Processing effort during reading texts in young adults: text simplification, readability assessment and preliminary eye-tracking data

Maria De Martino

University of Salerno, via Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy

English. The paper reports empirical data about the impact of text simplification procedures supported by readability assessment measures on processing effort during reading. Sixty-six Italian native undergraduate students read original and simplified versions of TV news texts and answered a comprehension question. Accuracy data, single word-based and sentence-based measures collected by means of an eye-tracker show that reading simplified texts requires less cognitive demands than their original versions.

Italiano. Il lavoro riporta dati empirici relativi all'efficacia dell'applicazione di metriche di leggibilità nella produzione di testi semplificati. Sessantasei studenti italiani madrelingua hanno letto versioni originali e semplificate dei lanci di notizie tratte da telegiornali e hanno risposto ad una domanda di comprensione. I dati sull'accuratezza e diverse metriche ottenute con l'impiego di un eye-tracker e calcolate al livello di singole parole e singole frasi mostrano che le versioni semplificate dei testi riducono significativamente il carico cognitivo dei lettori rispetto alle relative versioni originali.

Keywords
Text simplification, Readability, Eye-tracking

1. Introduction

Reading is a recent achievement in human evolution, but reading proficiency is considered an important component of success and life outcome [1]. The ability to read and understand with minimum effort depends on both reader characteristics (e.g., literacy, multilingualism, presence/absence of cognitive disorders) and text properties (e.g., length, topic, lexical and syntactic complexity, cohesion, coherence) [2]. One approach to improve inclusion by coping with disadvantage in reading skills is to match readers with texts appropriate to their reading abilities. This goal can be accomplished by exploiting readability formulas to predict the reading and comprehension difficulty of a text for a given target audience and, then, to obtain simplified, i.e. easy-to-understand, texts [3]. Different readability formulas are available for several languages [4] [5]. Many of them rely on text linguistic features such as lexical and syntactic features. Lexical features include the values of frequency, familiarity, imageability and age of acquisition of words within a text. Syntactic features include the complexity of syntactic structures such as the incidence of types of clauses and phrases in the text. Both lexical and syntactic features have been shown to impact cognitive demands in reading processes [6]. Hence, more recent readability formulas take into account measures of natural text processing and try to express readability in terms of cognitive processing effort [7]. Actually, reading texts is a complex behavior subserved by automated interactions between different cognitive processes: visual perception, attention, lexical access, working memory, semantic processing. All these processes are involved in the two main aspects of reading: visual information decoding and meaning construction (comprehension). One of the techniques that has been extensively used to study the reading behavior is the recording of eye movements [8]. During reading, the reader's eyes move from one position to the next in order to process different levels of information that can be extracted from words' visual form. Psycholinguists assume that eye movements during reading reflect different stages of language processing. Some movement have a perceptual function: saccades are rapid movements that shift the eye's focus between two fixed points and are necessary to bring the visual information into the zone of the visual field where acuity is best. Other movements have more complex functions. Fixations are short periods of steadiness of the eye on a word and their duration is a marker of the ease of accessing the meaning of the word and integrating this into the current sentence. Regressions are backward-directed saccades and are related to the necessity of the reader

CLiC-it 2023: 9th Italian Conference on Computational Linguistics, Nov 30 — Dec 02, 2023, Venice, Italy

mdemartino@unisa.it (M. De Martino)

to re-analyze previously explored portions of the text because of processing difficulties.

Gaze behavior during reading is exploited in several contexts and for different aims both in cognitive psychology and in Natural Language Processing (NLP) literature. For instance, eye-tracking metrics are exploited to unfold mechanisms of reading in L2 learners, typical and atypical readers [9] [10] [11] [12] [13] [14]. Corpora of eye-tracking data are available in many languages and are profitably used to implement language models that can predict human reading behavior (e.g., [15] [16] [17]).

A less investigated issue is to what extent the use of specific text simplification strategies reduces the processing effort as measured through eye-movements [18] [19] .

The aim of the present study is to test the impact of texts' simplification and readability assessment on young adults reading behavior through the analysis of eye-movements.

