

# ReadLet: an ICT platform for the assessment of reading efficiency in early graders

Marcello Ferro<sup>1</sup>, Claudia Cappa<sup>2</sup>, Sara Giulivi<sup>3</sup>, Claudia Marzi<sup>1</sup>, Franco Alberto Cardillo<sup>1</sup>, Vito Pirrelli<sup>1</sup>

<sup>1</sup>Institute for Computational Linguistics - National Research Council of Italy (ILC CNR) - [marcello.ferro@ilc.cnr.it](mailto:marcello.ferro@ilc.cnr.it)

<sup>2</sup>Institute for Clinical Physiology - National National Research Council of Italy (IFC CNR) - [claudia.cappa@cnr.it](mailto:claudia.cappa@cnr.it)

<sup>3</sup>Training and Learning Department - University of Applied Sciences and Arts of Southern Switzerland (DFA SUPSI) - [sara.giulivi@supsi.ch](mailto:sara.giulivi@supsi.ch)



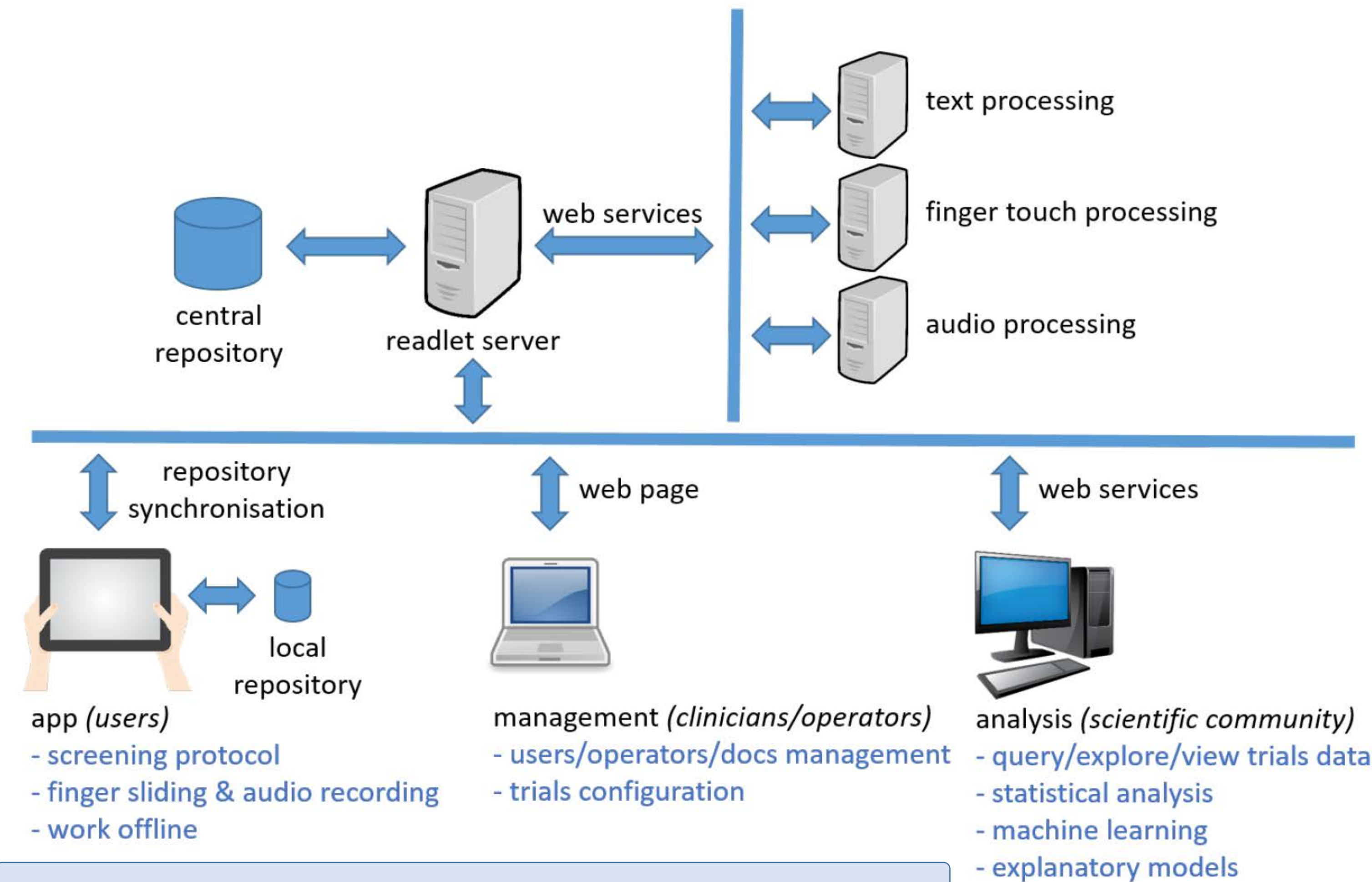
**ReadLet is an ICT platform specifically designed to provide accurate, evidence-based assessment of reading efficiency in early grade children. It is intended to offer an ecological, non-invasive protocol for extensive data elicitation, storage and analysis.**

