

Clustering Readers by Scanpath Similarity Reveals Variation in Processing Profiles

Michael Vrazitulis, Johan Hennert, Shravan Vasishth

IndiREAD Workshop, Saarbrücken, 2025-11-27

Outline

1 Motivation

2 Dataset

3 Clustering

4 Implications

1 Motivation

Two needs

1 Motivation

Two needs

Benchmark data
for battery of
controlled
experimental
designs

1 Motivation

Two needs

Benchmark data
for battery of
controlled
experimental
designs

Scanpath-level
analysis of
individual
differences
in reading

2 Dataset

GEPPU

*German Evaluation Benchmark for
Psycholinguistics from Potsdam University*

2 Dataset

GEPPU

げっぷ

burp

2 Dataset

GEPPU

月賦

*monthly
installment*

2 Dataset

GEPPU

*German Evaluation Benchmark for
Psycholinguistics from Potsdam University*

2 Dataset

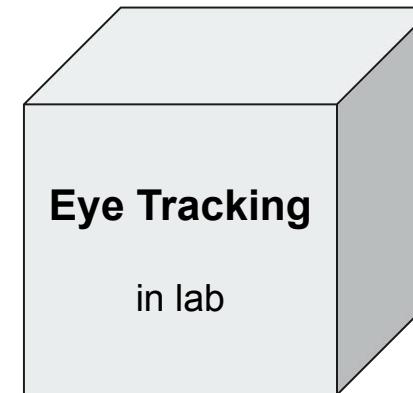
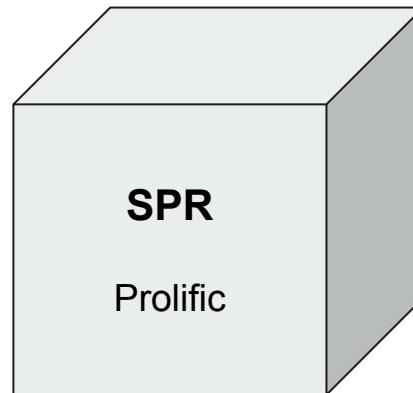
GEPPU

Like SAP
benchmark dataset
(Huang et al., 2024),
but for German

*German Evaluation Benchmark for
Psycholinguistics from Potsdam University*

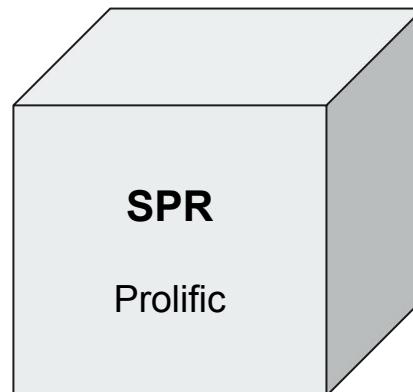
2 Dataset

GEPPU

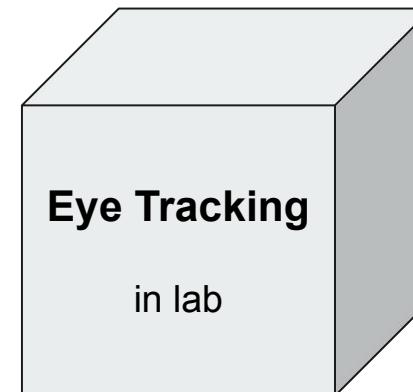


2 Dataset

GEPPU



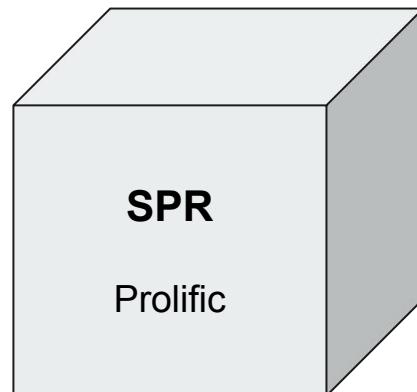
951 / 1100
done



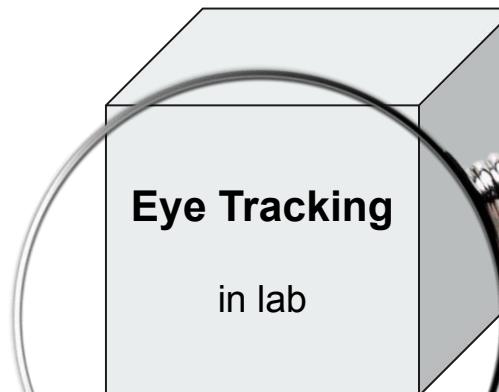
186 / ?
done

2 Dataset

GEPPU



951 / 1100
done



186 / ?
done

2 Dataset

SYAA (3×1): Syntax-Based Attachment Ambiguity

High-/Low-/Ambiguous-Attachment — closely replicating Logačev (2023)

SEAA (3×1): Semantics-Based Attachment Ambiguity

High-/Low-/Ambiguous-Attachment — German adaptation of Traxler et al. (2023)

2 Dataset

AGAT (2×2): Agreement Attraction in Grammatical Sentences

Singular-/Plural-Controller × Match/Mismatch — closely replicating Häussler (2009)

LOCO (2×2): Local Coherence

Coherent/Incoherent × Intervener/No-Intervener — closely replicating Paape & Vasishth (2016)

SBIN (2×2): Similarity-Based Interference

Subject-Cue [Yes/No] × Animacy-Cue [Yes/No] — closely replicating Schoknecht et al. (2025)

RCSO (2×2): Subject vs. Object Relative Clauses

Subject/Object × Double-/Single-Embedding — German adaptation of Hsiao & Gibson (2003)

2 Dataset

GPSD (2×2): Garden Paths From Subject-vs.-Direct-Object Ambiguity

Ambiguous/Unambiguous × S-O/O-S — closely replicating Meng & Bader (2000a)

GPSI (2×2): Garden Paths From Subject-vs.-Indirect-Object Ambiguity

Ambiguous/Unambiguous × Active/Passive — loosely replicating Meng & Bader (2000b)

GPCA (2×2): Garden Paths From Coordination Ambiguity

NP-/VP-Coordination × AP-/PP-Modifier — closely replicating Konieczny et al. (2000)

GPMI (2×2): Garden Paths From Modifier-vs.-Indirect-Object Ambiguity

Modifier/No-Modifier × Ambiguous/Unambiguous — closely replicating van Kampen (2001)

2 Dataset



GPSD (2x2): Garden Paths From Subject-vs.-Direct-Object Ambiguity

Ambiguous/Unambiguous \times S-O/O-S — closely replicating Meng & Bader (2000a)

GPSI (2x2): Garden Paths From Subject-vs.-Indirect-Object Ambiguity

Ambiguous/Unambiguous \times Active/Passive — loosely replicating Meng & Bader (2000b)

GPCA (2x2): Garden Paths From Coordination Ambiguity

NP-/VP-Coordination \times AP-/PP-Modifier — closely replicating Konieczny et al. (2000)

GPMI (2x2): Garden Paths From Modifier-vs.-Indirect-Object Ambiguity

Modifier/No-Modifier \times Ambiguous/Unambiguous — closely replicating van Kampen (2001)

2 Dataset

GPSD (2×2): Garden Paths From Subject-vs.-Direct-Object Ambiguity

- a. Alle wollten wissen, | welche Lokalpolitikerin | die amtierenden Minister | gelobt | hat, | als | die Pressekonferenz | begann.
All wanted to.know, | which.F.NOM/.ACC local.politician.F | the incumbent ministers | praised | has, | when | the press.conference | began.
- b. Alle wollten wissen, | welche Lokalpolitikerin | die amtierenden Minister | gelobt | haben, | als | die Pressekonferenz | begann.
All wanted to.know, | which.F.NOM/.ACC local.politician.F | the incumbent ministers | praised | have, | when | the press.conference | began.
- c. Alle wollten wissen, | welcher Lokalpolitiker | die amtierenden Minister | gelobt | hat, | als | die Pressekonferenz | begann.
All wanted to.know, | which.M.NOM local.politician.M | the incumbent ministers | praised | has, | when | the press.conference | began.
- d. Alle wollten wissen, | welchen Lokalpolitiker | die amtierenden Minister | gelobt | haben, | als | die Pressekonferenz | begann.
All wanted to.know, | which.M.ACC local.politician.M | the incumbent ministers | praised | have, | when | the press.conference | began.

2 Dataset

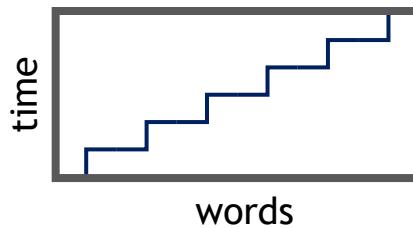
GPSD (2×2): Garden Paths From Subject-vs.-Direct-Object Ambiguity

- a. Alle wollten wissen, | welche Lokalpolitikerin | die amtierenden Minister | gelobt | hat, | als | die Pressekonferenz | begann.
All wanted to.know, | which.F.NOM/.ACC local.politician.F | the incumbent ministers | praised | has, | when | the press.conference | began.
- b. Alle wollten wissen, | welche Lokalpolitikerin | die amtierenden Minister | gelobt | haben, | als | die Pressekonferenz | begann.
All wanted to.know, | which.F.NOM/.ACC local.politician.F | the incumbent ministers | praised | have, | when | the press.conference | began.
- c. Alle wollten wissen, | welcher Lokalpolitiker | die amtierenden Minister | gelobt | hat, | als | die Pressekonferenz | begann.
All wanted to.know, | which.M.NOM local.politician.M | the incumbent ministers | praised | has, | when | the press.conference | began.
- d. Alle wollten wissen, | welchen Lokalpolitiker | die amtierenden Minister | gelobt | haben, | als | die Pressekonferenz | begann.
All wanted to.know, | which.M.ACC local.politician.M | the incumbent ministers | praised | have, | when | the press.conference | began.

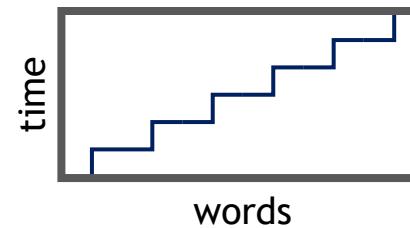
2 Dataset

Theoretical expectation (e.g., Frazier & Rayner, 1982)

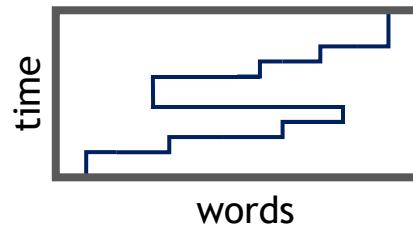
a.



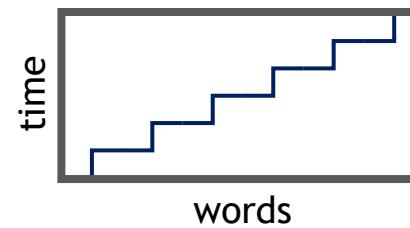
c.



b.