2. Method

We conducted an eye-tracking reading study to obtain objective and reliable measures of processing effort. The advantage of monitoring readers' eye movements during reading is that it is considered to be the experimental situation that better resembles natural reading [11] [20] [21].

2.1. Stimuli

2.1.1. Selection

The same materials as used in [22] were employed. They consisted of 18 texts of news scripts as read by Italian TV news anchors. Such news texts are usually short but often linguistically and stylistically complex and can be difficult to comprehend for non-native speakers and/or for people with low literacy, reading disorders or cognitive and intellectual disabilities. Hence, they constitute suitable materials to be employed in an empirical study on simplification.

2.1.2 Readability Assessment and Text Simplification

Each selected text in its original version (OV) underwent a readability assessment through the READ-IT tool [3] and then to a double manual simplification process that generated 2 simplified versions: S1 and S2. The simplification strategies used to implement S1 involved sentence splitting, changing passive voice to active, lexical and syntactic ambiguity resolution, low frequency and long words replacement.

In addition to the above mentioned simplification strategies, specific interventions on the semantic content were used in order to achieve the S2 versions. They were mainly focused on the temporal ordering of events and on reporting each factual event into a separate sentence (See [22] for further details and for examples of the original and simplified materials).

The obtained simplified versions were matched for text length calculated in number of words (average number of words: OV=56; S1=58; S2=58).

The text complexity measures obtained through the application of the READ-IT tool [3] revealed that the OV texts scores were significantly lower than the S1 and S2 ones (see Table 1: values shown in bold are the readability scores, values reported in parentheses are the p values of the t-tests comparing OV vs. S1 and S2).

The GULPEASE score indicates the readability of the texts: a higher GULPEASE score indicates higher readability of a text.

Table 1
Readability assessment: comparisons between OV vs. simplified versions

READ-IT scores	OV	S1	S2
Base	**36**	**11** (p<.005)	**6** (p<.001)
Lexical	**96**	**80** (p<.05)	**83** (p=.05)
Syntactic	**51**	**13** (p<.01)	**7** (p<.001)
Global	**80**	**34** (p<.001)	**21** (p<.001)
GULPEASE	**51**	**58** (p<.001)	**65** (p<.001)

The comparison between the two simplified versions, S1 vs. S2, revealed that S2 obtained better readability scores than S1 scores only on the global score ($t(17) = 215$, $p<.05$) and on the GULPEASE score ($t(17) = -5.08$, $p<.001$)[2] .

2.1.3. Implementation of the experimental text lists

The whole set of 54 texts (18 OV, 18 S1, 18 S2) was split into 3 lists. Each list included 1 of the 3 versions of each text and comprised 6 OV, 6 S1 and 6 S2. Participants were randomly assigned to a given list. This strategy allowed to ensure that each participant was presented only once with a given text in order to avoid putative effects of the repetition of materials. Each participant was administered with texts presented in a shuffled order.

2.2. Participants

Sixty-six undergraduate students from University of

[2] Further user-based readability scores on the materials used in the study are reported in [22] and provide information about the speakers' perceived difficulty of the linguistic formulation and of the topic of the texts and the perceived naturalness/acceptability of the Italian language used to generate the texts. The judgments of speakers did not reveal significant differences between OV, S1 and S2.

Salerno (45 females) were recruited; they voluntarily took part in the experiment. Their age ranged from 19 to 30 years (average = 23 years). They were all native speakers of Italian, had normal or corrected-to-normal vision and did not report history of reading, language, learning or neurological/psychiatric disorder.

2.3. Apparatus

The reading experiment was implemented and administered via Tobii Pro Lab 1.194 software.
A screen-based Tobii Pro X2-30 eye-tracker was connected to the monitor of an HP computer available at the LaPSUS laboratory (University of Salerno). The range of head movement allowed was within a three-dimensional range of 50 cm W × 36 cm H × 70 cm and the allowed operating distance from the monitor was within 40 and 90 cm.

2.4. Procedure

The experimental procedure consisted of different steps.

2.4.1. Calibration

The participants sat in front of the screen; the distance to the screen was adjusted. For calibration in the Tobii Pro Lab software, participants were required to keep their heads as still as possible and to look at a fixation point moving on the screen.