**Background** - Unlike learning to speak, which develops spontaneously and nearly effortlessly, learning to read requires conscious effort, dedication, focused attention, systematic instruction and corrective error feedback. We still know comparatively little about the basic mechanisms involved in learning to read fluently and efficiently. Besides, assessing reading skills is an extremely laborious and time-consuming task, which requires monitoring a variety of interlocked abilities, ranging from accurate word rendering, word-in-context reading fluency and lexical access, to linguistic comprehension [1], [2], and interpretation, management and inference of complex events in working memory [3], [4].

**Objective** - We intend to develop a "Reading Efficiency Model" (REM), which combines decoding accuracy and automaticity (fluency) with reading comprehension. This will be measured by the so-called "Reading Efficiency Parameter" [5] through individual tablet based test sessions that combine assessment of fluency and reading comprehension. Aspects of REM will be validated with ReadLet acquired data, and will be simulated with machine learning models.

## ReadLet: A platform for reading efficiency evaluation

- Centralized server (storage, management, configuration, text/touch/audio post-processing)
- Grant security and privacy
- Low cost and distributable endpoint (commercial tablet device)
- Web services for data access (operators, clinicians, researchers, 3<sup>rd</sup> party apps).



## Data post-processing

- Finger- and audio-to-text alignment over time
- Text linguistic annotation
- Token/n-grams frequency (SUBTLEX-IT [6])
- Part-of-speech tagging and readability analysis (READ-IT [7])
- Comprehension profile: questionnaire accuracy and timing
- Decoding profile: time spent on the whole text and on each letter/token/sentence/paragraph/page

