The Impact of Lexical Simplification by Verbal Paraphrases for People with and without Dyslexia

Luz Rello[1,2], Ricardo Baeza-Yates[1,3], and Horacio Saggion[2]

[1] Web Research Group, Dept. of Information and Communication Technologies,
Universitat Pompeu Fabra, Barcelona, Spain
[2] Natural Language Processing Research Group, Dept. of Information and
Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain
[3] Yahoo! Research Barcelona, Spain
luzrello@acm.org,rbaeza@acm.org,horacio.saggion@upf.edu

Abstract. Text simplification is the process of transforming a text into an equivalent which is easier to read and to understand, preserving its meaning for a target population. One such population who could benefit from text simplification are people with dyslexia. One of the alternatives for text simplification is the use of verbal paraphrases. One of the more common verbal paraphrase pairs are the one composed by a lexical verb (*to hug*) and by a support verb plus a noun collocation (*to give a hug*). This paper explores how Spanish verbal paraphrases impact the readability and the comprehension of people with and without dyslexia dyslexia. For the selection of pairs of verbal paraphrases we have used the *Badele.3000* database, a linguistic resource composed of more than 3,600 verbal paraphrases. To measure the impact in reading performance and understandability, we performed an eye-tracking study including comprehension questionnaires. The study is based on a group of 46 participants, 23 with confirmed dyslexia and 23 control group. We did not find significant effects, thus tools that can perform this kind of paraphrases automatically might not have a large effect on people with dyslexia. Therefore, other kinds of text simplification might be needed to benefit readability and understandability of people with dyslexia.

Keywords: Lexical simplification, verbal paraphrases, readability, understandability, eye-tracking, dyslexia.

1   Introduction

The goal of this paper is to present the impact of lexical simplification through verbal paraphrases in readability and understandability for people with and without dyslexia.

Dyslexia has been defined as a specific reading disability [39] and as a learning disability [20]. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties

typically result from a deficit in the phonological component of language that is often unrelated to other cognitive disabilities. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge [20]. According to cognitive neuroscience studies, people with dyslexia find difficulties with functional [26] and short words [37]. Functional and short words are present in the verbal paraphrases (support verb plus a noun collocation, *dar un paseo, 'to go for a walk'* to be simplified by a lexical verb *pasear, 'to walk'*).

In this study, we distinguish between **readability** and **understandability**. Readability refers to the legibility of a text, that is, the ease with which text can be read (that is, the person can reproduce it even though does not understand it) while understandability refers to comprehensibility, the ease with which text can be understood. Since readability strongly affects text comprehension [5], sometimes both terms have been used interchangeably [21]. However, previous research with people with dyslexia has shown that both concepts need to be taken into consideration separately. For instance, in [31] the inclusion of graphical schemes in the text improved their readability in terms of reading speed, but had a negative effect on the comprehension for people with dyslexia. Moreover, for people with dyslexia, comprehension has been found to be independent of the lexical quality of the text. While errors in text affect negatively readability and understandability of people without dyslexia, they do not affect that much in people with dyslexia [29].

This research is motivated by (1) its novelty and (2) by the social relevance of its results. First, lexical complexity such as word frequency, verb complexity and lexical ambiguity has an effect on the readability and understandability for people with dyslexia [18] and without this condition [28]. In this study, we try to enrich previous findings exploring how practical examples of verbal paraphrases impact readability and understandability to find out whether lexical simplification systems targeted for people with dyslexia shall include verbal paraphrases. To measure readability we analyze eye movements of readers with and without dyslexia using eye tracking and for addressing reading comprehension, we used questionnaires with inferential questions. To the best of our knowledge, this is the first time that the effect of verbal paraphrases is measured in terms of readability and understandability for people with and without dyslexia using this methodology.

Second, since dyslexia is universal and frequent, people with dyslexia are a relatively large group of users. The Interagency Commission on Learning Disabilities [19] states that 10 to 17.5% of the population in the U.S.A. have dyslexia and between 7.5 to 11.8 % of the Spanish speaking population has dyslexia [30]. Also, dyslexic-related difficulties are shared by other groups with special needs such as low vision [16] and symptoms of dyslexia are common to varying degrees among most people [14]. Thus, the results of this research may be applicable to general usability problems and other target groups.