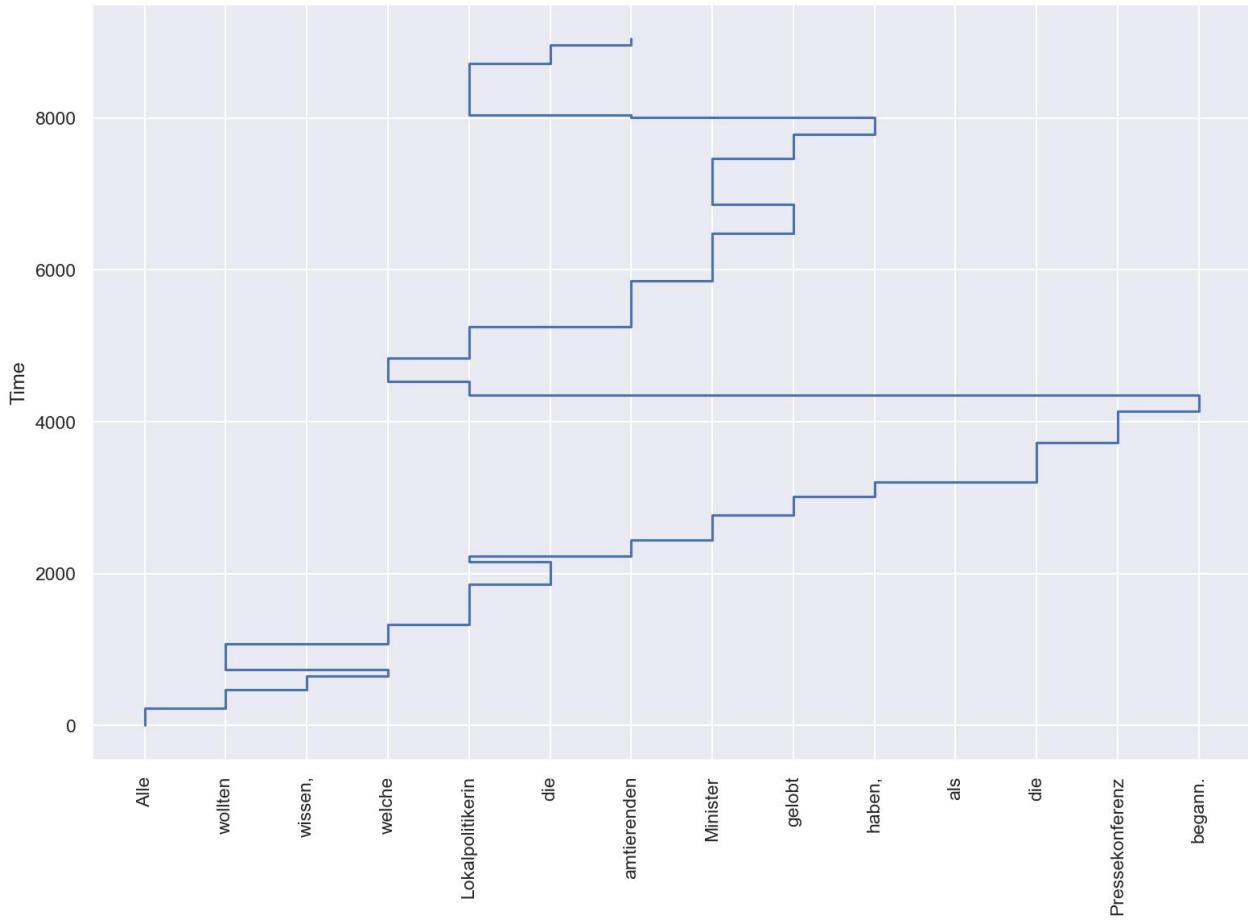
d.



2 Dataset

Empirical data
(subject S5):

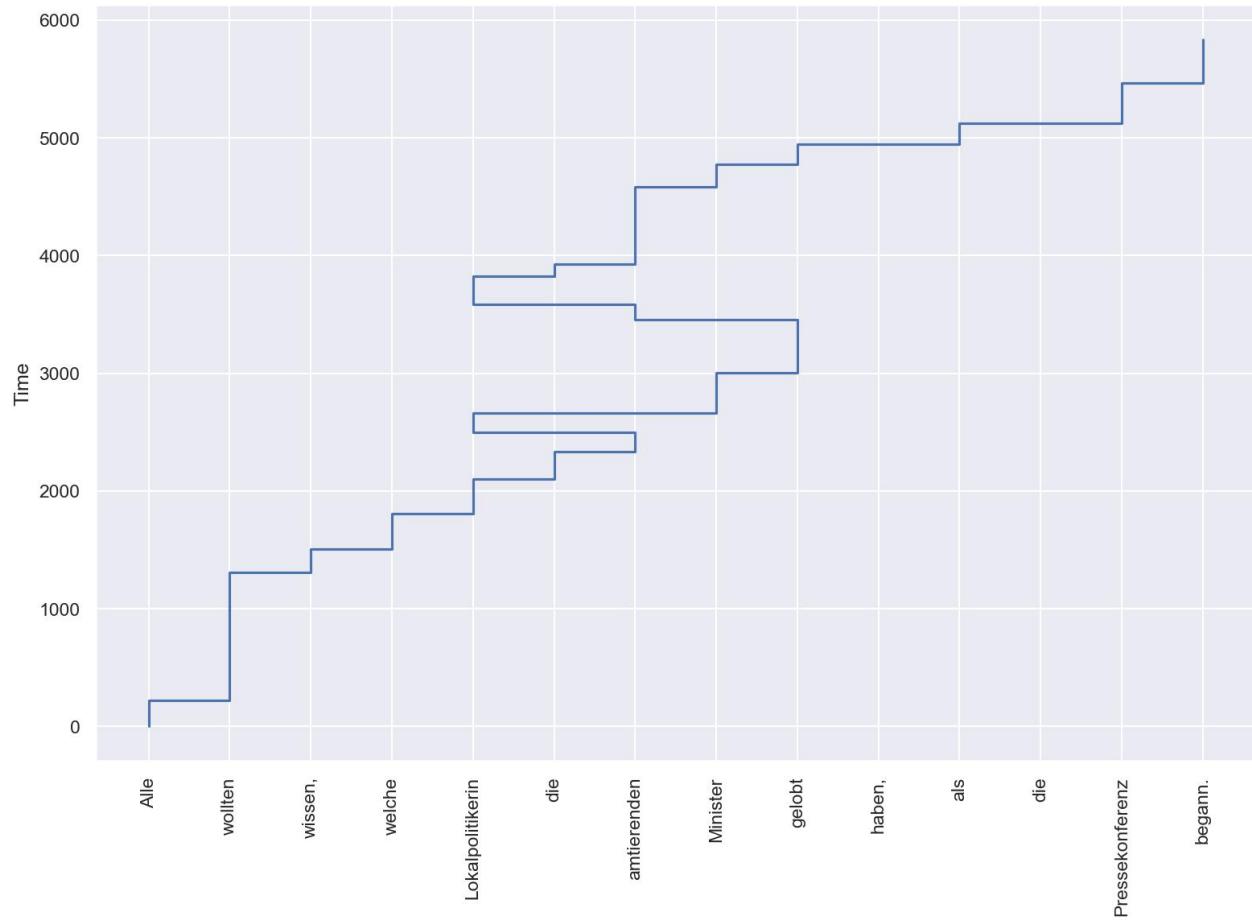
Critical condition b



2 Dataset

Empirical data
(subject **S17**):

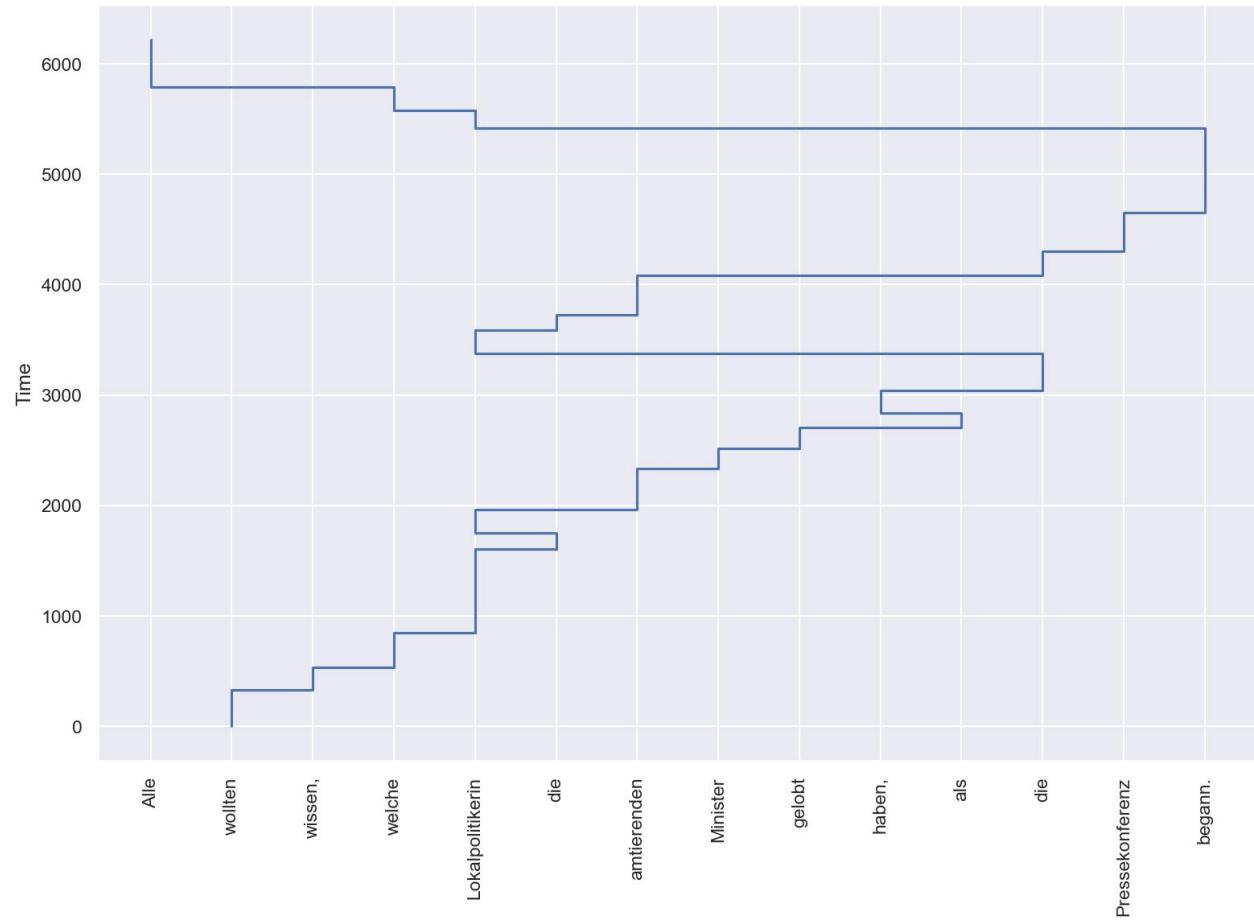
Critical condition b



2 Dataset

Empirical data
(subject **S18**):

