

5th International Conference on Software Development and Technologies for Enhancing  
Accessibility and Fighting Info-exclusion, DSAI 2013

## Online test tool to determine the CEFR reading comprehension level of text

Eric Velleman<sup>a\*</sup>, Thea van der Geest<sup>b</sup>

<sup>a</sup>Accessibility Foundation, Christiaan Krammlaan 2, 3571AX Utrecht, Netherlands

<sup>b</sup>University of Twente, P.O.Box 217, 7500AE Enschede, Netherlands

---

### Abstract

On the Common European Framework of Reference for Languages (CEFR) scale, the average reading comprehension level of the Dutch population is B1 and the average level of text provided by Dutch government organisations requires a considerably higher reading skills level (C1). This means that part of the population may have difficulty reading texts delivered to them by their own government. We built a simple and freely available online tool to give content editors an indication of the CEFR reading comprehension level of their texts. This paper describes the parameters of the tool and proposes a list of possible extensions to improve the quality and usability of the output.

© 2013 The Authors. Published by Elsevier B.V. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Selection and peer-review under responsibility of the Scientific Programme Committee of the 5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2013).

*Keywords:* Easy to Read; CEFR; E2R; Readability Tool; Text Comprehension; Online Reading Level Test Tool; RLTT

---

### 1. Introduction

On the Common European Framework of Reference for Languages (CEFR) scale<sup>1</sup>, the average reading comprehension level of the Dutch population is B1 while the average level of text provided by governments

---

\* E.Velleman. Tel.: +31 30 239 8270.

E-mail address: [evelleman@bartimeus.nl](mailto:evelleman@bartimeus.nl)

requires the much higher reading ability level of C1. This means that a large part of the Dutch population runs the risk of having difficulties with reading, understanding and using (online) texts delivered to them by their own government<sup>2</sup>. With the widespread use of the Web for daily needs, education and other public services, alternative multi-channel options are also slowly disappearing. This means that reading is becoming more and more important and there are fewer possibilities to work around it for people with low levels of reading skills.

Like in many countries, the Dutch government bodies and agencies, national, regional and local, are supposed to ensure that their websites meet the Web Content Accessibility Guidelines standard. This standard requires content to be “readable and understandable”<sup>3</sup>. For many authors of web content this standard is not obvious and not easy to comply with. That is why the Accessibility Foundation in the Netherlands has developed a simple and freely available online Reading Level Test Tool (RLTT)<sup>4</sup>. This tool provides text writers and content editors an indication of the CEFR reading comprehension level of their text.

The RLTT gives the author just a global indication of the reading comprehension level of a text, a limited output for revising and rewriting a text that does not have the desired reading comprehension level. However, RLTT is a free tool and comparative research in the Netherlands showed that RLTT effectiveness scores are equal or better than the scores of commercial tools in the market<sup>5,6</sup>. RLTT works on the basis of a corpus of word frequency lists that are directly related to the CEFR levels. Further, it uses the number of words per sentence, the number of words that are not in the frequency lists, the number of syllables per word, the number of personal and demonstrative pronouns and the number of prepositions per sentence. The different measurement elements of the tool have been calibrated using texts of which the reading comprehension level was established beforehand by specialists of a Dutch organization called Eenvoudig Communiceren, specialized in re-phrasing text to a lower reading comprehension level.

Since its launch late 2009, the tool has been used many times and has received positive comments from users and researchers. In this paper we give a more detailed description of the building blocks of the tool and their calibration and we address the possibilities proposed by users and researchers to extend the effectiveness and validity of the RLTT. We aim at developing the tool in such ways that it can be included in Content Management Systems used by governments in the Netherlands, and thus offer better support for the implementation of the Web Content Accessibility Guidelines in the Netherlands when it comes to easy to read content on the web.

### *1.1. WCAG 2.0*

The Web Content Accessibility guidelines (WCAG2.0) require that text content should be “readable and understandable”. There is a specific Success Criterion about reading level that continues: “When text requires reading ability more advanced than the lower secondary education level after removal of proper names and titles, supplemental content, a version that does not require reading ability more advanced than the lower secondary education level is available. (Level AAA)”<sup>3</sup>. The explanation states that content should be written as clearly and simply as possible. It provides a link to the International Standard Classification of Education<sup>7</sup> for an explanation of reading levels and a number of sufficient and advisory techniques to check for conformance.

