

Identifying tasks using keystroke frequencies and patterns

Rianne Conijn and Menno van Zaanen
m.a.conijn@uvt.nl

Outline

- Frequencies of keystrokes (Conijn & Van Zaanen, 2017)
- Patterns of keystrokes (ongoing)

Introduction

Keystrokes

- Writer characteristics / typing characteristics
 - Used for writer identification
- Writing processes/cognitive load
 - What we want

Introduction – Copy task

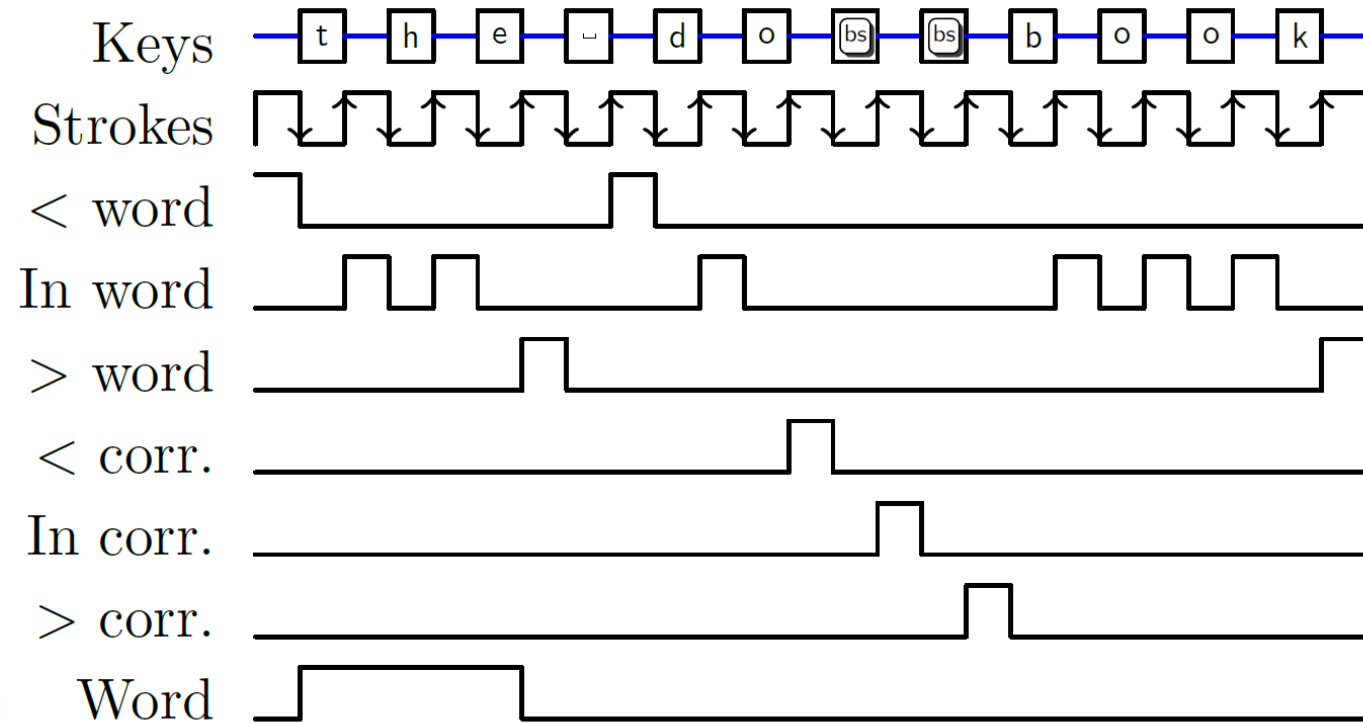
- Assumed to relate to typing skills only
- Differences between writing task – copy task = writing process?
- Here: Compare copy task with free-form task

Method – Dataset

- Villani keystroke data set (Tappert et al., 2009)
- 36 writers
- 338 copy texts & 412 free-form texts (>650 characters)

Method – Keystroke features

- 1) Interkey intervals
- 2) Corrections (#corrections, %words corrected)
- 3) Word time