This paper is organized as follows. Next section covers the related work while Section 3 covers lexical simplification by using verbal paraphrasing. In Section 4,

we present our experimental methodology while in Section 5 we show the results of it. We end in Section 6 with some concluding remarks and our future work.

2 Related Work

Related work to our study belong to different fields: (a) natural language processing (NLP) literature about paraphrases and their use in lexical simplification, and (b) experimental psychology studies which takes into account the impact of language complexity in reading comprehension and performance of people with dyslexia.

In **NLP** a paraphrase is an alternative surface form in the same language expressing the same semantic content as the original form [24]. The use of automatic methods for generating paraphrases has been successfully applied for text simplification among other NLP tasks. For instance, in [21] paraphrasing is used to remove difficult syntactic structures for deaf learners of written English and Japanese. Paraphrasing methods were applied to simplify newspaper texts for people with aphasia [10, 11] and Down syndrome [33] as well as to simplify online information for people with aphasia [13].

Text complexity and dyslexia also has been studied in **experimental psychology**. Word frequency, verb complexity and lexical ambiguity are related to the processing time of words [28, 34]. Hyönä and Olson measure the effect of word length and word frequency in relation with eye fixation patterns and show that low frequency and long words present longer gaze durations and more re-inspections in both, readers with and without dyslexia [18]. In that work, the analysis is focused on target words [18] while we measure the whole text and the integration of target words in the overall text. The rationale behind this is that readability and understandability pertain to longer segments of texts [17]. Comprehension in people with dyslexia was studied in correlation with syntax complexity including long sentences with complex structures [35], the sentence context [25], or the word fluency [12], among others.

However, there are no studies for Spanish which approach readability and comprehension of people with dyslexia taking into consideration one common verbal paraphrasing pair [2] used for lexical simplification. That is, the pair composed of a lexical verb (*abrazar, 'to hug'*) and by a support verb plus a noun collocation (*dar un abrazo, 'to give a hug'*).

3 Lexical Simplification by Verbal Paraphrases

Under 18% of manual simplification operations made by experts in newspaper articles are lexical changes [6]. One of the most common simplification solutions done manually in Spanish is the substitution of the combination of the support verb and a deverbal noun by the corresponding verb alone [15]. That is, *dar un paseo, 'to go for a walk'* by *pasear, 'to walk'* or *dar un abrazo, 'to give a hug'* by *abrazar, 'to hug'*. Although these kind of lexical simplifications are frequent in manual simplifications, their automatic computational process is still challenging

[15]. Thus, there are specific linguistic resources developed for such tasks, such as the *Badele.3000* database [3].

Badele.3000 is a database that contains more than 3,600 high frequency Spanish nouns and 2,800 high frequency Spanish verbs, including 23,000 collocations made from the combinations of both kinds of words. The paraphrase pairs consisting of a verb and a verb-noun collocation were manually extracted [4]. As *Badele.3000* was created manually by an expert, the linguistic validity of the paraphrases pairs used in our study is guaranteed.

The selected pairs of synonymic paraphrases are composed of a support verb plus a noun collocation and a lexical verb. According to the manual simplifications [15], the lexical verb alone is considered to be simpler; for instance:

[−simple] Sus lectores *tenían confianza* en ella.
　　　　　　Her readers *had trust* in her.
[+simple] Sus lectores *confiaban* en ella.
　　　　　　Her readers *trusted* her.

According to cognitive neuroscience studies, it would also be expected that people with dyslexia might find more difficult to read the [−simple] option since they have more frequent errors with functional [26] and short words [37]. However, from a linguistic point of view it is not clear which option is simpler.

Linguists agree in differencing lexical words and functional words [23]. Lexical words have a lexical meaning which is less ambiguous than the grammatical meanings expressed by functional words. Functional words are prepositions, pronouns, auxiliary verbs, conjunctions, among others. Support verbs have been considered as functional words because they are semantically empty, for instance verb *dar, 'to give'* is a support verb in *dar un abrazo, 'to give a hug'*.[4]

Since functional words do not have a lexical representation their processing is different than lexical words [8]. There are still many open questions about the difference levels of word processing by the human brain. However, in the case of dyslexia a special emphasis have been made for errors in functional words [26]. To the best of our knowledge, there is no formal explanation behind errors in functional words. They could be due to their nature (i.e. lack of lexical content) or could be simply due to the fact that higher errors rates are observed for shorter words [37].

