Logging Keystrokes in Writing by English Learners

Georgios Velentzas, Andrew Caines, Rita Borgo, Erin Pacquetet, Clive Hamilton, Taylor Arnold, Diane Nicholls, Paula Buttery, Thomas Gaillat, Helen Yannakoudakis, Nicolas Ballier

LREC-COLING 2024

Introduction

In this work we **collect** and **analyse** data representing the essay writing process from start to finish, by recording every **keystroke** from multiple writers participating in our study.

We describe our **data collection** methodology, the **characteristics** of the resulting dataset, namely **KUPA-KEYS**, and the **assignment of proficiency levels to the texts**

Introduction: Why Keylog data?

- A great deal of information about the writing process can be captured
- We can potentially detect **when learners are struggling** with their **writing**
- We might potentially enable **supportive interventions** to **aid the learner**
- Insights into **linguistic creativity** and **language complexity** in **production**
- Detect **malpractice in education or assessment settings** - Generative AI

Introduction: KUPA-KEYS Dataset

We have compiled a dataset of **texts**, **keystroke logs** and **metadata** for public release, which can be used for any of the aforementioned purposes, including research around **automated essay assessment**.

It contains a copied text and creative essay written in English by **1,006 crowdsourced participants**, both native speakers and non-native speakers of the language.

The **K**ing's College London & **U**niversité **Pa**ris Cité Keys (**KUPA-KEYS**) dataset is publicly available for non-commercial use* and our research code for data collection is open-source

* https://huggingface.co/datasets/ALTACambridge/KUPA-KEYS

Related Work: Other Datasets

Keystroke datasets generated from participants transcribing a **fixed-text** are more common (Allen, 2010; Giot et al., 2012) with the largest being reported in Dhakal et al. (2018) consisting of 136 million keystrokes.

On the contrary, datasets that include free-text are not as extensive, with one of the earliest ones being **Clarkson I** (Vural et al., 2014), comprised of 840,000 keystrokes produced by 39 participants

One of the largest datasets available is described in Sun et al. (2016), known as the **Buffalo dataset**, comprising 2.14 million keystrokes from **148** participants (both transcription and free-text included)

Related Work: Comparisons

Murphy et al. (2017) provides the **Clarkson II** dataset, where keystroke data were collected in an uncontrolled environment from 103 subjects, yielding a total of 12.6 million keystrokes.

Our dataset is distinct from the ones above because we focus on essay writing by both learners and native speakers of English, with both text-copy and free-text composition.

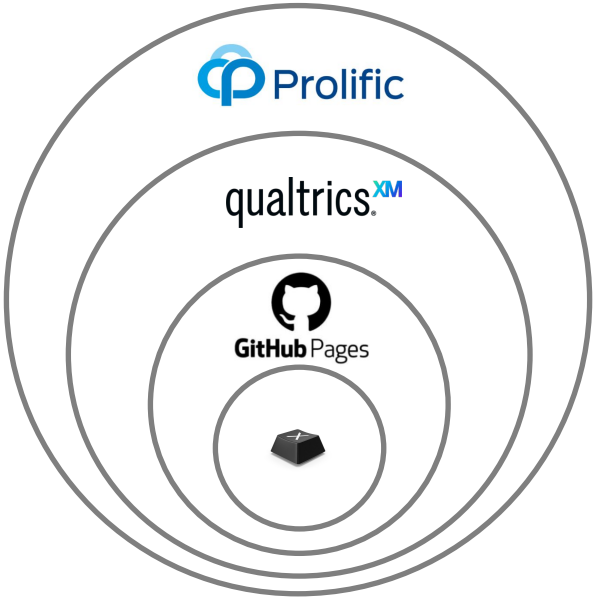
A new Kaggle shared task was launched in October 2023*. KUPA-KEYS is smaller, but includes logs from a transcription task, our essays are scored on a greater scale, and the alphanumeric characters are preserved.

* https://www.kaggle.com/competitions/linking-writing-processes-to-writing-quality

Data Collection: Protocol

- Recruited 1045 participants through Prolific crowdsourcing platform
- Participants were directed to Qualtrics for data collection
- After metadata collection they were redirected to our GitHub pages
- JavaScript keylogger in each page captured their keystroke data

Data Collection: Protocol



After the survey, participants were redirected from Qualtrics to our text authoring site

Participants were required to complete two writing tasks: a **copy-text** task & an **essay-writing** task

Data Collection: Copy-Text task

- 300 word excerpt from Steve Jobs' Speech at Stanford University
- Contains 197 distinct English Digraphs
- Extensive enough to be used for user-specific baselines
- Used by Sun et al. 2016 (Buffalo Dataset) - Data Augmentation

Data Collection: Essay-Writing task

- "Just for fun" prompts from the English learning platform Write & Improve*
- We chose "just for fun" prompts as opposed to level-specific prompts
- They are deliberately creative
- They are suitable across different proficiency levels
- We selected ones which tend not to elicit personal information
- Each participant was randomly allocated to one prompt (10 total)

* https://writeandimprove.com

Data Collection: Essay-Writing task (examples)

A Special Place.
If you could be anywhere in the world right now, where would you choose to be? Describe the place. Why do you want to be there?

