



This project is partially funded by the European Commission under the Seventh (FP7-2007-2013) Framework Programme for Research and Technological Development. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

D7.2: Benchmark report (results of piloting task)

Author: Dr Vesna Jordanova, Richard Evans, and Arlinda Cerga Pashoja

Affiliation: LNFT, WLV

Date: 14th April 2014

Document Number: FIRST_D7.2_20140414

Status/Version: Approved v 1.2

Distribution Level: Public

<i>Project Reference</i>	287607
<i>Project Acronym</i>	FIRST
<i>Project Full Title</i>	A Flexible Interactive Reading Support Tool
<i>Distribution Level</i>	Public
<i>Contractual Date of Delivery</i>	31st October 2012
<i>Actual Date of Delivery</i>	14 th April 2014
<i>Document Number</i>	FIRST_D7.2_20140414
<i>Status & Version</i>	Approved (v1.2 modified to address evaluators' comments from the second review meeting)
<i>Number of Pages</i>	40
<i>WP Contributing to the Deliverable</i>	WP7
<i>WP Task responsible</i>	Vesna Jordanova
<i>Authors</i>	Vesna Jordanova, Richard Evans, Arlinda Cerga Pashoja

<i>Other Contributors</i>	Eduard Barbu, Sandra Freire, Ana González, Martin Keats, Elena Lloret, Eugenio Martinez, Paloma Moreda, Sanja Stajner, Antoneta Shishkova
<i>Reviewer</i>	Constantin Orasan
<i>EC Project Officer</i>	Magdalena Szwochertowska
<i>Keywords:</i>	Benchmarking, texts, simplification, autism
<p><i>Abstract:</i></p> <p>Benchmarking involves structured comparisons to help define and implement best practice. Current practices of text simplification for people with Autistic Spectrum Disorders (ASD) involve manual processing of written information. Best practices are delivered by experienced special needs teachers and carers who have good understanding and knowledge about autism and its associated reading difficulties. The objective of the benchmarking exercise was to determine how long it might take for a special needs teacher or a professional carer to simplify written documents satisfactorily for adults with ASD so that a comparison might be drawn with the “Open Book” software as an alternative, digital service. The benchmarking task aims to provide a comparative measure of speed and accuracy provided by experienced professionals. In this way the software writers would know the degree of speed and accuracy that users of the software would have to exceed in order to justify its use. We recruited 15 professionals (special needs teachers and professional carers) with experience of working with people with ASD. They were given five different texts that they simplified manually to make them comprehensible for people with ASD. A total of 75 different texts covering a wide range of topics were analysed. The numbers of words in the original and simplified versions were recorded and compared as were the individual times taken as recorded by the teachers and carers who volunteered to take part in the study.</p> <p>Professionals took just over 54 minutes in average to simplify a text of 250-350 words long. Simplifying text for adult users took 23 minutes, professionals working with children texts spent 97 minutes and those addressing both target groups spent 43 minutes. They worked at the average speed of 12 words per minute.</p> <p>For all three languages, the use of established readability metrics demonstrated differences in the readability of the original and simplified versions of the texts. In several cases, these differences were statistically significant.</p>	

Contents

1. Introduction	4
2. Objectives	5
3. The benchmarking exercise	5
4. Method	5
4.1 Benchmarking exercise key method points	5
4.2 Recruitment of Teachers and Care-workers	6
4.3 Identification of texts for simplification	6
4.4 Text simplifications; Identification and Resolution of Obstacles	8
4.5 Use of aids in simplification of texts	13
4.6 Data gathering	13
4.7 Statistical methods	13
5. Results	14
5.1 Text simplifications	14
5.2 Choice of readability indices	16
5.3 Readability analysis	18
5.4 Unexpected results	26
5.5 Analysis of times taken to simplify texts	27
6. Discussion	28
6.1 Readability	28
6.2 Text simplification times	29
6.3 Limitations of the research method	31
6.4 Strengths of the research method	32
7. References	33
Appendix 1	37

1. Introduction

Communication needs of people with learning difficulties including people with Autistic Spectrum Disorders (ASD) have been addressed by different means. One of the essential methods used is by making information easier to understand. Different terminologies have been used to describe this process such as inclusive communication, total communication, easy read etc. The most common term used at present in the UK is ‘accessible information’ which has been defined by Mander (2008) as *‘A supportive process of making information easier for people with learning disabilities, that firstly involves simplifying the linguistic message and secondly conveying the simplified message indifferent mode(s) of communication, i.e. not just the written word or spoken message’* (1).

Providing accessible information for people with learning difficulties is not only considered best practice, but is also set within a legal framework in the UK. Firstly, the Disability Discrimination Act (1995) (2) instructs that ‘reasonable adjustments’ need to be made to improve access for people with disabilities and the Human Rights Acts Article 10 (1998)(3) sets out the right to information as does the Communication Bill of Rights(1992)(4). The Mental Capacity Act (1998) (5) emphasizes the need for ‘practicable steps’ to be taken to support individual communication needs with reference to the decision making process. Further significant documents such as Valuing People (DoH, 2001) (6) and Valuing People Now (DoH, 2009) (7) state that *‘the government expects organisations working with learning disabled people to develop communication policies and produce and disseminate information inaccessible formats’*. The legal basis for the European Union have also been provided by a series of documents such as: Article 13 of the European Treaty, which instructs the European Council to *‘take appropriate action to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age and sexual orientation’*; the Charter of Fundamental Rights and by the Commission communication ‘Towards a barrier free Europe for people with disabilities’ (European Commission 2000a).

Benchmarking is the process of comparing one's performance metrics to best practices. Dimensions typically measured are quality, time and cost (8). Learning through benchmarking can provide assurance as to the effectiveness of existing arrangements as well as identifying improvement points. Improvements mean doing things better, faster, and cheaper in the future (9).

The NHS Benchmarking Network (10) uses the following definition as a simple and accessible statement of what benchmarking is and why we do it: “Benchmarking is the use of structured comparisons to help define and implement best practice”. In the process of benchmarking, the best existing practice in the field is identified, and then it’s compared to the results and processes of those studied to one's own results and

processes. In this way, it is learned more about performance and, more importantly, the processes that explain why these practices are successful.

Current practices of text simplification for people with ASD involve manual processing of written information by those who work with this vulnerable group. Best available practices of text simplification relate to the work carried out by experienced special needs teachers and professional carers who have good understanding and knowledge about ASD and its associated reading difficulties. In their routine work they spend considerable time adapting various documents and making them more accessible for readers with autistic spectrum disorders. From an applied perspective, the goal is to implement a system for automatic text simplification that can perform the wide range of lexico-syntactic revisions that teachers and carers perform in their routine practice.

2. Objectives

The objective of this benchmarking exercise is to determine how long it might take for a special needs teacher or a professional carer to simplify written information satisfactorily for adults and children with ASD so that a comparison might be drawn with the “Open Book” software soon to be built as an alternative, digital service.

3. The benchmarking exercise

Following the gathering of the comprehension test data (WP2) which will help the technical partners to determine what areas to concentrate on when building the new program, the next task is to provide a comparative measure of speed and accuracy provided by experienced professionals. In this way the software writers would know the degree of speed and accuracy that users of the software would have to exceed in order to justify its use.

4. Method

4.1 Benchmarking exercise key method points

- Identification of simplification strategies that improve reading for people with ASD.
- Recruitment of five professionals (special needs teachers and health professionals) with experience of working with people with ASD in each clinical centre in UK, Spain and Bulgaria.
- Identification of 25 separate pieces of text suitable for simplification for adults with ASD in the UK and Spain, and for children in Spain and Bulgaria.
- Random allocation of 5 texts to each of the 5 professionals in each country.

- Professionals given instructions to simplify the texts to improve reading comprehension
- Manual text simplification was carried out by each professional and the time required for each text was recorded.
- Data gathering included the 25 original and simplified texts along with the 25 recorded times from the professionals in all three countries.
- Original and simplified versions of the texts were compared in terms of word length to ascertain the degree to which simplification expands or compresses the amount of information in a document. The time taken by professionals to simplify the texts was also obtained in the UK.
- Analysis of text readability was carried out by our technical partners from Wolverhampton University.
- Findings overall and for individual countries were obtained.
- Comparisons were made between countries.
- Benchmarks were identified.

4.2 Recruitment of Teachers and Care-workers

Recruitment was pragmatic, based on existing contacts with professionals. However, we attempted to have a good representation of different professionals that carry out text simplifications in real life for people with ASD. We also tried to attain a good balance regarding age (26-56 years old), education/qualifications and length of experience. The overall sample of professionals consisted of seven teachers, six psychologists, one speech and language therapist and one psychiatrist. All recruited professionals, were sent five different texts, identified locally by random allocation. They were asked to complete the work and return the original texts, the simplified versions and their timing figures to the UK centre where analysis was conducted. UK professionals worked with an adult audience in mind, while as the Bulgarian colleagues addressed children only. Professionals in Spain aimed their work towards both children and adults.

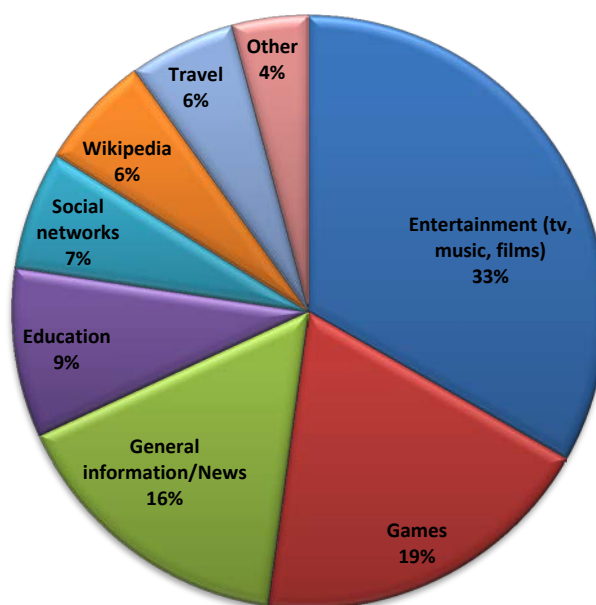
4.3 Identification of texts for simplification

Text identifications were carried out based on an attempt to keep a right balance between internal and external validity of the benchmarking exercise. While the use of a limited number of texts uniformly between the three participating centres would have provided excellent internal validity, this would have rendered the study not generalizable to the outside world. Therefore, different texts were used in order to address participating sites' differences in target age-groups, language and culture.