## High-level processing

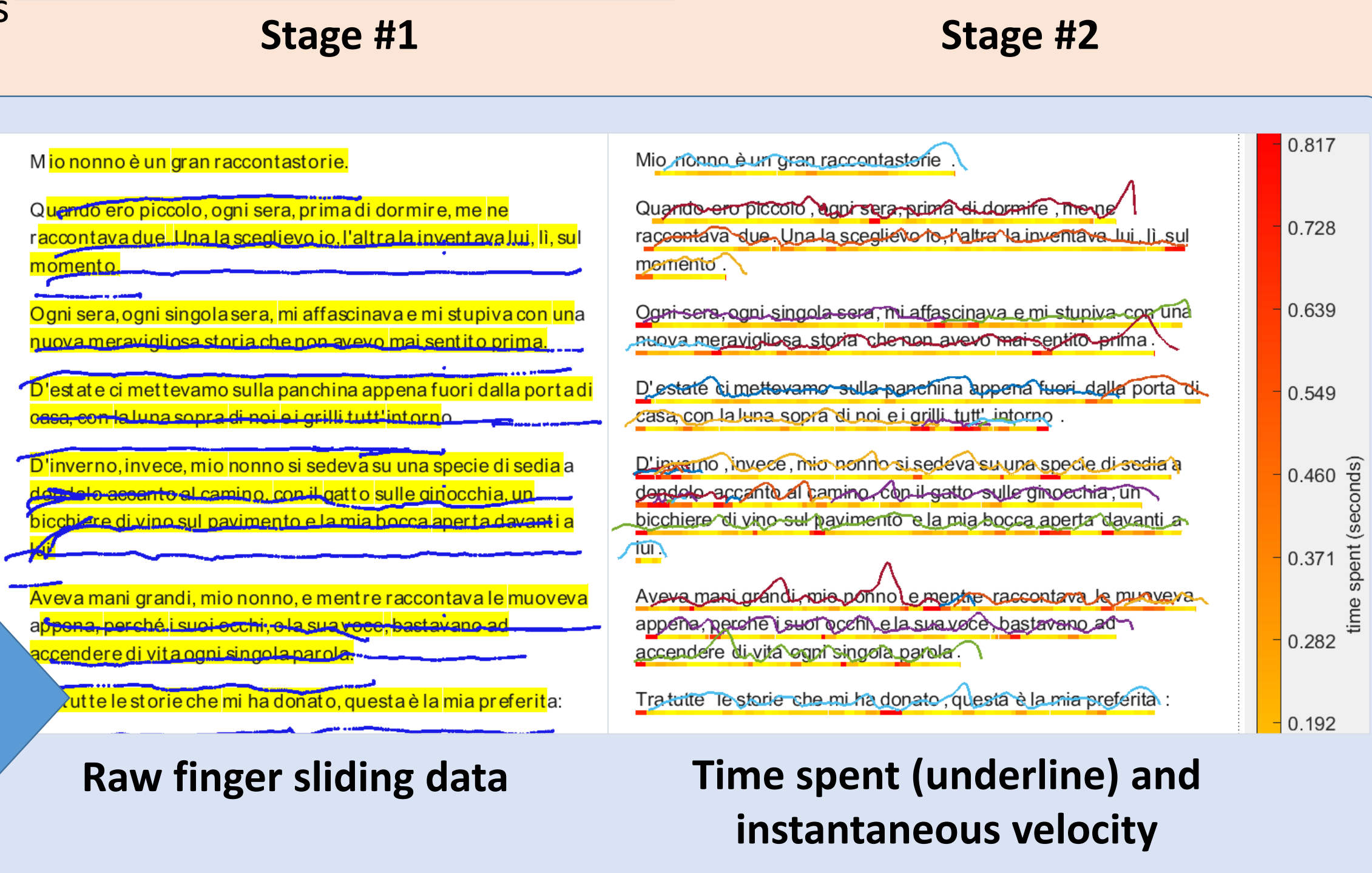
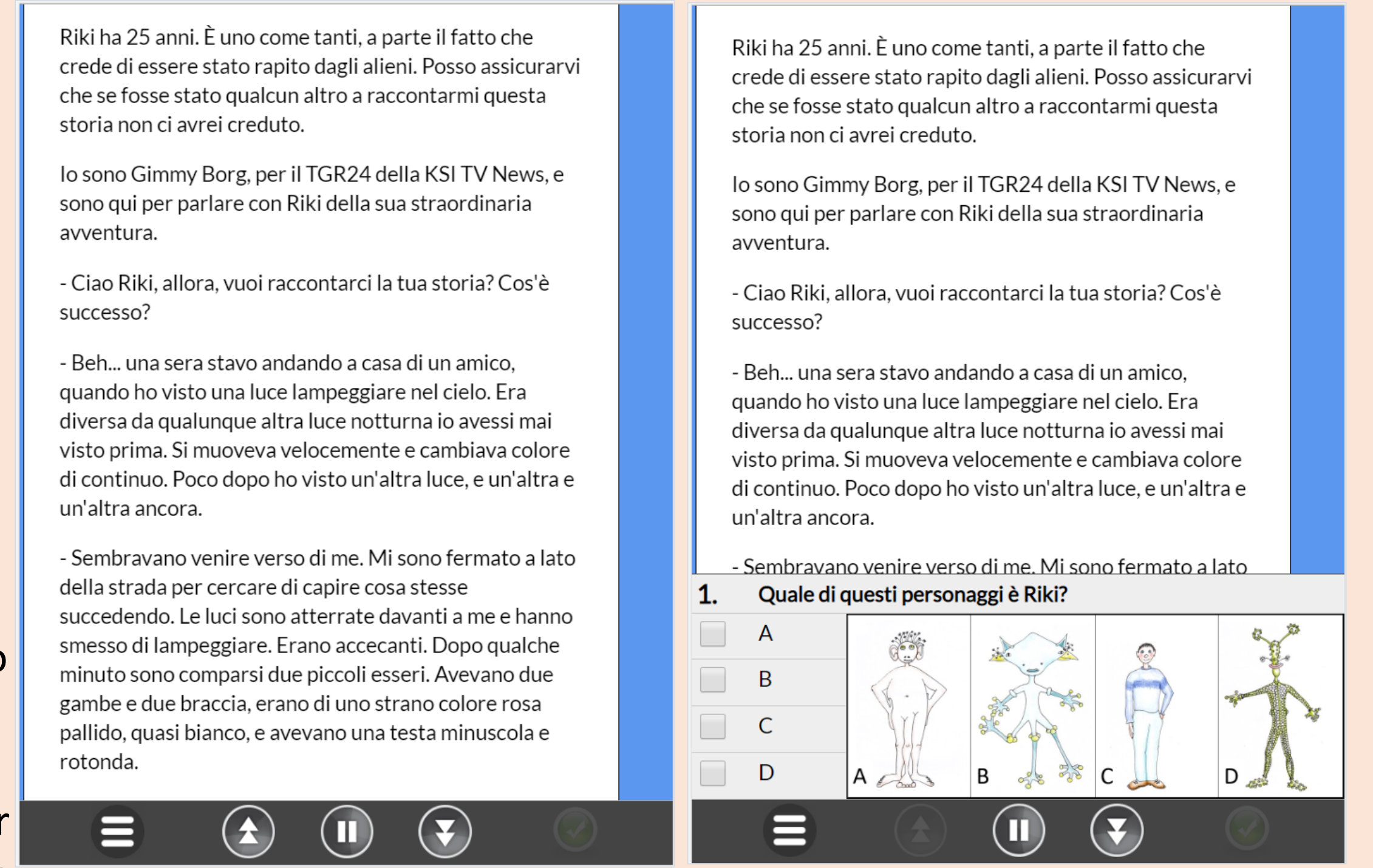
By anchoring finger sliding data on written text, we can associate sliding speed fluctuations with annotated linguistic structures, and with speed, rhythm and prosodic contours of the acoustic signal.