Unforgettable.
Write a short story with the title 'Unforgettable'. Your story must have a beginning, a middle and an end. The end must be surprising.

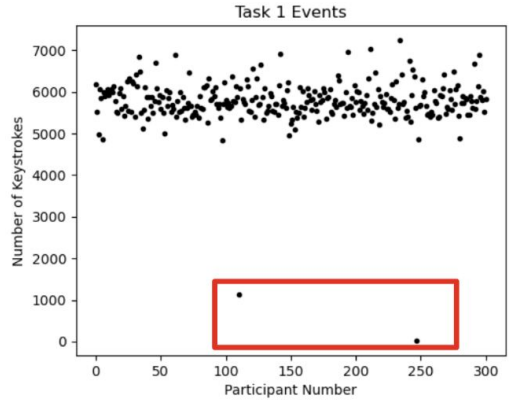
Data Collection: Participants

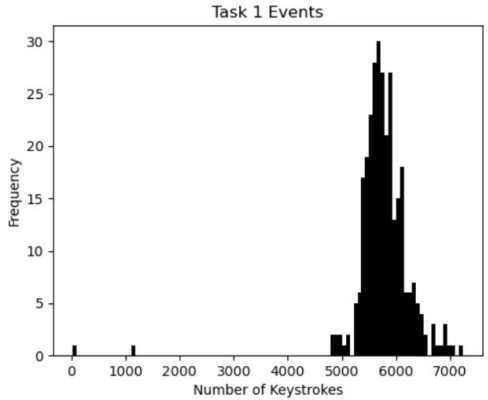
- After a pilot study we found that it took ~30 min for completion (total)
- We paid each participant £7.50 GBP/ $9.20 USD for the full survey
- Participants spent on average 33.9 min [copy: 9.6 min, essay: 13.4 min]
- Essay threshold: initially 250 words - reduced to 150 to attract beginners
- We recruited 1,045 participants in three phases of crowdsourcing

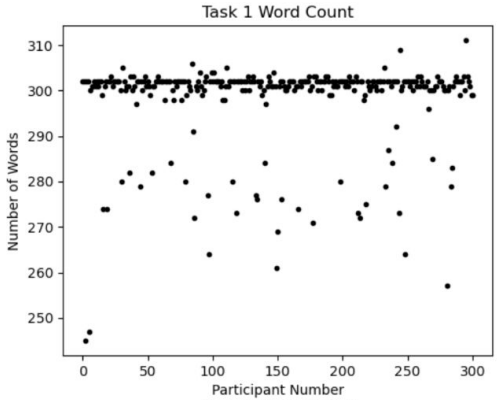
Data Collection: Common Rejection Reasons

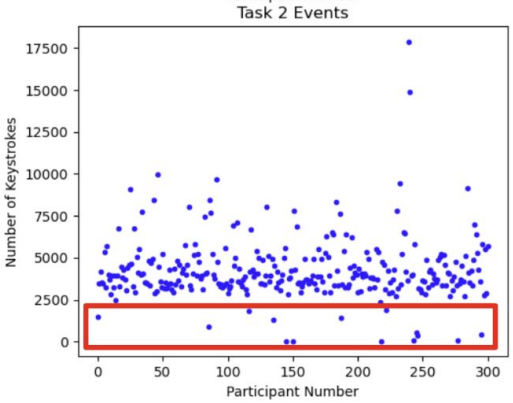
- Essays less than 80% of the minimum stated word length
- Copied and pasted text from external sources (even for copy-text task!)
- Typed in another language and translated at the end
- Most participants were asked to withdraw their submission before rejecting

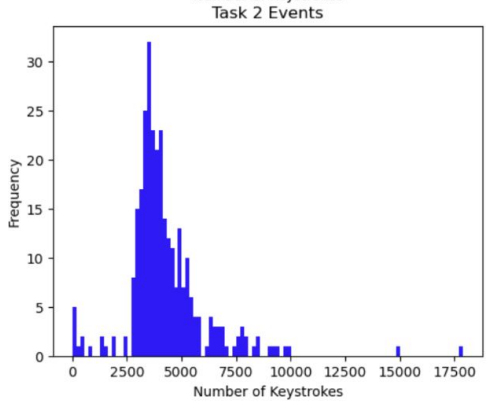
Data Collection: Common Rejection Reasons

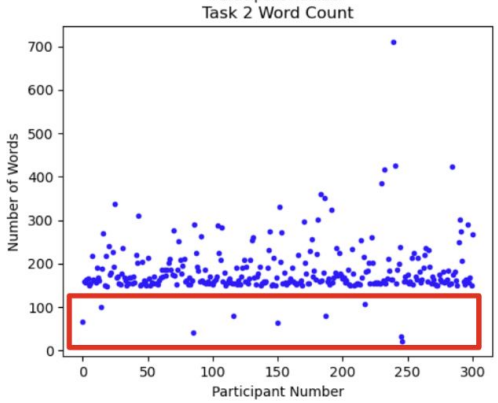












Data Processing

- Automatically graded each essay with W&I API
- Three qualified human assessors graded each essay
- 35 participants were rejected by human assessors
- Reasons: off-topic, offensive, potentially distressing
- Removed 4 more from post-processing (tablet/mobile)
- Therefore, the public release features 1006 essays

