

CLCM - A Linguistic Resource for Effective Simplification of Instructions in the Crisis Management Domain and its Evaluations

Irina Temnikova, Constantin Orasan, Ruslan Mitkov

Research Institute in Information and Language Processing, University of Wolverhampton
Stafford Street, Wolverhampton, UK
Irina.Temnikova@gmail.com, C.Orasan@wlv.ac.uk, R.Mitkov@wlv.ac.uk

Abstract

Due to the increasing number of emergency situations which can have substantial consequences, both financially and fatally, the Crisis Management (CM) domain is developing at an exponential speed. The efficient management of emergency situations relies on clear communication between all of the participants in a crisis situation. For these reasons the Text Complexity (TC) of the CM domain needed to be investigated and showed that CM domain texts exhibit high TC levels. This article presents a new linguistic resource in the form of Controlled Language (CL) guidelines for manual text simplification in the CM domain which aims to address high TC in the CM domain and produce clear messages to be used in crisis situations. The effectiveness of the resource has been tested via evaluation from several different perspectives important for the domain. The overall results show that the CLCM simplification has a positive impact on TC, reading comprehension, manual translation and machine translation. Additionally, an investigation of the cognitive difficulty in applying manual simplification operations led to interesting discoveries. This article provides details of the evaluation methods, the conducted experiments, their results and indications about future work.

Keywords: controlled languages, text simplification, text simplification evaluation methodologies

1. Introduction

Due to the increasing number of emergency situations which are affecting the world, the field of Crisis Management (CM) has been developing rapidly. The efficient management of emergency situations relies on clear communication between all of the participants in such situations. Unfortunately, previous research has shown that crisis management documents exhibit a large number of high text complexity issues. This is particularly problematic in emergency situations where human comprehension under stress is hindered (Winerman, 2009). For this reason, we argue that CM documents need Text Simplification (TS). In this paper, we define text simplification as a research area whose aim is to improve text comprehension through re-writing of particular text complexity phenomena, while preserving the information content of the original text. Some of the text simplification approaches rely on controlled languages which are sets of predefined linguistic restrictions at various text levels (e.g. lexical, syntactic and discourse) which should be used while writing a text.

This article presents a new linguistic resource in the form of Controlled Language (CL) guidelines for simplification in the crisis management domain. Given the importance of correct communication in the field and the fact that inaccurate communication can have fatal consequences, a multi-perspective evaluation was carried out in order to assess the quality of the texts produced using these guidelines and find out how difficult it is for humans to apply them. The article is structured as follows: Section 2 provides a brief overview of the related work in the crisis management and text simplification domains. The linguistic resource is described in Section 3 and evaluated in Section 4. The paper finishes with conclusions and a discussion of possible applications of the resource, including plans for future work.

2. Natural Language Processing (NLP) in the Crisis Management Domain

Although several CM computer systems were developed, the contribution of NLP components in the crisis management domain is limited. The most common uses of NLP in crisis management involved applying information extraction to detect emergency events on the web (Corvey et al., 2010; Ireson, 2008) or employ text mining techniques to ensure epidemic surveillance on the basis of clinical notes (Conway et al., 2009).

In contrast, the area of text simplification has received more attention from the NLP community who developed automatic and semi-automatic methods for simplifying texts. The existing text simplification approaches that rely on NLP have never addressed the CM domain, but instead they simplify texts for a variety of low-skilled readers (Canning, 2002; Gasperin et al., 2009) or as an input for other NLP applications (Chandrasekar et al., 1996; Siddharthan, 2003).

In addition to automatic and semi-automatic text simplification approaches, researchers have proposed manual methods which rely on controlled languages. The state-of-the-art approaches in controlled languages address mainly simplification of technical documentation (ASD, 2005; Kuhn, 2009) or texts for low-skilled readers like Basic English (Ogden, 1930) or the Plain English Campaign¹. Only a few CL approaches have been developed for the Crisis Management domain and were either restricted to a different language, like French (Renahy and et al, 2010) or to a well-defined CM area (e.g. aeronautics, police) (ASD, 2005; Johnson et al., 1993). Due to their domain- and document-specificity controlled languages cannot usually be easily exported to other domains. In addition, quite often these controlled

¹www.plainenglish.co.uk. Last accessed on March 21st, 2012

languages are confidential which means it is difficult to find many details about them and use them. The next Section presents the CL for English developed specially for the Crisis Management domain, and more specifically for emergency instructions.