## Protocol

- Stage#1: Text reading (decoding efficiency)
  - 2-4 pages short story
  - Silent reading vs reading aloud
  - Words vs nonwords
- Stage#2: Questionnaire (comprehension efficiency)
  - ~15 questions (tagged, possibly hyperlinked to text)

## Data acquisition

- Device-dependent text layout (used to exactly reconstruct the text as it was read by the child)
- Touchscreen timestamped data (finger sliding): 60-120Hz (12-24 touch events when reading at 5 syllables per second)
- Microphone audio data (voice): 48KHz, 16bit, stereo, 128kbps MP3
- Questionnaire timestamped filling data

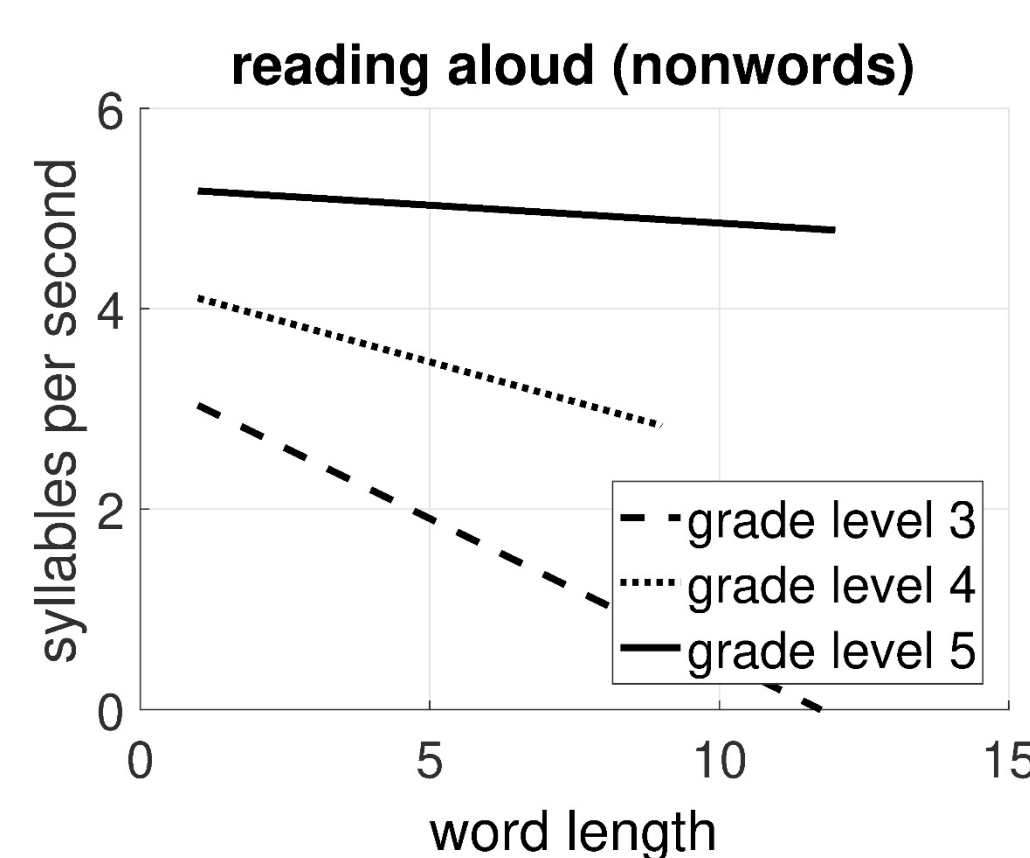


## Measurement campaign

- ~150 children
- ~200 Italian speaking area of Switzerland
- ~200 Tuscany, Italy
- ~15 Fez, Morocco (French and Arabic languages)
- ~50% F/M
- 3<sup>rd</sup> -5<sup>th</sup> grade level
- ~500 trials
- 1/3 listening (3 pages) + questionnaire (15 questions)
- 1/3 reading aloud (1 word page + 1 nonword page)
- 1/3 silent reading (3 pages) + questionnaire (15 questions)
- 15 stories and questionnaires (9 Italian, 3 French, 3 Arabic)