Considering it is impossible to identify different texts that are identical in terms of word count and language complexity, it was decided that each professional would be presented with five different texts, each with a word length of between 250 and 350 words. This method was also employed to aid rater reliability, as speed of text simplification will be affected by text complexity and varies between texts. The texts were independently selected by the research teams in the UK, Spain and Bulgaria. Criteria for selecting text themes were based on the cohort (UK, Spain and Bulgaria) outcomes from the requirement survey (WP2).

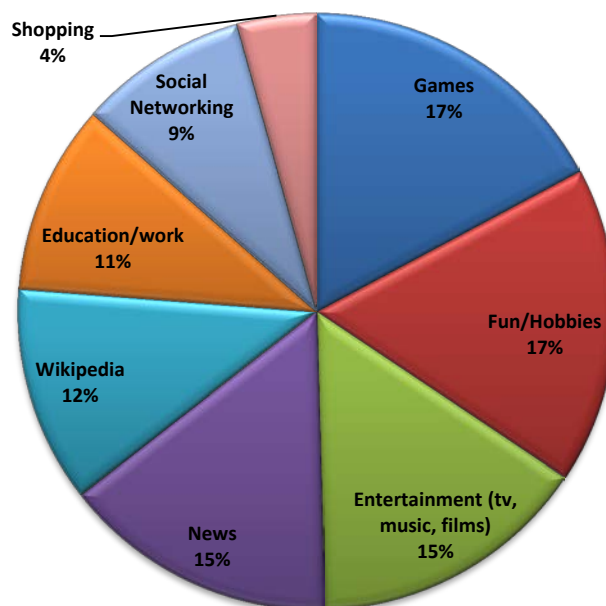
Children and adults that participated in the user requirement survey (WP2) identified preferred internet topics. Children participants (Figure 1) seem to use the internet mostly for entertainment (TV, music, film and games). However, they also read news or general information and utilise it for school work. Adult survey participants (Figure 2) reported that they use the internet for games, hobbies, entertainment as well as to read news and for education/work purposes. Figure 1 and Figure 2 provide a full description of these preferences.

Figure 1 Preferred internet topics for children with ASD in Spain and Bulgaria *



** Children in the UK are not being investigated therefore not included in the survey

Figure 2 Preferred internet topics for adults with ASD in the UK and Spain[†]



Consequently attention was focused to identify suitable texts on the following areas:

- News
- Health
- General culture
- Literature
- Education material

Considering the interface of Facebook, YouTube and eBay it was not viable to identify suitable texts from these sources. Therefore they were excluded and the rest of the websites in Figure 1 and Figure 2 were used for text identification. Researchers were also mindful to identify texts containing examples of comprehension obstacles such as: phraseological units, polysemic words, anaphors, abbreviations, acronyms, temporal concepts, and specialized terms.

4.4 Text simplifications; Identification and Resolution of Obstacles

In the absence of international or national guidelines for producing accessible texts for people with ASD, professionals were instructed to use best clinical practice to simplify each text. Further instructions were

[†] Adults in Bulgaria are not being investigated therefore not included in the survey

compiled aimed to make sure the text simplifications were appropriate for people with ASD; improved reading comprehension for the target group; and simplifications were carried out on all relevant obstacles.

Identification and resolution of text obstacles relevant to ASD reading comprehension were guided by two important elements:

- The literature review conclusions
- User requirement outcomes (WP2)

➤ Literature review

Several studies show that people with Asperger Syndrome have excellent phonetic decoding ability but poor comprehension (Myles 2002, O'Connor and Klein 2004, Miniscalco et al 2009, Smith, 2011; Hambley, 2011, Walters, 2011). Similar results were found by others, such as Huemer and Mann (2009), as well as Hambley (2011) who compared reading accuracy vs. reading comprehension in population with ASD.

Traditionally the difficulty with reading comprehension has been related with the cognitive profile of the readers especially with their inability to comprehend the perspective of others (Baron-Cohen 2001) and a lack of Central Coherence (Norbury 2005, Happé 2010). Recently it has also been related with a poor executive function (Sesma et al 2009). Gold et al (2010) suggest that differences in linguistic information processing cause difficulties in novel **metaphor comprehension** in ASD, who show limited language abilities compared to control participants. Thus, several studies confirm that people with ASD struggle to understand metaphors and make inferences about social scripts (Dennis, Lazenby and Lockyer, 2001; Heimann, Nelson, Tjus, and Gillberg, 1995; Beversdorf et al. 1998).

Although people with ASD often have severe problems with pragmatic aspects of language, little is known about the pragmatic reasoning. This was one of the conclusions of the study developed by Jolliffe and Baron-Cohen (2000) when they studied the comprehension of **homographs** and inferences. Same conclusions were drawn by Gold, Faust and Goldstein (2010) after analysing the semantic integration during metaphor comprehension.

The difficulties in processing verbal information are a great challenge to people with ASD throughout life. Even though reading accuracy improves with age in high functioning children with ASD, they continue to struggle with many linguistic phenomena such as **homographs, multiple meaning words, phrases and metaphors** (Huemer and Mann 2010; Jones et al. 2009; Prior and Hall 1979). A study comparing the cognitive and language development of children with ASD and children with Developmental Receptive Language Disorder shows that despite the improvements shown by the Autism group on standardized

language tests over time, their overall level of communicative competence in adulthood was generally far poorer than in the language group (Howlin, 2000).

Strategies to improve reading skills in Autistic Spectrum Disorder

In order to improve reading comprehension for people with ASD, different authors suggest several strategies based on findings from various studies. For example, the importance of using computer assisted tasks or instruction to improve phonological and reading scores has been suggested by authors for almost two decades (Heimann et al 1995, Tjus et al 1998, Moore and Calvert 2000, Coleman et al 2005). O'Connor and Klein (2004) studied a sample of 25 children with ASD exploring the effects of three facilitating strategies for improving the reading comprehension: (1) answering pre-reading questions, (2) completing close sentences embedded in the text and (3) resolving anaphora cuing by identifying. They concluded that text comprehension was improved when **anaphors were explained**, an **abstract of the text** was provided and **graphics were used**.

➤ **User requirements**

Functional requirements specifications and user preference survey (D2.1 and D2.2) guided further the simplification strategies employed by our professionals. Table 1 presents main obstacles and resolutions that were identified by people with ASD in the UK, Spain and Bulgaria as having a higher impact on reading comprehension.

Obstacle	Resolution
Multiple copulative coordinated clauses	Substitute with sentences divided by periods.
Long sentences	Sentences < 15 words
Semicolon and suspension points	Avoid the use of semicolon and suspension points
Brackets and non common punctuation marks (&,% , /...)	Avoid non common punctuation marks
Improper grammar	Correct grammar
Polysemy	Avoid using easier synonym. Detect and highlight when domain is not clear
Phraseological units (idioms, Lexicalized metaphors)	Substitute by a simple word. Highlight when substitution is not possible
	Provide simple definitions to explain phraseological units
	Substitute infrequent words with simpler synonym
	Provide simple definitions to explain infrequent words
Slang	Substitute infrequent slang with simpler synonym
	Provide simple definitions to explain slang
Infrequent acronyms and abbreviations	Expand infrequent acronyms and abbreviations
Temporal adjectives	Disambiguate temporal adjectives
Anaphors	Resolve all types of anaphors when possible. Leave anaphors with low resolution confidence level.
Non lexicalized metaphors	Provide idea of inferred meaning when possible and highlight
Long paragraphs	Divide long paragraphs

Table 1 Obstacles and resolutions identified by user survey

All the examples given above confirmed literature review findings and complimented them with real world suggestions. Based on the above a document was composed (Table 2), which gave specific instructions to professionals on how to simplify the given texts.

Instructions to professionals: How to Simplify Texts

- You need to time yourself from the moment you start reading each text to the moment you finish the simplification. This should be done separately for each text.
- Should you need to interrupt the work, pause the timer and restart timing once you restart simplification task.
- Please **do not read the texts beforehand**.
- Read the instructions below carefully. Do not start simplifications if you are not clear or sure about any of the instructions (or parts of them). Please contact me if you do not understand or need clarification about the instructions below before starting the task.

After you have read the text carefully please:

1. **Detect infrequent words and substitute them with simpler synonyms or definitions (in case no simpler synonyms exist).**
 2. **Identify figurative expressions such as idioms and metaphors and replace with simpler words or definitions.**
 3. **Identify jargons or specialised terms and replace with specific definitions.**
 4. **Identify phraseological units and polysemic words and replace with specific definitions.**
 5. **Identify and divide long paragraphs.**
 6. **Detect long sentences and divide them into shorter easier to understand chunks.**
 7. **Rewrite complicated sentences to make them easier to understand.**
 8. **Identify and resolve anaphora.**
 9. **Replace abbreviations and acronyms with full definitions.**
 10. **Use bullets points if necessary to break down the text in easier parts.**
- Email back both the original and simplified versions of each text.
 - At the bottom of each text please make a note of the how long it took for the text to be simplified (in minutes).

Dictionary of key terms

- **Anaphora** is: A rhetorical term for the repetition of a word or phrase at the beginning of successive clauses (i.e. 'Ben went to work to find John sitting at his desk. **He** was not happy about this'. Although we deduct that Ben [not John] was not happy, people with ASD can get confused. Therefore replace 'he' with Ben -> Ben went to work to find John sitting at his desk. **Ben** was not happy about this')
- **Abbreviations:** Mr. (Mister), Prof. (Professor), op. (opus), mm (millimeters)
- **Metaphors:** A metaphor is a figure of speech that describes a subject by asserting that it is, on some point of comparison, the same as another otherwise unrelated object e.g. *Broken heart, light of my life, feel blue, rollercoaster of emotions etc*
- **Acronyms:** FBI->Federal Bureau of Investigation; NATO-> North Atlantic Treaty Organization, laser, scuba....
- **Phraseological units:** *bring about, fold away etc*
- **Idioms:** *Cry my Eyes Out, A Picture Paints a Thousand Words, Slap on the Wrist, Against The Clock, Break A Leg etc*
- **Temporal concepts:** *previously, currently, finally etc*
- **Specialized terms:** jargons such *asecophysiology etc*

Table 2 Instructions given to professionals regarding text simplifications

4.5 Use of aids in simplification of texts

Professionals used on-line dictionaries and thesauruses to help with the simplification process. Furthermore it should be noted that professionals in Spain incorporated images from the web into their simplification of their chosen texts to improve comprehension for the intended readership.