3. The Controlled Language in Crisis Management (CLCM)

3.1. Origins of CLCM

The Controlled Language in Crisis Management was developed in the context of the MESSAGE Project², an EU-funded project with four partner universities. The project was coordinated by Centre Tesnière, Besançon, France who have fifteen years of CL experience in the CM domain. The aim of the project was to transfer the CL for French developed by the coordinators (Renahy and et al, 2010) to the languages of the partners' countries (English, Spanish and Polish) in close collaboration with domain specialists. LiSe consisted of a collection of predefined rules for writing simple documents and addressed different document types (protocols for specialists and alert messages), application domains (medical, aeronautics, and police) and readers (specialists and non-specialists). CLCM was developed by adapting LiSe's guidelines from French to English in collaboration with UK CM specialists and on the basis of a corpus analysis of English CM documents, psycholinguistic studies about human text comprehension, and information from the English grammar. The CLCM guidelines presented in this paper focus on the simplification of a specific CM document type: emergency instructions for the general, non-specialist population. The choice of the document type was motivated by two reasons: the non-specialist population is the weakest link in CM communication and previous analysis of a corpus of CM documents showed that they would greatly benefit from simplification using a controlled language.

3.2. Description of CLCM

From the point of view of purpose, CLs can be human-readers-oriented, machine-oriented or mixed-purpose. CLCM is a mixed-purpose CL designed mainly to improve human comprehension of written text in emergency situations, but it can also be used to ensure good translation results. It is different from the formal-logic-based CLs by having more free-text rules and also different from the human-only CLs, as it has more formal rules (constituted by a reference number, definition, and incorrect and correct examples).

CLCM relies on several types of simplification rules: prescriptive, proscriptive, construction, interpretation and paraphrasing rules listed in special guidelines. The CLCM guidelines follow the LiSe guidelines for protocols for specialists and thus contain a description of the rules notation, general rules that apply to the whole document, rules for specific document elements, definitions of the allowed syntactic structures, and an example of a re-written

text. They feature thirty-five pages of over eighty rules for TS of emergency instructions for the general population. The CLCM guidelines adapt the LiSe guidelines to instructions for the general population by adding

- a small dictionary of grammatical terms (in case the writer has no linguistic training),
- a list of forbidden syntactic structures,
- a list of lexical rules, a list of forbidden lexical expressions,
- a domain dictionary (in case the writer is not a domain specialist),
- a step-by-step re-writing example, and
- changes to the rules format, which makes the guidelines more accessible

An example of a rule is provided in Figure 1, while an example of a simplification on the basis of the CLCM simplification rules in Figure 2.

In_L_05: Keep preposition and verb together in phrasal verbs.	
Switch the lights off.	Switch off the lights.
Explanation: Preposition and verb separated by many words create difficulties for both non-native speakers and machine translation engines.	

Figure 1: Example of a CLCM rule.

Figure 1 shows that each rule has a unique reference number (e.g. *In.L.05*), which describes to which document-type and document element it belongs and what type of rule it is (in this case, lexical). This notation allows easy browsing among the rules. The table contains both the original, unsimplified snippet of text, and how it should be simplified. The rule is explained below the examples.

Original	Simplification
If you suspect there is something embedded, take care not to press on the object. Instead press firmly on either side of the object and build up padding around it before bandaging to avoid putting pressure on the object itself.	<p>How to treat severe bleeding</p> <p>If you suspect there is an embedded object:</p> <ol style="list-style-type: none"> 1. Avoid pressing on the embedded object. 2. Do the following actions simultaneously: <ul style="list-style-type: none"> • Press firmly on either side of the embedded object. • Build up padding around the embedded object. <p>Explanation: In order to avoid putting pressure on the object itself.</p> 3. Bandage the wound.