W&I score	CEFR level
0	–
1	A1.i
2	A1.ii
3	A2.i
4	A2.ii
5	B1.i
6	B1.ii
7	B2.i
8	B2.ii
9	C1.i
10	C1.ii
11	C2.i
12	C2.ii
13	C2.ii

Dataset Description: KUPA-KEYS

We conducted a thorough data cleaning process to streamline the generation of three primary tables, conveniently saved in CSV format

KUPA-KEYS-META

- Metadata
- Demographic information
- Post-processing data
- Human markers' evaluations
- Automarker evaluations
- Original prompts
- Final submitted text

KUPA-KEYS-TASK-1

- copy-text task keylog events

KUPA-KEYS-TASK-2

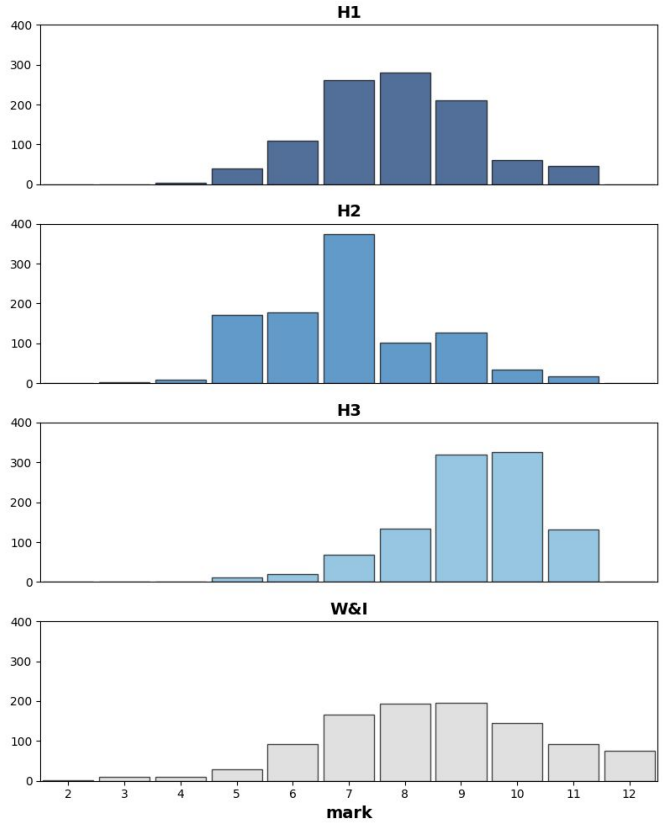
- essay-writing task keylog events

Dataset Description: KUPA-KEYS

KUPA-KEYS-TASK-<i>

id	time	type	key	key_code	alt_key	ctrl_key	meta_key	shift_key	is_repeat	range_start	range_end	text
xa2	563.4	down	'I'	KeyI	-	-	-	True	-	0	0	-
xa2	564.7	capture	-	-	-	-	-	-	-	-	-	'I'
xa2	564.7	input	-	-	-	-	-	-	-	1	1	'I'
xa2	691.6	up	'I'	KeyI	-	-	-	True	-	1	1	-
xa2	708.0	up	'Shift'	ShiftLeft	-	-	-	-	-	1	1	-
xa2	708.3	down	' '	Space	-	-	-	-	-	1	1	-
xa2	709.6	input	-	-	-	-	-	-	-	2	2	' '
xa2	835.6	up	' '	Space	-	-	-	-	-	2	2	-

Inter-annotator Agreement



H1: evenly distributed
H2: more strict
H3: more lenient
W&I: evenly + lenient

A2	B1	B2	C1	C2
4	111	549	323	19
0.4%	11%	54.6%	32.1%	1.9%

AC = 0.854

$$P_e = \frac{1}{g-1}\sum_{k=1}^{g}\pi_k(1-\pi_k)$$

$$\pi_k = \frac{1}{|\mathcal{A}|}\sum_{a\in\mathcal{A}} p_{a,k}$$

(Gwet 2002; Yannakoudakis and Cummins 2015)

Inter-annotator Agreement

Spearman's rank correlations

	H1	H2	H3	W&I
H1	–	0.633	0.711	0.574
H2	0.633	–	0.567	0.563
H3	0.711	0.567	–	0.514
W&I	0.574	0.563	0.514	–
Avg	0.639	0.588	0.598	0.550

Strong and significant correlations on the whole. Judgements of the human assessors correlate with each other more than they do with the automarker