4.6 Data gathering

The original and simplified texts were received back to the UK team (FIRST_D7.2_V3_annex). The numbers of words in the original and simplified versions were noted and compared as were the individual times taken as recorded by the professionals who volunteered to take part in the study.

Language obstacles were recorded and each modification was noted down.

4.7 Statistical methods

All data were entered onto SPSS (Version 18.0) for data analysis. Distribution and central tendency (such as Standard Deviations (SD) for normally distributed continuous data) were examined and simple descriptive statistics used to describe the samples. Comparisons were made by using paired t-tests.

4.8 Choice of readability indices

The indices used to assess the readability of texts were selected on the basis of their ability to be derived reliably using consistent automatic methods and their ability to provide information on linguistic phenomena likely to pose obstacles to reading comprehension for people with autism. Eight language independent metrics were applied to texts of all three languages (BG, EN, and ES), as was one language specific one. Their exploitation is justified by relevance to the potential user requirements presented in D7.2. This distinguishes them from existing indices that pertain to readability for different types of reader (such as high school students or readers with different types of reading comprehension impairment). The selected indices are detailed in Section 5.2.

5. Results

5.1 Text Simplifications

Texts used for the benchmarking task were inspected manually as well as electronically. Manual inspections were focused towards aspects that could not be identified electronically such as metaphors, idioms, specialised terms etc. (Table 3)

Simplification Operations	Spanish Texts (frequency)	English Texts (frequency)	Bulgarian Texts (frequency)
Synonyms An entity usually a noun or an adjective is replaced with its less complex synonym	64	83	67
Sentence Splitting A long sentence is split in shorter sentences or in bullet list.	40	70	70
Definition A difficult term is explained using a explanation extracted from Wikipedia, a dictionary or Web	34	65	--
Near – Synonymous	33	21	36
Image A concept is illustrated by an image	27	N/A	29
Explanation The sentence is rewritten with different words	24	43	116
Anaphora Resolving the anaphor to its antecedent	17	--	1
Deletion Parts of the sentence are removed	17	58	69
Coreference resolution	9	3	--
Syntactic Operation A transformation on the syntactic parse trees	9	42	73
Metaphor A metaphor is explained	8	30	19
Summarization It gives the gist of the sentence or the paragraph	3	10	1

Table 3. Manually inspected simplification operations

Metaphors and idioms were identified and replaced with easier to understand explanation or definitions e.g:

- Text 21, Bulgarian: '*не си дават сметка*' Changed to -> '**не осъзнават**'.
- Text 9, Spanish: '*... fuertes precipitaciones que han dejado ocho muertos en Murcia*' Changed to -> '**... fuertes precipitaciones por las que murieron ocho personas en Murcia...**'
- Text 5, English: '*..as world leaders struggle to find a way to resolve the war that has left the Security Council hopelessly deadlocked.*' Changed to -> '**..The world leaders still haven't decided on what needs to be done. Therefore Security Council cannot take steps to resolve the issue.**'

Long sentences were broken up into shorter sentences e.g. :

- Text 1 English: '*Ultra-thin electronics that dissolve inside the body have been devised by scientists in the US and could be used for a range of medical roles.*' Changed to -> '**Scientists in America devised very thin electronics that dissolve inside a body. The device can be used in medicine**'.
- Text 22, Bulgarian: '*Освен, че трябва да си купиш колело, което се забелязва лесно, трябва да се облечеш в цветове, които се виждат от далеч... и никакви веещи се поли!*' Changed to --> '**Купи си колело. То трябва да се вижда отдалече. Трябва да си облечен в ярки цветове. Ярките цветове ще се виждат отдалече. Не носи широки поли.**'
- Text 16, Spanish: "*... uno de los guitarristas más inventivos y aventureros de la actualida, poseedor de una técnica asombrosa...*" Changed to -> "**... uno de los guitarristas mas inventivos y aventureros. Posee una técnica maravillosa...**"

Infrequent or rare words were replaced with easier to understand, more commonly used, synonyms e.g :

- Text 7, Bulgarian: '*Медоносните пчели са невероятни създания*' Changed to -> '**Пчели, които правят мед се наричат медоносни пчели.**'
- Text 16, Spanish: '*... sorprenden a la audiencia...*' Changed to -> '**... sorprenden al público...**'
- Text 6, English: '*..the English governess quarrelled with the housekeeper..*' Changed to -> '**..An English woman who was paid to look after the children argued with the woman who looked after the house. ..**'

Specialised terms or jargons were explained e.g:

- Text 23, Bulgarian: '*..микроорганизми..*' Changed to -> '**..много малки живи организми..**'
- Text 2, English: '*Previous studies suggest that the most prominent biological responses are changes in distribution, phenology and productivity.*' Added-> '**Phenology is the study of plant and animal life cycle events and how these are influenced by seasonal and changes in climate.**'

Anaphora resolutions were also made in several texts:

- Text 11, Spanish: ‘... *antes de hablar con nuestros hijos. Son los amos de las excusa*’ Changed to -> ‘... **antes de hablar con nuestros hijos. Los adolescentes ...**’
- Text 16, Bulgarian: ‘*В двора имаше череши, която цъфтеше с бели цветове. Човекът имаше и голям червен петел.*’ Changed to -> ‘**В двора на къщата му имаше череши. Черешата цъфтеше с бели цветове.**’
- Text 3, English: ‘They flew by propeller plane to the beautiful island of KohSamui, where they went kayaking, before taking a boat to the neighbouring island of Koh Tao, where they drank pink cocktails overlooking the sea.’ Changed to -> ‘**They flew there by propeller plane. Then they went kayaking on the beautiful island of KohSamui. They also took a boat to the island of Koh Tao. There they enjoyed the sea-view and drank cocktails.**’

Acronyms were explained e.g:

- Text 24, English: ‘... *Ram...*’ Changed to -> ‘**..Random Access Memory..**’
- Text 6, Bulgarian: ‘*ООН*’ Changed to -> ‘**Организацията на Обединените Нации. Това е организация между народите за мир. Нарича се със съкратено име.**’
- Text 23, English: ‘...*LGBT activism.*’ Changed to -> ‘**...lesbian, gay, bisexual, and transgender activism..**’

5.2 Choice of readability indices

The eight language independent readability indices used to assess the texts used in benchmarking are:

- Commas index**, which indicates the average syntactic complexity of the sentences occurring in the texts (relevant to user requirements UR301-303 and UR307[‡]). It is calculated as the ratio of 100 multiplied by the number of commas occurring in a text to 10 multiplied by the number of words occurring in the text. Accessible texts are expected to have smaller values for the *commas index*.
- Index of words with three or more syllables**, which indicates the length of the words in the texts measured in syllables (relevant to user requirements UR316, UR318, UR319). In previous research, the syllable length of words in documents has been considered an indicator of readability (Kitson, 1921; Gray and Leary, 1935; Coleman, 1965; Coleman, 1971; Flesch, 1948; Swanson, 1948; Farr et al., 1951; Gunning 1952; Fry, 1977; McLaughlin, 1969; Caylor et al., 1973; Kincaid et al., 1975, etc). LT exploiting multilingual LaTeX hyphenation patterns was developed in order to obtain approximations of the numbers of syllables per word. The index is calculated as the ratio of 100 multiplied by the number of words with 3 or more syllables in a text to 10 multiplied by the total number of words in the text. Accessible texts are expected to have smaller values for the *index of words with three or more syllables*.

[‡] The full set of user requirements is detailed in project deliverable D2.2.

- c. **Index of words in sentences**, which indicates the numbers of words in sentences in the texts (relevant to user requirement UR309). This index is calculated as the ratio of the number of words in a document to the number of sentences in that document. Accessible texts are expected to have smaller values for the *index of words in sentences*.
- d. **Index of word diversity**, which indicates the number of different words used in a document (marginally relevant to user requirements related to ambiguity in meaning, as the greater the number of different words occurring in a text, the greater the likelihood that one or more of them will constitute obstacles to reading comprehension). This index is calculated as the ratio of the number of word types in a document to the number of word tokens in that document. Accessible texts are expected to have smaller values for the *index of word diversity*.
- e. **Pronoun index**, which indicates the percentage of words in the document that are anaphoric pronominal references (relevant to user requirements UR418 and UR419). This index is calculated as the ratio of 100 multiplied by the number of pronouns in a document to the number of words in that document. Accessible texts are expected to have smaller values for the *pronoun index*.
- f. **Metaphor index**, which indicates the percentage of sentences in the document that contain phraseological units and non-lexicalised metaphors (relevant to user requirements UR402, UR410, UR422, UR425, and UR507-508). This index is calculated as the ratio of 100 multiplied by the number of phraseological units and non-lexicalised metaphors in a text to the number of sentences in that text. Accessible texts are expected to have smaller values for the *metaphor index*.
- g. **Passive verb index**, which indicates the percentage of sentences in the text that are in the passive voice (relevant to user requirement UR313). LT was implemented to detect the occurrence of passive verbs on the basis of a set of part of speech patterns. The passive verb index is calculated as the ratio of 100 multiplied by the number of sentences in a text that contain a passive verb to the total number of sentences in that text. Accessible texts are expected to have smaller values for the *passive verb index*.
- h. **Polysemic word index**, which indicates the percentage of words in the document that are semantically ambiguous (relevant to user requirements UR401, UR425, UR504, UR505, and UR511). LT exploiting WordNet for English, EuroWordNet for Spanish and the Bulgarian portion of BalkaNet for Bulgarian was implemented to acquire the number of different senses of each word in a text. The Polysemic word index is calculated as the ratio of 100 multiplied by the number of words with two or more senses in a text to the total number of words in that text. Accessible texts are expected to have smaller values for the *polysemic word index*.

LT was developed to acquire values for indices a-e and g-h automatically. Values of index f (*metaphor index*) were obtained by manual inspection of the original and more accessible versions of each text used in benchmarking.