The level of this specific Success Criterion is AAA. Yet, difficult and incomprehensible text will exclude large parts of the population from access to public information and services. Governments and many other parties that offer information and services for the public at large limit themselves increasingly to offering their content and services through web pages and web applications. Other channels, like paper forms, written documentation or face-to-face contact, are rapidly becoming unavailable. The shift to online self-service information systems does not only affect people with low levels of reading literacy, but might also lead to info-exclusion of people with cognitive disabilities that influence reading skills, people with hearing disabilities, people who are blind and use a screenreader to listen to text, and many others.

The Success Criterion includes a number of informative (sufficient) techniques that can be used to check whether a piece of text is readable and understandable. The most relevant one is General Technique 153 (G153): “Make the text easier to read”. Other techniques include providing a text summary, visual illustrations, or a spoken version. Technique G153 first describes a large number of ways to reduce the complexity of text. It then proposes the following test procedure to measure the readability of the text: “(1) Measure the readability of the text and then (2) check whether the text requires reading ability less advanced than the lower secondary education level. Then check if #2 is true”.

This approach to measuring the guideline and the underlying Success Criterion may not be self-explanatory for many content providers and editors of government websites. The guidelines do not specify how to measure the reading comprehension level of a text. This is where RLTT offers a preliminary solution to help gain a quick and very general indication of the reading comprehension level of a text.

The Web Content Accessibility Guidelines are referenced directly by the EU Directive on the accessibility of public sector bodies’ websites<sup>8</sup> and the UN Convention on the rights of persons with disabilities<sup>9</sup>. Many countries in the world use them as the *de facto* standard, for example in the form of section 508 in the US<sup>10</sup>, the Referentiel in France<sup>11</sup> and the Web Guidelines in the Netherlands<sup>12</sup>.

## 1.2. CEFR

The Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR) is the result of twenty years of work by the Council of Europe and “provides a transparent, coherent and comprehensive basis for the elaboration of language syllabuses and curriculum guidelines, the design of teaching and learning materials”, and how to evaluate them. “It is used in Europe but also in other continents and is now available in 39 languages”<sup>13,14</sup>.

CEFR basically describes six levels of language skills: A1, A2, B1, B2, C1 and C2. It encompasses a number of separate scales, including one for reading comprehension levels. The CEFR views users and learners of a language as members of a society who may wish to accomplish tasks in a given set of circumstances, in a specific environment and within a particular field of action<sup>1</sup>. The RLTT uses CEFR to indicate the reading comprehension level from technically measurable text elements, thus taking a somewhat limited approach that is applied by other text tools in the Netherlands as well<sup>6</sup>.

The CEFR overall reading comprehension level scale for written text (reception written)<sup>1</sup> distinguishes six skill levels, ranging from very basic reading skills (A1) to near-native written language fluency (C2).

A1: Can understand very short, simple texts a single phrase at a time, picking up familiar names, words and basic phrases and rereading as required.

A2: Can understand short, simple texts on familiar matters of a concrete type which consist of high frequency everyday or job-related language. Can understand short, simple texts containing the highest frequency vocabulary, including a proportion of shared international vocabulary items.

B1: Can read straightforward factual texts on subjects related to his/her field and interest with a satisfactory level of comprehension. About 95% of the Dutch population can understand texts of this level<sup>15</sup>.

B2: Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low-frequency idioms.

C1: Can understand in detail lengthy, complex texts, whether or not they relate to his/her own area of speciality, provided he/she can reread difficult sections. About 60% of the Dutch population has difficulty understanding a text of this level<sup>15</sup>.

C2: Can understand and interpret critically virtually all forms of the written language including abstract, structurally complex, or highly colloquial literary and non-literary writings. Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.

There is a large number of different scales for different situations like for public announcements, for reading correspondence etc. Each separate scale refers to particular aspects, elements, contexts, processes, etc. distinguished within the model<sup>1</sup>.

### *1.3. Situation in the Netherlands*

In the Netherlands the group of people with a low level of reading literacy is estimated to be around 10 percent of the population. That means that even in a country with an excellent educational system like The Netherlands, an estimated 1,5 million people<sup>16</sup> is considered “functionally low-literate”. For this group, their literacy level is insufficient to use written information that is of interest to them, to fully participate in society, to realize their own goals and to develop their knowledge and competences<sup>2</sup>.