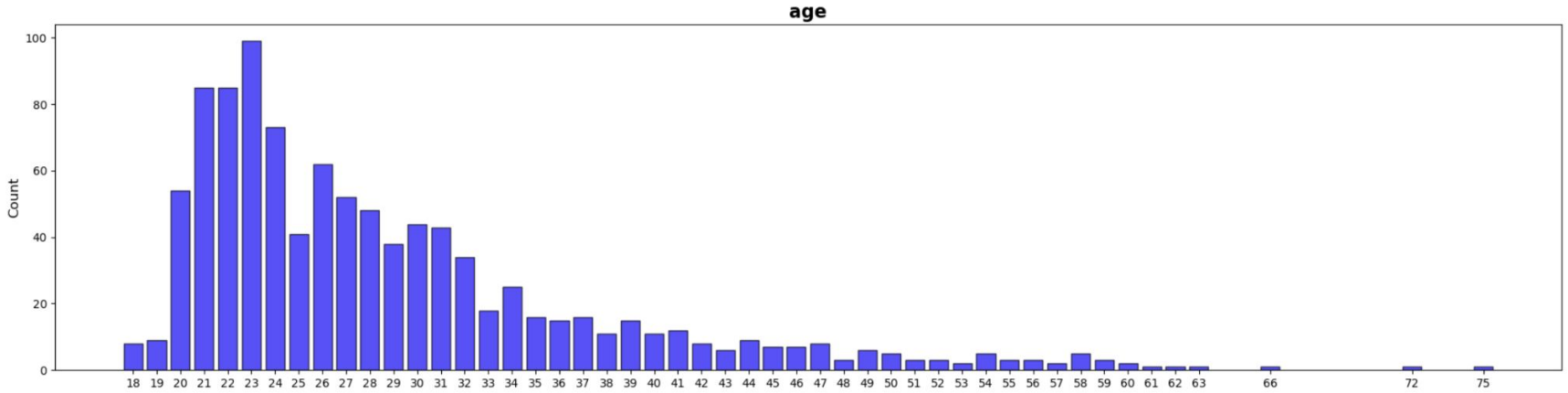
Inter-annotator Agreement

RMSD

	H_{avg}	H1	H2	H3
H1	0.622	-	-	-
H2	1.288	1.487	-	-
H3	1.371	1.695	2.586	-
W&I	1.543	1.708	2.241	1.770

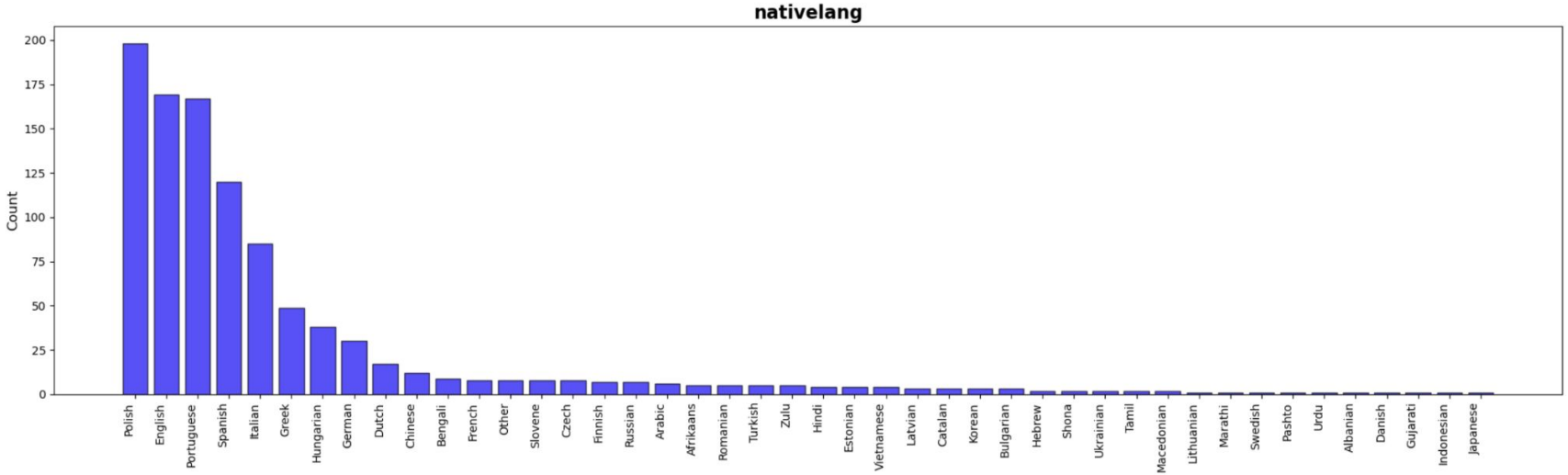
H1 is closest to the mean of human marks and has the lowest deviation from other markers including the automarker. H2 is involved in the highest RMSD values.