Figure 2: CLCM simplification example.

As can be seen in Figure 2, the original passage³ is given in the left column, whilst the simplified version of the passage is displayed in the right column.

The main TC issues addressed by CLCM are:

- discourse and text-level complexity (non-logical and non-chronological information order, unclear text structuring)

²<http://clg.wlv.ac.uk/projects/Message/> (last accessed on March 13th, 2012.)

³Passage taken from www.redcross.org "How to treat severe bleeding", last accessed on March 11th, 2012)

- lexical complexity and syntactic complexity (e.g. technical terms, figurative language, passive voice, negation)
- lexical and syntactic text ambiguity (e.g. unclear anaphora, non-transparent abbreviations, vague quantifiers, unclear attachment, garden path sentences)

The simplified documents follow a predefined structure composed of easily identifiable elements. The allowed elements are: a title, subsection titles, conditions, actions to be taken, comments (e.g. explanations), and lists of items. Some of these elements are compulsory (e.g. title, instructions), whereas some are not (e.g. comments).

The **title** is very important, as it specifies the topic of the whole document. It is essential to separate it graphically from the other elements, and to make it short and meaningful. If the original document contains large enough specific situations, they should be placed into separate sections, explicitly marked by specific titles, which follow similar rules to the title of the document.

The **instructions** contain the main information, i.e. they list the actions to be undertaken during an emergency. The **conditions** are important because they specify under which circumstances certain actions need to be carried out. For this reason, actions corresponding to a condition are listed under it and indented. The lists visually improve the understanding of enumerated items. The lists are also optional, except when there are enumerations of more than two elements.

The **comments** are the least important elements of the simplified text. There are two types of comments: the ones which can be put in the beginning of a document, such as a **definition** of technical terms in the document, the **target audience**, or a **reference** to another document; and those which can be put after a condition or an instruction, such as the **aim**, its **explanation**, any **exceptions** that should be considered, or an **example**. Particular types of comments also include the **warnings**, which can be written in order to warn about a dangerous situation, and which have a higher priority than the other types of comments.

Each of the document elements is restricted by a set of different rules (general, formatting, syntactic, punctuation, and lexical). The general rules (**G-rules**) define the purpose of the simplification and of the document, its structure, its contents, and the ways to order and group information. They also impose restrictions concerning the language of the whole document. Examples of G-rules are provided in Figure 3.

The Formatting rules (**F-rules**) define how to graphically present the document, including the formatting between and inside the different document sub-parts, specifying blank lines, font style, font size, indentation, etc. An example of an F-rule is: "In_F_03: Separate each block of instructions with a new line." The syntactic rules (**S-rules**) impose domain-independent and domain-dependent restrictions on the syntactic realization of the phrases and sentences in the simplified documents. Examples of S-rules can be seen in Figure 4.

Rule restrictions	Example
Purpose of the simplification	In_G_00: Preserve the meaning and the information content of the original document.
Purpose of the document	In_G_01: This type of document is mainly aimed at the general audience (not specialists).
Document structure	In_G_03: The document should contain easily identifiable parts: (title of the document, sections, titles of sections and sub-sections, compulsory actions to be always done, conditions, instructions, lists, and comments).
How to group and order information	In_Cd_G_01: If you have 2 or more alternative conditions: Start with the most specific one. End with the most general one.

Figure 3: Examples of G-rules.

Rule restrictions	Example
Domain-independent restrictions	In_S_01: Use only the allowed syntactical structures. In_S_02: Avoid the forbidden syntactical structures on p.29.
Concrete, domain-dependent restrictions	In_S_03: Avoid demonstrative pronouns. In_S_15: If an adjective determines more than one noun: Repeat the adjective. In_Li_S_01: Use only NP with indefinite articles as elements of the list.

Figure 4: Examples of S-rules.

The lexical rules (**L-rules**) provide restrictions at the lexical level, and, like the syntactic rules, can impose restrictions of a domain-independent and of a concrete, domain-dependent nature. Examples of L-rules can be seen in Figure 5.

