

# Does finger-tracking point to child reading strategies?

**Claudia Marzi**  
ILC - CNR Pisa

claudia.marzi@ilc.cnr.it

**Anna Rodella**  
Università di Pisa

a.rodella@studenti.unipi.it

**Andrea Nadalini**  
ILC - CNR Pisa

andrea.nadalini@ilc.cnr.it

**Loukia Taxitari**  
ILC - CNR Pisa

loukia.taxitari@ilc.cnr.it

**Vito Pirrelli**  
ILC - CNR Pisa

vito.pirrelli@ilc.cnr.it

## Abstract

The movement of a child’s index finger that points to a printed text while (s)he is reading may provide a proxy for the child’s eye movements and attention focus. We validated this correlation by showing a quantitative analysis of patterns of “finger-tracking” of Italian early graders engaged in reading a text displayed on a tablet. A web application interfaced with the tablet monitors the reading behaviour by modelling the way the child points to the text while reading. The analysis found significant developmental trends in reading strategies, marking an interesting contrast between typically developing and atypically developing readers.

## 1 Introduction

Recent experimental evidence in visual perception analysis (Lio et al., 2019) shows that eye movements and finger movements strongly correlate during scene exploration, at both individual and group levels. In Lio *et al.*’s (2019) experiment, subjects are invited to explore a blurred image displayed on a touchscreen by moving their fingers on the display. Picture areas that are located immediately above the touch point of the subject’s finger on the screen are automatically shown in high resolution, thereby simulating the subject’s central (foveal) vision. The experiment proves that the subjects’ image-exploring patterns in the two modalities (optical and tactile) are highly congruent. The result is not surprising. A familiar context which exploits this synergistic behaviour is when

children are learning to read. Despite the undoubtedly different dynamics of the two types of text exploration, finger-pointing to text helps children learn to look at print, and supports critical early reading behaviours: directional movement, attention focus, and voice-print match (Mesmer and Lake, 2010; Uhry, 2002).

ReadLet (Ferro et al., 2018a; Ferro et al., 2018b) is a web application with a tablet front-end, designed to support online monitoring of silent and oral reading abilities through *finger-tracking*. Finger-tracking consists of recording the time series of touch events on the tablet screen where a child is reading a short story, while the child is pointing to the text with the index finger of her dominant hand.<sup>2</sup> Preliminary analyses of our finger-tracking data (Pirrelli et al., 2020) highlighted a diminishing influence of word frequency and word length on reading time as readers get older and more proficient (from 3<sup>rd</sup> to 6<sup>th</sup> grades). With increasing exposure to written words, differences in tracking time between high and low frequency words gradually tend to decrease, suggesting a ceiling effect in the entrenchment of both high- and low-frequency lexical representations in long-term memory (Zoccolotti et al., 2009). Similarly, word length was found to significantly interact across grades. Younger readers show increasing difficulty with longer words, with a steeper time increment for word length > 6, while older readers are slowed down when words are longer than 8 letters. This integrates previous evidence (De Luca et al., 2008), confirming that not even the most experienced readers can avoid the slowing down effect of word length.

The two-fold interaction of word frequency and word length with grade levels strongly suggests that Italian children use a *lexical route* to decod-

Copyright ©2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

<sup>2</sup><https://www.readlet.it/>: see section *publications* for materials related to this paper.

ing even at early stages of their reading development, despite the transparency of Italian orthography (Bates et al., 2001; Davies et al., 2013). It also suggests that young readers can make use of sublexical information, whenever they are confronted with words that are not contained in their orthographic lexicon. This developmental dynamic would account for the stronger sensitivity of less skilled readers to both lexical frequency and word length. As lexical information increases with age, with rarer and longer words finding their way into the reader’s orthographic lexicon, the reader makes an increasingly prominent use of lexical information and an increasingly sparser use of sublexical information.

In this paper, we provide a finer-grained quantitative analysis of the finger-tracking profile of typically developing readers, offering further evidence that their reading strategy results from an optimal, interactive combination of both lexical and sublexical information. The evidence is compared with the finger-tracking profile of difficult readers. To provide a more realistic developmental profile of these effects, we restrict our focus on nouns only, which are less likely to be skipped while finger-tracking, and present a narrower range of variability in both length and frequency.