Data Analysis: Demographics - Age



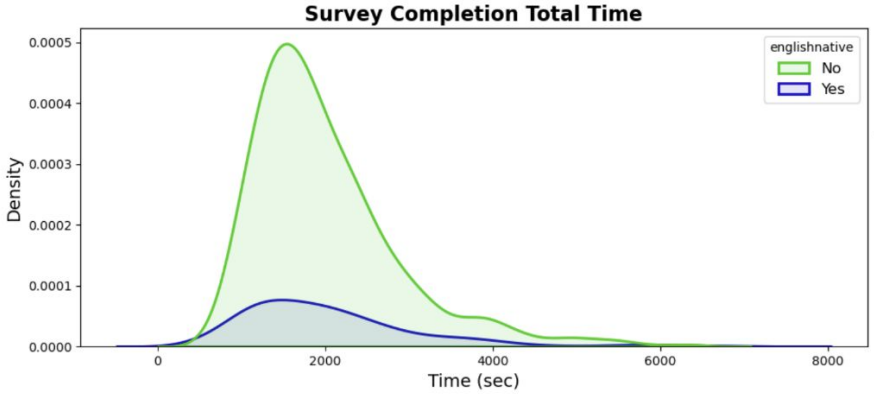
Age distribution: 25% of the participants between 18 and 23 years old, median age 26 years, 25% of the participants above 32 years old.

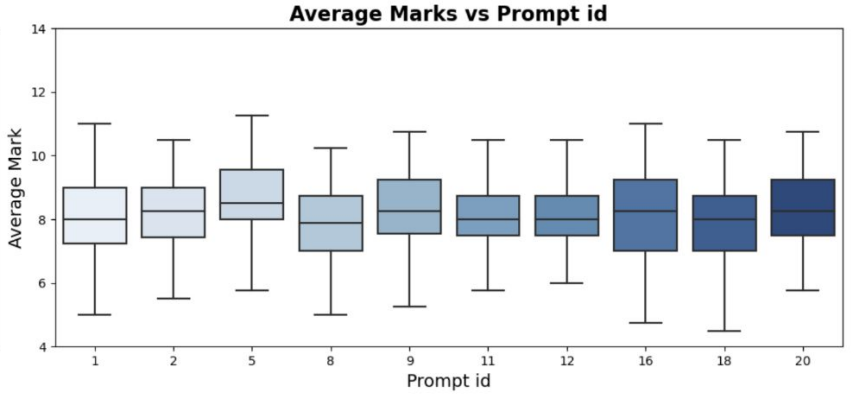
Data Analysis: Demographics - Native Language

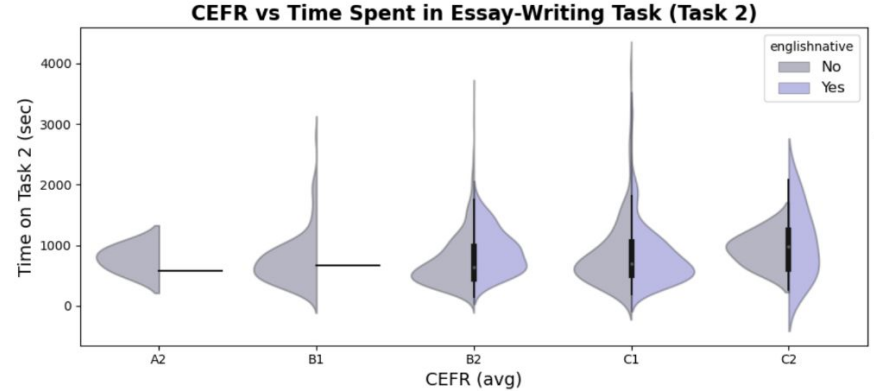


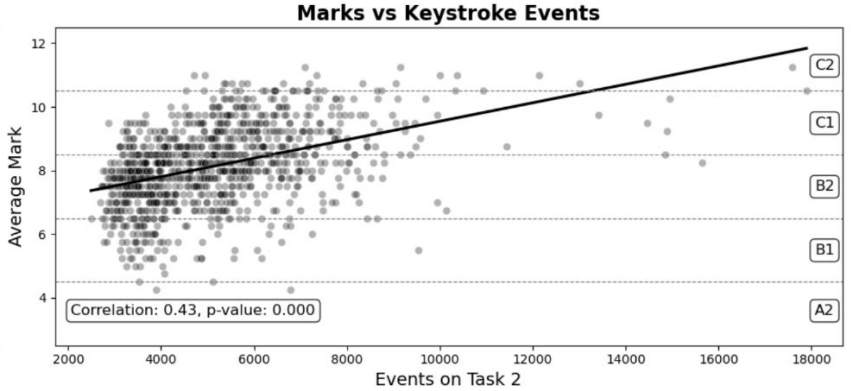
Polish was the most commonly reported native language among our participants, constituting 20% of the responses. It was closely followed by English and Portuguese, each comprising 17% of the total.

Data Analysis: Survey Completion Duration









Data Analysis: Survey Completion Duration

$$\text{K-S} = \sup_x |F_a(x) - F_b(x)|$$

We compared the **survey completion duration between native English speakers (NS) and non-native English speakers (NNS)** using the Kolmogorov-Smirnov (K-S) test.