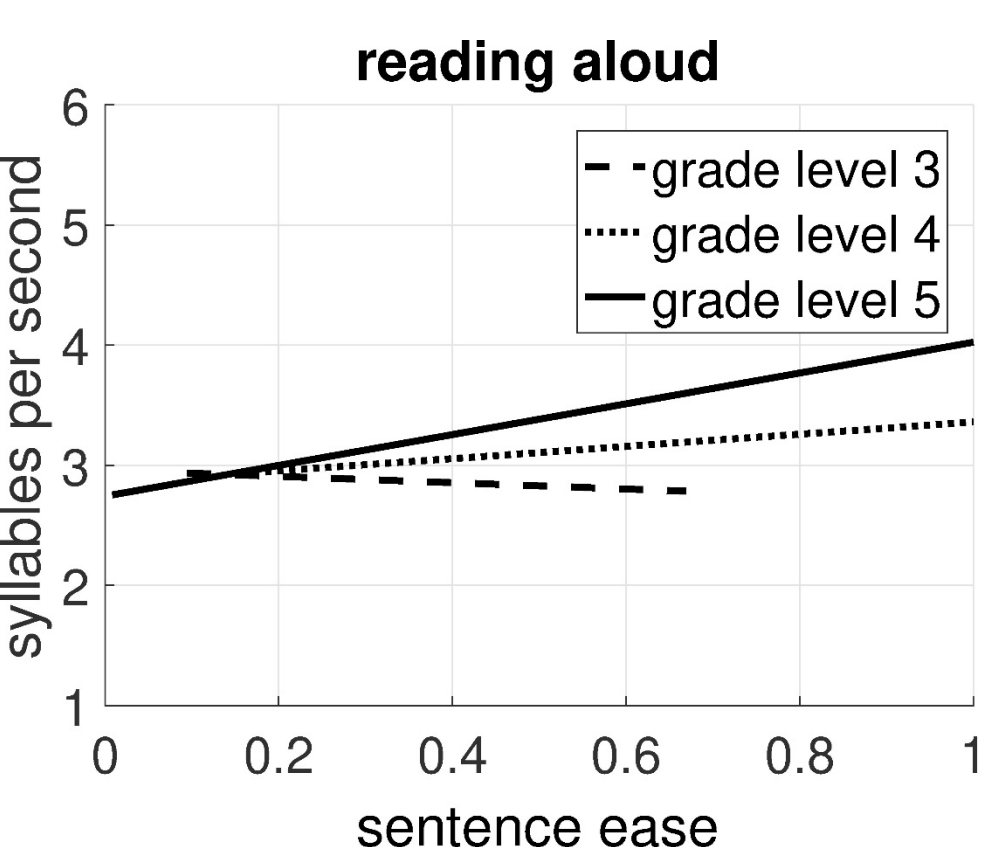
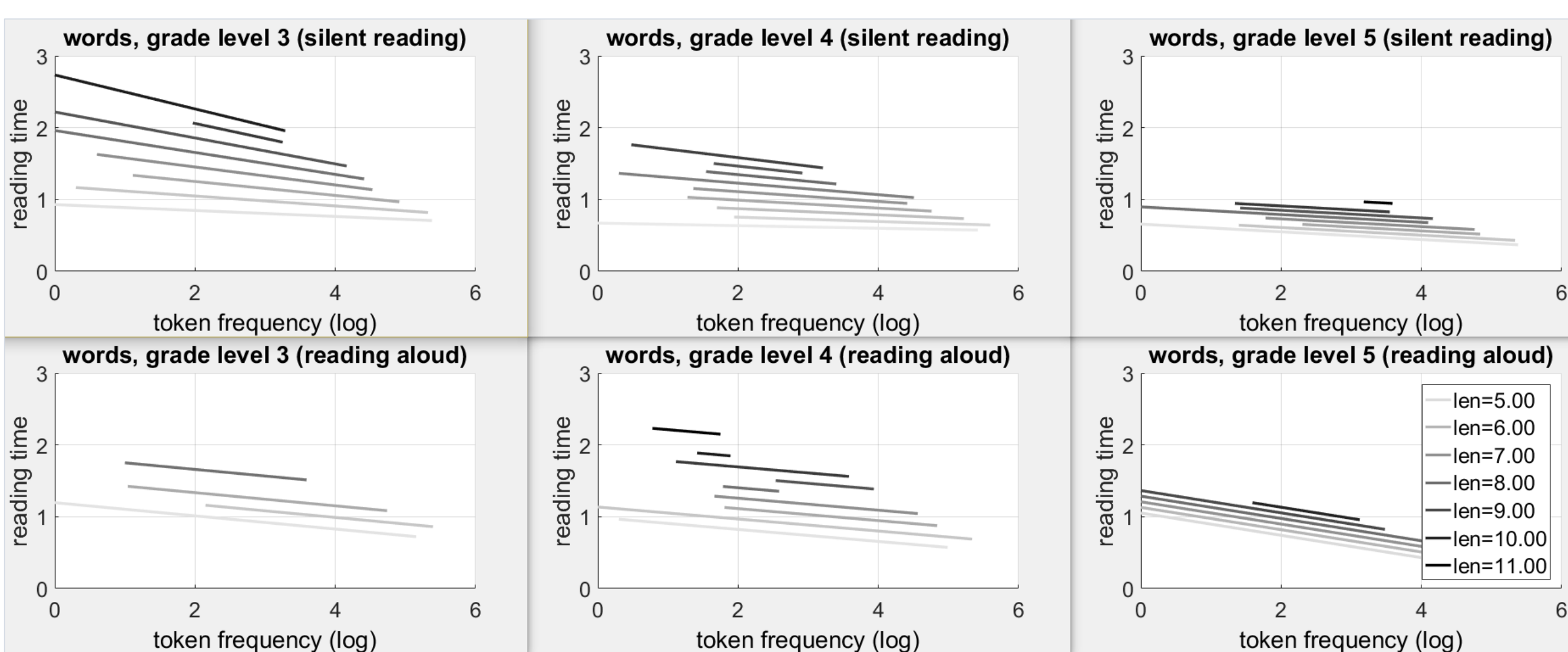
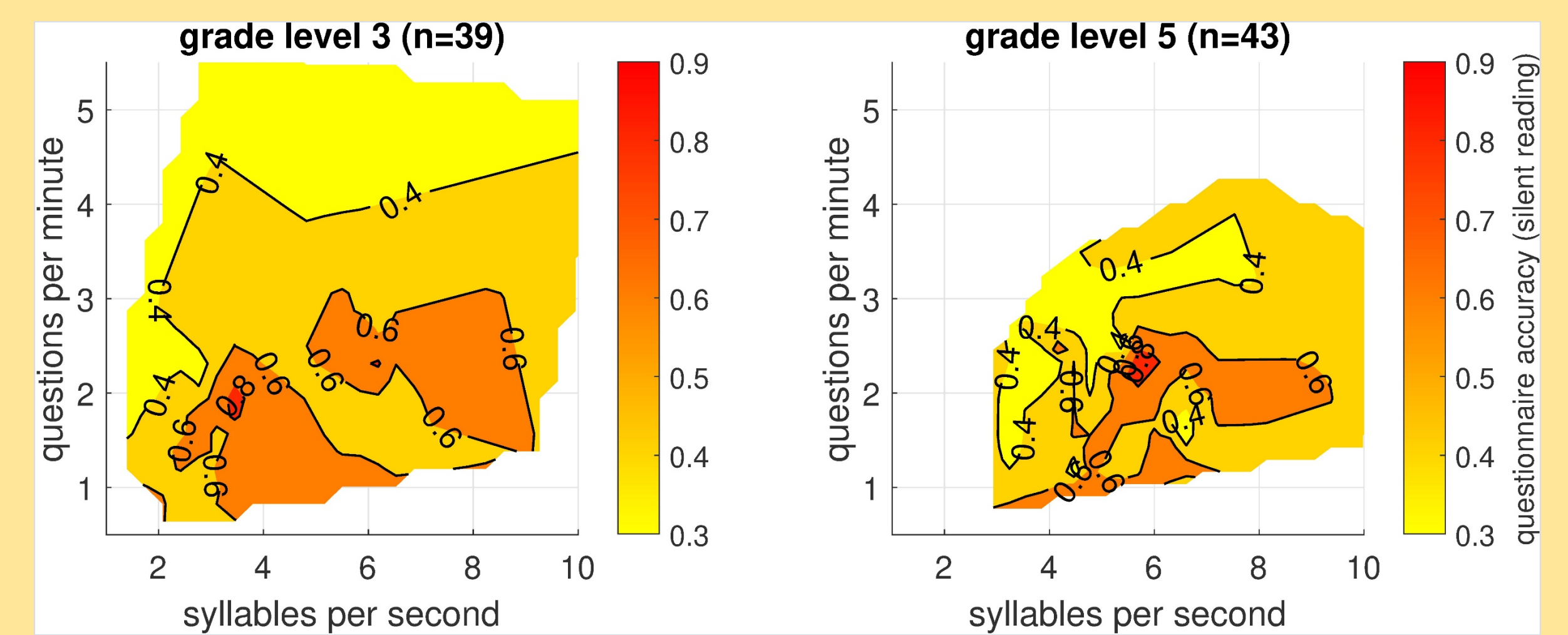
LME modeling shows finger sliding time significantly ( $p < 0.0001$ ) correlates with:

- Word length (+)
- Token frequency (-)
- Grade level (-)
- Reading aloud (+) / silent reading (-)
- Nonwords (+) / words (-)
- Text reading ease (-) ( $p < 0.01$ )



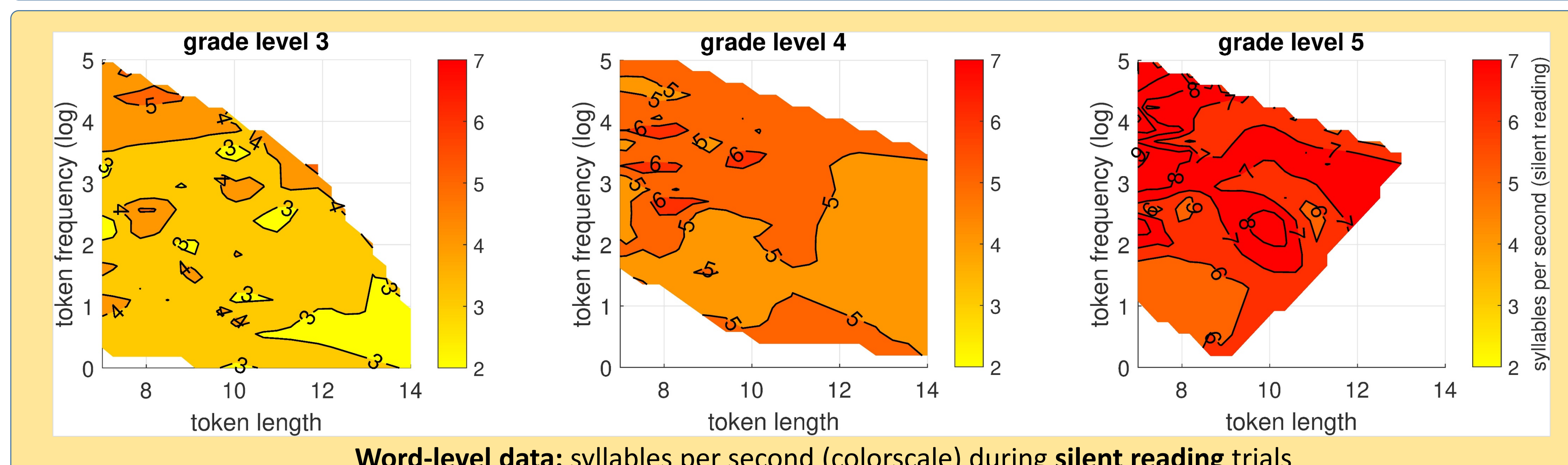
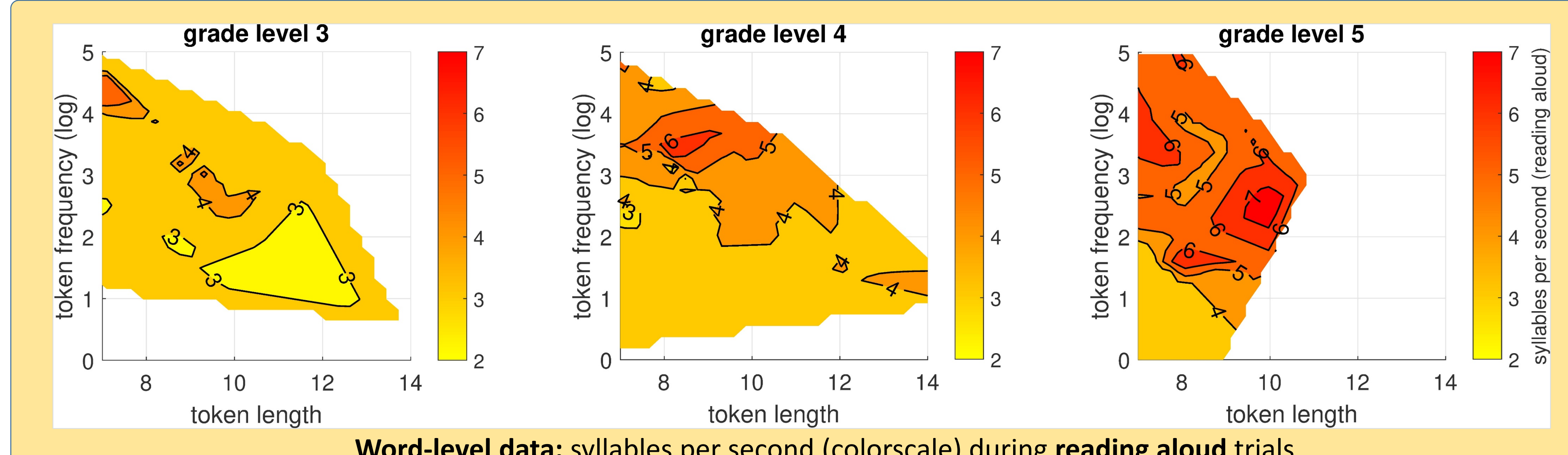
## Document-level data