Method – Analyses

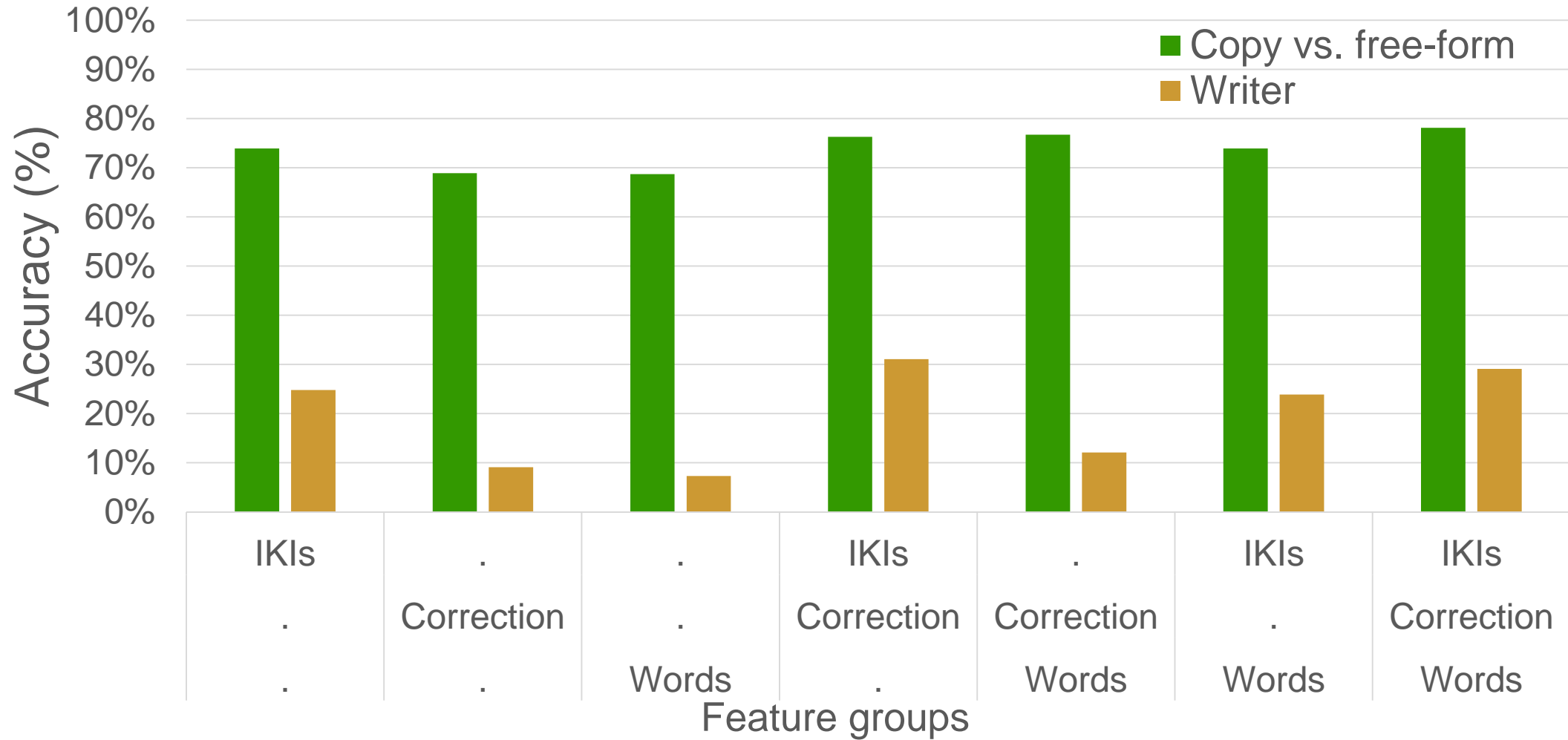
- Statistical analyses
- Classification (SVM)
 - Copy versus Free-form text
 - Writer

Results – Statistical analyses

Free-form versus copy task:

- More keystrokes
- More corrections
- Higher % words with corrections
- Longer total time
- Longer interkey intervals before words
- Shorter interkey intervals within words

Results – Classification



Discussion

- Frequencies of keystroke differ between copy task and free-form task
- Classification contains little writer/typing information
- What about patterns of keystrokes?