The calculated K-S statistic was found to be 0.10, with a p-value of 0.10, suggesting **no significant difference** between NS and NNS in terms of the time required to complete the survey.

Data Analysis: Marks

Weak correlation between the **time spent on the essay task** and the **average mark received**, both from human markers and the automarker (**r=0.09**, $\rho = 0.003$)

Negative correlation between the **time they spent on the copy-text task** and the **average mark received on the essay-writing task**, suggesting that fast typists could generally achieve higher marks (**r=−0.26**, $\rho < 0.001$)

A strong correlation was evident between the **number of words** and the **average CEFR score** (**r=0.51**, $\rho < 0.001$), aligning with expectations based on previous studies

Conclusion & Future Work

- We introduce the new KUPA-KEYS dataset which includes keystroke data from 1,006 participants
- Participants completed two tasks: a copy-text task of 300 words, and an essay-writing task responding to one of 10 prompts
- The dataset also includes metadata about the individual participants, such as age, location, level of English and other languages known
- We annotated the essays with proficiency assessments from both human assessors and a pre-trained automarker
- We found a decent level of agreement amongst these assessors and initial analyses revealed some interesting results

Conclusion & Future Work

The dataset carries the potential for **further analyses of typing patterns**, indications of complex word and character sequences, and identification of **hierarchical structures in the writing process** per Ballier et al. (2019) and Leijten et al. (2019)

Conclusion & Future Work

We note that keystroke data can empower writing support for authors if we can successfully identify **when writers are struggling with linguistic constructions** (Conijn et al., 2021). This support could be in the form of writing suggestions, a chatbot or dictionary look up tools.

Conclusion & Future Work

Other future work includes the use of **features** derived from keystroke data to **enhance Transformer-based assessment models** (Mizumoto and Eguchi, 2023)

Conclusion & Future Work

Finally, the continuing challenge of **generative AI text detection** is acknowledged in recent literature (Krishna et al., 2023; Sadasivan et al., 2023), signifying a necessity for increased endeavour in this domain.

Our exploration may potentially facilitate the advent of novel research towards **generative-AI text detection** based on **event-based information including keystrokes**.

References

Pin Shen Teh, Andrew Beng Jin Teoh, Connie Tee, and Thian Song Ong. 2011. A multiple layer fusion approach on keystroke dynamics. *Pattern Analysis and Applications*, 14(1):23–36.

Ivor Timmis. 2002. Native-speaker norms and International English: a classroom view. *ELT Journal*, 56:240–249.

Vishaal Udandarao, Mohit Agrawal, Rajesh Kumar, and Rajiv Ratn Shah. 2020. On the inference of soft biometrics from typing patterns collected in a multi-device environment. In *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*, pages 76–85.

Esra Vural, Jiaju Huang, Daqing Hou, and Stephanie Schuckers. 2014. Shared research dataset to support development of keystroke authentication. In *IEEE International joint conference on biometrics*, pages 1–8. IEEE.

Helen Yannakoudakis and Ronan Cummins. 2015. Evaluating the performance of automated text scoring systems. In *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 213–223.

Mo Zhang, Jiangang Hao, Chen Li, and Paul Deane. 2016. Classification of writing patterns using keystroke logs. In *Quantitative psychology research*, pages 299–314. Springer.

Alejandro Acien, Aythami Morales, Ruben Vera-Rodriguez, Julian Fierrez, and John V. Monaco. 2020. TypeNet: Scaling up keystroke biometrics. In *2020 IEEE International Joint Conference on Biometrics (IJCB)*.

Svenja Adolphs. 2005. "I don't think I should learn all this" – a longitudinal view of attitudes towards 'native speaker' English. In Claus Gnutzmann and Frauke Intemann, editors, *The globalisation of English and the English language classroom*, pages 119–131. Tübingen: Narr Verlag.

Jeffrey D Allen. 2010. *An analysis of pressure-based keystroke dynamics algorithms*. Ph.D. thesis, Southern Methodist University.

Cem Alptekin. 2002. Towards intercultural communicative competence in ELT. *ELT Journal*, 56:57–64.

Veerle M Baaijen and David Galbraith. 2018. Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, 36(3):199–223.

Yukino Baba and Hisami Suzuki. 2012. How are spelling errors generated and corrected? a study of corrected and uncorrected spelling errors using keystroke logs. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 373–377, Jeju Island, Korea. Association for Computational Linguistics.

Nicolas Ballier, Erin Pacquetet, and Taylor Arnold. 2019. Investigating Keylogs as Time-Stamped Graphemics. In *Proceedings of Graphemics in the 21st Century, Brest 2018*, pages 353–365.

Ritwik Banerjee, Song Feng, Jun Seok Kang, and Yejin Choi. 2014. Keystroke patterns as prosody in digital writings: A case study with deceptive reviews and essays. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1469–1473.

Evgeny Chukharev-Hudilainen. 2019. Empowering automated writing evaluation with keystroke logging. In *Observing writing*, pages 125–142. Brill.

Evgeny Chukharev-Hudilainen, Aysel Saricaoglu, Mark Torrance, and Hui-Hsien Feng. 2019. Combined deployable keystroke logging and eyetracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, 41(3):583–604.