Rule restrictions	Example
Domain-independent restrictions	In_L_01: Choose the words in accordance with the lexical rules on p. 31. In_L_04: If possible: Use the alternative expressions in the dictionary on p.33.
Concrete, domain-dependent restrictions	In_L_06: If possible: Avoid acronyms and abbreviations. In_I_L_01: If possible: Use discourse connectives (e.g. first, second, next, then, finally).

Figure 5: Examples of L-rules.

The final type of rules are the punctuation rules (**P-rules**), which impose restrictions on the use of punctuation marks in the document. An example of a P-rule is "In_T_P_01: Avoid any punctuation signs at the end of the titles. ". The number of different types of rules per document elements is not evenly distributed and varies between zero and seventeen rules per type per kind of element. Most of the rules describe the main writing of the document, while the sections describing specific document elements contain fewer rules (between four and ten). The CLCM rules can be adapted to other types of documents, such as language learning materials, medical leaflets, or technical manuals. The next section evaluates the controlled language described here.

4. CLCM Evaluations

Due to the fact that this linguistic resource is aimed at the Crisis Management field, where accuracy and quality of the produced texts are crucial, an extensive evaluation of CLCM from several different perspectives was conducted. The evaluation of this linguistic resource follows the evaluation methodology proposed by Hirschman and Mani (2001) for assessing the output of NLP systems. Namely, we conducted an intrinsic evaluation of the resource

(Section 4.1.) followed by an assessment of the impact of the resource on extrinsic tasks (Section 4.2.). In addition, evaluation of its acceptability with end-users is presented in Section 4.3.

4.1. Intrinsic Evaluation of CLCM

The first experiment aimed to estimate the level of TC in CM documents. Although there has been enough research in measuring TC, such as the classic readability formulae (DuBay, W. H., 2004) and modern approaches (McNamara et al., 2010), none of it is tailored for the CM domain, where human comprehension is different (Kiwan et al., 1999) and thus a specific set of high text complexity issues must be examined.

The TC evaluation presented in this section employed a set of original (considered “complex”) texts and their simplified versions, produced manually by six linguists. The evaluation method first required identifying the TC issues, which, if high, affect human comprehension under stress. The high TC issues were identified on the basis of state-of-the-art psycholinguistic findings regarding the process of reading (Harley, 2008) and taking into account stress situations. Table 1 provides the results of the following measures of text complexity: Average Sentence Length (ASL), Average Word Length (AWL), Lexical Diversity (LD) and Average Number of Word Senses (ANWS), Proportion of Coordination Markers (PCM), Proportion of Subordination Markers (PSM), Proportion of Relative Clause Markers (PRCM), Proportion of Ambiguous Quantifiers (PAQ) and Proportion of Personal and Possessive Pronouns (PPPP).

High TC issues	Original text	Simplifications
ASL	15.922	10.799
AWL	5.327	5.389
LD	0.976	0.961
ANWS	8.478	8.811
PCM	0.063	0.036
PSM	0.040	0.039
PRCM	0.014	0.009
PAQ	0.010	0.005
PPPP	0.048	0.042

Table 1: Measures of text complexity for complex and simplified texts.

As can be seen, the first column of Table 1 shows the list of TC measures followed by their values for the original text in column 2 and for the simplified text in column 3. If CLCM decreases text complexity, then the numbers in column 2 should be higher than the numbers in column 3. As can be seen, this is true for all TC issues except for AWL (Average Word Length) and ANWS (Average Number of Word Senses) which are the other way around. In addition, although all of the other values are statistically significant with 95% confidence, the difference between the Proportions of Subordination Markers in the original and simplified texts is not statistically significant. The higher value of AWL for the simplified text can be explained by the fact that at the time of simplification

no dictionary of alternative terms was available and thus the linguists could not consult any resource for replacing long words with appropriate shorter synonyms. The higher value of ANWS for the simplified text can be explained by the fact that while replacing more technical terms with common synonyms, these new synonyms are more ambiguous according to Zipf’s law (Zipf, 1949). As can be seen, however, overall the results show a positive impact of the CLCM simplification on TC.