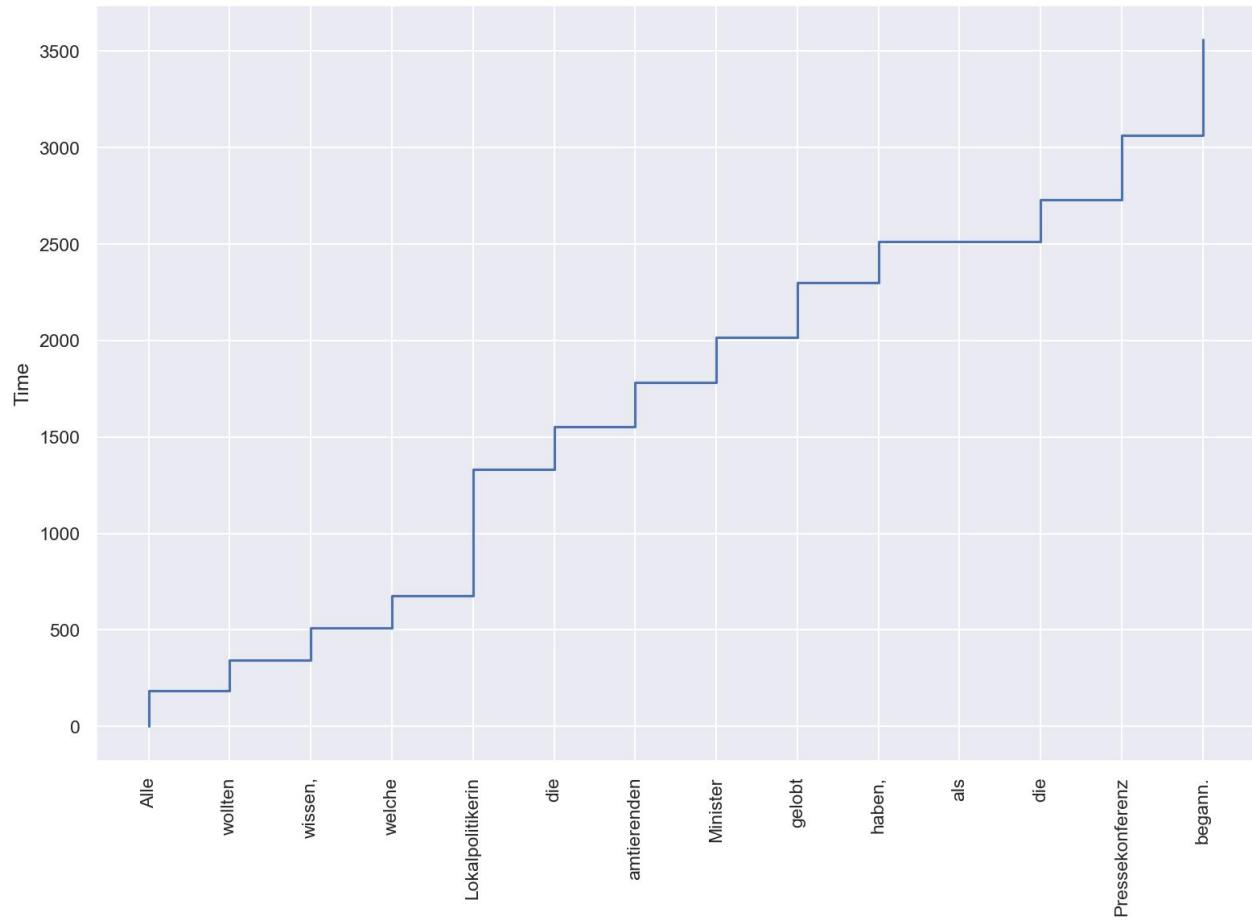
Critical condition b



2 Dataset

Empirical data
(subject **S41**):

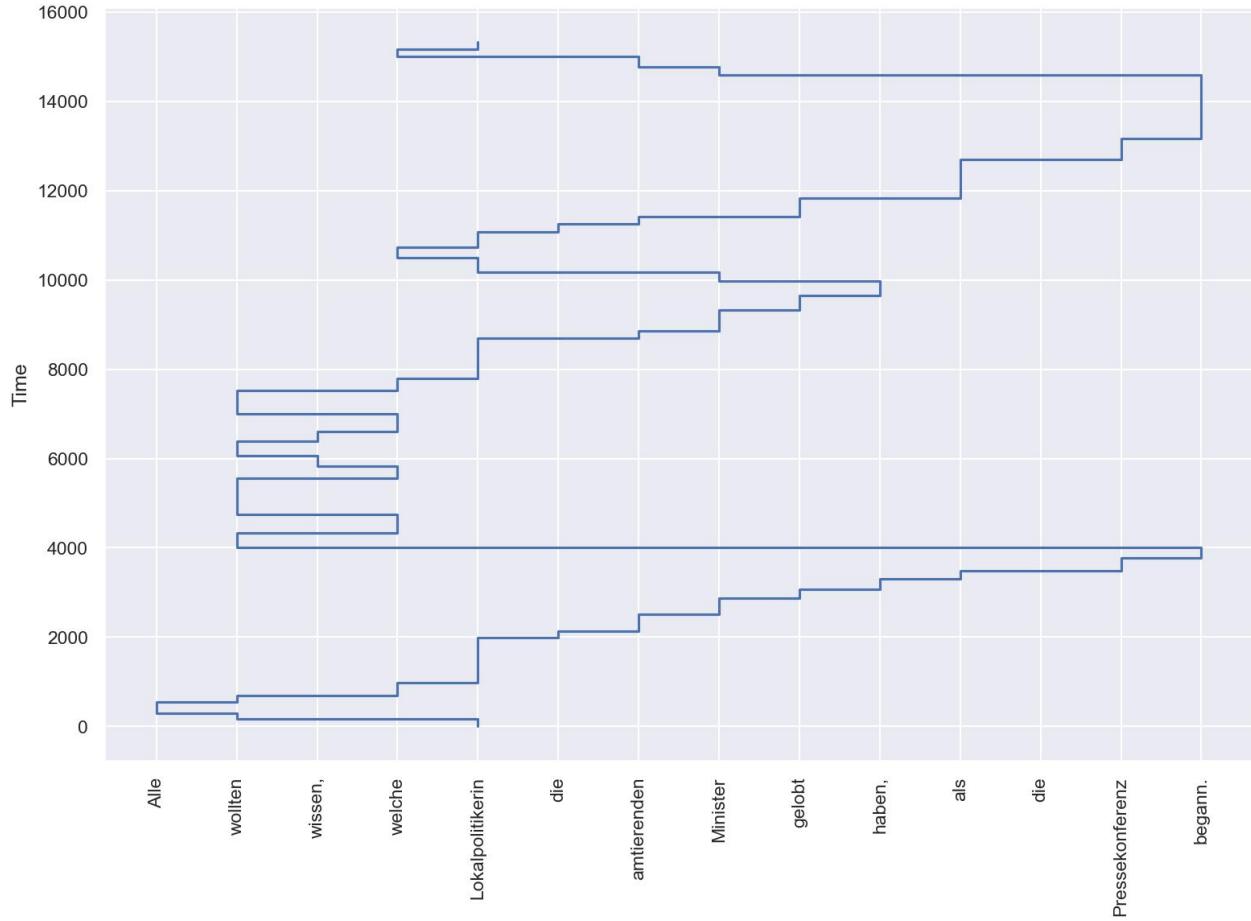
Critical condition b



2 Dataset

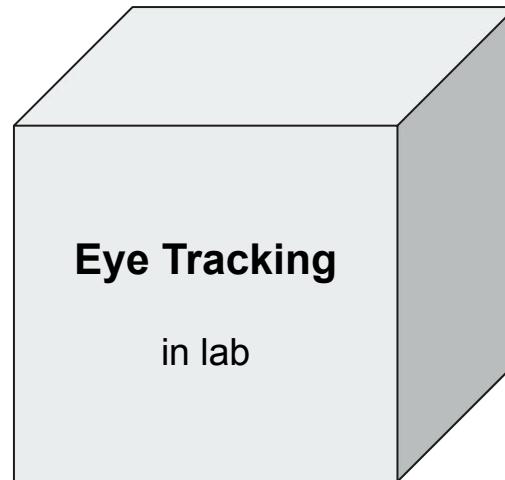
Empirical data
(subject **S149**):

Critical condition b



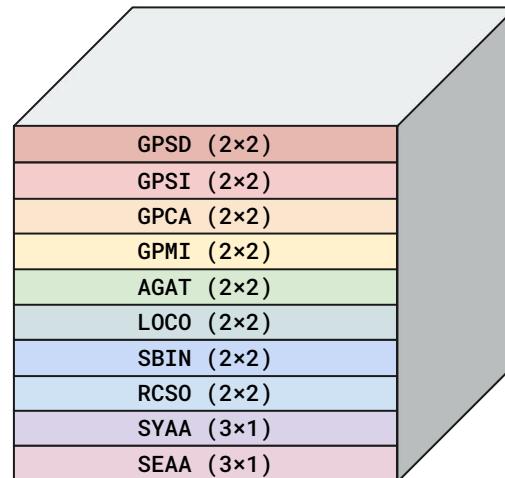
2 Dataset

GEPPU



2 Dataset

GEPPU



2 Dataset

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
a	a	a	a	a	a	a	a	a	a
b	b	b	b	b	b	b	b	b	b
c	c	c	c	c	c	c	c	c	c
d	d	d	d	d	d	d	d	d	

38 conditions

2 Dataset

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
....
....
....
....

3 data points per subject and condition
(Latin square)

3 Clustering

Quantify **spatio-temporal similarity** between scanpaths:

Scasim (von der Malsburg & Vasishth, 2011; 2013)

- Based on Needleman–Wunsch algorithm (edit distance)
- But also sensitive to fixation duration and visual angle

3 Clustering

R package: [github . com/tmalsburg/scanpath](https://github.com/tmalsburg/scanpath)

Quantify **spatio-temporal similarity** between scanpaths:

Scasim (von der Malsburg & Vasishth, 2011; 2013)

- Based on Needleman–Wunsch algorithm (edit distance)
- But also sensitive to fixation duration and visual angle