Furthermore, research by Bohnenn shows that 27,1 percent of the population cannot read above IALS level 2, which is more or less equivalent to CEFR level B1<sup>17</sup>. This group can be classified as the “less advanced than the lower secondary education level” used in WCAG. Making web content easy to read and understand is therefore a crucial and critical condition for effective communication between government and citizens<sup>2</sup>.

The government cannot exclude some groups of citizens from its information and services. Yet, there is a mismatch between the reading skills levels of the average citizen and the texts the government send out to the average citizen. Content editors wonder whether they are writing at the right level when publishing text online or in print, but they lack good and easy to use instruments that can help them to assess the reading comprehension level of the texts they are producing. In addition, many producers of content are not experienced and trained text writers, but contribute content for the website because they are the specialist for a particular topic. RLTT offers all content providers, trained writers as well as incidental content producers, a solution that is easy to use and that can quickly give them an indication of the reading comprehension level of the text they are working on. The tool was calibrated in cooperation with the specialists of Eenvoudig Communiceren, who created specific reference texts for all the reading skills levels of the CEFR.

## **2. The Reading Level Test Tool**

The Reading Level Text Tool (RLTT) is a free and easy to use online tool. It was created as part of a short project for the Revalidatiefonds (Rehabilitation Fund), which included a series of twelve presentations for Dutch municipalities about low literacy in the Netherlands. During the presentations, content editors indicated their need for a tool that could give them an indication of the reading comprehension level of their texts. Although the project had only a limited budget of 6000 euro, it rendered a free, online RLTT that is currently also used by Microsoft in their special accessibility support website in the Netherlands<sup>18</sup>, called “Kleed je site uit” (see figure 1, below). The modest purpose of the tool is to create awareness among content providers of the difficulty level of their text and to give them a first indication of the ease-of-reading of a text draft.

The tool allows the text writer to upload unformatted text that is then compared to a corpus of words and other sets of measurable elements. The corpus has been put together in collaboration with Eenvoudig Communiceren (Simple Communication), specialists in reading comprehension level in the Netherlands. They provided the reference texts for the different CEFR reading levels that we needed to calibrate the tool.

The different parameters that are used by the tool to analyse text for their ease-of-reading and calculate an indication of the reading comprehension level are:

- Average number of simple words. We have a corpus of words that is checked for level by specialists from their panels of users and experience in re-translating texts to lower CEFR levels. The average is calculated as the number of simple words from the corpus of words divided by the total number of words (minus the names and terms that are indicated by the user or presumed by the tool if no user input is detected). The corpus of words contains words from the A1 reading comprehension level of CEFR. This is the most extensive part of the RLTT.
- Average number of words per sentence.
- Average number of pronouns in a sentence.
- Average number of syllables per word. This is calculated using the number of syllables and dividing that by the number of words (minus the names and terms input by the user or presumed by the tool if no user input is detected). The syllables are determined using a rather simple method of searching for the following combinations in the words and counting them as syllables:
  - Three letters: aai, eau, eeu, eui, oei, ooi
  - Two letters: aa, ai, au, ee, ei, eu, ie, oe, oo, ou, ui, uu, oi, ue, ae
  - One letter: a, e, i, o, u, y
 The tool first checks for the three letters and then subsequently goes to two and one letter not counting previously counted letter combinations.
- Average number of prepositions per sentence. This is calculated by dividing the number of prepositions in the text by the number of sentences. The list currently contains a list of 91 Dutch prepositions with which the RLTT has been calibrated.
- Number of names and terms. We follow the WCAG proposal to ignore proper names in the text. WCAG specifies that “proper names can be ignored or removed from the text before assessing whether it meets the reading ability requirement”<sup>3</sup>. The tool offers the option to count names (including brand names that are considered to be well-known) and asks for user input on the total number of names. If there is no user input, the tool will itself try to find names through a simple algorithm.

WCAG also proposes to remove titles from the test<sup>3</sup>. At the moment, this is not necessary for the RLTT analysis, because the user of the tool submits unformatted text and it is not yet possible to select complete pages including titles. The tool can however warn users not to include titles into the text that is tested.