Current work: Patterns in keystrokes

- Differences in keystrokes in one task
- Where do students struggle?
- Improve instruction/feedback
- Real-time feedback

Analyses of patterns in keystrokes

Patterns of what?

- Interkey intervals?
- Within/Before/after word?
- With/without pauses? Pause length = 2 sec?

How to extract patterns?

1) n-grams of interkey intervals (Van Zaanen & Gaustad, 2010; Conijn et al., *under review*)

- How many ordinal categories?
- What is the length of subsequence (n)?

Max. categories:

... – 262 – 160 – 3340 – 156 – 145 – 400 – 3289 – 2345 – 123 – 189 – ...

3 categories:

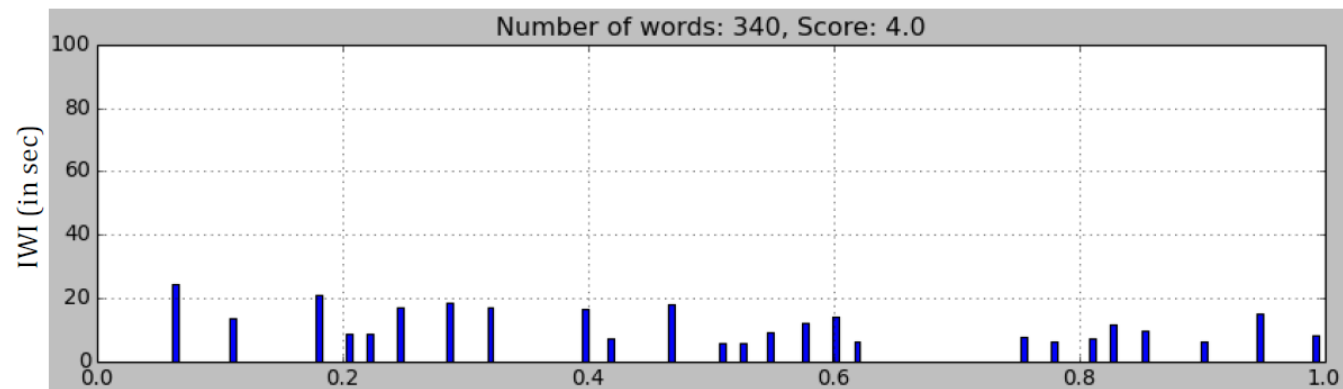
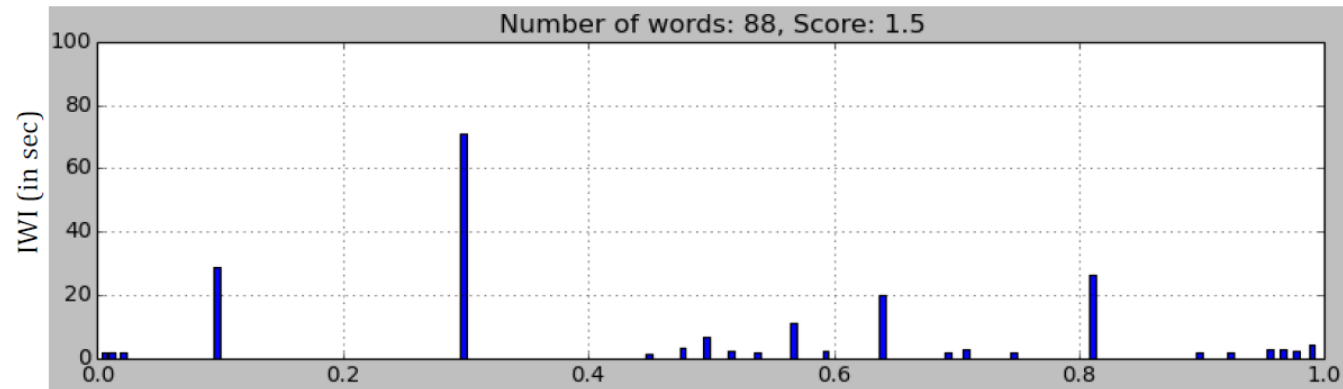
... – Med – Low – High – Low – Low – Med – High – High – Low – Low – ...

n = 4

How to extract patterns?