2.4.2. Training and warm up trials

The participants were presented with a slide displaying written instructions about the reading task. Then, a training phase was administered. Participants were requested to perform 2 warm up trials: they were shown a text on the screen and were instructed to read it silently at their own pace of comprehension and to press any key of the PC keyboard to move to the subsequent slide. After reading the text, they were asked to use the mouse to select the correct response of a true–false question presented on the screen. Then, a second trial was administered.

2.4.3. Experimental Session

After the warm up trials, each participant was administered 1 of the 3 lists made up with 18 experimental trials.

2.5. Dependent variables

Different measures were analysed in the current study.

2.5.1. Global reading time

The global reading time was recorded.

2.5.2. Accuracy

Answers to the true-false question were recorded and analysed as indicators of the comprehension of the texts.

2.5.3. Eye-tracking metrics

The Tobii Pro X2-30 hardware and software equipment (Tobii Pro Lab 1.194) provides a large number of eye-tracking measures. However, the current study reports the most commonly measures found to be related to text difficulty and cognitive demand [23].
The following measures were analysed for word-based and sentence-based areas of interest (AOI):

- Number and Duration of Fixations: fewer and shorter fixations are supposed to be associated with lower reading effort.
- Number and Duration of Visits: visits can be defined as the number of times that the reader's eye move towards a given AOI with either progressive or regressive saccades. The entry and exit saccades are excluded. The number and duration of visits indicate that specific portions of the text receive specific amount of attentional and linguistic resources to be processed.
- Regression path duration: it describes the time that elapses between a first fixation on an AOI to the moment when gaze is directed away from that region to the right. Thus, it includes time spent re-reading earlier parts of the text before the reader is ready to proceed with the rest of the text.
- Re-reading duration: it corresponds to the regression-path duration minus first-pass duration and it is assumed to reflect strategic, controlled processes involved in reading comprehension.

3. Results

The results of reading times, accuracy rates and eye-tracking metrics were analyzed through a series of ANOVAs.

3.1.1. Whole text reading times and accuracy data

No significant effect was found on global reading time; this result replicates the findings of [22] obtained in a different experimental setting.
On the contrary, ANOVA on accuracy data (Figure 1) showed a significant effect of simplification ($F(2,1185) = 3{,}8094$, $p=.02$). LSD post hoc tests revealed that questions to S2 were responded significantly better than OV ($p=.008$). The difference between S1 and OV was only marginally significant ($p=.052$), while the difference between S1 and S2 was not significant ($p=.46$).

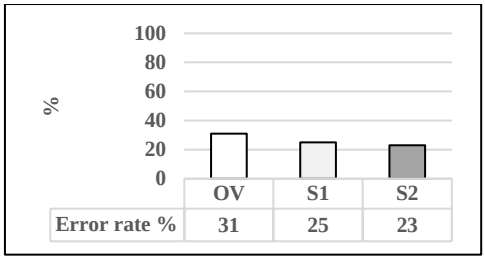

Figure 1: Error rate for OV, S1 and S2

3.1.2. Word-based eye-tracking data

The mean and standard deviation values obtained for word-based and sentence-based eye-tracking metrics are reported in Tables 2 and 3.

Table 2
Word-based eye-tracking metrics

WORD-BASED METRICS		OV	S1	S2
Number of Fixations	mean	2,11	2,0	2,0
	ds	1,59	1,5	1,47
Duration of Fixations (ms)	mean	413,94	393,5	389,46
	ds	370,97	342,49	340,43
Number of Visits	mean	1,75	1,71	1,69
	ds	1,13	1,08	1,07
Duration of Visits (ms)	mean	430,96	408,77	404,67
	ds	398,37	368,13	364,58
Regression-Path Duration (ms)	mean	413,78	391,46	389,28
	ds	1416,05	1195,97	1174,15
Re-Reading Duration (ms)	mean	285,36	259,77	254,21
	ds	1383,62	1167,14	1143,29

ANOVAs performed on word-based metrics showed a significant effect of simplification both on the number ($F(2, 42566) = 21,098$, $p < .001$) and the duration of Fixations ($F(2, 42566) = 19,733$, $p < .0001$). More specifically, the post hoc LSD tests revealed that the number and the duration of fixations is significantly higher in OV than in both S1 ($p < .001$) and S2 ($p < .001$). No significant difference was detected between S1 and S2.