On the other hand, word processing depends on the complexity of the morphological components of the word [9]. For instance, *paseo, 'walk'* is simpler than *pasear, 'to walk'* because it is composed by one lexeme while *pasear* is made by one lexeme plus one derivative morpheme *pasear = paseo + ar*. Since it is not trivial to access the complexity of the paraphrase pairs from a linguistic point of view, we take as our criteria the empirical analysis observed in manual simplifications performed by experts [15].

---

[4] However, Barrios [2] analyzed extensively the meaning of support verbs concluding that some of them are not fully empty.

4 Experimental Methodology

We designed one experiment which combines reading tests, comprehension tests and semi-structured interviews. Twenty three participants with dyslexia and a comparable control group undertook the experiment.

4.1 Participants

Twenty-three native Spanish speakers with a confirmed diagnosis of dyslexia took part in the study, twelve of whom were female and eleven male. Their ages ranged from 13 to 37, with a mean age of 20.74. Three of the participants were also diagnosed with attention deficit disorder. All participants were frequent readers; eleven read less than four hours per day, nine read between four and eight hours per day, and three participants read more than eight hours daily. Ten people were studying or already finished university degrees, eleven were attending school or high school and two had no higher education. All the participants were asked to bring their diagnoses to the experiment. Therefore, we can guarantee that the participant was diagnosed in an authorized center or hospital. A control group of 23 participants without dyslexia with the same age range and similar age average (20.91) also took part the experiment.

4.2 Design

The experiment is composed of: (1) a questionnaire designed to collect demographic information, (2) two reading tests with their corresponding target words, (3) two tests designed to control the comprehension, and (4) a semi-structured interview about their impression and opinions about the readability of the texts. The experiment followed a within-subjects design, so every participant contributed to each of the conditions, [+simple] and [−simple], in both experiments. The order of the conditions was counter-balanced to cancel out sequence effects, guaranteeing that the person never reads the same text twice (see Figure 1).[5]

With the reading tests we collect the quantitative data to measure readability, while with the comprehension tests we measure understandability. At the end, with the semi-structured interviews we gather information about the participant preferences.

We selected two very similar newspaper texts from the Spanish Simplex corpus [7]. To meet the comparability requirements among the texts, we slightly adapted the texts maintaining as much as possible the original text. To determine these comparability requirements we took into account the parameters that different complexity measures take into consideration [15]. Next, we present the characteristics shared by the texts of the experiment:

(a) They are about similar topics: a literature award (Text *María*) and a cinema award (Text *Alex*). See the Appendix for the texts used in the experiment.

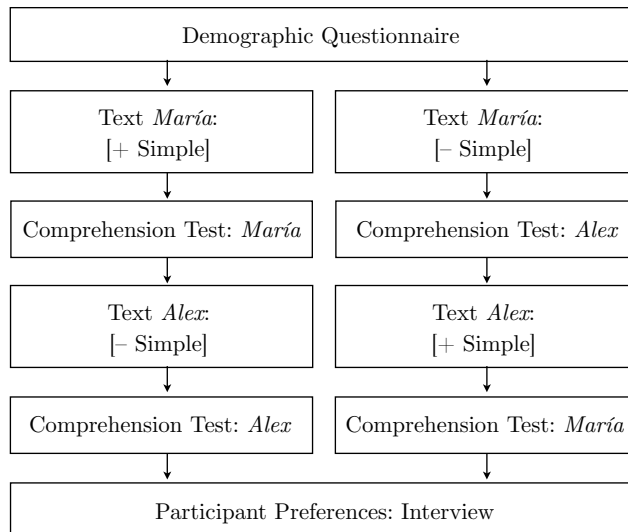[5] We do not need to consider the two texts in different order as they have similar text complexity.

Fig. 1. Structure of the experiment.

(b) They have the same number of target lexical substitutions: nine verbal para-
    phrase pairs [±simple]. See the Appendix for all the paraphrase pairs used
    in the experiment.
(c) They share the same genre: culture news.
(d) They have the same number of sentences per text, five sentences.
(e) They have the same number of words per text, 100 words.
(e) All the texts have a similar word length average ranging from 4.87 to 5.19
    letters per word.
(f) They contain the same number of named entities mentioned for the first
    time.
(i) The texts do not contain numerical expressions, foreign words or acronyms.