3 Clustering

Clustering scanpaths by **Scasim** can reveal distinct processing patterns

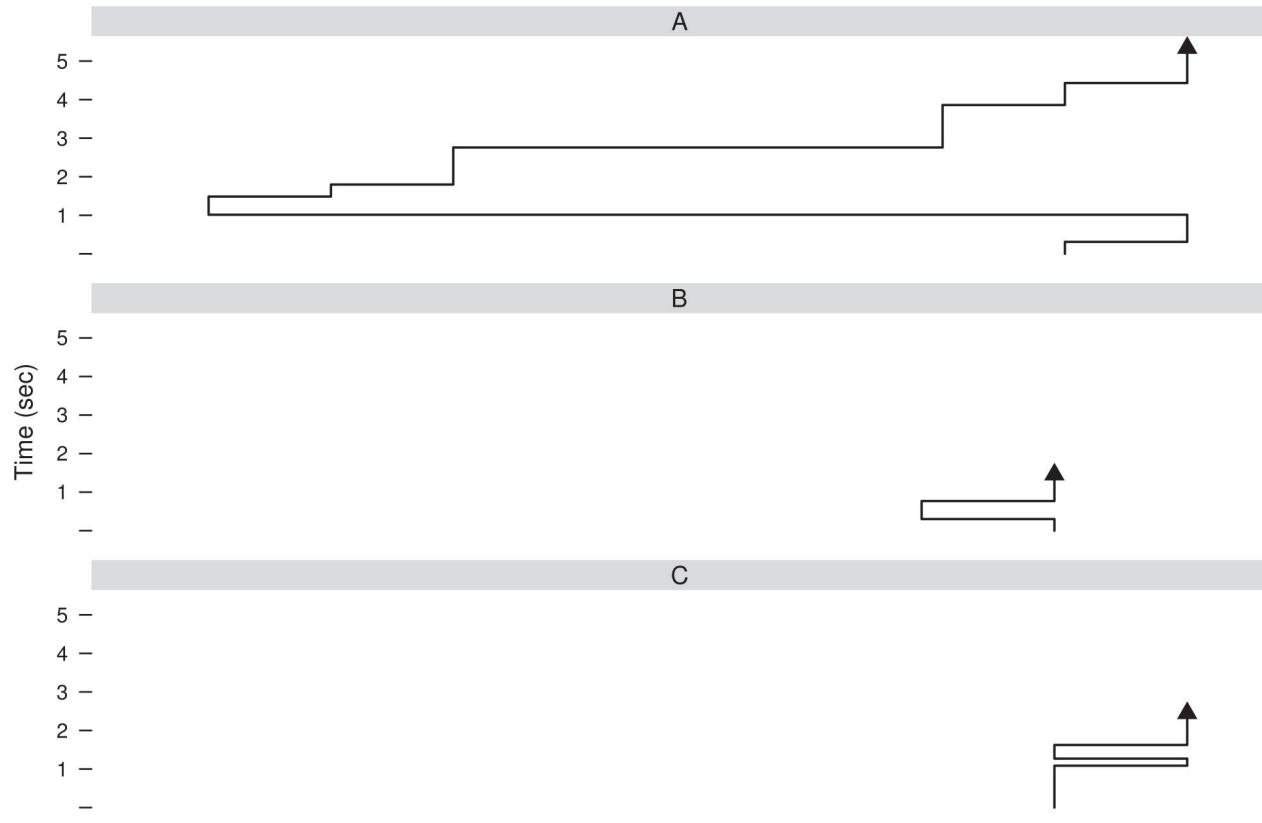


Figure 3 from von der Malsburg & Vasishth (2013)

3 Clustering

Adapted idea, here:

Cluster subjects based on average Scasim into reader types

“Cluster people, not scanpaths.”

3 Clustering

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
....
....
....
....

1) Determine **centroid scanpath** in each condition

3 Clustering

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
... -									
... -									
... -									
... -									

1) Determine **centroid scanpath** in each condition

3 Clustering

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•		

1) Determine **centroid scanpath** in each condition

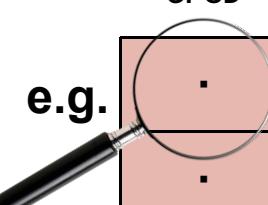
3 Clustering

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
·	·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·	·
·	·	·	·	·	·	·	·	·	·

2) Construct between-subject **distance matrix** per condition

3 Clustering

e.g.

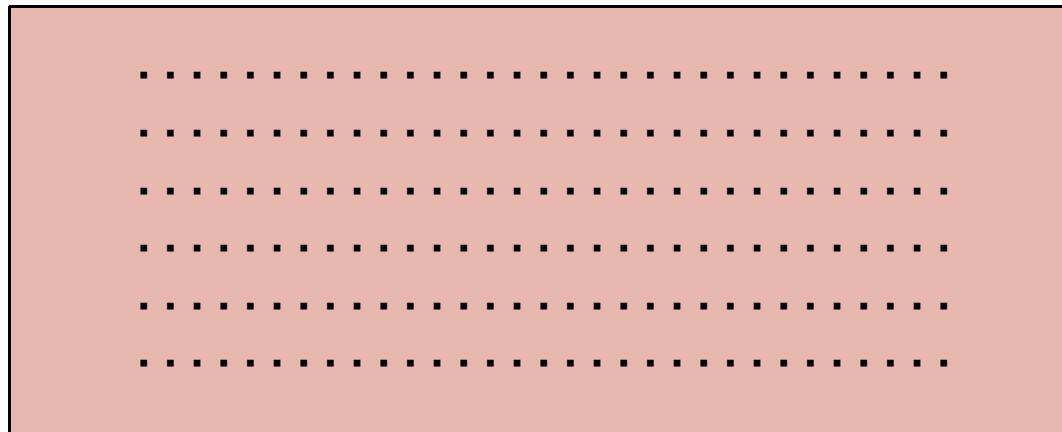


GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
.
.
.
.

2) Construct between-subject **distance matrix** per condition

3 Clustering

186 subjects



2) Construct between-subject **distance matrix** per condition

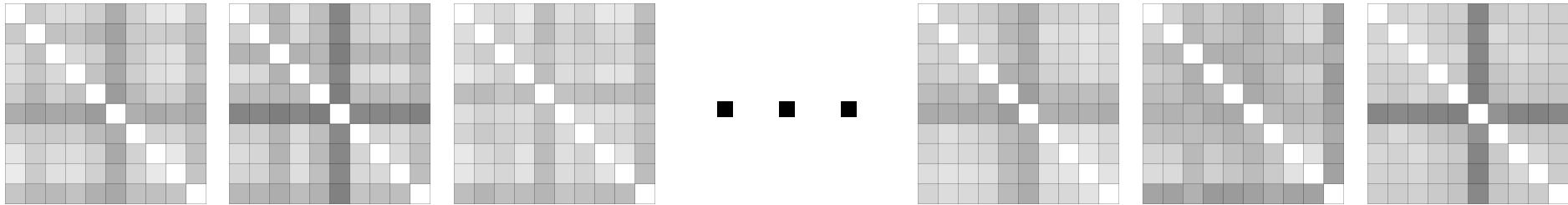
3 Clustering

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	...
S1											
S2											
S3											
S4											
S5											
S6											
S7											
S8											
S9											
S10											
:											

2) Construct between-subject **distance matrix** per condition

3 Clustering

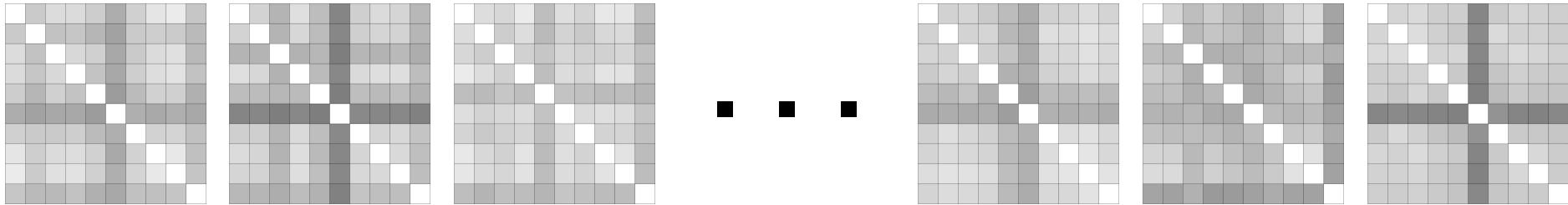
38 conditions



2) Construct between-subject **distance matrix** per condition

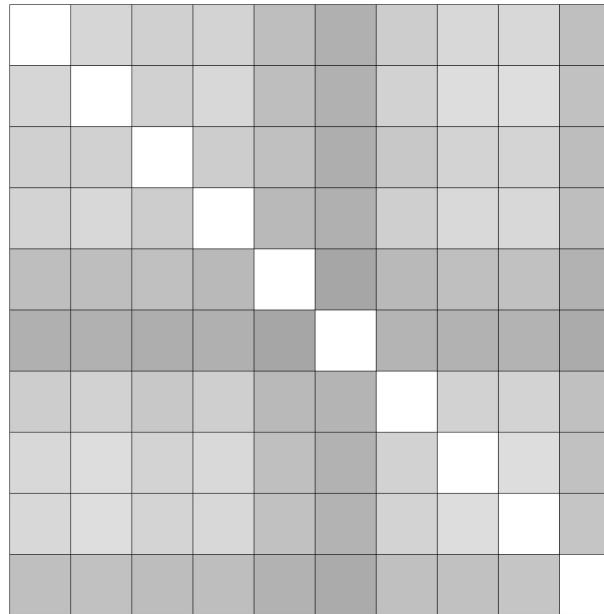
3 Clustering

38 conditions



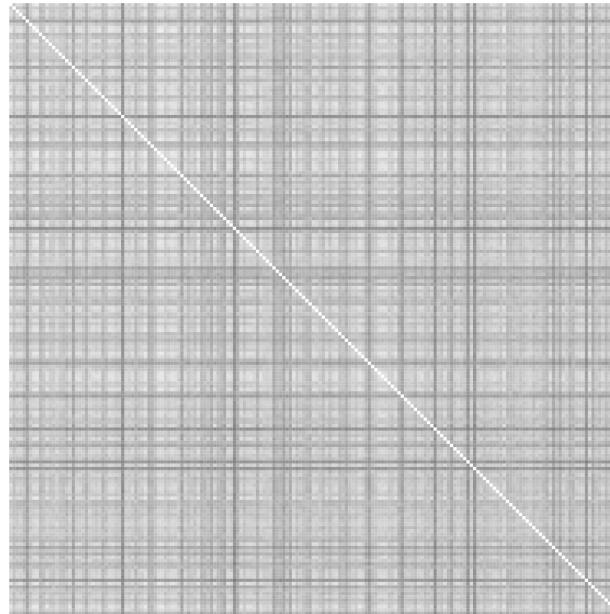
3) **Average** over 38 distance matrices

3 Clustering



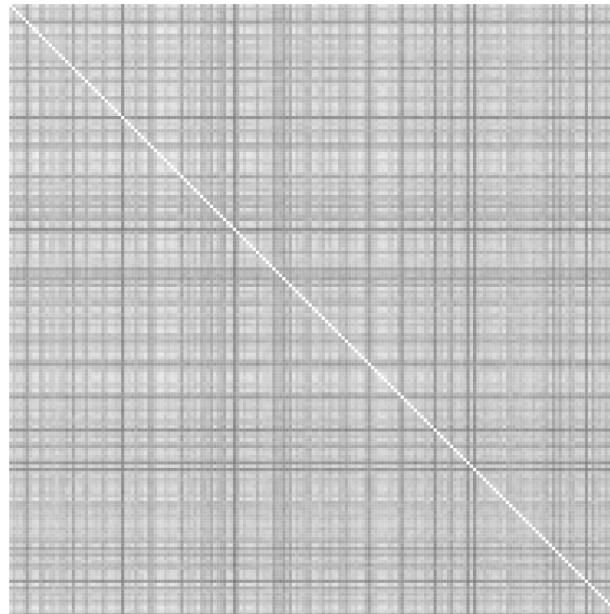
3) **Average** over 38 distance matrices