An effect of simplification was observed both on the number ($F(2, 42566) = 11,202$, $p < .001$) and the duration of visits ($F(2,42566) = 19,920$, $p < .0001$): the OV texts received significantly higher scores than S1 and S2, while S1 and S2 resulted equivalent.

Data on the re-reading scores showed that simplification elicited a slight tendency to the statistical significance ($F(2, 42560) = 2,5594$, $p = .07$); however, additional planned comparisons showed that the OV texts required a re-reading time significantly higher than the S2 texts ($p < .05$), but did not differ from S1 ($p = .08$).

The regression-path duration did not show any significant effect of simplification at the word-level.

Table 3
Sentence-based eye-tracking metrics

SENTENCE-BASED METRICS		OV	S1	S2
Number of Fixations	mean	20,05	14,94	15,35
	sd	17,34	11,68	9,32
Duration of Fixations (ms)	mean	3928,22	2906,84	2967,14
	sd	3660,62	2471,19	1949,63
Number of Visits	mean	3,58	3,50	4,39
	sd	2,58	2,55	2,22
Duration of Visits (ms)	mean	4975,30	3615,57	3686,45
	sd	4660,07	3109,65	2430,72
Regression-Path Duration (ms)	mean	3970,04	2904,92	2967,20
	sd	4793,85	3588,49	3048,10
Re-Reading Duration (ms)	mean	2249,18	1806,07	2122,96

A slightly different picture emerged from the analyses performed on sentence-based metrics. The ANOVAs revealed that the effect of simplification reached the statistical significance for all the collected eye-tracking metrics:

- Number of fixations: $F(2, 5692) = 163,15$, $p = .0000$;
- Duration of Fixations: $F(2, 5692) = 144,61$, $p = .0000$;
- Number of visits: $F(2, 5692) = 5,2883$, $p = .00507$;
- Duration of visits: $F(2, 5692) = 160,30$, $p = .0000$;
- Regression-path duration: $F(2, 5651) = 75,393$, $p = .0000$;
- Re-reading duration: $F(2, 5651) = 20,504$, $p = .00000$.

OV was found to be the version that required significantly higher processing effort when compared both to S1 and S2 ($p < .001$). In addition, S1 was found to be more demanding than S2 ($p < .05$). Only for the number of visits OV and S1 showed equivalent amount of processing effort ($p = .33$).

4. Conclusion

The paper investigates to what extent the application of text simplification strategies improves the readability of texts and reduces the reading processing effort as it emerges from cognitive indexes that are out of the awareness of the reader, i.e. eye movements patterns.

The preliminary data reported in the paper show that accuracy in comprehension questions increases significantly when texts undergo simplification procedures based on the reduction of the reader's

amount of processing inferences (i.e., event reordering or coreference chains explaining).

On the other hand, the physiological and cognitive measures related to the processing effort during reading are affected by simplification strategies that involve both the lexical-syntactic level and the content level.

Moreover, the metrics collected at the sentence-level and single-word level are found to be suitable and sensitive measures to detect respectively the efficacy of simplification procedures in modulating the strategic controlled processes involved in comprehension and the attentional and lexical processing effort during reading. Interestingly, the data were obtained by analyzing the performance of young adult skilled readers that are supposed to be less likely influenced by the readability of texts.

References

[1] Dehaene S. Reading in the brain: The science and evolution of a human invention. New York: Viking; 2009.

[2] Kuperman V, Matsuki K, Van Dyke JA. Contributions of reader-and text-level characteristics to eye-movement patterns during passage reading. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2018 Nov;44(11):1687. doi: 10.1037/xlm0000547

[3] Dell'Orletta F, Montemagni S, Venturi G. 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 2011 Jul (pp. 73-83).