Since the presentation of the text has an effect on reading speed of people with
dyslexia [32], we used the same layout for all the texts. We chose a recommended
font type, sans serif arial [1], unjustified text [27], a big size of 20 points, 62
characters per column, and recommended color and brightness contrast using a
black font with creme as background[6] [32].

To control the comprehension, after each text we designed a test including
inferential items related to the main idea. We did not include items referred to
details because they involve memory more than comprehension [36]. Each of the
items has three choices where one is correct, one is partially correct (normally
containing details), and one is incorrect. We gave 100, 50 and 0 points for each
type of answer, respectively, to compute a comprehension score

---

[6] The CYMK are creme (FAFAC8) and black (000000). Color difference: 700, Bright-
   ness difference: 244.

The test finishes with one semi-structured interview to learn the participant preferences. The participant was asked which text seemed to be more readable. After this, we asked face-to-face the reasons leading to the selected answer, which difficulties they encountered when reading the texts, and which options would they like to find to achieve a better understanding.

4.3   Equipment

The eye tracker used was the Tobii T50 [38, 17-inch TFT monitor] using a resolution of 1024x768 pixels. The eye tracker was calibrated for each participant and the light focus was always in the same position. The distance between the participant and the eye tracker was constant (approximately 60 cm. or 24 in.) and controlled by using a fixed chair.

4.4   Procedure

The sessions were conducted at Pompeu Fabra University and they took from 20 to 30 minutes each, depending on the amount of information given by the participant. In each session, the participant was alone with the interviewer (first author) in the quiet room prepared for the study, and had to do the following three steps: (1) answer a questionnaire designed to collect demographic information; (2) perform the eye-tracking experiment; and (3) answer the semi-open interview. For (2) the participant was asked to read the texts in silence and to complete the comprehension tests. We obtained 40 test samples out of the 44 possible that were successfully recorded.

4.5   Data Analysis

The software used for analyzing the eye tracking data was Tobii Studio 3.0 and the R 2.14.1 statistical software. The dependent variables used for the comparison of the text passages were the means of the fixation duration and the total duration of reading. Differences between groups and dependent variables were analyzed by means of matched-pairs, and two-way Student t-tests.
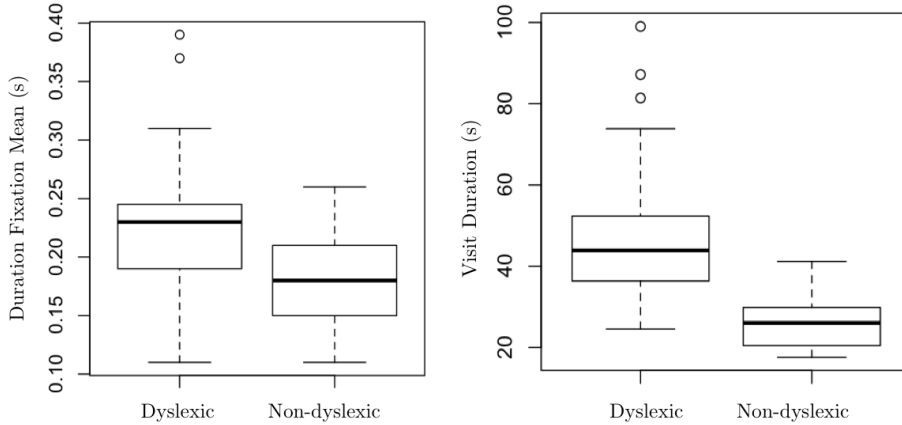
5   Results

To measure the impact of verbal simplification in readability we analyzed two variables derived from eye-tracking data: the average fixation duration and the total visit duration of the text passages. In general, shorter fixations are preferred to longer ones because according to previous studies [22, 28, 34], readers make longer fixations at points where processing loads are greater. Also, shorter reading durations are preferred to longer ones since faster reading is related to more readable texts [40]. We compare readability with understandability through the inferential items of the comprehension test which are assessed by the percentage of correct answers.

Table 1. Experimental results of the eye-tracking and comprehension user study for the texts using paraphrases (none of the differences are statistically significant).