### 2.1. Calculating the Reading Level Score

The tool calculates the RLTT level score on the basis of the elements that are present in the text and their calculated averages. Every element receives a score of 1 to 10. This is done by adding a cascading switch that scores from 1 (lowest) to 10 (highest). For example, the score for the average number of simple words per sentence is calculated by checking the words in the text against a corpus of simple words, and then calculate the proportion of simple words in the total number of words in a sentence (excluding proper names and terms that are indicated by the user).

Table 1. Score calculation for average number of simple words per sentence.

Outcome of calculation	Score
< 0.625	10
< 0.65	9
< 0.675	8
< 0.7	7
< 0.725	6
< 0.775	5

< 0.8	4
< 0.85	3
< 0.9	2
< 0.95	1

This type of calculation is made for every sentence in a text, and for the whole series of elements that are covered by RLTT. The tool determines the reading comprehension level score on the basis of the total score of all calculations for all the elements. For the final calculation of the reading comprehension level score, such as A2 or C1, the tool uses an algorithm that is based on earlier calibration of the tool with reference texts.

## 2.2. Determining the reading level from the elements

To determine the overall reading comprehension level score of the text, we have used prefabricated reference texts that were developed by specialists of the Eenvoudig Communiceren (Simple Communication) organisation in the Netherlands. They based the reference texts and the calculation of the CEFR reading comprehension level score on their experience with re-phrasing and adapting texts to make them easy-to-read for the various CEFR levels and on their experience with panels of representative users of all skills levels.

All over, 50 texts were created or adapted that represent the various reading skills levels. The 50 reference texts were used to calibrate the different elements of the RLTT. The calibration is done inside the tool after collecting all the elements.

The tool measures full scores (like A1, A2, etc.) or intermediate scores (A2/B1, B1/B2), to give the users of the tool a bit more insight in the score. One can find RLTT on the site of the Accessibility Foundation (accessibility.nl). On that site, the score is indicated both graphically and by presenting the overall score, e.g. A2. Web developers are encouraged to include the tool in their content management system or in their own tool and use the ease-of-reading scores delivered by RLTT.

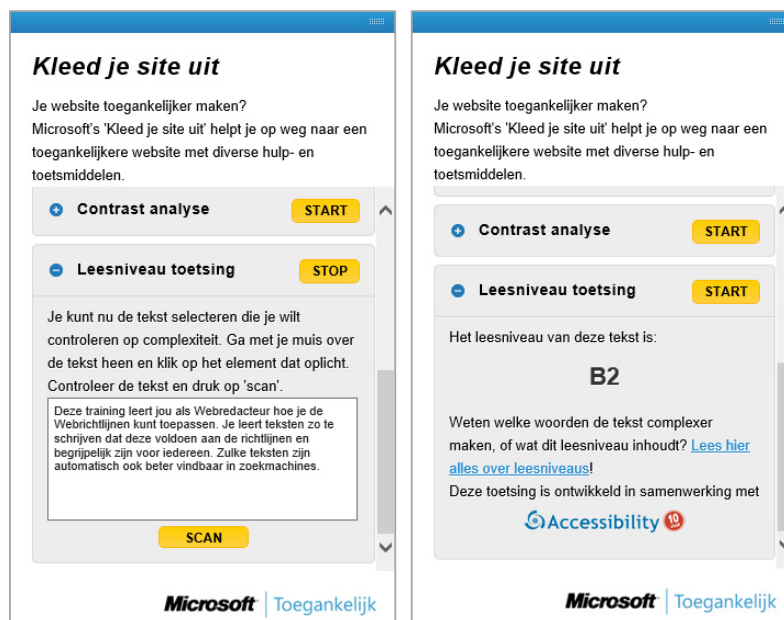


Fig. 1. (a) Microsoft use of RLTT: Importing text from the webpage; (b) outcome after pressing 'scan'.

Please note that the tool gives just an indication of the reading comprehension level; its primary function is raising awareness and creating support. The tool does not (yet) take elements like order of words in account, or grammatical complexity, personal or situational characteristics of the reader, attitudes like motivation and self-efficacy, etc. The tool itself has no built-in artificial intelligence or natural language processing facilities. For example, if the order of words in a sentence is completely changed, the tool will still render the same reading comprehension level score for that sentence.