[4] Crossley SA, Skalicky S, Dascalu M. Moving beyond classic readability formulas: New methods and new models. Journal of Research in Reading. 2019 Nov;42(3-4):541-61. doi:10.1111/1467-9817.12283

[5] Nahatame S. Text readability and processing effort in second language reading: A computational and eye-tracking investigation. Language learning. 2021 Dec;71(4):1004-43

[6] Keller TA, Carpenter PA, Just MA. The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing. Cerebral cortex. 2001 Mar 1;11(3):223-37. doi.org/10.1093/cercor/11.3.223

[7] Huckin TN. A cognitive approach to readability. InNew essays in technical and scientific communication 2019 Mar 8 (pp. 90-108). Routledge.

[8] Hyönä J, Kaakinen JK. Eye movements during reading. Eye movement research: An introduction to its scientific foundations and applications. 2019:239-74.

[9] Cop U, Drieghe D, Duyck W. Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. PloS one. 2015 Aug 19;10(8):e0134008. doi.org/10.1371/journal.pone.0134008

[10] Franzen L, Stark Z, Johnson AP. Individuals with dyslexia use a different visual sampling strategy to read text. Scientific reports. 2021 Mar 19;11(1):6449. doi.org/10.1038/s41598-021-84945-9

[11] Godfroid A. Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide. 2019.

[12] Prabha AJ, Bhargavi R. Predictive model for dyslexia from fixations and saccadic eye movement events. Computer Methods and Programs in Biomedicine. 2020 Oct 1;195:105538. doi.org/10.1016/j.cmpb.2020.105538

[13] Rayner K, Shen D, Bai X, Yan G, editors. Cognitive and cultural influences on eye movements. Taylor & Francis; 2023 May 31.

[14] Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological bulletin. 1998 Nov;124(3):372.

[15] Hollenstein N, Rotsztejn J, Troendle M, Pedroni A, Zhang C, Langer N. ZuCo, a simultaneous EEG and eye-tracking resource for natural sentence reading. Scientific data. 2018 Dec 11;5(1):1-3. doi.org/10.1038/sdata.2018.291

[16] Luke SG, Christianson K. The Provo Corpus: A large eye-tracking corpus with predictability norms. Behavior research methods. 2018 Apr;50:826-33. doi.org/10.3758/s13428-017-0908-4

[17] Siegelman N, Schroeder S, Acartürk C, Ahn HD, Alexeeva S, Amenta S, Bertram R, Bonandrini R, Brysbaert M, Chernova D, Da Fonseca SM. Expanding horizons of cross-linguistic research on reading: The Multilingual Eye-movement Corpus (MECO). Behavior research methods. 2022 Dec;54(6):2843-63.

[18] Rivero-Contreras M, Engelhardt PE, Saldana D. Do easy-to-read adaptations really facilitate sentence processing for adults with a lower level of education? An experimental eye-tracking study. Learning and Instruction. 2023 Apr 1;84:101731. doi.org/10.1016/j.learninstruc.2022.101731

[19] Rets I, Rogaten J. To simplify or not? Facilitating English L2 users' comprehension and processing of open educational resources in English using text simplification. Journal of Computer Assisted Learning. 2021 Jun;37(3):705-17. doi: 10.1111/jcal.12517

[20] Raney GE, Campbell SJ, Bovee JC. Using eye movements to evaluate the cognitive processes involved in text comprehension. JoVE (Journal of Visualized Experiments). 2014 Jan 10(83):e50780. doi: 10.3791/50780

[21] Roberts L, Siyanova-Chanturia A. Using eye-tracking to investigate topics in L2 acquisition and L2 processing. Studies in Second Language Acquisition. 2013 Jun;35(2):213-35. doi:10.1017/S0272263112000861

[22] De Martino M, Colella A. La produzione di testi semplificati di notiziari televisivi italiani destinati a persone con disturbi cognitivi acquisiti: un'integrazione tra metodi psicolinguistici e analisi automatiche (Implementing Simplified TV News Texts in

Italian for People with Acquired Cognitive Disorders: Psycholinguistic Methods and Automatic Analyses). In Proceedings of the Eighth Italian Conference on Computational Linguistics (CLiC-IT).CLiC-it 2021.

[23] Rayner K, Chace KH, Slattery TJ, Ashby J. Eye movements as reflections of comprehension processes in reading. Scientific studies of reading. 2006 Jul 1;10(3):241-55. doi.org/10.1207/s1532799xssr1003_3