Indices a-d are exploited in the readability index proposed by Garcia (2012). Indices e-h were added in order to achieve sensitivity to the occurrence of ambiguity in meaning (e, f, and h), and a particular problematic type of structural complexity (g). It would be possible to exploit these indices as terms in a novel formula to assess the readability of documents for people with ASD. However, this has not been done for two reasons. First, the relative significance of each linguistic factor for autistic readers has not yet been determined.

Second, combining the indices in one formula would obfuscate potential reasons for differences in accessibility.

One language specific index was used to assess the readability of texts in each of Bulgarian and English:

- a. **Verb temperature** (BG; Gerganov et al., 1976) is the proportion of tokens in the text that are verbs. In the work described in this report, the Tree Tagger[§] was used for tokenisation and identification of verbs in Bulgarian text. Information about the frequency of verbs in a text indicates the number of distinct clauses and events described. Accessible documents are expected to have smaller values of *verb temperature*.
- b. **Index of productive syntax** (EN; Scarborough, 1990) was developed to detect the presence of numerous syntactic constructions in natural language “to serve as a readily obtained summary scale of grammatical complexity that would be appropriate for the study of individual differences in [child] language acquisition”. In FIRST, values of this index were obtained by means of the *AC-IPSyn*^{**} distribution which detects and counts occurrences of various types of words and syntactic constituents (phrases and clauses) in texts. Accessible documents are expected to have smaller values for the *index of productive syntax*.

5.3 Readability Analysis

5.3.1 Bulgarian

	Sents	Words/Text	Char/Text	Words/Sent	Char/Word	SC Sign Usage
Original (μ)	18.56	345.08	3099.00	20.01	8.99	5.73
Simplified (μ)	30.64	333.64	2960.56	11.52	9.17	4.36
t-test	0.00	0.22	0.20	0.00	0.71	0.00

Table 4: Readability statistics (BG)

Table 4 displays a summary of readability information derived from the original Bulgarian documents and the simplified versions produced by participants in the benchmarking experiment. The first row contains the mean values of each characteristic calculated for the original forms of the documents. The second row contains this information for the manually simplified forms of the documents. The final row displays 2-tailed statistical significance of the paired t-test. Numbers in bold font indicate that the difference between the original and simplified texts is statistically significant ($\alpha=0.001$). The columns present different

[§] Available at <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/> with BG resources available at <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/bulgarian-par-linux-3.1.bin.gz> and description of the BG tag set available at <http://www.bultreebank.org/TechRep/BTB-TR03.pdf>.

^{**} Available at <http://www.hlt.utdallas.edu/~nisa/ipsyn.html>.

characteristics of the documents. These are the number of sentences in the document (*Sents*), the number of words in the document (*Words*), the number of characters (letters) in the document (*Chars*), the average sentence length (*Words/Sent*), the average word length (*Chars/Word*), and the percentage of tokens in the document that are signs of syntactic complexity^{††} (*SC Sign Usage*).

Metric	<u>Comma Index</u>		<u>Index of words with > 3 syllables</u>		<u>Index of words in sentences</u>		<u>Index of word diversity</u>		<u>Pronoun index</u>	
File	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP
bg.Text10.txt	0.66	0.44	2.47	2.65	20.28	12.74	0.62	0.60	7.67	5.81
bg.Text11.txt	0.49	0.47	3.22	3.24	18.28	15.74	0.59	0.60	10.94	10.70
bg.Text12.txt	0.66	0.31	3.75	3.36	17.71	13.63	0.58	0.57	9.30	9.79
bg.Text13.txt	0.53	0.41	3.26	3.15	24.36	13.78	0.60	0.63	15.54	14.20
bg.Text14.txt	0.51	0.11	3.36	3.34	16.39	15.20	0.37	0.37	2.71	5.00
bg.Text15.txt	0.60	0.26	4.72	3.58	32.17	9.05	0.61	0.57	3.89	8.14
bg.Text16.txt	0.94	0.37	1.85	1.94	16.92	9.00	0.51	0.48	16.75	15.69
bg.Text17.txt	0.42	0.11	1.89	2.10	9.34	7.00	0.53	0.52	11.83	10.78
bg.Text18.txt	0.63	0.22	1.82	2.10	15.85	7.33	0.56	0.57	13.32	12.98
bg.Text19.txt	0.67	0.45	1.73	1.86	15.00	11.64	0.61	0.63	15.33	14.78
bg.Text1.txt	0.56	0.38	2.66	2.67	15.39	12.12	0.57	0.60	14.12	14.92
bg.Text20.txt	0.65	0.54	2.39	2.22	15.33	12.76	0.58	0.57	11.41	12.16
bg.Text21.txt	0.64	0.42	3.14	3.24	24.92	15.24	0.59	0.61	8.36	8.49
bg.Text22.txt	0.75	0.29	2.66	2.52	15.25	9.12	0.59	0.59	12.79	11.29
bg.Text23.txt	0.46	0.70	3.92	4.11	18.00	7.52	0.64	0.63	6.21	2.53
bg.Text24.txt	0.79	0.34	3.36	2.19	30.75	15.48	0.47	0.43	6.78	13.20
bg.Text25.txt	0.90	0.49	2.46	2.44	22.31	11.60	0.61	0.59	10.92	8.33
bg.Text2.txt	0.81	0.37	3.05	2.50	19.28	12.00	0.59	0.53	10.37	12.64
bg.Text3.txt	0.67	0.32	2.24	2.06	18.05	9.69	0.59	0.57	17.20	15.63
bg.Text4.txt	0.76	0.40	2.48	2.58	14.68	10.94	0.53	0.55	16.35	16.94
bg.Text5.txt	0.84	0.44	3.34	2.48	20.19	12.25	0.59	0.55	6.81	11.66
bg.Text6.txt	0.68	0.26	2.45	2.56	29.58	12.45	0.56	0.56	10.70	7.25
bg.Text7.txt	0.46	0.07	3.12	3.22	23.07	9.50	0.61	0.59	7.51	11.84
bg.Text8.txt	0.54	0.18	2.42	3.13	23.40	9.38	0.61	0.63	10.26	5.51
bg.Text9.txt	0.51	0.35	2.81	2.34	23.64	12.82	0.56	0.61	6.34	8.16
μ	0.64	0.35	2.82	2.70	20.00	11.52	0.57	0.56	10.54	10.74

Table 5a: Readability of original and manually converted forms of BG texts used in benchmarking experiments

Values of the readability indices presented in Section 5.2 are presented in Tables 5a and 5b. They are broadly in line with expectation (exceptions in all three languages are discussed in Section 5.4). For manually converted versions of the BG texts, values of *comma index*, *index of words in sentences*, and *metaphor index*, and *passive verb index* were significantly^{††} smaller than those derived from the original versions ($\alpha < 0.03$ in

^{††} Punctuation, conjunctions, complementisers, and wh-words.

^{††} In this section, the adverb *significantly* will denote “with statistical significance” for a two-tailed t-test with maximum threshold for $\alpha=0.05$.

each case). Values of the *index of word diversity* were also smaller, but not significantly so ($\alpha=0.69$), for manually converted texts.

Metric	<u>Metaphor index</u>		Passive verb index		Polysemic word index		Verb temperature	
	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP
bg.Text10.txt	16.67	3.70	0.00	3.70	54.80	58.41	0.13	0.13
bg.Text11.txt	5.56	0.00	0.00	0.00	48.41	59.77	0.14	0.13
bg.Text12.txt	0.00	0.00	5.88	8.33	58.89	63.13	0.13	0.13
bg.Text13.txt	0.00	0.00	7.14	4.35	58.31	61.02	0.13	0.13
bg.Text14.txt	0.00	0.00	0.00	0.00	42.41	54.23	0.08	0.11
bg.Text15.txt	8.33	0.00	25.00	0.00	63.10	62.95	0.12	0.16
bg.Text16.txt	12.50	1.96	0.00	0.00	52.89	50.99	0.19	0.18
bg.Text17.txt	5.26	0.00	2.63	1.89	56.41	51.47	0.18	0.19
bg.Text18.txt	7.41	0.00	11.11	3.28	48.34	51.42	0.17	0.19
bg.Text19.txt	5.00	4.00	0.00	0.00	56.71	51.45	0.18	0.16
bg.Text1.txt	4.35	3.85	4.35	3.85	49.85	53.18	0.19	0.19
bg.Text20.txt	4.17	3.45	0.00	0.00	50.17	54.74	0.14	0.14
bg.Text21.txt	0.00	0.00	0.00	0.00	51.61	47.63	0.14	0.15
bg.Text22.txt	5.00	0.00	5.00	0.00	72.54	73.16	0.16	0.18
bg.Text23.txt	0.00	0.00	5.88	0.00	42.23	56.69	0.11	0.07
bg.Text24.txt	0.00	0.00	8.33	0.00	65.76	67.32	0.09	0.12
bg.Text25.txt	6.25	0.00	0.00	6.67	60.85	59.84	0.13	0.14
bg.Text2.txt	11.11	0.00	11.11	0.00	53.50	60.40	0.14	0.17
bg.Text3.txt	10.53	0.00	0.00	2.86	56.67	56.01	0.16	0.17
bg.Text4.txt	12.00	5.88	0.00	0.00	56.52	60.81	0.19	0.20
bg.Text5.txt	6.25	0.00	6.25	0.00	50.50	50.97	0.10	0.15
bg.Text6.txt	8.33	0.00	0.00	6.45	46.56	55.48	0.12	0.13
bg.Text7.txt	0.00	0.00	6.67	0.00	36.60	44.30	0.12	0.17
bg.Text8.txt	0.00	0.00	13.33	0.00	53.66	72.47	0.14	0.10
bg.Text9.txt	0.00	0.00	14.29	4.55	59.66	62.93	0.15	0.15
μ	5.15	0.91	5.08	1.84	53.88	57.63	0.14	0.15

Table 5b: Readability of original and manually converted forms of BG texts used in benchmarking experiments

5.3.2 English

	Sents	Words/Text	Char/Text	Words/Sent	Char/Word	SC Sign Usage
Original (μ)	12.40	281.40	1336.04	24.36	4.76	4.13
Simplified (μ)	20.28	300.72	1372.68	16.00	4.60	4.57
t-test	0.00	0.11	0.49	0.00	<u>0.01</u>	0.19

Table 6a Readability statistics 2 (EN)

Table 6a presents readability statistics for the original and simplified versions of English documents. Statistically significant differences between the two are indicated by values for *t-test* presented in bold type ($\alpha=0.001$) and underlined italics ($\alpha=0.05$). It can be noted that there is a statistically significant difference in word length (measured in characters) between the original and simplified versions of the texts. Inspection of the mean values of *Char/Word* in Table 6a indicates that the simplified versions of the texts contain words that are shorter than those in the original versions. There are significant differences in the lengths of sentences in the simplified versions of documents compared with those in the original versions, and in the number of sentences contained in those versions. Inspection of the mean values for these characteristics indicates that the simplified versions of the texts contain a higher number of sentences and that these sentences are of a shorter length (measured in words) than is the case for the original versions.