4.2. Evaluation of CLCM on extrinsic tasks

The evaluation of CLCM on extrinsic tasks consisted on testing the impact of the CLCM simplification output on reading comprehension and manual and machine translations. The experiments conducted for these purposes are described in the following sections.

4.2.1. Evaluation of the CLCM impact on reading comprehension

The second experiment is a large text-understanding experiment involving one hundred and four volunteers. The participants had to read in a limited time simple and complex emergency instructions texts in random order and to reply to multiple-choice questions after each of them. The texts used for the experiment were four different complex texts of the same length and similar TC levels, and their manual simplifications. The participants were shown complex and simplified texts in alternating randomized order. However, no participant was shown both the complex and simplified versions of the same text. Each participant had to read four texts in total: two complex and two simplified ones. The questions regarding each set of texts were the same to ensure comparability of the results. The order of questions and answers to choose from were randomized. The impact of the CLCM simplification was measured by comparing the proportion of correct answers given to the complex and simplified texts and by comparing the time necessary to provide correct answers for both kinds of texts. The participants were very diversified according to variables such as gender, age, profession and native language, but most of them were university students or coming from the research community.

The overall results showed no clear indication of any positive or negative effect of the CLCM simplification on reading comprehension. This was due to the fact that the participants were highly competent readers able to deal with both complex and simple texts. However, analysis of the results revealed that some specific categories of readers did benefit from the CLCM simplification. When the native language is considered, Basque, Chinese, Indian and Vietnamese native speakers were helped by the simplified version, whilst native speakers of Germanic, Romanic or Slavic languages were not. Readers with high reading skills, such as lawyers, linguists or translators, benefited less from using simplified texts than students and NLP researchers. The same was noticed when comparing women vs. men. Male participants, especially native English speakers, could answer more questions on the basis of the simplified texts than when a complex text was shown to them.

Figure 6 shows the impact of the CLCM simplification on the time taken by male and female participants to provide the correct answers. The line with the circle represents the complex text, while the wider line with the triangle - the simplified text.

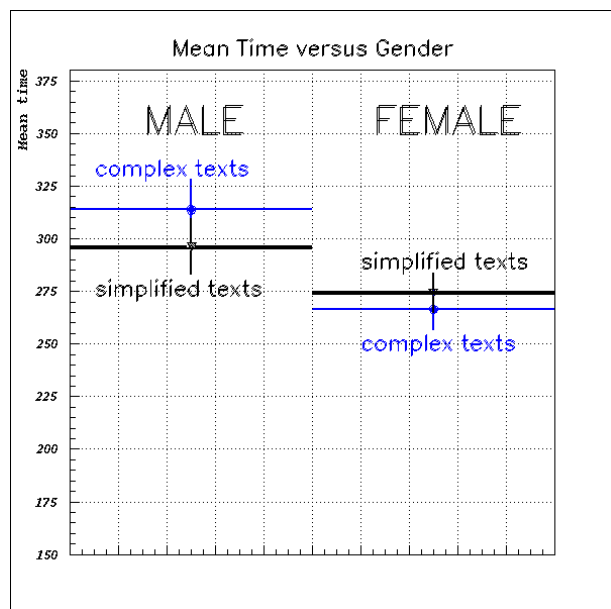


Figure 6: Time taken to provide correct answers for female and male participants.

Figure 6 clearly shows that male participants provide correct answers faster to the questions after the simplified text, while the female participants were the other way around. The shorter time that female participants require in giving correct answers to both texts may be explained by women's better reading skills (Lietz, 2006), whilst the poor impact of text simplification on women could be explained by the fact that they need to collect more information about the environment before making a decision (Sanz et al., 2007). This is not immediately possible with simplified texts because they have shorter sentences. Figure 7 shows another interesting discovery, that among all the language groups, the native speakers of Basque benefited most from text simplification. They were able to correctly answer more questions on the basis of simplified texts than when they were shown complex texts. Analyzing the performance of Basque speakers was motivated on one hand by the fact that there was a large group of participants with this native language and on the other that this is non an Indo-European language.