3 Clustering



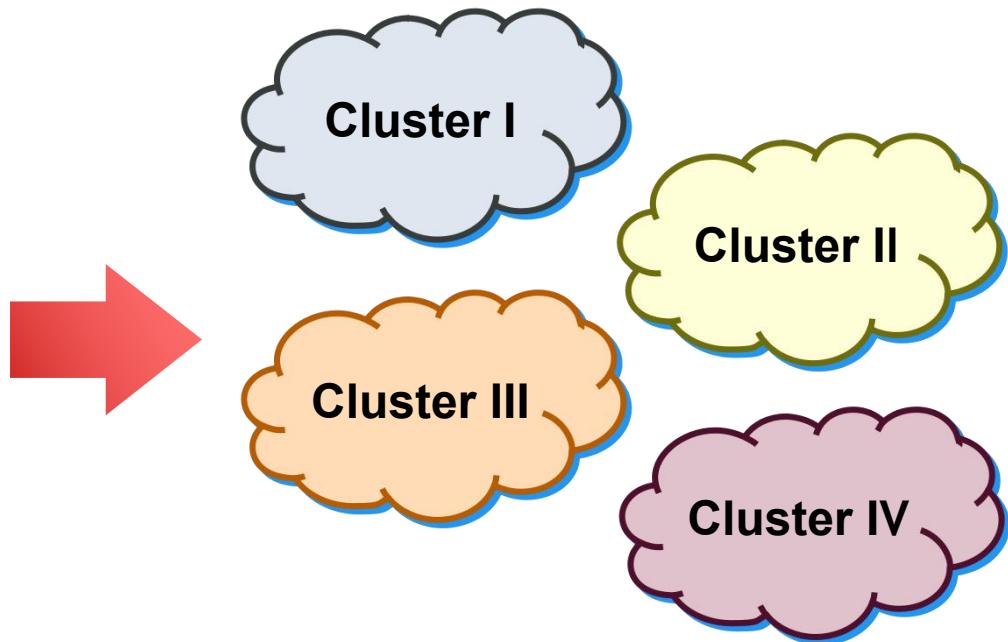
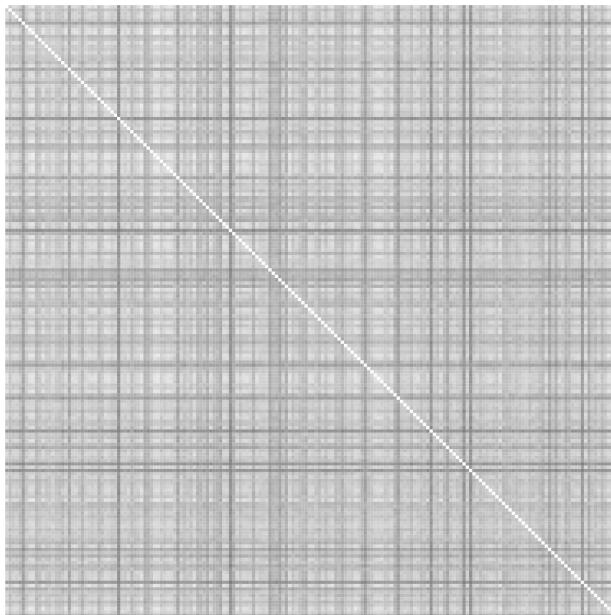
3) **Average** over 38 distance matrices

3 Clustering



4) Multidimensional scaling and Gaussian mixture modeling

3 Clustering

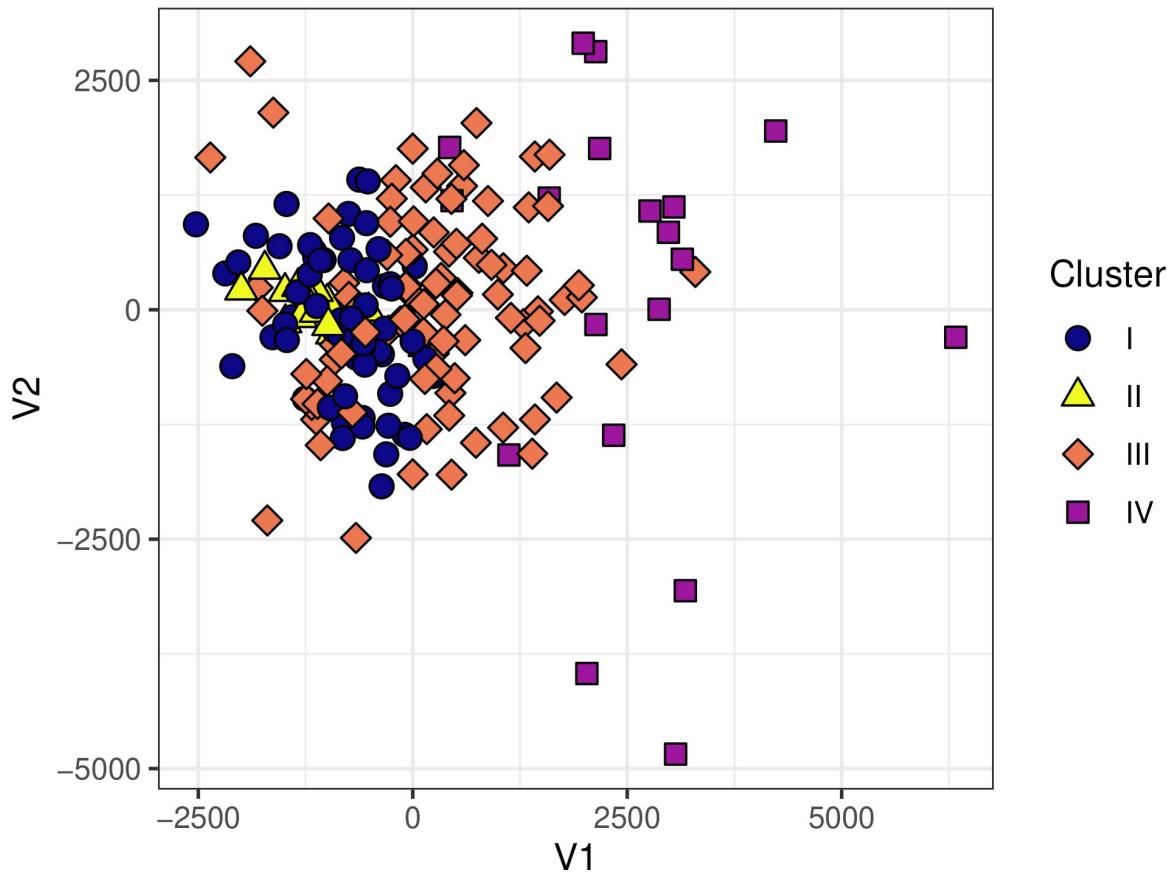


4) Multidimensional scaling and Gaussian mixture modeling

3 Clustering

2D projection

Scanpath Similarity Map of Participants

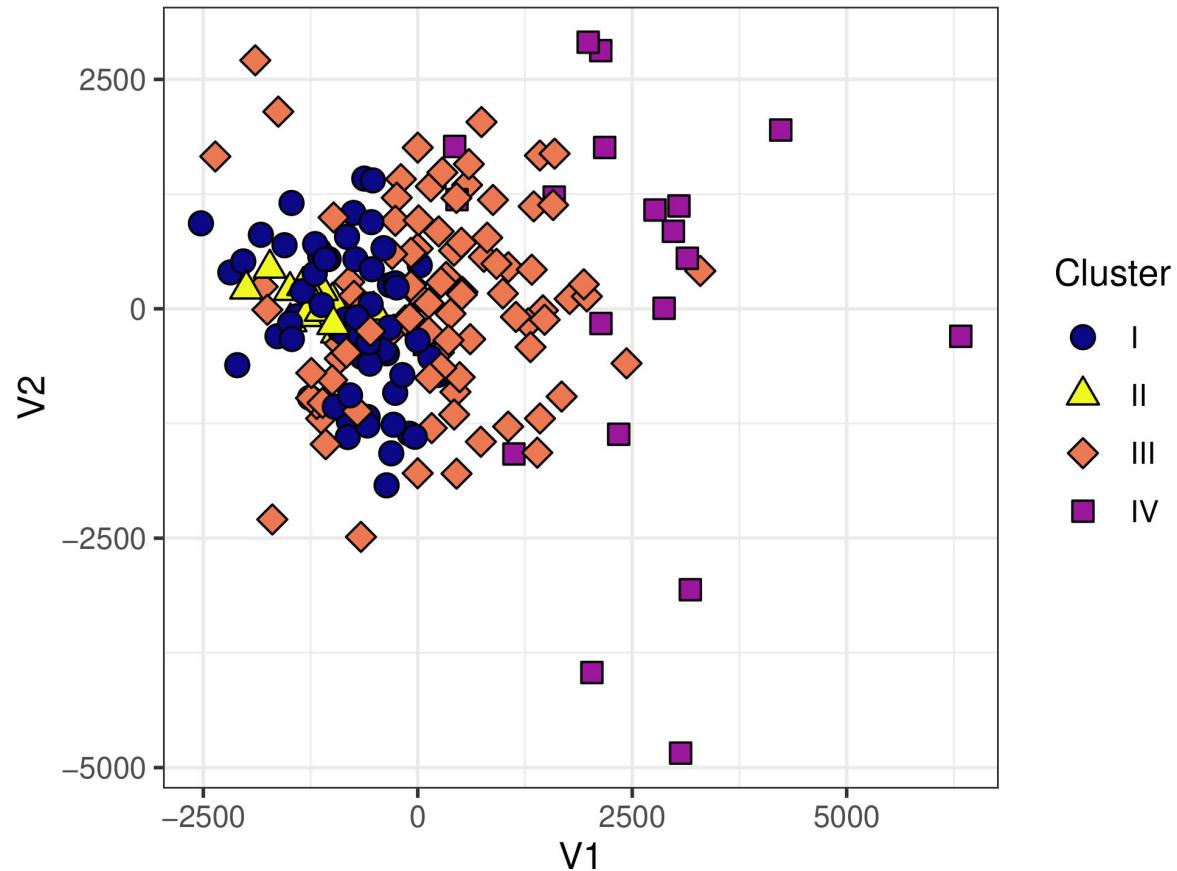


3 Clustering

I	32%
II	7%
III	50%
IV	11%

2D projection

Scanpath Similarity Map of Participants



3 Clustering

GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
a	a	a	a	a	a	a	a	a	a
b	b	b	b	b	b	b	b	b	b
c	c	c	c	c	c	c	c	c	c
d	d	d	d	d	d	d	d	d	