## 2 ReadLet

ReadLet is a tablet-based application that combines an objective assessment of a child’s reading fluency and comprehension skills with careful collection of large-scale behavioural data, and quantitative modelling of the specific factors affecting reading development. It leverages an ICT infrastructure with a cloud-based back-end exposing a battery of web services acting as an interface between the central repository and the users. The ReadLet front-end is an ordinary tablet, where short stories are displayed for children to read, either silently or aloud. In both cases, the child is asked to finger-point to the text while reading. Texts are displayed on a 10” tablet screen in Lato font 17pt in black against a white background. During each reading session, the behaviour of the child is captured through large streams of time-aligned signals including voice recording, time-stamped finger-tracking patterns, reading time and question-answering time. Data are automatically captured and sent to a centralised server for post-

processing, where audio and finger-tracking time series are aligned with the text. Recorded and post-processed data are exposed through a set of web services offered by the cloud server.

## 3 The Data

For our present goal, we focus on reading data of 237 children, sampled from entire classes ranging from 3<sup>rd</sup> through to 6<sup>th</sup> school grades, in Italian and Italian-speaking Swiss schools.<sup>3</sup> Participants included both *typically* developing readers (N=214) and children screened and reported in schools as *atypically developing readers* (N=23), but who did not receive a clinical diagnosis. Eight short stories were created for the pilot study, one for each of the four school grades, and for each experimental condition (silent and aloud reading). Children were asked to read a story while finger-tracking the text. After reading in the silent condition only, children were asked a few multiple-choice questions, to ascertain they actually carried out the task.

Texts were automatically annotated for part-of-speech, word token frequency, and word typicality (measured as either the size of the word’s lexical neighbourhood, or *N*-size, or the mean Levenshtein distance from its top 20 neighbouring words, or *Old20* (Yarkoni et al., 2008)).

For each child, in both reading conditions, we calculated the *token tracking time* as the total time spent in finger-tracking each word token while reading. To ensure reliability and precision in the alignment of finger-tracking data with the text being read, we selected reading trials with  $\geq 75\%$  of finger-tracked text pages. From the original set of tokens making up the 8 short stories, we selected 97 lemmas for 109 noun tokens, by intersecting our data with age of acquisition and imageability assessments by Italian speakers (Montefinese et al., 2014; Montefinese et al., 2019). In the resulting data sample, word frequency<sup>4</sup> is observed

<sup>3</sup>Data come from a pilot data collection carried out by Claudia Cappa and Sara Giulivi in schools of Southern Tuscany and Canton of Ticino in 2018 and 2019 for the *AEREST* project, funded by the Department of Teaching and Learning of the University of Applied Sciences and Arts of Southern Switzerland. Data were encrypted and pseudonymized locally, to then be delivered through a *https* protocol to a centralized server, where they are accessed with authentication credentials and dedicated web services. A background history of children including information about any clinical and/or psychological disorders was provided by the schools.

<sup>4</sup>Token frequency is given as the natural logarithm of occurrences in the Subtlex corpus (<http://cr.ugent.be/subtlex->

to vary between  $\text{min}=5.61$  and  $\text{max}=11.77$ , and word length between 4 and 10 letters ( $\text{median}=5$ ,  $\text{mean}=5.62$ ,  $\text{sd}=1.40$ ).

#### 4 Typical and atypical reading development

The main goal of the *ReadLet* project is to propose and validate an ICT methodology for assessing the typical reading development of children in Italian schools. In this section we focus on the finger-tracking behaviour of typically developing children engaged in reading a short text. The idea is to provide evidence that finger-tracking patterns exhibit lexical effects that are well-established in the reading literature: namely word frequency, word length and word similarity (or *N-size*).<sup>5</sup>

Figure 1 shows the effects of word frequency across grade levels, in both aloud (left panel) and silent (right panel) reading, for typically developing readers. A linear mixed model fitting token tracking time as a function of reading type, word frequency and grade levels shows shorter tracking times in reading more frequent words. The model also highlights a significant interaction between years of schooling and word frequency, with facilitation effects getting smaller for older graders, particularly in silent reading. The difference in facilitation rate between the two reading tasks is not statistically significant.