Figure 7 shows the comparison of the four sets of complex/simplified texts for Basque native speakers. The vertical axis contains the number of correctly answered questions, while the horizontal one - the specific texts. The higher columns of the simplified text show that there is a higher number of correct answers to the questions about it. What can be noticed is that sometimes (Set 1) the proportion is almost double, while sometimes (Set 4), the correct answers of the simplified text are 100%.

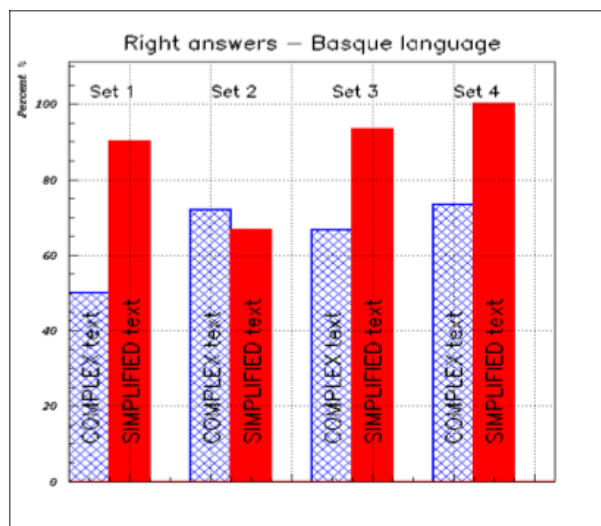


Figure 7: Proportion of correctly answered questions for native speakers of Basque.

4.2.2. Evaluation of the impact of CLCM on Manual and Machine Translation

An extrinsic evaluation was carried out in order to demonstrate the impact of the CLCM simplification on translation tasks. The two tasks chosen were Manual Translation (ManT) and Machine Translation (MT), as they are considered to be important for the efficient communication of emergency instructions. The experiment involved 25 professional translators with seven working languages and the publicly available statistical MT engine Google Translate⁴. The experiment consisted of splitting a complex text and a simplified text into a set of sentences and then asking the translators to manually translate half of the sentences and post-edit the MT-translated versions of the other half of sentences. In both settings, a special web interface was used.

In order to achieve an objective evaluation, a different approach from the existing ManT evaluation approaches, which are mostly manual (House, 2001; Hale and Campbell, 2002) was followed. Specifically, the impact of the CLCM simplification on manual translation was evaluated by calculating the time employed to manually translate texts, which can provide an objective measure of the ease of translation. The results of the manual translation experiment showed rather positive results for the simplified text, but the statistical significance of the results was not satisfactory (i.e. in most of the cases only 85%).

As the experiment did not rely on alternative MT translations, the evaluation of the impact of CLCM on MT could not employ BLEU (Papineni et al., 2002) or TER (Snover et al., 2006). In our evaluation, we measured and compared the post-editing effort for automatically translating the complex and simplified sets of sentences. The assumption behind this evaluation was that if the CLCM simplification has a positive impact on MT, then the post-editing cost of the simplified sentences should be lower than the post-editing cost of the complex sentences.

⁴<http://translate.google.com/>

Similarly to (Krings, 2001) and (O'Brien, 2005), the post-editing effort was evaluated from three perspectives: temporal, technical and cognitive points of view. The temporal evaluation consisted in comparing the times spent to manually post-edit the two MT output texts, the technical evaluation compared the edit distance between the MT output and the post-edited versions of the two texts. The results of this evaluation showed completely different results to the experiment that used ManT. The temporal and technical evaluation of the impact of CLCM on MT has shown a clear improvement for post-editing the simplified text, which was statistically significant with 98-99% confidence, using directional z-test. In addition, the average normalized time employed to post-edit the text was much lower than the time employed to manually translate the same text, which confirmed the discovery of C. M. de Sousa et al. (2011) that MT+post-editing is faster than manual translation from scratch. The simplified text showed larger difference (post-editing was 34% faster than translating) than the complex text (post-editing was 17% faster) in the comparison of the two tasks. Both differences of the means were statistically significant with 98-99% confidence, which showed that the CLCM simplification had a decisively positive impact on the more important of the two translation tasks.