38 conditions

3 Clustering

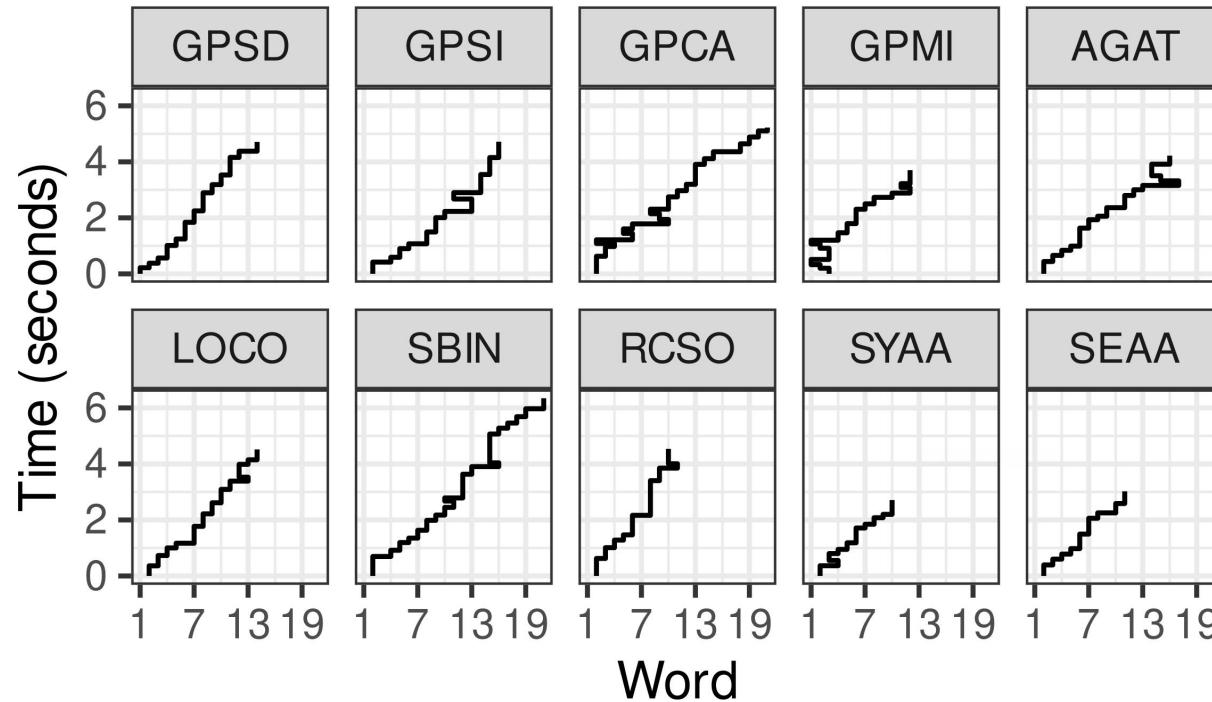
GPSD	GPSI	GPCA	GPMI	AGAT	LOCO	SBIN	RCS0	SYAA	SEAA
						a			
b	b				b				
			c	c				c	c
		d					d		

10 critical conditions

3 Clustering

Centroid of Cluster I

A fast and linear reader (Cluster I)

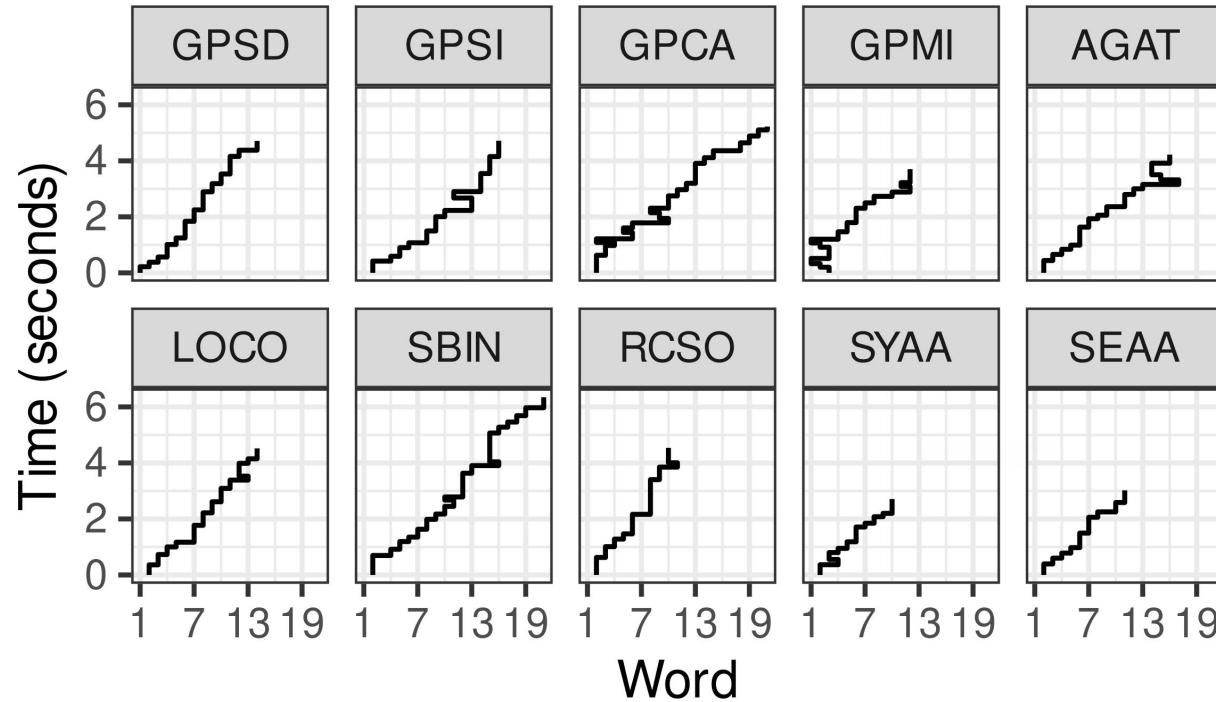
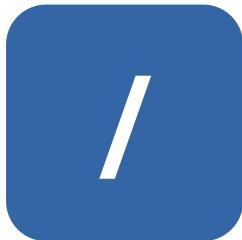


3 Clustering

Centroid of Cluster I

A fast and linear reader (Cluster I)

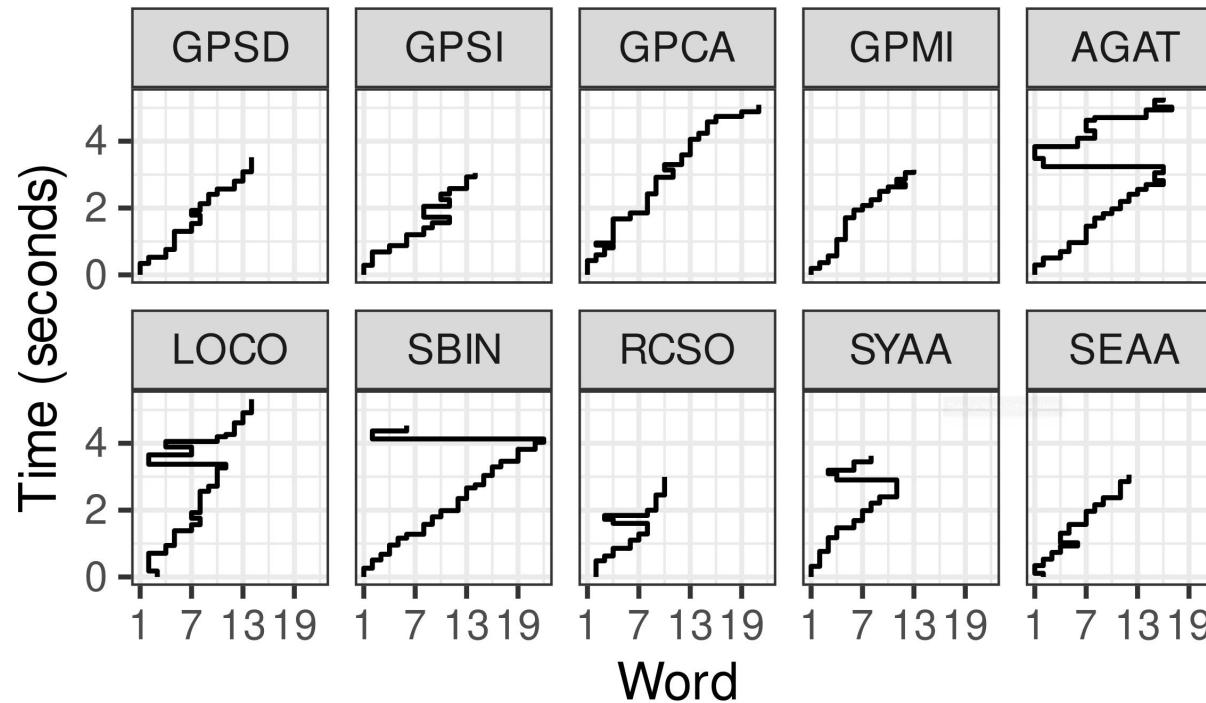
mnemonic:



3 Clustering

Centroid of Cluster II

A fast occasional re-reader (Cluster II)



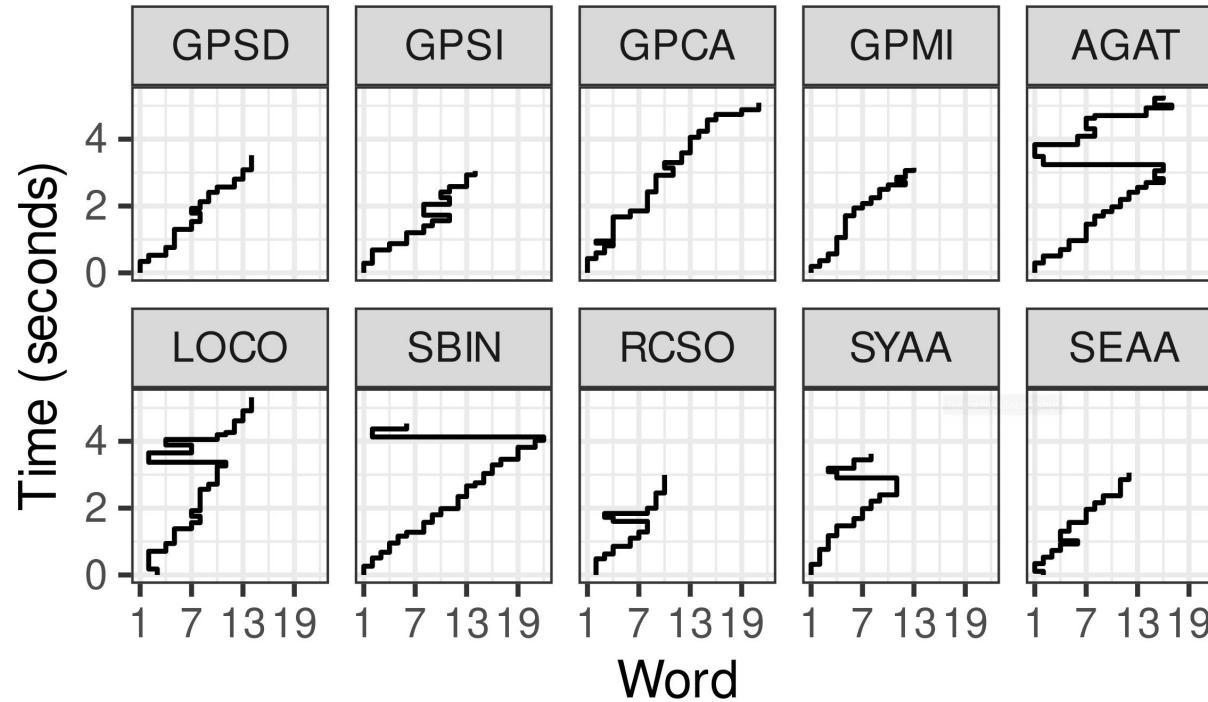
3 Clustering

Centroid of Cluster II

A fast occasional re-reader (Cluster II)

mnemonic:

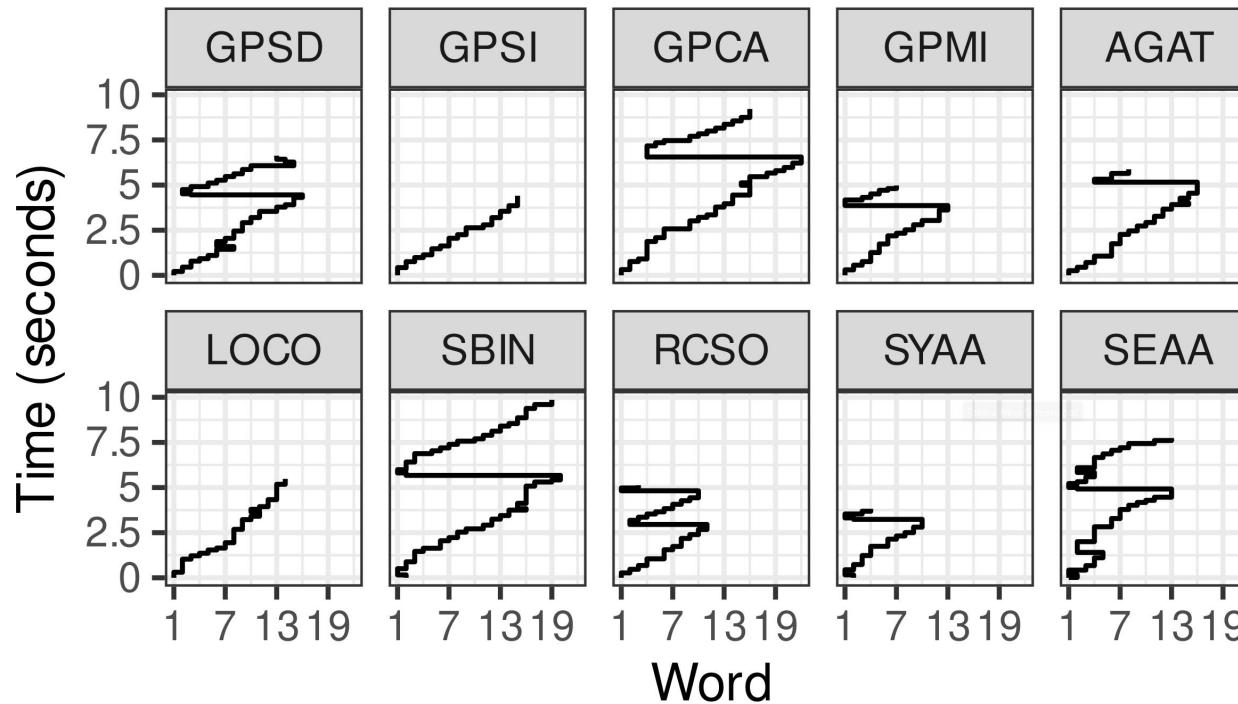
7



3 Clustering

Centroid of Cluster III

An average regular re-reader (Cluster III)



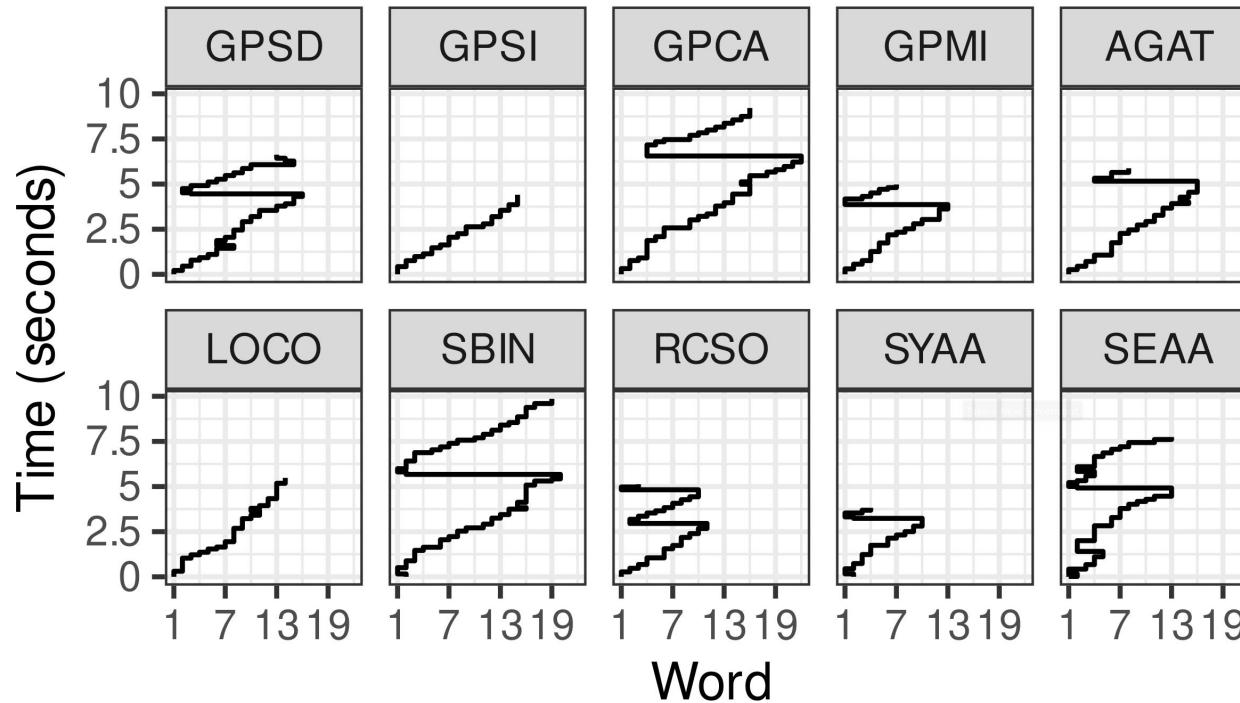
3 Clustering

Centroid of Cluster III

An average regular re-reader (Cluster III)

mnemonic:

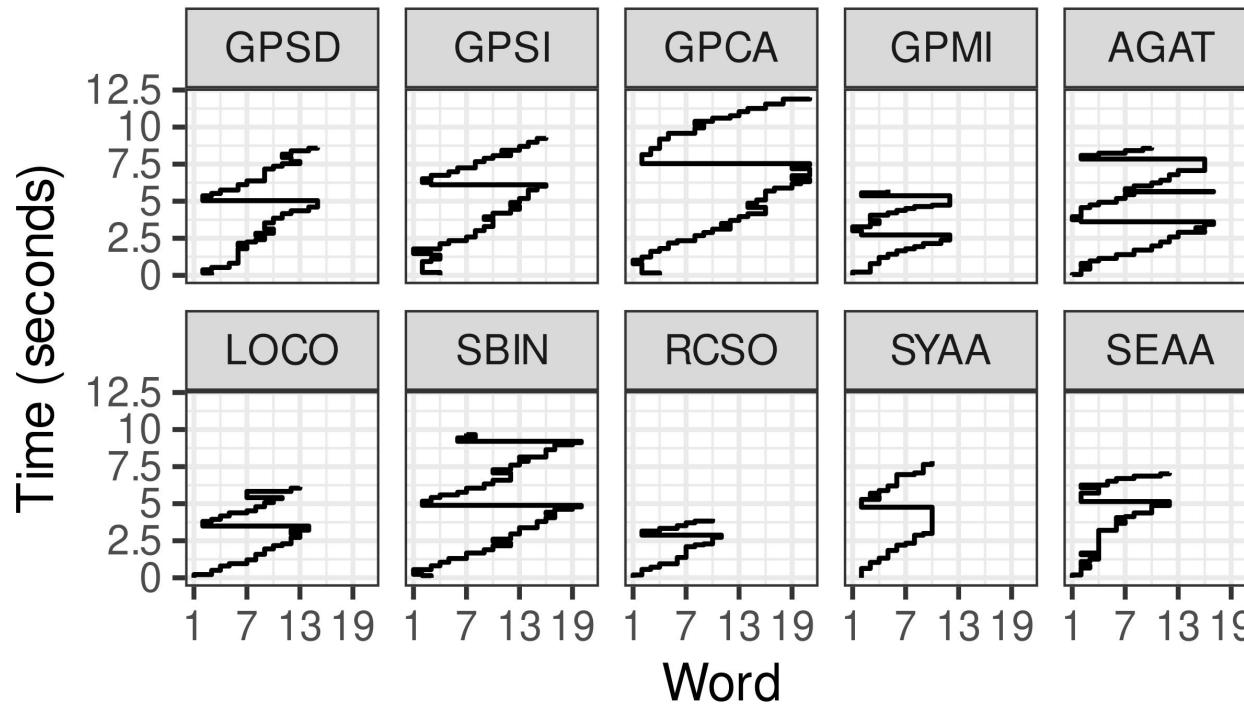
S



3 Clustering

Centroid of Cluster IV

A slow excessive re-reader (Cluster IV)



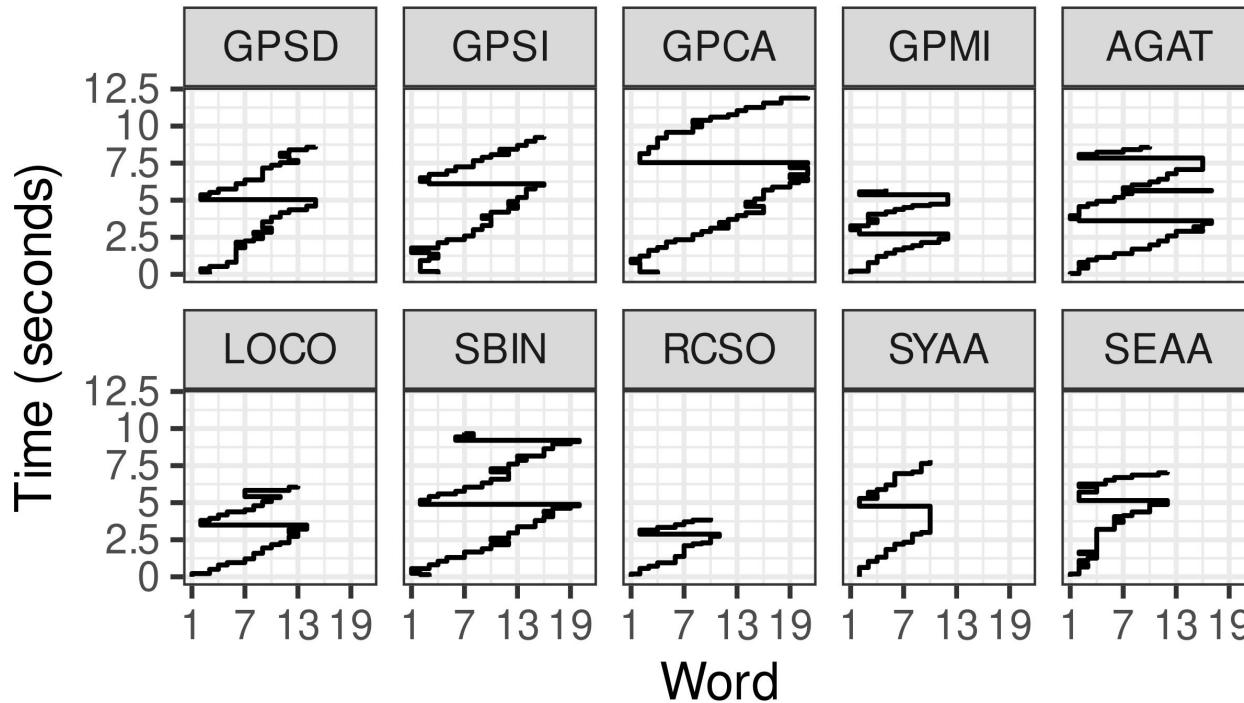
3 Clustering

Centroid of Cluster IV

A slow excessive re-reader (Cluster IV)

mnemonic:

3



3 Clustering

Comprehension question accuracies (95% Crl)

- / **Cluster I:** [77.2%, 82.5%]
- 7 **Cluster II:** [71.0%, 81.5%]
- S **Cluster III:** [82.2%, 85.3%]
- 3 **Cluster IV:** [77.2%, 84.7%]

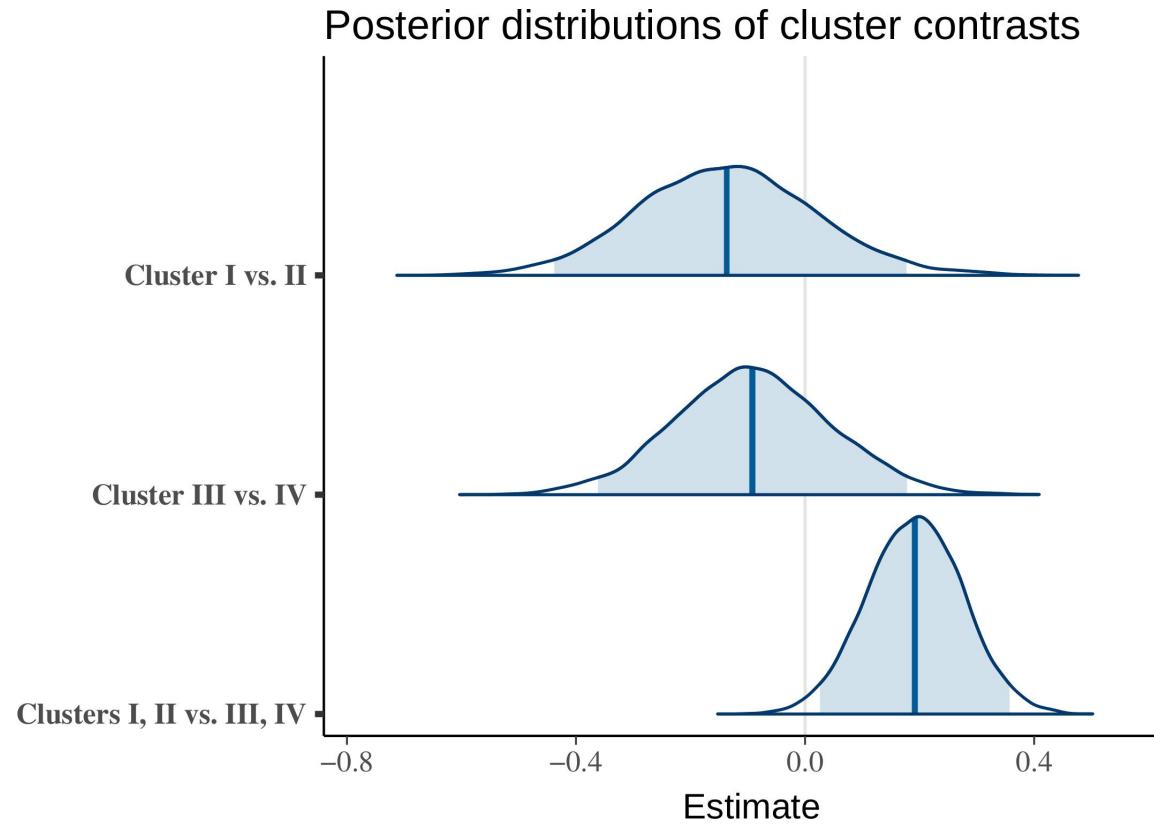
3 Clustering

Comprehension question accuracies (95% Crl)

fast	/	Cluster I: [77.2%, 82.5%]
	7	Cluster II: [71.0%, 81.5%]
slow	S	Cluster III: [82.2%, 85.3%]
	3	Cluster IV: [77.2%, 84.7%]

3 Clustering

Beta regression
on question
accuracies



4 Implications

Methodological:

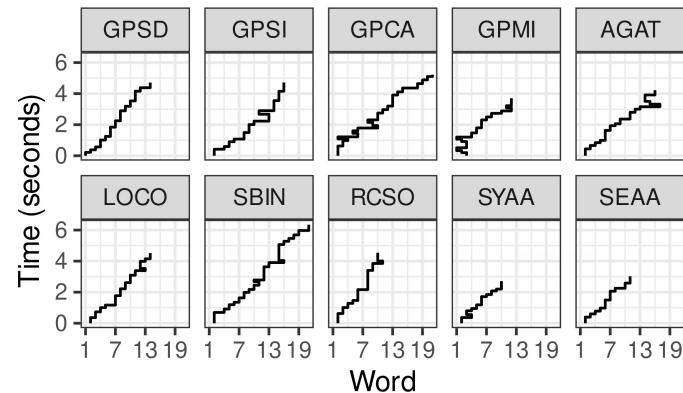
- Analysis of **raw scanpaths** can unveil differences that aggregate measures hide
- Such differences may manifest **between subjects** (i.e., as individual differences)

4 Implications

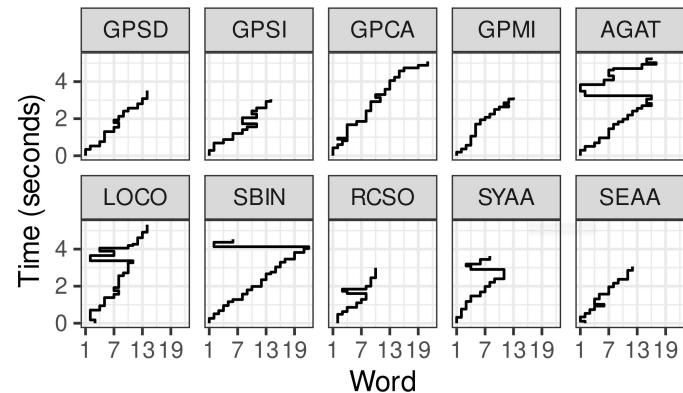
Present results:

- Different subjects show **distinct reading patterns**
 - More / fewer full regressions (/-, 7-, S-, or Ζ-shaped)
 - Overall reading speed
- Patterns might reflect different **task strategies** (Rayner, 1998)
 - Systematic re-reading only due to question anticipation?
- Higher response accuracy in slow and regressive clusters
 - Indicates **speed–accuracy trade-off** (Wickelgren, 1977)

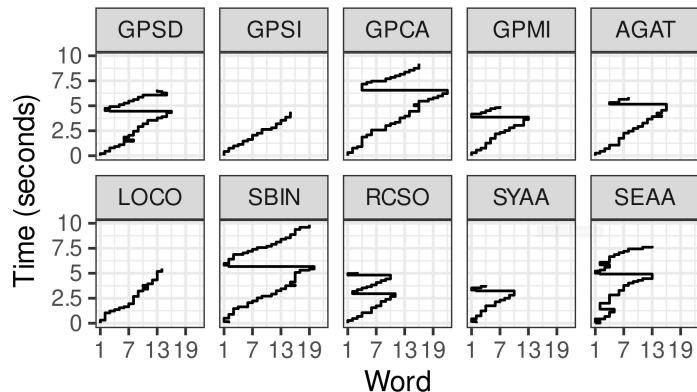
A fast and linear reader (Cluster I)



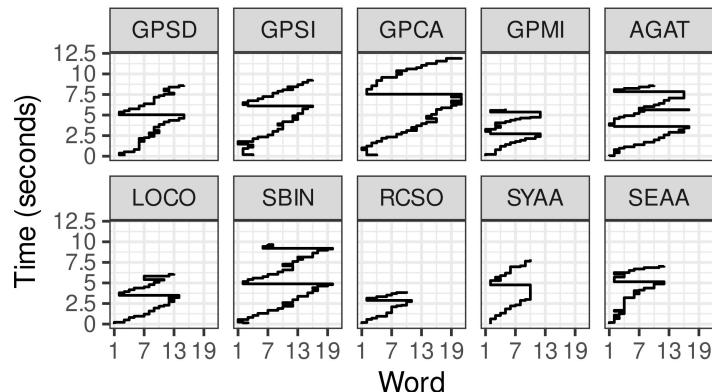
A fast occasional re-reader (Cluster II)



An average regular re-reader (Cluster III)



A slow excessive re-reader (Cluster IV)



References [1 / 2]

- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210.
- Häussler, J. (2009). *The Emergence of Attraction Errors During Sentence Comprehension*. PhD thesis, University of Konstanz.
- Hsiao, F. and Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, 90(1), 3–27.
- Huang, K. J., Arehalli, S., Kugemoto, M., Muxica, C., Prasad, G., Dillon, B., & Linzen, T. (2024). Large-scale benchmark yields no evidence that language model surprisal explains syntactic disambiguation difficulty. *Journal of Memory and Language*, 137, 104510.
- Konieczny, L., Hemforth, B., and Scheepers, C. (2000). Head position and clause boundary effects in reanalysis. In Hemforth, B. and Konieczny, L., editors, *German Sentence Processing*, pages 247–278. Springer.
- Logačev, P. (2023). The role of underspecification in relative clause attachment: Speed-accuracy tradeoff evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 49(9), 1471.
- Meng, M. and Bader, M. (2000a). Mode of disambiguation and garden-path strength: An investigation of subject–object ambiguities in German. *Language and Speech*, 43(1), 43–74.
- Meng, M. and Bader, M. (2000b). Ungrammaticality detection and garden path strength: Evidence for serial parsing. *Language and Cognitive Processes*, 15(6), 615–666.
- Paape, D. and Vasishth, S. (2016). Local coherence and preemptive digging-in effects in German. *Language and Speech*, 59(3), 387–403.

References [2 / 2]

- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372
- Schoknecht, P., Yadav, H., & Vasishth, S. (2025). Do syntactic and semantic similarity lead to interference effects? Evidence from self-paced reading and event-related potentials using German. *Journal of Memory and Language*, 141, 104599.
- Traxler, M. J., Pickering, M. J., and Clifton Jr, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39(4), 558–592.
- van Kampen, A (2001). *Syntaktische und semantische Verarbeitungsprozesse bei der Analyse strukturell mehrdeutiger Verbfinalsätze im Deutschen: Eine empirische Untersuchung*. PhD thesis, Free University of Berlin.
- von der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis?. *Journal of Memory and Language*, 65(2), 109-127.
- von der Malsburg, T., & Vasishth, S. (2013). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes*, 28(10), 1545-1578.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41(1), 67-85.
- TODO frazier rayner 1982