Rianne Conijn, Emily Dux Speltz, and Evgeny Chukharev-Hudilainen. 2021. Automated extraction of revision events from keystroke data. *Reading and Writing*, pages 1–26.

Vivian Cook. 1999. Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33:185–209.

Scott A. Crossley, Jennifer L. Weston, Susan T. McLain Sullivan, and Danielle S. McNamara. 2011. The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication*, 28(3):282–311.

Vivek Dhakal, Anna Maria Feit, Per Ola Kristensson, and Antti Oulasvirta. 2018. Observations on typing from 136 million keystrokes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Anna Maria Feit, Daryl Weir, and Antti Oulasvirta. 2016. How we type: Movement strategies and performance in everyday typing. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 4262–4273.

Julian Fierrez, Javier Galbally, Javier Ortega-Garcia, Manuel R Freire, Fernando Alonso-Fernandez, Daniel Ramos, Doroteo Torre Toledano, Joaquin Gonzalez-Rodriguez, Juan A Siguenza, Javier Garrido-Salas, et al. 2010. Biosecurid: a multimodal biometric database. *Pattern Analysis and Applications*, 13(2):235–246.

David Galbraith and Veerle Baaijen. 2019. Aligning keystrokes with cognitive processes in writing. In E. Lindgren and K. Sullivan, editors, *Observing writing: Insights from keystroke logging and handwriting*, pages 306–325. Leiden: Brill.

Romain Giot, Mohamad El-Abed, and Christophe Rosenberger. 2009. GREYC keystroke: a benchmark for keystroke dynamics biometric systems. In *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6.

Romain Giot, Mohamad El-Abed, and Christophe Rosenberger. 2012. Web-based benchmark for keystroke dynamics biometric systems: A statistical analysis. In *2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 11–15. IEEE.

Nahuel González, Enrique P. Calot, Jorge S. Ierache, and Waldo Hasperué. 2021. On the shape of timings distributions in free-text keystroke dynamics profiles. *Heliyon*, 7(11):e08413.

Adam Goodkind and Andrew Rosenberg. 2015. Muddying the multiword expression waters: How cognitive demand affects multiword expression production. In *Proceedings of the 11th Workshop on Multiword Expressions*, pages 87–95, Denver, Colorado. Association for Computational Linguistics.

Daniele Gunetti and Claudia Picardi. 2005. Keystroke analysis of free text. *ACM Transactions on Information and System Security (TISSEC)*, 8(3):312–347.

Kilem Gwet. 2002. Inter-rater reliability: dependency on trait prevalence and marginal homogeneity. *Statistical Methods for Inter-Rater Reliability Assessment Series*, 2:1–9.

Syed Zulkarnain Syed Idrus, Estelle Cherrier, Christophe Rosenberger, and Patrick Bours. 2013. Soft biometrics database: A benchmark for keystroke dynamics biometric systems. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, pages 1–8. IEEE.

Pilsung Kang and Sungzoon Cho. 2015. Keystroke dynamics-based user authentication using long and free text strings from various input devices. *Information Sciences*, 308:72–93.

M. Karnan, M. Akila, and N. Krishnaraj. 2011. Biometric personal authentication using keystroke dynamics: A review. *Applied Soft Computing*, 11(2):1565–1573. The Impact of Soft Computing for the Progress of Artificial Intelligence.

Zixuan Ke and Vincent Ng. 2019. Automated essay scoring: A survey of the state of the art. In *IJCAI*, volume 19, pages 6300–6308.

Kalpesh Krishna, Yixiao Song, Marzena Karpinska, John Wieting, and Mohit Iyyer. 2023. Paraphrasing evades detectors of AI-generated text, but retrieval is an effective defense. arXiv:2303.13408.

John Leggett, Glen Williams, Mark Usnick, and Mike Longnecker. 1991. Dynamic identity verification via keystroke characteristics. *International Journal of Man-Machine Studies*, 35(6):859–870.

Mariëlle Leijten, Eric Van Horenbeeck, and Luuk Van Waes. 2019. Analysing keystroke logging data from a linguistic perspective. In E. Lindgren and K. Sullivan, editors, *Observing writing: Insights from keystroke logging and handwriting*. Leiden: Brill.

Hilbert Locklear, Sathya Govindarajan, Zdeňka Sitová, Adam Goodkind, David Guy Brizan, Andrew Rosenberg, Vir V. Phoha, Paolo Gasti, and Kiran S. Balagani. 2014. Continuous authentication with cognition-centric text production and revision features. In *IEEE International Joint Conference on Biometrics*.

Cerstin Mahlow. 2015. Learning from Errors: Systematic Analysis of Complex Writing Errors for Improving Writing Technology. In *Language Production, Cognition, and the Lexicon*, pages 419–438.

Cerstin Mahlow, Malgorzata Anna Ulasik, and Don Tuggener. 2022. Extraction of transforming sequences and sentence histories from writing process data: a first step towards linguistic modeling of writing. *Reading and Writing*, pages 1–40.