### 3. Effectiveness of the tool

Since its launch late 2009, the tool has been used more than 325.000 times, including the use through other parties than the Accessibility Foundation by means of a software API. The tool is used more than 350 times per day in the Netherlands, mostly by government agency officials. They often use the tool a few consecutive times, to test and re-test a text they are writing. The RLTT page does not offer a feedback option, but every now and then we received emails from organizations and individuals with specific requests and comments. User for example requested that we edited the accompanying text for ease-of-reading (was level B2), proposed changes for the tool engine (both to add elements to the tool and to add information about the results of a test), or just shared their enthusiasm about the tool.

“an original and high-profile tool that helps in the struggle against low literacy”  
(Jury report of the National Literacy Award 2009)

The effectiveness of the tool has been studied by Jansen and Boersma in their article “*Meten is weten?*”<sup>6</sup>. In their study, They compare RLTT to two other (commercial) tools available in the Netherlands. Although the authors indicate that their evidence is not very strong, they consider RLTT to be the most effective of the three tools.

Jansen and Boersma indicate a number of shortcomings<sup>6</sup>, some of which apply to the three tested tools. First, they conclude that for the calibration of RLTT, no real users where consulted but the work was done by specialists of a publishing house specializing in CEFR reading comprehension levels (*Eenvoudig Communiceren*). However, the texts provided by the experts of the company did originate from documents that have been tested with panels of users, contrary to what Jansen and Boersma stated.

Second, the reading comprehension levels as implemented in the three tools go beyond CEFR intentions. The Common European Framework is about much more than just text comprehensibility. Jansen and Boersma find the tools limited in their coverage of CEFR. According to them, there is no support for the claim that one single text can have a CEFR reading comprehension level score.

Finally, Jansen and Poersma put the tools to the test with twenty different texts and found that the three tools failed to recognize the A1 and A2 level texts in their sample. Also it showed that the three tools assigned different reading comprehension levels to the submitted texts.

If we consider the limited size of the project and the funding, the considerable usage numbers of the tool, its free, online availability and its ease-of-use, its modest aim to create awareness, and the outsiders assessment that it is as least as effective as currently available commercial reading comprehension level tools, we must conclude that the RLTT is fulfilling its task to create awareness and at the same time help content developers to get an indication of the reading comprehension level of their text. The tool is also open for future development.

#### 4. Future development of RLTT

We are very much aware of the limitations of RLTT as identified by Kraf and by Jansen and Boersma<sup>5,6</sup>. We think that what we have now is a promising start for further development. We want to do more research on the effectiveness and validity of the tool and find good ideas in the comments of researchers and users. We specifically would like to work with panels of representative users with various reading skills to extend the possibilities of the tool and to validate its validity. We would like to study in more detail the factors that influence the experienced difficulty of texts.

Besides the research plans above, we hope to develop RLTT further in a number of ways.

1. Add more and different corpus texts and lists, specifically for government information. We intend to look at existing government texts and analyze them with panels of real users. The outcome will be use to build up new corpus libraries for the tool. There are already corpus libraries available and the University of Twente has done relevant research for the Dutch Tax Administration, that can be used as input.
2. Add corpus libraries for different target groups who have trouble with difficult texts. We could distinguish between target groups, for example people who read Dutch as a foreign language, people with forms of aphasia due to brain impairments, severe forms of dyslexia, and relate target groups to different content domains, such as government information, health information, tax information.
3. Add more input from the user panels and judges when examining the texts that are provided for the tool. This can include the use of eye tracking. We would also like to include people with disabilities in these tests such as people with partial eyesight in different age groups (there is an ongoing project about eye-tracking at Bartiméus for people who are partially sighted).
4. Extend the API in such a way that the tool can be used in Content Management Tools and by other parties who want to make their content providers continuously aware of the need to provide comprehensible content.
5. Extend the output of the tool to give more insight in the results. Currently the tool just indicates the reading comprehension level score, e.g. A1 or B2. In the API we have already built in the possibility to identify the elements that lead to a higher reading comprehension level score but this option needs further work.
6. The experiences of users of the RLTT could provide input for a European Plain Language Act, comparable to the American Plain Language Act of 2010.

With these steps, we aim at making the free tool even more valid and effective, using the input from users and researchers to extend the features of RLTT. We envision that the tool is included on a regular basis in the Content Management Systems that are used by government bodies in the Netherlands. In that way we give better support to the implementation of the Web Content Accessibility Guidelines in the Netherlands when it comes to easy to read on the web.