- High performance areas found in the {questionnaire accuracy; syllables per second; questions per minute} space
- Low-speed areas (left-bottom) as well as high-speed areas (right-top) predicts low comprehension accuracy



## Concluding remarks

We believe that technology cannot and should not supplant the role and professional judgement of teachers and therapists in helping children with reading and learning difficulties. Having said that, ReadLet protocols and tools can effectively support daily decisions and education/intervention management, and offer an example of effective introduction of adequate ICT tools into school curricula and daily teaching.



## References

- L. E. Cutting and H. S. Scarborough, "Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured", Scientific studies of reading, vol. 10, no. 3, pp. 277-299, 2006.
- R. Padovani, "La comprensione del testo scritto in età scolare. una rassegna sullo sviluppo normale e atipico", Psicologia clinica dello sviluppo, vol. 10, no. 3, pp. 369-398, 2006.
- M. A. Gernsbacher and M. Faust, "Skilled suppression", in Interference and inhibition in cognition. Elsevier, 1995, pp. 295-327.
- K. Nation and M. J. Snowling, "Factors influencing syntactic awareness skills in normal readers and poor comprehenders", Applied psycholinguistics, vol. 21, no. 2, pp. 229-241, 2000.
- C. Cappa, S. Giulivi, and C. Muzio, "L'efficienza di lettura: l'integrazione dell'abilità di comprensione del testo con quella della velocità di lettura", in Proceedings of the XXV AIRIPA National Conference, 2016.
- D. Crepaldi, S. Amenta, P. Manderla, E. Keuleers, and M. Brysbaert, "SUBTLEX-IT. Subtitle-based word frequency estimates for Italian", in Proceedings of Annual Meeting of the Italian Association for Experimental Psychology, Rovereto, 10-12 September 2015.
- F. Dell'Orletta, S. Montemagni, and G. Venturi, "Read-it: Assessing readability of Italian texts with a view to text simplification", in Proceedings of the second workshop on speech and language processing for assistive technologies. Association for Computational Linguistics, 2011, pp. 73-83.