed. by Nicoletta Calzolari et al. Torino, Italia: ELRA and ICCL, pp. 14717–14728

Gustave Cortal and Alain Finkel (2025). "Formalizing Style in Personal Narratives". In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. Ed. by Christos Christodoulopoulos et al. Suzhou, China: Association for Computational Linguistics, pp. 7322–7337. ISBN: 979-8-89176-332-6

References

References I

Bonard, Constant and Gustave Cortal (2024). "Improving Language Models for Emotion Analysis: Insights from Cognitive Science". In: *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*. Ed. by Tatsuki Kuribayashi et al. Bangkok, Thailand: Association for Computational Linguistics, pp. 264–277. DOI: 10.18653/v1/2024.cmcl-1.23.

Bruner, Jerome (1990). *Acts of Meaning*. Acts of Meaning. Cambridge, MA, US: Harvard University Press. ISBN: 978-0-674-00360-6.

Cortal, Gustave (2024). "Sequence-to-Sequence Language Models for Character and Emotion Detection in Dream Narratives". In: *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. Ed. by Nicoletta Calzolari et al. Torino, Italia: ELRA and ICCL, pp. 14717–14728.

References II

Cortal, Gustave and Alain Finkel (2025). "Formalizing Style in Personal Narratives". In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. Ed. by Christos Christodoulopoulos et al. Suzhou, China: Association for Computational Linguistics, pp. 7322–7337. ISBN: 979-8-89176-332-6.

Cortal, Gustave, Alain Finkel, et al. (2023). "Emotion Recognition Based on Psychological Components in Guided Narratives for Emotion Regulation". In: *Proceedings of the 7th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Ed. by Stefania Degaetano-Ortlieb et al. Dubrovnik, Croatia: Association for Computational Linguistics, pp. 72–81. DOI: `10.18653/v1/2023.latechclfl-1.8`.

Halliday, M.A.K. et al. (2014). *An Introduction to Functional Grammar*. 0th ed. Routledge. ISBN: 978-1-4441-1908-4. DOI: `10.4324/9780203783771`.

References III

Micheli, Raphaël (2013). "Esquisse d'une typologie des différents modes de sémiotisation verbale de l'émotion". In: *Semen. Revue de sémio-linguistique des textes et discours* 35. ISSN: 0761-2990. DOI: 10.4000/semen.9795.

Scherer, Klaus R. (2022). "Theory Convergence in Emotion Science Is Timely and Realistic". In: *Cognition and Emotion* 36.2, pp. 154–170. ISSN: 0269-9931. DOI: 10.1080/02699931.2021.1973378.