	Kincaid	ARI	CLI	Flesch	Fog	Lix	SMOG	Syll/Word
Original (μ)	11.21	13.16	12.23	58.14	14.26	48.84	11.81	1.51
Simplified (μ)	7.21	8.22	11.28	71.85	10.13	36.85	9.53	1.46
t-test	0.00	0.00	<u>0.01</u>	0.00	0.00	0.00	0.00	<u>0.02</u>

Table 5b Readability statistics 1 (EN)

Table 6b presents information derived from English documents using resources available exclusively for that language. This includes values for traditional readability metrics computed for the original forms of documents included in the benchmarking experiment together with information on their manually simplified versions. The columns present the values obtained for these documents of the Kincaid Formula (*Kincaid*), Automated Readability Index (*ARI*), Coleman-Liau Formula (*CLI*), Flesch Reading Ease Formula (*Flesch*), Fog Index (*Fog*), Lix Formula (*Lix*), and SMOG Grading (*SMOG*). In addition to this, the column *Syll/Word*

indicates the average number of syllables per word in the documents. These values were obtained using an automatic method for syllabification based on LaTeX hyphenation patterns.^{§§}

METRIC	<u>Comma index</u>		Index of words with > 3 syllables		<u>Index of words in sentences</u>		<u>Index of word diversity</u>		Pronoun index	
File	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP
Melt in the body electronics	0.18	0.00	1.35	1.06	17.00	9.40	0.31	0.27	2.94	3.19
Climate change	0.43	0.12	2.62	1.37	15.50	12.05	0.27	0.24	1.08	3.32
Forrest wedding	0.74	0.04	0.81	0.48	19.28	10.00	0.33	0.29	8.15	10.43
Leeds vs. Wigan	0.35	0.08	0.58	0.42	22.36	8.50	0.30	0.24	2.88	3.36
Syrian news	0.34	0.00	1.74	1.35	16.28	10.79	0.28	0.25	1.71	2.70
Blood booster	0.44	0.10	0.55	0.54	24.91	11.25	0.26	0.24	3.28	3.81
DNA profiling	0.63	0.21	1.54	1.35	17.81	10.67	0.30	0.28	3.86	4.17
iPhone 5	0.40	0.13	1.02	1.29	16.83	11.04	0.29	0.27	3.63	3.88
Movie paradoxes	0.42	0.04	0.72	0.92	14.57	8.70	0.31	0.31	7.84	8.81
To Kill a Mockingbird	0.50	0.16	1.35	1.44	20.07	11.30	0.30	0.28	6.76	8.52
Cannabis	0.51	0.11	1.72	0.90	18.47	14.83	0.25	0.25	1.27	0.75
Drones over Pakistan	0.87	0.13	1.60	0.65	19.13	11.30	0.30	0.19	1.39	4.04
Xray plants	0.38	0.28	1.56	1.38	22.81	18.74	0.29	0.28	0.82	1.69
Pulp Fiction	0.77	0.43	1.80	1.45	17.65	12.63	0.33	0.30	3.00	3.30
Tom Sawyer	0.59	0.09	0.39	0.25	18.06	14.04	0.28	0.25	4.56	7.12
Casual Vacancy	0.47	0.44	1.32	1.24	18.65	11.92	0.31	0.30	3.47	5.03
Fall North	0.61	0.55	0.51	0.40	12.56	13.31	0.32	0.30	9.24	8.38
Lady Gaga	0.59	0.40	1.23	1.17	25.30	13.65	0.31	0.30	5.93	7.69
PCs Features	0.35	0.30	1.09	1.42	11.18	9.91	0.26	0.25	5.43	4.90
Rock Sound Mag	0.40	0.29	0.43	0.50	21.36	10.06	0.25	0.23	10.70	11.11
Melt in the body electronics	0.23	0.03	1.33	0.93	14.61	11.86	0.28	0.22	3.04	7.83
Anna Karenina	0.61	0.09	0.68	0.37	28.09	12.96	0.27	0.25	8.41	9.88
iPhone 5	0.39	0.14	1.02	0.41	16.00	12.82	0.29	0.20	3.62	5.05
James Joyce	0.77	0.13	0.58	0.41	16.12	14.59	0.30	0.21	4.38	7.11
Rock Sound Mag	0.39	0.20	0.46	0.65	17.88	15.39	0.25	0.18	10.53	9.65
μ	0.49	0.18	1.12	0.89	18.50	12.07	0.29	0.25	4.72	5.83

Table 7a: Readability of original and manually converted forms of EN texts used in benchmarking experiments

Inspection of Tables 7a and 7b reveals that for manually converted versions of the EN texts, values of *comma index*, *index of words in sentences*, *index of word diversity*, and *metaphor index* were significantly smaller than those derived from the original versions ($\alpha < 0.01$ in each case). Values of *index of words with more than 3 syllables* and *passive verb index* were smaller for manually converted versions of the EN texts, but not significantly so.

^{§§} This method underestimates the average number of syllables per word in a text due to the hyphenation of short multi-syllable words such as *again*.

METRIC	<u>Metaphor index</u>		Passive verb index		Polysemic word index		Index of productive syntax	
	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP
File								
Melt in the body electronics	5.00	0.00	60.00	30.00	46.47	47.34	72.00	72.00
Climate change	16.67	0.00	16.67	15.00	57.35	55.19	70.00	66.00
Forrest wedding	7.14	2.17	21.43	8.70	42.22	40.00	74.00	76.00
Leeds vs. Wigan	46.43	3.57	21.43	7.14	42.81	45.80	66.00	67.00
Syrian news	33.33	4.17	16.67	12.50	54.27	53.28	77.00	79.00
Blood booster	50.00	25.00	54.55	28.57	53.28	49.52	74.00	76.00
DNA profiling	21.88	14.81	43.75	51.85	44.56	48.26	83.00	82.00
iPhone 5	41.67	21.43	27.78	25.00	49.83	50.49	80.00	79.00
Movie paradoxes	42.86	5.00	0.00	3.33	46.41	51.34	78.00	79.00
To Kill a Mockingbird	50.00	22.22	7.14	7.41	46.98	49.84	73.00	79.00
Cannabis	2.94	2.78	5.88	16.67	50.00	53.56	73.00	70.00
Drones over Pakistan	40.00	1.04	13.33	30.43	53.48	57.50	67.00	85.00
Xray plants	16.67	7.89	56.25	31.58	48.77	49.16	82.00	83.00
Pulp Fiction	41.18	33.33	35.29	29.17	47.67	49.17	71.00	67.00
Tom Sawyer	17.65	4.35	0.00	21.74	43.00	44.27	73.00	77.00
Casual Vacancy	50.00	10.00	11.76	12.00	45.43	47.32	76.00	77.00
Fall North	36.00	19.23	0.00	0.00	42.99	43.93	84.00	83.00
Lady Gaga	60.00	25.00	10.00	15.00	49.41	51.28	66.00	68.00
PCs Features	10.71	8.11	10.71	8.11	50.80	51.77	74.00	74.00
Rock Sound Mag	67.86	13.24	14.29	5.88	48.16	51.46	75.00	77.00
Melt in the body electronics	2.78	3.57	44.44	17.86	49.05	47.59	77.00	73.00
Anna Karenina	18.18	0.00	9.09	8.00	38.51	39.20	80.00	76.00
iPhone 5	39.47	19.12	26.32	11.76	50.00	48.62	78.00	83.00
James Joyce	23.53	14.81	17.65	18.52	52.19	50.25	79.00	79.00
Rock Sound Mag	55.88	12.12	11.76	0.00	49.34	48.82	74.00	80.00
μ	31.91	10.92	21.45	16.65	48.12	49.00	75.04	76.28

Table 7b: Readability of original and manually converted forms of EN texts used in benchmarking experiments

5.3.3 Spanish

	Sents	Words/Text	Char/Text	Words/Sent	Char/Word	SC Sign Usage
Original (μ)	14.00	283.44	1372.56	21.51	4.85	6.93
Simplified (μ)	17.16	280.00	1349.32	18.63	4.85	7.13
t-test	0.04	0.40	0.36	0.02	0.48	0.40

Table 8 Readability statistics (ES)

Table 8 presents statistics on the original and simplified versions of Spanish documents for criteria used to estimate their readability. The first row contains the mean values of each characteristic calculated for the original forms of the documents. The second row contains this information for the manually simplified forms of the documents. The final row displays 2-tailed statistical significance of the paired t-test (obtained via use of the SPSS package).

Tables 9a and 9b display the values of readability indices calculated for ES texts.^{***} The figures presented there are ratios, not percentages. It can be observed that values of the metaphor and words in sentences indices of simplified versions of the texts are significantly smaller than those calculated from the original versions. Values of passive verb index were also smaller, but not significantly so ($\alpha=0.28$) for manually converted texts.

^{***} In Tables 9a and 9b, the statistics are presented as decimals rather than percentages.