Measure (ave. ± std.dev.)	[+simple]	[−simple]
	Group with Dyslexia	
Fixation Duration	0.229 ± 0.063	0.226 ± 0.054
Visit Duration	44.403 ± 17.225	47.425 ± 14.610
Correct Answers	67.5%	67.5%
	Group without Dyslexia	
Fixation Duration	0.180 ± 0.040	0.178 ± 0.039
Visit Duration	25.172 ± 5.482	27.825 ± 6.993
Correct Answers	75%	77.5%

All our results are given in Table 1. As expected, comprehension for people with dyslexia is slightly lower than those for people without dyslexia.

First, we studied the differences between participants with dyslexia and control group. The average fixation duration of people with dyslexia (0.228 ± 0.058) was significantly higher than for people without dyslexia (0.179 ± 0.039), with $t(80) = 4.4583$, and $p < 0.001$ (see Figure 2).

Fig. 2. Box plots for the average fixation and the total duration for the two groups.

The results for fixation duration do not show statistical significance, because we obtain $t(40) = 0.1613$, $p < 0.873$. The same happens with visit duration with $t(40) = 0.1753$, $p < 0.862$.

To estimate the likelihood that we missed revealing an existing effect of verbal paraphrases on the mean of fixation durations, we calculated the achieved statistical power. Given a p-value of 0.873, an effect size of 0.052 (Cohen's d), and a sample size of 40, the achieved power is 0.880. Hence, the probability of not committing a Type II Error is 88%, that is, the likelihood that an unrevealed effect exists is only 12%.
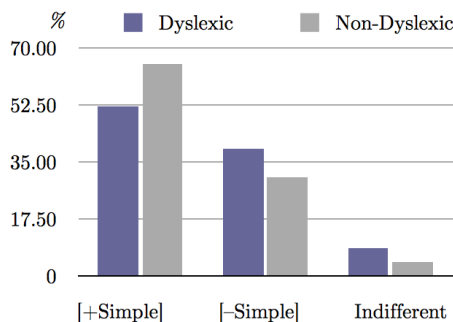
Fig. 3. Preferences of the two groups.

Second, we studied the visit duration time. The statistical results were similar to the ones for the average duration time and hence corroborated the negative finding. The larger range of values for visit time in the group with dyslexia compared with the control group probably indicates the individual variations in reading methods which would make correlations between individuals difficult to observe. On the other hand, the results of the semi-open interview did not matched the analysis of the quantitative variables as shown in Figure 3. That is, the perception of the people is that the simpler text was indeed simpler, although they did not read faster.

6   Conclusions

In this paper we have studied the impact of verbal paraphrases in lexical simplification for people with and without dyslexia.

We chose to study these kind of paraphrases because of two reasons. First, there are already linguistic resources for NLP including these type of Spanish paraphrases [3], which can serve as a starting point. Second, according to cognitive neuroscience studies, this kind of verbal simplification might be especially suitable for people with dyslexia because they find difficulties with functional [26] and short words [37].

The effect of the verbal paraphrases is concluded to be insignificant. Our results are negative in the sense that verbal paraphrases neither improved readability nor understandability in our experiment. However, we can argue a few reasons that may explain this result, implying that further research is needed. The main reason might be that the impact of paraphrasing may depend on the complexity of the text (e.g. in more complex texts verbal paraphrases might be beneficial). Another reason is that the impact is small and hence we need larger texts and a larger number of people to detect it.

As people with dyslexia do have problems with reading most texts, including simple texts, as shown by our results, a more promising line of future research is studying more complex techniques to perform lexical simplification. For example, other types of paraphrasing or synonym substitution.