#### References

- [1] Council of Europe. *The Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)*. Strasbourg: COE; 2000. Available: [http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key\\_reference/Overview\\_CEFRscales\\_EN.pdf](http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key_reference/Overview_CEFRscales_EN.pdf); Last accessed 20 July 2013.
- [2] Fouarge D, Houtkoop W, Van der Velden R. *Laaggeletterdheid in Nederland*. 's-Hertogenbosch: Expertisecentrum Beroepsonderwijs; 2011.
- [3] Caldwell B, Cooper M, Guarino Reid L, Vanderheiden G. *Web Content Accessibility Guidelines (WCAG) 2.0*. Boston: W3C; 2008. Available: <http://www.w3.org/TR/2008/REC-WCAG20-20081211/>. Last accessed 22 July 2013.
- [4] Accessibility Foundation. *Accessibility Online Leesniveautoetstool*. Available: <http://www.accessibility.nl/kennisbank/tools/leesniveautool/leesniveaus>. Last accessed 22 July 2013.
- [5] Kraf R, Lentz L, Pander Maat H. *Drie Nederlandse instrumenten voor het automatisch voorspellen van begrijpelijkheid. Een klein consumentenonderzoek*. Tijdschrift voor Taalbeheersing; 2011. 33 (3) p. 249-265.



- [6] Jansen C, Boersma N. *Meten is weten? over de waarde van de leesbaarheidsvoorspellingen van drie geautomatiseerde Nederlandse meetinstrumenten*. Tijdschrift voor Taalbeheersing; 2013. 35 (1), p47-62.
- [7] UNESCO. *International Standard Classification of Education*; 1997. Available: [http://www.unesco.org/education/information/nfsunesco/doc/isced\\_1997.htm](http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm). Last accessed 20 september 2013.
- [8] European Commission. *Directive of the European Parliament and of the Council on the accessibility of public sector bodies' websites*. Brussels: EC; 2012. Available: [http://ec.europa.eu/information\\_society/newsroom/cf/dae/document.cfm?doc\\_id=1242](http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1242). Last accessed 22 July 2013.
- [9] UN CRPD. *Convention on the Rights of Persons with Disabilities*. 2006. Available: <http://www.un.org/disabilities/convention/conventionfull.shtml>. Last accessed 22 July 2013.
- [10] US Access Board. *Information and Communication Technology (ICT) Standards and Guidelines*. US; 2000. Available: <http://www.access-board.gov/guidelines-and-standards/communications-a-it/about-the-ict-refresh/draft-rule-2011>. Last accessed 1 march 2012.
- [11] DGME. *Référentiel Général d'Accessibilité pour les Administrations RGAA*. Paris; 2009. Available: <http://references.modernisation.gouv.fr/rgaa-accessibilite>. Last accessed 20 March 2013.
- [12] Berkouwer G, Fase A, Hay S, Rooij R de, Velleman E, Willems K. *Webrichtlijnen versie 2*. Den Haag; 2011. Available: <http://versie2.webrichtlijnen.nl/norm/20110701/>. Last accessed 1 july 2013.
- [13] Council of Europe. *Education and Language, Language Policy. Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*. Strasbourg: COE; 2012. Overviewpage of website. Available: [http://www.coe.int/t/dg4/linguistic/Cadre1\\_en.asp](http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp). Last accessed 22 September 2013.
- [14] Council of Europe. *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)*. Strasbourg: COE; 2009.
- [15] BureauTaal. (2013). Wat is eenvoudig Nederlands. Available: <http://www.bureautaal.nl/wat-is-eenvoudig-nederlands-1.html>. Last accessed 22 September 2013.
- [16] OESO. (2000). Literacy in the information age: Final report of the international adult literacy survey (IALS). Paris.
- [17] Bohnenn, E., Ceulemans, C., Guchte, C, van de., Kurvers, J., Tendeloo, T, van. (2004). Laaggeletterd in de Lage Landen, Hoge prioriteit voor beleid. Den Haag: Nederlandse Taalunie.
- [18] Microsoft. (2013). Kleed je site uit. Available: <http://www.microsoft.com/netherlands/toegankelijk/kleedjesiteuit/>. Last accessed 22 september 2013.