METRIC	<u>Comma index</u>		Index of words with > 3 syllables		<u>Index of words in sentences</u>		Index of word diversity		Pronoun index	
File	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP
Adolescentes	0.74	0.43	2.83	2.95	20.21	17.94	0.53	0.50	6.71	7.21
Ahorrar energia	0.39	0.14	4.08	4.01	28.40	17.63	0.48	0.48	5.28	5.32
Artes escénicas	0.73	0.72	3.27	3.07	35.43	30.60	0.52	0.49	2.42	2.61
Aventu leer	0.74	0.36	2.53	2.47	14.73	18.05	0.50	0.47	10.49	7.48
BLANCANIEVES	0.88	0.75	2.60	3.01	25.00	18.83	0.53	0.50	3.60	3.98
COLAPSO EN LAS CIUDADES	0.52	0.19	2.54	2.69	30.70	31.25	0.42	0.45	2.93	3.73
CONVOCATORIA DE EXAMENES	0.39	0.30	3.57	3.43	25.50	18.44	0.53	0.46	7.06	7.23
Concierto de Marc Ribot	1.01	0.40	2.60	2.01	30.80	18.63	0.55	0.45	3.57	3.69
Cuento corto.literatura	0.54	0.21	2.17	1.69	25.09	15.76	0.45	0.46	3.62	5.44
Emergencias Sanitarias	0.59	0.41	3.19	3.30	30.40	27.63	0.39	0.40	2.96	4.98
InstruccionesLavadoEnSeco	0.45	0.37	2.68	3.04	22.08	19.29	0.56	0.55	6.97	5.19
JUSTIN BIEBER	0.40	0.41	2.85	2.82	14.06	19.00	0.51	0.42	7.51	9.77
LA MITOLOGIA GRIEGAlite	1.20	0.27	3.14	2.92	23.58	13.07	0.49	0.48	3.18	4.37
Lluvias torrenciales	0.50	0.63	3.09	3.33	25.90	25.50	0.56	0.53	4.63	4.31
LorcaLIT	0.66	0.68	3.09	2.87	24.73	25.10	0.48	0.41	1.84	6.37
CONTRATO DE ARRENDAMIENTO	0.37	0.30	3.30	3.35	33.38	24.45	0.54	0.41	6.74	8.18
Moby Dick	0.58	0.35	2.43	2.45	18.69	16.33	0.52	0.44	10.29	6.71
Oro negro	1.00	0.70	3.55	3.75	23.25	13.59	0.42	0.43	5.38	2.01
SANCIONADMINISTRATIVA	0.40	0.04	3.80	3.43	39.14	17.69	0.60	0.58	5.11	3.89
Tim Burton	0.53	0.48	3.06	2.90	29.43	26.57	0.53	0.52	2.43	0.54
VUELVE TOLKIEN	0.89	0.90	1.98	2.19	31.00	33.29	0.55	0.43	5.65	5.58
children health	0.45	0.57	3.66	3.56	24.09	13.86	0.47	0.38	5.66	3.35
children literature	0.52	0.18	2.49	2.62	23.21	15.12	0.50	0.43	12.62	8.91
física cuantica	1.10	0.49	3.65	3.17	31.33	30.90	0.30	0.21	6.03	5.18
instrucciones juego	0.69	0.38	3.01	3.01	22.23	15.73	0.49	0.43	4.50	6.57
μ	0.65	0.42	3.00	2.96	26.09	20.97	0.49	0.45	5.49	5.30

Table 9a: Readability of original and manually converted forms of ES texts used in benchmarking experiments

METRIC	<u>Metaphor Index</u>		Passive verb index		Polysemic word index	
	ORIG	SIMP	ORIG	SIMP	ORIG	SIMP
File						
Adolescentes	0.14	0.00	0.36	0.06	0.42	0.46
Ahorrar energia	0.00	0.00	0.70	0.50	0.37	0.40
Artes escénicas	0.00	0.00	0.00	0.20	0.34	0.35
Aventu leer	0.18	0.05	0.18	0.10	0.40	0.43
BLANCANIEVES	0.00	0.00	0.20	0.08	0.32	0.34
COLAPSO EN LAS CIUDADES	0.10	0.00	0.10	0.33	0.31	0.35
CONVOCATORIA DE EXAMENES	0.00	0.00	0.80	0.67	0.33	0.33
Concierto de Marc Ribot	0.40	0.00	0.10	0.06	0.33	0.37
Cuento corto.literatura	0.09	0.00	0.27	0.14	0.42	0.43
Emergencias Sanitarias	0.00	0.00	0.30	0.63	0.24	0.22
InstruccionesLavadoEnSeco	0.00	0.00	0.69	0.57	0.38	0.36
JUSTIN BIEBER	0.00	0.00	0.17	0.29	0.36	0.36
LA MITOLOGIA GRIEGAlite	0.00	0.00	0.17	0.25	0.37	0.42
Lluvias torrenciales	0.00	0.00	0.60	0.50	0.29	0.27
LorcaLIT	0.36	0.00	0.18	0.40	0.26	0.29
CONTRATO DE ARRENDAMIENTO	0.00	0.00	0.75	0.64	0.28	0.30
Moby Dick	0.31	0.00	0.46	0.10	0.40	0.45
Oro negro	0.25	0.09	0.50	0.18	0.42	0.48
SANCIONADMINISTRATIVA	0.00	0.00	0.71	0.38	0.38	0.43
Tim Burton	0.14	0.00	0.14	0.00	0.33	0.35
VUELVE TOLKIEN	0.00	0.00	0.50	0.86	0.31	0.29
children health	0.18	0.00	0.55	0.25	0.36	0.40
children literature	0.07	0.04	0.36	0.27	0.39	0.44
física cuantica	0.00	0.00	0.44	0.60	0.37	0.37
instrucciones juego	0.23	0.10	0.08	0.10	0.41	0.40
μ	0.10	0.01	0.37	0.33	0.35	0.37

Table 9b: Readability of original and manually converted forms of ES texts used in benchmarking experiments

5.4 Unexpected results

When considering unexpected results, it is important to recall that the reading comprehension profiles of autistic readers are heterogeneous. The text conversion process undertaken by carers is likely to be similarly heterogeneous, regardless of their access to the writing guidelines presented in Table 2 (Section 4.4).

Although they are not statistically significant, several unexpected results can be observed in the statistics presented in Tables 5a-b, 7a-b and 9a-b:

- In 48% of BG, 84% of EN, and 56% of ES text pairs, the *pronoun index* reveals that manually converted versions of the texts contain more pronominal anaphoric references than the original versions, indicating that they are less accessible for end users. A comparison of corresponding text pairs revealed that this phenomenon is caused by the manual rewriting of long sentences as sequences of shorter sentences. Carers tended to pronominalise references to avoid explicit repetition of

information in different sentences. Given that such new pronouns refer to recently mentioned and salient concepts, it is expected that these pronominal references do not pose obstacles to reading comprehension.

- In 52% of the BG, 72% of the EN, and 68% of ES text pairs, the *polysemic word index* reveals that manually converted versions of the texts contain a greater proportion of semantically ambiguous words than the original versions do. This observation is consistent with the findings of previous related work (Stajner et al., 2014). It is argued that the identification of ambiguous words alone contributes little to the assessment of text accessibility. Instead, it is necessary to consider the specific sense of each word in a text and to consult a sense frequency distribution derived from a representative corpus of texts of the same category. When an ambiguous word is used with a sense that matches the most frequently occurring sense of that word derived from the corpus, then it poses no obstacle to reading comprehension. When the ambiguous word is used with a sense that matches less frequent senses of that word derived from the representative corpus, then it has a negative impact on accessibility. Unfortunately, resources enabling an assessment of the prevalence of word senses in natural language and enabling word sense disambiguation are currently insufficient.
- Values computed for *verb temperature* (BG) indicate that the manually converted versions of texts introduce more clauses and facts than the corresponding original versions, implying that they are less accessible. However, in many cases this additional information is of an explanatory nature, which may contribute to an overall improvement in the accessibility of those versions.
- In EN texts, values of the *index of productive syntax* imply that the manually converted versions of the document are less accessible, containing a greater variety of syntactic constructions than the original versions of the texts. Inspection of the text *Drones over Pakistan*, for which manual conversion led to the most substantial increase in the *index of productive syntax*, revealed that the main cause of the change was the introduction of new explanatory text providing more detail about the use and characteristics of entities mentioned in the text.

Five of the texts used in benchmarking for English were also assessed in terms of user requirement UR312 (improper grammar). This examination revealed that the process of manually converting the five texts into a more accessible form led to the introduction of 13 typos and errors. It is suggested that the methodology for manual conversion supported by the Open Book interface include a revision step to promote correction of manually introduced errors.

5.5 Analysis of times taken to simplify texts

The professionals took just over 55 minutes on average to simplify a text between 250-350 words long (Table 10). They worked at the average speed of 8 words per minute. The speed of simplification was calculated by dividing the number of words on each original text with the time taken to simplify it, as recorded by each professional. Means were initially calculated for each professional. Means for each centre were calculated from the average means for each professional in corresponding centres.

Professionals' speed ranged from 1 word per minute (Prof 13 in Bulgaria, SameSim Triangles) to 26 words per minute (Prof 7 in Spain, text title 'Lluvias torrenciales'). The overall time taken to simplify one single text ranged from 15 minutes (Prof 3 in the UK, 'Rock Sound Mag' text) to 180 minutes (Prof 13 in Bulgaria, text title 'SameSim Triangles'). Please see Appendix 1 for a breakdown of all outcomes for each professional in each participating centre.

Participating Centres	Mean times taken to simplify a text of 250-350 words (in minutes)	Words produced per minute
UK (adults)	27	11
Spain (adults & children)	43	9
Bulgaria (children)	97	4
Overall	55.6	8

Table 10 Outcomes from text-simplification process

6. Discussion

6.1 Readability

Pairwise comparison of the original and simplified versions of the Bulgarian texts shows that simplified versions contain a significantly ($\alpha=0.001$) greater number of sentences, fewer words per sentence, and a greater number of signs of syntactic complexity per word than the original versions.

For English, there was no significant difference in the average number of signs of syntactic complexity per word in the original and simplified versions of the texts. The average length of the original and simplified versions (measured both in words and in characters) also did not differ significantly. The latter finding was unexpected as previous research has shown that automatic syntactic simplification leads to the generation of simplified documents with greater word lengths than the originals (Siddharthan, 2004).^{†††} Pairwise comparison of the original and simplified EN texts using all of the traditional readability metrics (Kincaid, ARI, CLI, Flesch, Fog, Lix, SMOG) indicates that there are statistically significant ($\alpha=0.001$) differences in the readability of the simplified versions and the original versions of the texts. Inspection of the mean values of these metrics implies that the manually simplified versions are easier to read than the corresponding

^{†††} The simplified documents are characterised by a larger number of shorter sentences than the originals. In Siddharthan's (2004) study, this characteristic led to an overall increase in the total word length of simplified documents.

originals. To illustrate, the mean Flesch Reading Ease Formula for the original texts indicates suitability for readers at a minimum age of 15-16. By contrast, the simplified texts are suitable for readers at a minimum age of 12-13. It can be noted that there is a statistically significant difference in word length (measured in syllables) between the original and simplified versions of the texts. Inspection of the mean values of *Syll/Word* in Table 6 indicates that the simplified versions of the texts contain words that are shorter than those in the original versions.