2) Vectors of interkey interval duration + location (Zhang et al., 2016)

- How to create vectors of equal length?
- How to compare vectors? (Cosine similarity, ...)

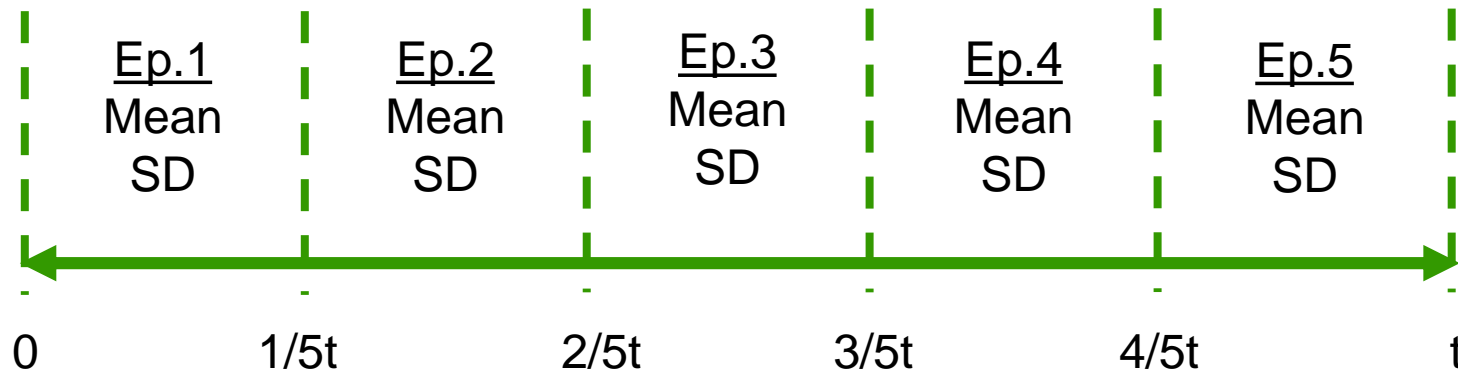


Taken from
Zhang et al.
(2016), Figure 2

How to extract patterns?

3) Frequencies of interkey intervals per episode (Tillema et al., 2011)

- How many episodes?



Conclusion

- Frequencies are useful
- Patterns might provide even more information
 - Many choices need to be made
 - Use data mining to reduce search space
 - Use theory to reduce search space

Questions?

Identifying tasks using keystroke frequencies and patterns

Rianne Conijn and Menno van Zaanen
m.a.conijn@uvt.nl

References

- Conijn, R., Knight, S., Martinez-Maldonado, R., & Van Zaanen, M. (*under review*). Discovering Sequential Patterns of Critical Rhetorical Moves in Students' Academic Writing. In *Proceedings of the Eighth International Conference on Learning Analytics & Knowledge*. Sydney, Australia: ACM.
- Conijn, R., & Van Zaanen, M. (2017). Identifying writing tasks using sequences of keystrokes. In *Benelearn 2017: Proceedings of the Twenty-Sixth Benelux Conference on Machine Learning* (pp. 28–35). Eindhoven, The Netherlands.
- Tappert, C. C., Villani, M., & Cha, S.-H. (2009). Keystroke biometric identification and authentication on long-text input. *Behavioral Biometrics for Human Identification: Intelligent Applications*, 342–367.
- Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2011). Relating self reports of writing behaviour and online task execution using a temporal model. *Metacognition and Learning*, 6(3), 229–253.
- Van Zaanen, M., & Gaustad, T. (2010). Grammatical inference as class discrimination. *Grammatical Inference: Theoretical Results and Applications*, 245–257.
- Zhang, M., Hao, J., Li, C., & Deane, P. (2016). Classification of Writing Patterns Using Keystroke Logs. In L. A. van der Ark, D. M. Bolt, W.-C. Wang, J. A. Douglas, & M. Wiberg (Eds.), *Quantitative Psychology Research: The 80th Annual Meeting of the Psychometric Society, Beijing, 2015* (pp. 299–314). Cham: Springer International Publishing.