References

1. Al-Wabil, A., Zaphiris, P., Wilson, S.: Web navigation for individuals with dyslexia: an exploratory study. Universal Access in Human Computer Interaction. Coping with Diversity pp. 593–602 (2007)
2. Barrios, M.A.: El dominio de la funciones léxicas en el marco de la Teoría Sentido-Texto. Estudios de Lingüística del Español (ELiEs) 30 (2010)
3. Barrios, M.A., Aguado de Cea, G., Ramos, J.A.: Enriching a lexicographic tool with domain definition problems and solutions. In: Sierra, G., Pozzi, M., Torres, J. (eds.) 1st International Workshop on Definition Extraction, RANLP 09. INCOMA Ltd, Shoumen, Bulgaria (2009)
4. Barrios Rodríguez, M., Rello, L.: False paraphrase pairs in Spanish for verbs and verb + noun collocations. Procesamiento del Lenguaje Natural 46, 107–112 (2011)
5. Barzilay, R., Elhadad, N., R. McKeown, K.: Inferring strategies for sentence ordering in multidocument news summarization. Journal of Artificial Intelligence Research 17, 35–55 (2002)
6. Bott, S., Rello, L., Drndarevic, B., Saggion, H.: Can Spanish be simpler? LexSiS: Lexical simplification for Spanish. In: Proceedings of the 24th International Conference on Computational Linguistics (Coling 2012). Mumbay, India (December 2012)
7. Bott, S., Saggion, H.: Text simplification tools for Spanish. In: Proceedings of the eighth international conference on Language Resources and Evaluation (LREC 2012). ELRA, Istanbul, Turkey (May 2012)
8. Caramazza, A.: How many levels of processing are there in lexical access? Cognitive neuropsychology 14(1), 177–208 (1997)
9. Caramazza, A., Laudanna, A., Romani, C.: Lexical access and inflectional morphology. Cognition 28(3), 297–332 (1988)
10. Carroll, J., Minnen, G., Canning, Y., Devlin, S., Tait, J.: Practical simplification of English newspaper text to assist aphasic readers. In: Proceedings of the AAAI-98 Workshop on Integrating Artificial Intelligence and Assistive Technology. pp. 7–10. Citeseer (1998)
11. Carroll, J., Minnen, G., Pearce, D., Canning, Y., Devlin, S., Tait, J.: Simplifying text for language-impaired readers. In: Proceedings of EACL. pp. 269–270 (1999)
12. Cutting, L., Materek, A., Cole, C., Levine, T., Mahone, E.: Effects of fluency, oral language, and executive function on reading comprehension performance. Annals of dyslexia 59(1), 34–54 (2009)
13. Devlin, S., Unthank, G.: Helping aphasic people process online information. In: Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility. pp. 225–226. ACM (2006)
14. Dixon, M.: Comparative study of disabled vs. non-disabled evaluators in user-testing: dyslexia and first year students learning computer programming. Universal Acess in Human Computer Interaction. Coping with Diversity pp. 647–656 (2007)
15. Drndarevic, B., Saggion, H.: Towards automatic lexical simplification in Spanish: an empirical study. In: Proceedings of the NAACL HLT 2012 Workshop Predicting and improving text readability for target reader populations (PITR 2012) (2012)
16. Evett, L., Brown, D.: Text formats and web design for visually impaired and dyslexic readers-clear text for all. Interacting with Computers 17, 453–472 (July 2005)
17. Huenerfauth, M., Feng, L., Elhadad, N.: Comparing evaluation techniques for text readability software for adults with intellectual disabilities. In: Proceedings of the

11th international ACM SIGACCESS conference on Computers and accessibility. pp. 3–10. ACM (2009)

18. Hyönä, J., Olson, R.: Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. Journal of Experimental Psychology: Learning, Memory, and Cognition 21(6), 1430 (1995)

19. Interagency Commission on Learning Disabilities: Learning Disabilities: A Report to the U.S. Congress. Government Printing Office, Washington DC, U.S. (1987)

20. International Dyslexia Association: Definition of dyslexia: `http://interdys.org/DyslexiaDefinition.htm` (2011), Based in the initial definition of the Research Committee of the Orton Dyslexia Society, former name of the IDA, done in 1994.

21. Inui, K., Fujita, A., Takahashi, T., Iida, R., Iwakura, T.: Text simplification for reading assistance: A project note. In: Proceedings of the second international workshop on Paraphrasing-Volume 16. pp. 9–16. Association for Computational Linguistics (2003)

22. Just, M., Carpenter, P.: A theory of reading: From eye fixations to comprehension. Psychological review 87, 329–354 (1980)

23. Lyons, J.: Semantics, vol. 2. Cambridge Univ Pr (1977)

24. Madnani, N., Dorr, B.J.: Generating phrasal and sentential paraphrases: A survey of data-driven methods. Computational Linguistics 36(3), 341–387 (2010)

25. Nation, K., Snowling, M.: Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. Child Development 69(4), 996–1011 (1998)

26. Patterson, K., Marshall, J., Coltheart, M.: Surface dyslexia: Neuropsychological and cognitive studies of phonological reading. Lawrence Erlbaum Associates, London (1985)