Statistically significant differences between Spanish documents in their original and manually simplified forms were only observed in criteria related to document length (measured in sentences) and sentence length (*Words/Sent*). There was no significant difference between documents in their original and manually simplified forms with regard to document length (measured in words or characters) or average word length. As noted for English documents, the two different versions of Spanish documents do not differ significantly with regard to the percentage of vocabulary comprising signs of syntactic complexity.

Direct observation of the simplified versions of the Spanish documents indicates that the manual simplification process applied to the Spanish texts was qualitatively different from that applied to the Bulgarian and English ones. Intra-sentential operations such as word replacements, the provision of definitions for metaphors and technical words, and non-linguistic changes such as the insertion of images were more evident in the simplified documents prepared for Spanish. By contrast, there was less evidence of “free” paraphrasing of sentences and re-ordering of sentences. The Spanish simplification operations were noted to include sentence deletion, splitting of complex sentences into sequences of simple sentences. This is reflected in the differences between values of readability criteria calculated.

6.2 Text simplification times

Professionals in the UK worked almost three times faster than their counterparts in Bulgaria, but comparable to the professionals in Spain (Figure 3). Apart from obvious language differences, these disparities may also be due to different simplification techniques used for children and adults, especially around substituting complex text with pictures instead of other words. Thus, professionals in Bulgaria most often use bullet points and try to present the information in clear, even schematic (Text No.24) way. They also removed sentences which they reported as not essential for understanding the meaning of the text. Pictures were widely used in the Spanish sample but not in the UK and Bulgaria.

UK professionals, on the other hand, worked with an adult, high-functioning autism audience in mind. The work was guided by comments that survey cohort participants made such as: *‘I do come across versions of ‘Easy Read’ documents which are just far too basic and take important information out altogether. I think many higher-functioning people on the spectrum find them quite insulting to our intelligence!’*

Avoiding the use of different words to refer to the same subject was thought to be important in the UK. This would avoid confusion: people with autism prefer to have an anchor word to focus upon so it might have to be repeated in later paragraphs to provide the required continuity of comprehension for the individual reading the text. Sometimes it was felt necessary to go as far as to change around the sentence order, rather than just alter a few words, so as to make better sense for someone with ASD.

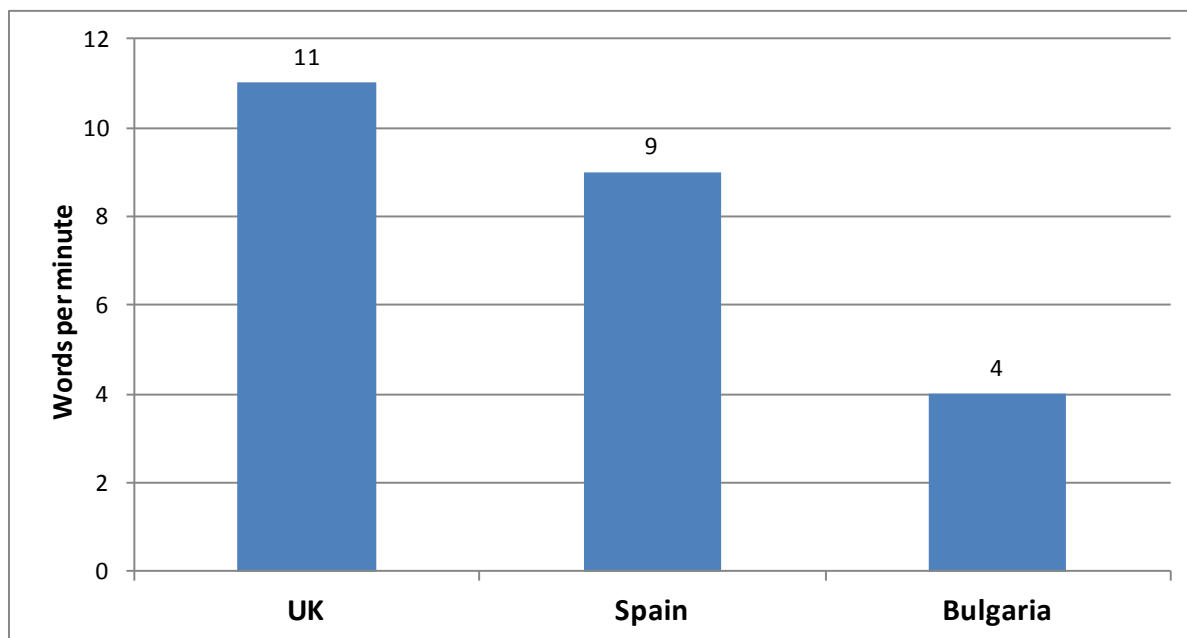


Figure 3 Spread of manual simplification for each site

Longest simplification times were recorded in **Bulgarian**. For example, a Maths lesson for ‘Same and Similar Triangles’ (text 14, Bulgarian) took 180 minutes (3 hours) to be simplified. In this case the special education teacher who carried out the simplification carried out extra work otherwise not necessary for adult texts. Some of the main simplification steps for text 14, Bulgarian are presented below:

- added definitions that are not presented in the original text
- break up long sentences into shorter ones
- used bold fonts to underline the most important ideas
- **Made up graphs** to visualize each new statement in the lesson

- Added those graphs to the simplified version

Obviously, the last two steps took a considerable period of time to be completed.

Other texts that took long time to adapt in Bulgarian were texts from the news domain, which children with ASD find difficult to understand due to the journalistic writing style. The most time spent, however, is where additional explanations were decided as necessary. For example in text 6 from News domain, the teacher explains who the main characters of the specific article are and provides some historical information for them. The teacher had to:

- **research the characters**
- **collect information from different sources including online**
- Collate the information into useful chunks
- Include this information in the simplified document

In view of the fact that the professionals in Spain and Bulgaria simplified texts for children, their timings tend to be generally longer than those recorded by the UK team, who worked with texts suitable for adults. The professionals used online dictionaries and thesaurus, which reflects their everyday practice and approach to manual text simplification for people with ASD. Professionals working with autistic children use visual aid more often than those who work with adults.

Findings from the benchmarking trials will be useful in appropriate customisation of the prototypes developed in the project.

6.3 Limitations of the research method

The size of the cohort of participants used for areas of web usage was limited to just 40 in each centre and the size of the cohort of teachers and care-workers used for benchmarking was limited to just five. We were relying upon the benchmarking professionals to time themselves, while working on their own at home or their places of work. These were therefore not strictly laboratory-controlled conditions.

By the very nature of syntax and meaningful context we were understandably not able to find 25 completely different pieces of text of exactly the same length for assessment of before-and-after-simplification comparison and timings.

Inter-rater reliability demonstrates consistency or agreement among different professionals carrying out a specific task. In our case this would mean that simplification operations and times would be similar for all professionals through all texts and languages. However, in reality inter-rater reliability is expected to be low in a linguistic based task between professionals from different countries. Essential linguistic characteristics in grammar, morphology and syntax make direct comparisons between languages invalid. Inter-rater reliability is not very high even within same language group professionals, when an unstructured test is used. Thus, one person's interpretation of a simplified text will always be different from another's so there may be some inconsistency in the results. This may also be affected by the amount of experience and training each one has had in dealing with people with ASD, length of involvement, personal differences, perhaps contributing to any possible discrepancies in interpretation. For example, text 13, in English (I-phone) was simplified by professional 2 and professional 5. The outcomes were different and professional 5 took almost half the time (29 minutes) professional 2 did (41 minutes).

The use of established readability metrics to assess the readability of the texts can be criticised as these metrics were not designed to assess the comprehensibility of documents from the perspective of readers with ASD. Instead, they provide values that can be mapped to US reading grade levels in general populations. We make the implicit assumption that there is a correlation between the reading grade level of a text and the difficulty that it poses to comprehension by end users in the project. In the research described in this report, readability assessments derived from established metrics were supplemented with information from indices selected on the basis of their relevance to specific reading comprehension difficulties of people on the autistic spectrum. It may be possible, on the basis of psycholinguistic studies, to combine these indices as terms in readability formulae relevant to the needs of people with ASD.

6.4 Strengths of the research method

The advantage of using the data gathered from the participants of the on-line questionnaire (WP2) for the benchmarking exercise meant that we had valuable data as guidance of favoured web usage and preferred websites used by adults/children with ASD.

Having 25 different texts interpreted by 5 individuals in each centre, ensured that we were making as wide a selection as possible for the small number of professionals being asked to volunteer their services.

Spreading the benchmarking exercise across three different countries with distinctive linguistic structures should give a wider, more appropriate outcome for the software-writers who themselves are building a cross-lingual translation programme for the FIRST project.

The assessment of readability of the documents involved in the benchmarking trials provides a linguistically oriented comparison between documents in their original and simplified forms. This comparison will partially inform design of the functionality of the prototypes developed during the course of the project.