Emanuele Maiorana, Himanka Kalita, and Patrizio Campisi. 2021. Mobile keystroke dynamics for biometric recognition: An overview. *IET Biometrics*, 10(1):1–23.

Danielle S. McNamara, Scott A. Crossley, Rod D. Roscoe, Laura K. Allen, and Jianmin Dai. 2015. A hierarchical classification approach to automated essay scoring. *Assessing Writing*, 23:35–59.

Atsushi Mizumoto and Masaki Eguchi. 2023. Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2):100050.

John V Monaco, Ned Bakelman, Sung-Hyuk Cha, and Charles C Tappert. 2012. Developing a keystroke biometric system for continual authentication of computer users. In *2012 European Intelligence and Security Informatics Conference*, pages 210–216.

John V Monaco, Gonzalo Perez, Charles C Tappert, Patrick Bours, Soumik Mondal, Sudalai Rajkumar, Aythami Morales, Julian Fierrez, and Javier Ortega-Garcia. 2015. One-handed keystroke biometric identification competition. In *2015 International Conference on Biometrics (ICB)*, pages 58–64. IEEE.

John V Monaco, John C Stewart, Sung-Hyuk Cha, and Charles C Tappert. 2013. Behavioral biometric verification of student identity in online course assessment and authentication of authors in literary works. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8.

John V. Monaco and Charles C. Tappert. 2017. Obfuscating keystroke time intervals to avoid identification and impersonation. arXiv:1609.07612.

Jugurta R Montalvão Filho and Eduardo O Freire. 2006. On the equalization of keystroke timing histograms. *Pattern Recognition Letters*, 27(13):1440–1446.

Mozilla Foundation. 2023. Mozilla WebDocs: High precision timing. https://developer.mozilla.org/en-US/docs/Web/API/Performance_API/High_precision_timing. Accessed: 2023-10-17.

Christopher Murphy, Jiaju Huang, Daqing Hou, and Stephanie Schuckers. 2017. Shared dataset on natural human-computer interaction to support continuous authentication research. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 525–530. IEEE.

Alen Peacock, Xian Ke, and Matthew Wilkerson. 2004. Typing patterns: A key to user identification. *IEEE Security and Privacy*, 2(5):40–47.

Robert Phillipson. 1992. *Linguistic Imperialism*. Oxford University Press, Oxford, UK.

Barbara Plank. 2016. Keystroke dynamics as signal for shallow syntactic parsing. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 609–619.

Barbara Plank. 2018. Predicting authorship and author traits from keystroke dynamics. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 98–104, New Orleans, Louisiana, USA. Association for Computational Linguistics.

Dadi Ramesh and Suresh Kumar Sanampudi. 2022. An automated essay scoring systems: a systematic literature review. *Artificial Intelligence Review*, 55(3):2495–2527.

Giorgio Roffo, Cinzia Giorgetta, Roberta Ferrario, Walter Riviera, and Marco Cristani. 2014. Statistical analysis of personality and identity in chats using a keylogging platform. In *Proceedings of the 16th International Conference on Multimodal Interaction*, pages 224–231.

Vinu Sankar Sadasivan, Aounon Kumar, Sriram Balasubramanian, Wenxiao Wang, and Soheil Feizi. 2023. Can AI-generated text be reliably detected? arXiv:2303.11156.

Moritz Jonas Schaeffer, Michael Carl, Isabel Lacruz, and Akiko Aizawa. 2016. Measuring cognitive translation effort with activity units. In *Proceedings of the 19th Annual Conference of the European Association for Machine Translation*, pages 331–34195.

John C Stewart, John V Monaco, Sung-Hyuk Cha, and Charles C Tappert. 2011. An investigation of keystroke and stylometry traits for authenticating online test takers. In *2011 International Joint Conference on Biometrics (IJCB)*, pages 1–7.

Giuseppe Stragapede, Paula Delgado-Santos, Ruben Tolosana, Ruben Vera-Rodriguez, Richard Guest, and Aythami Morales. 2023. Mobile keystroke biometrics using transformers. In *2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*.

Yan Sun, Hayreddin Ceker, and Shambhu Upadhyaya. 2016. Shared keystroke dataset for continuous authentication. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6. IEEE.

Charles C Tappert, Mary Villani, and Sung-Hyuk Cha. 2010. Keystroke biometric identification and authentication on long-text input. In *Behavioral biometrics for human identification: Intelligent applications*, pages 342–367.

Thank you!

any questions? please email helen.yannakoudakis@kcl.ac.uk or nicolas.ballier@u-paris.fr or andrew.caines@cl.cam.ac.uk

This work was supported by a research grant from Université Paris Cité and King's College London, under the ANR grant ANR-18-IDEX-0001, Financement IdEx Université de Paris.
The ALTA Institute is supported by Cambridge University Press & Assessment.
Thanks to Ece Washbrook, Russell Moore, Souradj Mounien Dit Ravi, Øistein Andersen, Mark Elliott, and ELiT (English Language iTutoring).