27. Pedley, M.: Designing for dyslexics: Part 3 of 3 (2006), `http://accessites.org/site/2006/11/designing-for-dyslexics-part-3-of-3`

28. Rayner, K., Duffy, S.: Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. Memory & Cognition 14(3), 191–201 (1986)

29. Rello, L., Baeza-Yates, R.: Lexical quality as a proxy for web text understandability. In: The 21st International World Wide Web Conference (WWW 2012). Lyon, France (April 2012)

30. Rello, L., Baeza-Yates, R.: The presence of English and Spanish dyslexia in the Web. New Review of Hypermedia and Multimedia 8, 131–158 (2012)

31. Rello, L., Baeza-Yates, R., Saggion, H., Graells, E.: Graphical schemes may improve readability but not understandability for people with dyslexia. In: Proceedings of the NAACL HLT 2012 Workshop Predicting and improving text readability for target reader populations (PITR 2012) (2012)

32. Rello, L., Kanvinde, G., Baeza-Yates, R.: Layout guidelines for web text and a web service to improve accessibility for dyslexics. In: International Cross Disciplinary Conference on Web Accessibility (W4A 2014). ACM Press, Lyon, France (April 2012)

33. Saggion, H., Martínez, E., Etayo, E., Anula, A., Bourg, L.: Text simplification in Simplext. Making text more accessible. Procesamiento de Lenguaje Natural 47, 341–342 (2011)

34. Sereno, S., Rayner, K.: Measuring word recognition in reading: eye movements and event-related potentials. Trends in Cognitive Sciences 7(11), 489–493 (2003)

35. Simmons, F., Singleton, C.: The reading comprehension abilities of dyslexic students in higher education. Dyslexia 6(3), 178–192 (2000)

36. Sinatra, R., Stahl-Gemake, J., Berg, D.: Improving reading comprehension of dis-abled readers through semantic mapping. The Reading Teacher 38(1), 22–29 (1984)
37. Sterling, C., Farmer, M., Riddick, B., Morgan, S., Matthews, C.: Adult dyslexic writing. Dyslexia 4(1), 1–15 (1998)
38. Tobii Technology: Product description Tobii 50 Series (2005)
39. Vellutino, F., Fletcher, J., Snowling, M., Scanlon, D.: Specific reading disability (dyslexia): What have we learned in the past four decades? Journal of child psy-chology and psychiatry 45(1), 2–40 (2004)
40. Williams, S., Reiter, E., Osman, L.: Experiments with discourse-level choices and readability. In: Proceedings of the 9th European Workshop on Natural Language Generation (ENLG-2003). Budapest, Hungary (2003)

Appendix

The text and the corresponding paraphrases pairs used are shown below.

Text María: Se [premia/otorga un premio] premia a Ana María Matute

Sus lectores [confiaban/tenían la confianza] en ella a pesar de que la humildad de Ana María no [ambicionara/tuviera ambición de] más premios. Tras [aparecer/hacer aparición] en las quinielas como la principal aspirante, finalmente, el Ministerio de Cultura [ha galardoneado/otorgó el galardón] con el Premio Cervantes a la escritora. Ana María Matute [ha contribuido/ha hecho una contribución] a la literatura española con novelas y relatos aunque también [ha atendido/ha prestado atención] al público más joven con cuentos para niños. Ana María tenía diez años cuando [comenzó/dio comienzo] la Guerra Civil Española. Luciérnagas fue su primera obra premiada, pero la [censuraron/impusieron censura] censuraron y no fue publicada hasta años más tarde.

Text Alex: Se [premia/otorga un premio] a Álex de la Iglesia

El Ministerio de Cultura [concedió el/hizo la concesión del] Premio Nacional de Cinematografía al director Álex de la Iglesia. Este premio del Instituto Nacional de la Cinematografía y de las Artes Audiovisuales [contribuye/hace una contribución] a [recompensar/dar una recompensa] a la aportación más sobresaliente en el ámbito cinematográfico español [manifestado/puesta en manifiesto] a través de una obra du-rante el año. En casos excepcionales como éste también se [reconoce/ofrece un re-conocimiento] a una trayectoria profesional. El jurado [valoró/dió valor] a la trayecto-ria profesional de álex de la Iglesia, que [enriquecido/ha aportado riqueza] al lenguaje de nuestro cine, así como su gran labor por poner [acercar/poner más cerca] el cine español a la sociedad.