Figure 2 compares the developmental patterns of token tracking time of typically (left panel) and atypically (right panel) developing readers, modelled as a linear function of word token frequency and grade level. The two patterns exhibit a clear facilitatory effect of token frequency on reading speed, confirming that frequency makes reading consistently easier for both typical and atypical populations of young readers, who appear to entertain the same *lexical* reading strategy. However, only in typically developing children the effect tends to diminish across grade levels, with slopes getting less steep as grade levels increase (Figure 2, left panel).<sup>6</sup>

A similar overall pattern is shown in Figure 3, where the sensitivity to word length of typical

it/) plus one. For our set of noun data the mean frequency is 9.45 ( $\text{sd}=1.61$ ).

<sup>5</sup>All figures in the section show regression plots of the interaction of main effects, using the *ggplot* function.

<sup>6</sup>Regression slopes for 4<sup>th</sup> and 5<sup>th</sup> grades are not statistically different from 3<sup>rd</sup> grade, but there is a significant difference when comparing slopes for 3<sup>rd</sup> and 6<sup>th</sup> grades.

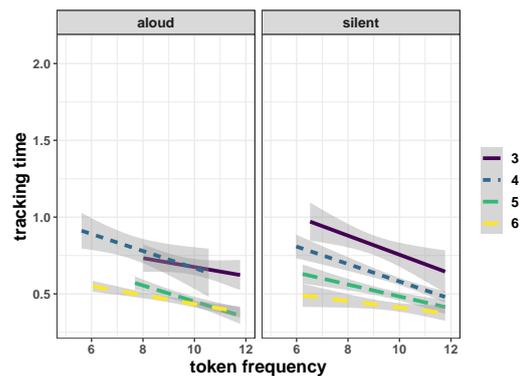


Figure 1: Linear interaction of token frequency and grade level (from 3<sup>rd</sup> to 6<sup>th</sup> level), for typically developing readers in both aloud and silent reading. Shaded areas indicate 95% confidence intervals.

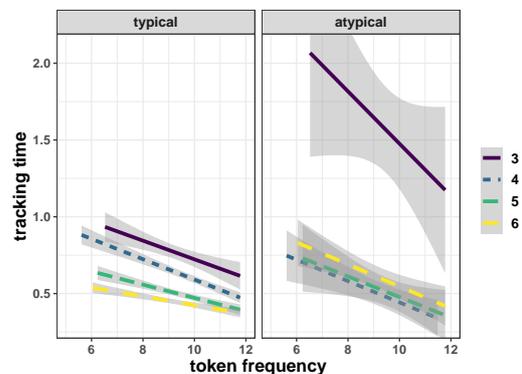


Figure 2: Linear interaction of token frequency and grade level (from 3<sup>rd</sup> to 6<sup>th</sup> level), for typically and atypically developing readers.

readers is contrasted with the same effect in atypical readers. In both groups, children take more time to read longer words, but only typically developing children exhibit a less strong sensitivity to word length as grade level increases. The statistical significance of this interaction disappears in atypical readers, with the only exception of 3<sup>rd</sup> graders, compared with all remaining graders.

Figure 4 shows how grade levels interact with *N-size* in affecting aloud (left panel) and silent (right panel) reading time. The dominant effect is facilitatory, with a clear incremental advantage in reading times for words with a high number of neighbours. Words are finger-tracked more quickly when they belong to more dense neighbourhoods, and this facilitatory effect is stronger for younger (3<sup>rd</sup> and 4<sup>th</sup> grade) than older (5<sup>th</sup>

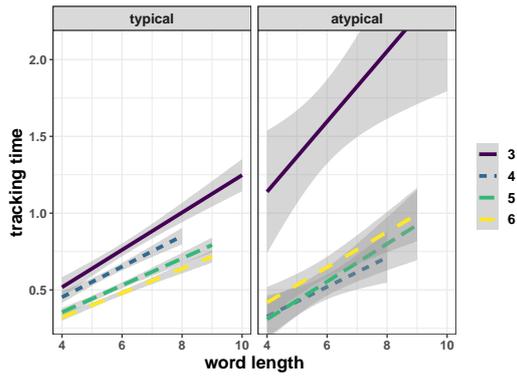


Figure 3: Linear interaction of word length and grade level (from 3<sup>rd</sup> to 6<sup>th</sup>) for typically and atypically developing readers.

