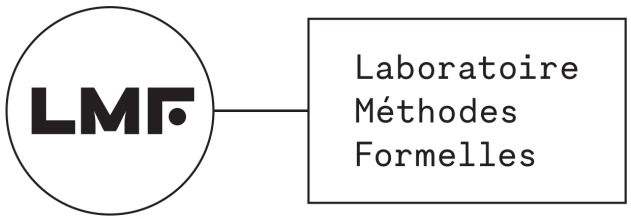
Formalizing Style in Personal Narratives

Gustave Cortal and Alain Finkel



Introduction

Research question: How is subjective experience communicated in narratives?

We use narratives to express our representations of reality and make sense of the world

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

Problem: Style is an intuitive notion; we need an operational definition

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

Research task: Formalize style as *patterns of linguistic choices that encode subjective experience*

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

Categorizing Linguistic Features

Our categorization is grounded in *systemic functional linguistics*: language represents experience through *processes, participants and circumstances*

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]$_{\text{Possessed}}$

Table: Processes with their participants.

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
3. Each narrative is mapped to a symbolic sequence using an alphabet based on identified features

Case Study on Dream Narratives

We apply our framework to dream narratives as they possess a narrative structure and represent attempts to communicate subjective experience

We use DreamBank, a database of more than 27,000 narratives with 72 series of dreamers

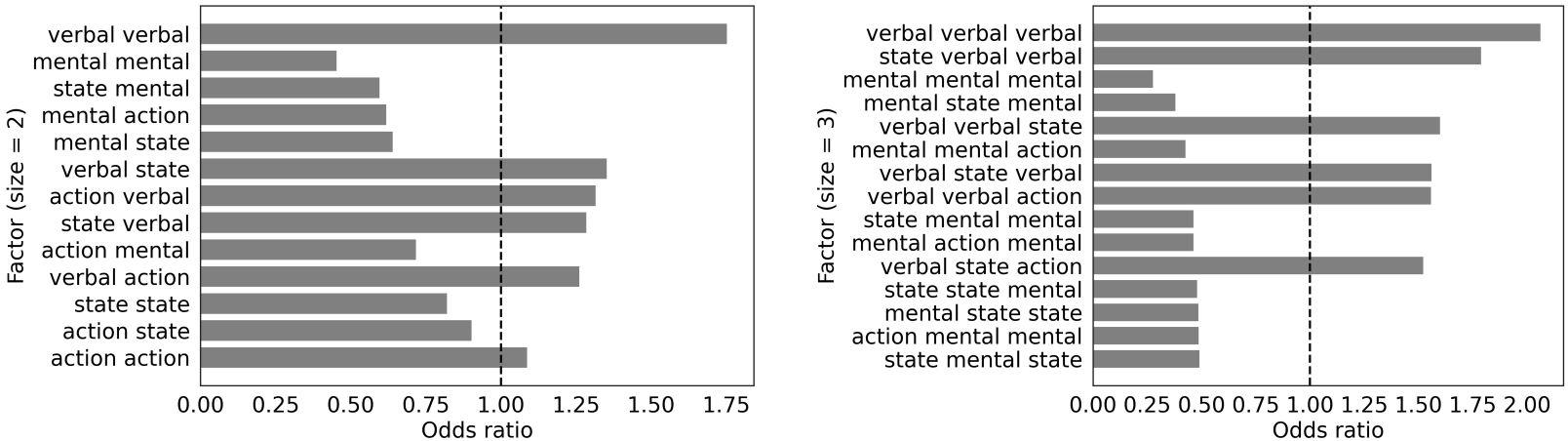
We analyze five series of dreamers: long-term blind dreamers (n=361), *ed* (a widower, n=139), *izzy* (a teenager, n=1091), *merri* (an artist, n=202), and *viet* (a Vietnam War veteran with PTSD, n=566)

We construct a *norm* (n=720) to compare how each series deviates from a hypothetical average dreamer

To identified features, we perform in-context learning with Llama 3 8B

Results on the Vietnam 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 *viet* 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 Vietnam War veteran

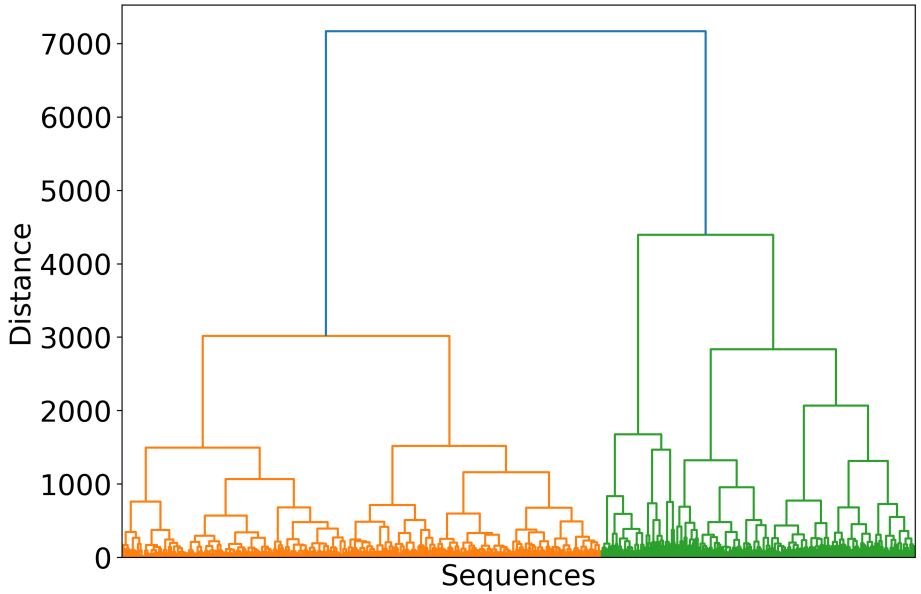


Figure: Dendrogram with Ward linkage and cosine similarity

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

Two templates: a highly action-oriented structure or a more varied structure alternating between state and action processes

Conclusion

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

Perspectives

1. Authorship profiling: identifying signature patterns (*e.g.*, distinctive substrings) that characterize an author's unique way of constructing narratives
2. Style-conditioned narrative generation: generating narratives from a sequence of linguistic features
3. Applying methods from complexity science and formal language theory: analyzing subsequences, using complexity measures to quantify redundancies, etc.