7. References

1. <http://www.accessibleinfo.co.uk/information/index.php>
2. Office of Public Sector Information. London. (2005). Disability Discrimination Act
3. Office of Public Sector Information (1998). London: Human Rights Act
4. American Speech and Hearing Association (1992) Communication Bill of Rights
5. NAS: London
6. Beversdorf, D.Q., Anderson J.M., Manning S.E., Anderson S.L., et al. (1998). The effect of semantic and emotional context on written recall for verbal language in high functioning adults with autism spectrum disorder. *Journal of Neurology, Neurosurgery and Psychiatry*, 65. (5) 685-692.
7. Caylor, J. S., T. G. Sticht, L. C. Fox, and J. P. Ford. (1973) Methodologies for determining reading requirements of military occupational specialties: Technical report No. 73-5. Alexandria, VA: Human Resources Research Organization.
8. Coleman, E. B. (1965) "On understanding prose: some determiners of its complexity." NSF Final Report GB-2604. Washington, D.C.: National Science Foundation.
9. Coleman, E. B. (1971) "Developing a technology of written instruction: some determiners of the complexity of prose." In *Verbal learning research and the technology of written instruction*, eds. E. Z. Rothkopf and P. E. Johnson. New York: Teachers College Press, Columbia University, pp. 155-204.
10. Coleman, M. and Liao, T. L. (1975) A Computer readability formula designed for machine scoring, *Journal of Applied Psychology*, Vol. 60: 283-284
11. Dennis, M., Lazenby, A.L., Lockyer, L. (2001). Inferential Language in High-Function Children with Autism. *Journal of autism and developmental disorders*, 31(1): 47-54.
12. Department of Health. London. (2008). The Mental Capacity Act Deprivation of Liberty safeguards
13. Department of Health. London. (2001). Valuing People
14. Department of Health. London. (2009). Valuing People Now: A New Three-Year Strategy for Learning Disabilities

15. Benchmarking for Competitive Advantage (1994). Robert J Boxwell Jr, New York: McGraw- Hill. pp. 225
16. Camp, R. (1989). The search for industry best practices that lead to superior performance
17. Farr, J. N., J. J. Jenkins, and D. G. Paterson. (1951) "Simplification of the Flesch Reading Ease Formula." Journal of applied psychology 35, no. 5:333-357.
18. Flesch, R. (1948) "A new readability yardstick." Journal of Applied Psychology 32:221-233.
19. Fry, E. B. (1977) "Fry's readability graph: Clarifications, validity, and extension to level 17." Journal of Reading 21, no. 3:242-252.
20. Gerganov E., Mateeva A. and Kristev B. (1976). Measurement of readability of written Bulgarian. In: "Использование математических моделей и электронных вычислительных машин в лингвистике". 1976. София, Издательство Болгарской академии наук, p. 331-334.
21. Gold, R.; Faust, M. and Goldstein, A. (2010). Semantic integration during metaphor comprehension in Asperger syndrome. Brain & Language 113, 124–134.
22. Gerganov, E., Mateeva, A. (1978) Experimental Research of the Frequency of the Bulgarian Language, In the proceedings *Contemporary problems of the native language education*, Sofia, 1978
23. Gray, W. S. and B. Leary. (1935) What makes a book readable. Chicago: Chicago University Press.
24. Gunning, R. (1952) The technique of clear writing, New York: McGraw-Hill.
25. Heimann, M., Nelson, K.E., Tjus, T., and Gillberg, C. (1995). Increasing reading and communication skills in children with autism through an interactive multimedia computer program. Journal of Autism and Developmental Disorders, 25, 459-480.
26. Howlin, P., Mawhood, L. and Rutter, M. (2000), Autism and Developmental Receptive Language Disorder—a Follow-up Comparison in Early Adult Life. II: Social, Behavioural, and Psychiatric Outcomes. Journal of Child Psychology and Psychiatry, 41: 561–578.
27. <http://www.nhsbenchmarking.nhs.uk>
28. Huemer, S.V., Mann, V. (2010). A Comprehensive Profile of Decoding and Comprehension in Autism Spectrum Disorders. Journal of Autism and Developmental Disorders, 40:485–493.
29. Jolliffe, T. and Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger syndrome. Can global coherence be achieved? A further test of central coherence theory. Psychological Medicine, 30, 1169-1187.
30. Kincaid, J. P., R. P. Fishburne, R. L. Rogers, and B. S. Chissom. (1975) Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. CNTECHTRA Research Branch Report 8-75.
31. Kitson, H. D. (1921) The mind of the buyer. New York: Macmillan.
32. McLaughlin, G. H. (1969) "SMOG grading - a new readability formula," Journal of reading 22:639-646.

33. Miniscalco, C., Dahlgren, A. (2010). Basic reading skills in Swedish children with late developing language and with or without autism spectrum disorder or ADHD. *Research in Developmental Disabilities*, 31 (5), 1054-1061.
34. Moore M. Calvert S. (2000). Brief report: Vocabulary acquisition for children with autism: Teacher or computer instruction. *Journal of Autism and Developmental Disorders*. ;36,304-311
35. Myles et al. (2002). Analysis of reading skills in individuals with Asperger syndrome. *Focus on autism and other developmental disabilities*, 17, 44-47.
36. Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology*, 90, 142–171.
37. O'Connor, I.M. and Klein, P.D. (2004). Exploration of Strategies for Facilitating the Reading Comprehension of High-Functioning Students with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 34, (2), 115-127.
38. Sesma, H.W.; Mahone, E.M.; Levine, T.; Eason, S.H.; Cutting, L.E. (2009). The contribution of executive skills to reading comprehension. *Sesma HW - Child Neuropsychology* - 01-MAY-2009; 15(3): 232-46 MEDLINE®.
39. Siddharthan, A. (2004) *Syntactic simplification and text cohesion*. Technical Report UCAM-CL-TR-597, University of Cambridge.
40. Smith, G. C. (2011). Variability in language and Reading in High-Functioning Autism. En Mohammad-Reza Mohammadi, *A comprehensive book on Autism Spectrum Disorder*, 4; (pp.63-84).
41. Štajner, S., Mitkov, R. and Corpas Pastor, G. (2014) 'Simple or not simple? A readability question'. To appear in N. Gala, R. Rapp, and G. Bel-Enguix (eds), *Recent Advances in Language Production, Cognition and the Lexicon*, Springer.
42. Swanson, C. E. (1948) "Readability and readership: A controlled experiment." *Journalism quarterly* 25:339-343.
43. Tjus, T., Heimann, M., and Nelson, K. E. (1998) Gains in Literacy through the Use of a Specially Developed Multimedia Computer Strategy: Positive Findings from 13 Children with Autism. *Autism*, June 1998; vol. 2, 2: pp. 139-156.
44. Walters, K.A. (2011). Improving reading comprehension and social skills with high functioning autism: an elementary school intervention. Project (M.A., Child Development (Applied Settings) California State University, Sacramento.

Appendix 1

FIRST	United Kingdom				
Text No.	Name	Original ^{***} word count	Simplified word count	Time taken in minutes	Words produced per min
1	'Melt in the body' electronics	164	180	30	5
2	Climate change	272	235	50	5
3	Forrest wedding	259	221	28	9
4	Leeds v Wigan	298	227	32	9
5	Syrian news	286	251	34	8
Prof 1	Means	255.8	223	35	7
11	Blood booster	258	308	16	16
12	DNA Profiling	286	279	25	11
13	iPhone 5	298	290	41	7
14	Movie Paradoxes	296	254	18	16
15	To Kill a Mockingbird	270	297	25	11
Prof 2	Means	281.6	285.6	25	12
16	Cannabis	306	262	26	12
17	Drones over Pakistan	277	500	45	6
18	Fossilized plants	353	341	28	13
19	Pulp Fiction	295	296	21	14
20	Tom Sawyer	303	317	19	16
Prof 3	Means	306.8	343.2	27.8	11
21	Casual Vacancy	302	265	22	14
22	Fall North	306	326	18	17
23	Lady Gaga	252	267	12	21
24	PCs Features	299	347	20	15
25	Rock Sound Mag	291	319	15	19

^{***} Word counts on these tables were calculated through Microsoft Word by the CNWL researcher. The team in Wolverhampton, however, used a different method, which requires that the documents be converted into plain text initially, and then the word count is carried out.

Prof 4	Means	290	304.8	18	17
1	'Melt in the body' electronics	253	324	26	10
6	Anna Karenina	301	321	29	10
13	IPhone 5	299	422	29	10
18	James Joyce	274	392	33	8
25	Rock Sound Mag	292	494	29	10
Prof 5	Means	283.8	390.6	29	10
UK	Overall means	283.6	309.4	27	11

FIRST	Spain				
Professionals	Text Title	Original word count	Simplified word count	Time taken in minutes	Words produced per min
	Ahorrar energía	289	264	23	13
Prof 6	Artes escénicas	251	129	28	9
	Instruc Lavado	285	205	21	13
	Lorca	267	253	39	7
	Tim Burton	208	189	26	8
	Means	260	208	27	9
	Blancanieves	252	233	21	12
Prof 7	Convocatoria de EXÁMENES	257	169	17	15
	Justin Bieber	252	224	30	8
	Lluvias torrenciales	259	127	10	26
	Vuelve Tolkien	251	235	16	16
	Means	254	198	19	14
	Adolescentes	283	307	27	10
Prof 8	Aventura de leer	324	361	43	8
	Colapso en las ciudades	311	384	50	6
	Cuento corto	276	331	42	7
	SANCIÓN Administrativa	280	289	48	6
	Means	295	334	42	7

	Concierto Marc Ribot	310	280	62	5
Prof 9	Emergencias Sanitarias	309	228	58	5
	Física cuántica	317	319	75	4
	Modelo de Contrato	269	267	99	3
	Moby Dick	276	316	86	3
	Means	296	282	76	4
	Instrucciones juego	289	507	55	5
Prof 10	LA MITOLOGÍA GRIEGA	284	406	40	7
	Literatura infantiles	322	417	46	7
	Oro negro	279	347	73	4
	Salud Niños	266	405	39	7
	Means	288	416	51	6
Spain	Overall means	279	288	43	9

FIRST	Bulgaria				
Professionals	Text Title	Original word count	Simplified word count	Time taken in minutes	Words produced per min
	ArgueTeacher	291	258	40	7
Prof 11	Fatigue	288	277	100	3
	Love	287	268	110	3
	Offence	300	298	45	7
	Stress	263	280	120	2
	Means	286	276	83	3
	Anan_Freeman	286	296	180	2
Prof 12	Bee Brain	286	244	180	2
	BgWriter BBCaward	291	292	40	7
	Ikea	273	221	120	2
	USA world news	294	315	45	7
	Means	286	274	113	3
	CommunSkills	275	236	120	2
Prof 13	Habitat	248	271	120	2

	SelfImage	298	259	60	5
	SameSimTriangles	239	327	180	1
	TextPublicCommun	323	264	150	2
	Means	277	271	126	2
	CherryRedHearts	336	369	100	3
Prof 14	DonkeyAdvice	284	281	100	3
	HappyPrince	340	348	100	3
	LionMouse	250	236	40	6
	Yotso	308	302	45	7
	Means	304	307	77	4
	ComputerGames	252	215	40	6
Prof 15	FiveAdvicesBicycle	262	240	120	2
	Probiotics	259	266	40	6
	SexOrientation	294	282	120	2
	Water	296	273	120	2
	Means	273	255	88	3
Bulgaria	Overall means	284.92	277	97	4