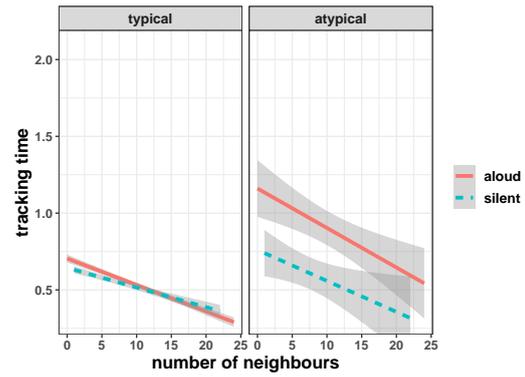


Figure 5: Linear interaction of  $N$ -size and reading task (aloud and silent) for typical and atypical readers.

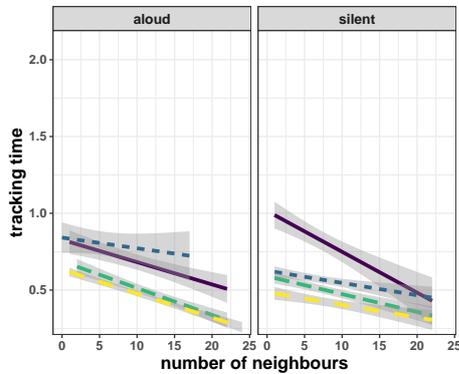


Figure 4: Linear interaction of  $N$ -size and grade level (from 3<sup>rd</sup> to 6<sup>th</sup>) for aloud and silent reading of typical readers.

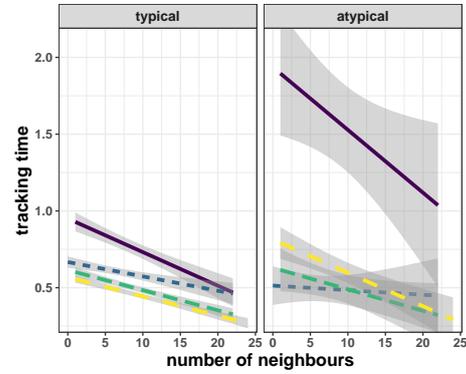


Figure 6: Linear interaction of  $N$ -size and grade level (from 3<sup>rd</sup> to 6<sup>th</sup>) for typically and atypically developing readers.

and 6<sup>th</sup> grade) readers. No significant difference is found in the interaction between reading type and  $N$ -size in typical readers (Figure 5, left panel). Atypical readers show equal slopes in both aloud and silent reading, but different intercepts, which capture the additional processing demands of concurrent articulation (Figure 5, right panel). This evidence suggests a *sublexical* reading strategy that relies on orthographic similarity: words that are not read lexically (because they are too long or less frequent), are read by decoding and combining the smaller parts they share with other neighbouring words. Fitting a mixed model with  $N$ -size, frequency ranges, and grade levels, as variables predicting the token tracking time, and with subjects as random effect, we find that all predictors and interactions are highly significant for typically developing children (Figure 6).

The behaviour of atypically developing read-

ers does not replicate the trend of typical readers. First, the tracking time of atypical readers is more strongly - and significantly - affected by token frequency, when compared with the typical tracking time of their age-matched peers. This is especially true for the youngest readers in our sample (3<sup>rd</sup> graders in the right plot of Figure 2). In addition, sensitivity to frequency appears to persist with age, as there are no significant differences in the facilitatory effect of frequency across later grade levels. This suggests a delay in developing and integrating lexical information. A nearly identical developmental pattern is replicated with  $N$ -size effects (Figure 6, right panel): younger children read words in denser neighbourhoods more easily, taking advantage of the recurrent sublexical parts shared by neighbouring words. Once more, no significant developmental pattern is observed across grade levels, as atypical readers do

not appear to be able to increasingly rely on lexical reading as they get more experienced (Figure 2). Finally, their reading time is persistently slowed down by longer words, suggesting a difficulty in memorizing and making them accessible through the lexical route (Figure 3, right panel).

## 5 General discussion

Facilitatory effects of lexical frequency on reading reaction time have been reported for Italian children (Barca et al., 2006; Burani et al., 2002) as well as adults (Barca et al., 2002; Burani et al., 2007). The effects are argued to reflect the working of the lexical route in dual-route models of reading (Coltheart et al., 2001): word items are accessed in the reader's orthographic lexicon, to then be pronounced after their full phonological code is retrieved. The faster reading of high-frequency lexical items thus reflects the well-established sensitivity of lexical access to word frequency. In Italian, the systematic nature of letter-to-sound mapping rules makes the operation of a sublexical reading strategy a reliable alternative to the lexical route. However effective, sublexical reading is nonetheless less efficient, since it requires the on-line, serial decoding of a word by its parts (e.g. *n*-grams or syllables). We conjecture that Italian children optimize reading efficiency at early stages of their reading practice, through dynamic integration of sublexical and lexical reading. Whenever possible, they resort to word-sized orthographic information in their lexicon (e.g. short and frequent words), and make it up for missing orthographic items through sublexical information.

Such an opportunistic strategy is in keeping with the idea that early readers strive, through reading practice, to “chunk” letter *n*-grams into longer orthographic units. Chunked units are stored and made accessible in the readers' lexicon, where they are associated with their fully specified phonological code. The length of stored items is a function of their frequency, and the reader's processing efficiency, reading practice and age. Our data confirm that this strategy is consistently adopted by typically developing readers in both silent and aloud reading, suggesting that the influence of lexical frequency is not confined to the retrieval and planning stage of the word phonological code, but appears to extend beyond response initiation, to affect full articulation of the code (Balota and Yap, 2006).

This strategy remains in operation through reading development, as shown by the decreasing tracking times of 6<sup>th</sup> graders as a function of lexical frequency (Figure 2, left panel). Nonetheless, the impact of word frequency is less strong in older readers, whose orthographic lexicon makes room for increasingly rarer (and longer) words. Also atypical readers appear to use a similar “chunking” strategy, but their developmental pattern fails to show a clear interaction between grade level and frequency. In Figure 2 (right panel), 3<sup>rd</sup> graders show a robust word frequency effect, but the diminishing role of frequency on tracking time across grades turns out not to be significant. This suggests that atypical readers have problems with developing orthographic representations for rarer (and longer) words, and they are not quite as successful as typical readers in optimally integrating lexical and sublexical information.

This interpretation is supported by the analysis of two other lexical effects on child reading development: word length and neighbourhood size (*N*-size). As expected, longer words elicit longer response latency and reading duration, but the effect is bigger for younger, typically developing readers compared to older ones (Figure 3, left panel), and for atypical readers compared to their age-matched peers (Figure 3, right panel). The use of sublexical information and serial *n*-gram decoding appears to be more prominent in younger and atypical readers than in the older and more skilled group of readers. Once more, the effect can be argued to reflect the absence of fully specified orthographic representations for longer words in the lexicon of less skilled readers, and a related difficulty in building up complex orthographic chunks.

Facilitatory effects of *N*-size on reading time are reported for atypical Italian readers by Marinelli et al. (2013), who, however, found no significant facilitation in age-matched typical readers. They argue that atypical readers over-rely on co-activation of word neighbours during reading to make it up for their poorly entrenched lexical representations. Conversely, access to individual lexical representations by typical readers is fast enough to make *N*-size effects hardly detectable. Our data are consistent with Marinelli et al.'s evidence, but integrate it in two important respects. First, the speeding-up influence of *N*-size is detected in both aloud and (for the first time to our knowledge) silent reading of Italian, with no

significant difference between the two (Figure 4 and Figure 5, left panel). This supports an interpretation of the  $N$ -size effect as having an impact on both phonological planning and overt articulation. Secondly, our data show that the effect is not limited to the reading pace of younger and atypical readers, as observed by Marinelli et al., but it also holds for typically developing readers (Figure 6), with an interesting modulation by grade level. This is mainly due to our focus on nouns, which include longer and less frequent words, for which  $N$ -size effects are known to be stronger and easier to detect (Davies et al., 2013). Finally, the diminishing impact of  $N$ -size for increasing grade levels confirms a sparser use of the sublexical route by more skilled readers, who are equipped with a richer and more efficient orthographic lexicon.

To sum up, typical and atypical readers alike strive to optimally integrate lexical and sublexical input patterns while reading, using the former whenever possible for efficient decoding, and the latter as a fall-back strategy, whenever the lexical route fails. This dynamic, however straightforward, has non-trivial consequences. In a developmental perspective, the orthographic lexicon gets richer with practice, boosted by an age-driven improvement of children's global ability in information processing (Zoccolotti et al., 2009), which makes longer and rarer words easier to store. As a result, the dynamic balance is shifted towards lexical reading. Conversely, atypical readers find it more difficult to develop and store detailed mappings between orthographic and phonological sequences, as confirmed by their greater sensitivity to frequency and length effects (Figures 2 and 3) and by a prolonged, larger effect of  $N$ -size on finger-tracking (Figure 5, right panel).

## 6 Concluding remarks

We provided evidence that finger-tracking data of reading children can highlight congruent developmental patterns in the acquisition of literacy skills. We only replicated established benchmark effects reported in the psycholinguistic literature on decoding transparent orthographies. Nonetheless, to our knowledge, this is the first time that finger-tracking patterns are shown to significantly correlate with more established reading data.

Unsurprisingly, typically developing readers were shown to read at a faster rate than atypical readers. Our comparative analysis shows that both

groups of readers are sensitive to the same lexical effects, but that atypical readers rely on an impoverished lexicon. We take this evidence to show that although the two groups adopt the same strategy, they differ in their global ability in serial information processing, which has a boosting influence on lexical development and reading speed.

Despite our promising results, one could legitimately wonder why we propose using finger-tracking as a proxy of a more established technology such as eye-tracking. Portability and task ecology are our strongest arguments. ReadLet can be used in almost any environment with no data-acquisition specialist or invasive, anxiety-provoking equipment. This has practical consequences for research in education, computer science, human cognition and medical sciences. Our architecture supports highly parallel and distributed processes of data acquisition, which can be delivered in real time to research, clinical and education centers as terminals for data harvesting and quantitative analysis. Large-scale studies can be conducted, paving the way to more generalizable results than ever in the past. In addition, the possibility to take single-subject measurements on more occasions and in different environments makes finger-tracking evidence usable not only in group studies but also for individual diagnostic purposes. Furthermore, the fine-grained, multimodal evidence of different signal streams which are aligned with time and with linguistically annotated texts provides invaluable training data for artificial neural networks and classification algorithms designed to solve engineering problems or simulate neurophysiological correlates of cognitive tasks. Last but not least, we know that reading probes are commonly used for monitoring progress in reading fluency and text comprehension (Miura Wayman et al., 2007), but take huge time and effort to collect. The use of a tablet for extended reading enables deriving this information unobtrusively and continuously, wherever the child fancies reading, even at home.

## Acknowledgments

This work is supported by PRIN grant 2017W8HFRX *ReadLet: reading to understand. An ICT driven, large-scale investigation of early grade children's reading strategies* (2020-22), from the Italian Ministry of University and Research.

## References

- David A Balota and Melvin J Yap. 2006. Attentional control and the flexible lexical processor: Explorations of the magic moment of word recognition. *From inkmarks to ideas: Current issues in lexical processing*, 229.
- Laura Barca, Cristina Burani, and Lisa S Arduino. 2002. Word naming times and psycholinguistic norms for italian nouns. *Behavior Research Methods, Instruments, & Computers*, 34(3):424–434.
- Laura Barca, Cristina Burani, Gloria Di Filippo, and Pierluigi Zoccolotti. 2006. Italian developmental dyslexic and proficient readers: Where are the differences? *Brain and Language*, 98(3):347–351.
- Elizabeth Bates, Cristina Burani, Simona D’Amico, and Laura Barca. 2001. Word reading and picture naming in italian. *Memory & Cognition*, 29(7):986–999.
- Cristina Burani, Stefania Marcolini, and Giacomo Stella. 2002. How early does morpholexical reading develop in readers of a shallow orthography? *Brain and language*, 81(1-3):568–586.
- Cristina Burani, Lisa S Arduino, and Laura Barca. 2007. Frequency, not age of acquisition, affects italian word naming. *European Journal of Cognitive Psychology*, 19(6):828–866.
- Max Coltheart, Kathleen Rastle, Conrad Perry, Robyn Langdon, and Johannes Ziegler. 2001. Drc: a dual route cascaded model of visual word recognition and reading aloud. *Psychological review*, 108(1):204.
- Robert Davies, Javier Rodríguez-Ferreiro, Paz Suárez, and Fernando Cuetos. 2013. Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography. *Reading and Writing*, 26(5):721–738.
- Maria De Luca, Laura Barca, Cristina Burani, and Pierluigi Zoccolotti. 2008. The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology*, 21(4):227–235.
- Marcello Ferro, Claudia Cappa, Sara Giulivi, Claudia Marzi, Franco Alberto Cardillo, and Vito Pirrelli. 2018a. ReadLet: an ICT platform for the assessment of reading efficiency in early graders. page 61, Edmonton, Alberta (Canada), 25-29 September, 2018. 11th International Conference on the Mental Lexicon.
- Marcello Ferro, Claudia Cappa, Sara Giulivi, Claudia Marzi, Ouaphae Nahli, Franco Alberto Cardillo, and Vito Pirrelli. 2018b. Readlet: Reading for understanding. In *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, pages 1–6.
- Guillaume Lio, Roberta Fadda, Giuseppe Doneddu, Jean-René Duhamel, and Angela Sirigu. 2019. Digit-tracking as a new tactile interface for visual perception analysis. *Nature Communications*, 10(5392):1–13.
- Chiara Valeria Marinelli, Daniela Traficante, Pierluigi Zoccolotti, and Cristina Burani. 2013. Orthographic neighborhood-size effects on the reading aloud of italian children with and without dyslexia. *Scientific Studies of Reading*, 17(5):333–349.
- Heidi Anne E. Mesmer and Karen Lake. 2010. The role of syllable awareness and syllable-controlled text in the development of finger-point reading. *Reading Psychology*, 31(2):176–201.
- Miya Miura Wayman, Teri Wallace, Hilda Ives Wiley, Renáta Tichá, and Christine A Espin. 2007. Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education*, 41(2):85–120.
- Maria Montefinese, Ettore Ambrosini, Beth Fairfield, and Nicola Mammarella. 2014. The adaptation of the affective norms for english words (anew) for italian. *Behavior research methods*, 46(3):887–903.
- Maria Montefinese, David Vinson, Gabriella Vigliocco, and Ettore Ambrosini. 2019. Italian age of acquisition norms for a large set of words (itaoa). *Frontiers in psychology*, 10:278.
- Vito Pirrelli, Claudia Cappa, Davide Crepaldi, Viola Del Pinto, Marcello Ferro, Sara Giulivi, Claudia Marzi, Andrea Nadalini, and Loukia Taxitari. 2020. Tracking the pace of reading with finger movements. Canada, 16-18 October, 2020. Words in the World International Conference.
- Joanna K. Uhry. 2002. Finger-point reading in kindergarten: The role of phonemic awareness, one-to-one correspondence, and rapid serial naming. *Scientific Studies of Reading*, 6(4):319–342.
- Tal Yarkoni, David Balota, and Melvin Yap. 2008. Moving beyond coltheart’s N: A new measure of orthographic similarity. *Psychonomic bulletin & review*, 15(5):971–979.
- Pierluigi Zoccolotti, Maria De Luca, Gloria Di Filippo, Anna Judica, and Marialuisa Martelli. 2009. Reading development in an orthographically regular language: Effects of length, frequency, lexicality and global processing ability. *Reading and Writing*, 22(9):965–992.