Further on, in order to ensure a more fine-grained examination of the changes applied to the MT output during post-editing, a new cognitive evaluation method was developed. The existing cognitive evaluation approaches such as relying on think-aloud feedback, the number of alternative post-editing versions and manual ratings of translations for post-editing difficulty (Krings, 2001; O'Brien, 2005; C. M. de Sousa et al., 2011) were not used because they were considered too subjective and difficult to apply. The proposed method consists of comparing the proportion of cognitively difficult-to-correct MT errors in the complex and simplified texts. For more details see Temnikova (2010). The assumption was that if CLCM has a positive impact on MT, then the proportion of cognitively difficult-to-correct MT errors in the complex text will be higher than the proportion of cognitively difficult-to-correct MT errors in the simplified text. The investigation of the cognitive effort to post-edit texts in three out of the seven languages (Bulgarian, Russian and Spanish) also showed that the errors corrected by post-editors in the simplified text require low cognitive effort to be corrected, comparing with those of the complex text. A detailed evaluation of this can be found in (Temnikova, 2010).

4.3. Evaluation of the CLCM acceptability with end-users

Finally, the difficulty of applying the CLCM rules and the user requirements for a CLCM writing aid were investigated. The experiment involved six computational linguists who manually simplified four emergency instruction texts of a total of two thousands words, according to the CLCM guidelines. Several evaluations have been carried out on the produced results. First, a comparative TC analysis (already described in Section 4.1.1.) of the complex and simplified texts

was conducted. Second, a comparison of the different manual simplifications of the same complex text, similar to the CNA analysis, was conducted. Third, the time to read the guidelines and simplify the text was measured and finally, the difficulty of applying concrete simplification operations and the necessity to implement some of them automatically was investigated by asking the participants to fill in a questionnaire.

The results of the comparison of the different simplified versions of the same complex text showed that there were differences in terms of rendering the same text complexity issue, text structure and discourse order of segments. Measurement of the time revealed that although a learning effect was observed, manual simplification takes a substantial amount of time (the best simplifying speed was 127 chars/min). All the aforementioned findings lead to the conclusions that manual simplification is time consuming and involves substantial cognitive effort. In light of this, the implementation of a tool assisting with emergency instructions simplification is essential.

The users' responses to the questionnaire showed that the most difficult rules to apply were those concerning re-writing lexical and syntactic ambiguity, negation, passive voice and unclear anaphora. This finding was not surprising, as these operations are cognitively difficult, especially for some groups of readers. At the same time, a tool may not be able to help much with these operations as they are also challenging for NLP applications (Jurafsky and Martin, 2008). More details about the results of this experiment can be found in (Temnikova, 2011).

5. Possible Applications, Conclusions and Future Work

This article has presented a new linguistic resource for achieving an effective text simplification of written emergency instructions in English. A detailed evaluation from multiple perspectives was carried out in order to validate its quality. The evaluation results have shown a positive impact of the resource on a variety of tasks, such as reduction of text complexity, enhancement of text comprehension under time pressure, and improvement of manual and machine translation. The intrinsic evaluation also showed that manual simplification is highly time- and cognitive effort-consuming, and thus that implementation of a CLCM authoring aid is imperative.

Possible applications of this resource are a CL text authoring aid or a TS engine, which could benefit from the pre-defined text simplification rules. The resource can be used as a basis for implementing a Text Simplification (TS) engine or text authoring aid, as well as can be also applied to other domains and document types.

The planned future work comprises many different directions. The evaluation results could be used as the starting point to make improvements to the CLCM guidelines. The evaluation methodology could be also refined to make it more precise. The planned enhancements include a revised version of reading comprehension evaluation, expanding the set of high TC issues in the respective analysis, analysis of a more diversified sample of the population in order to include readers with lower

reading skills, employing HTER (Snover et al., 2009) for MT evaluation. Collection of a larger sample of data for both the ManT and the MT evaluations, as well as initial implementation of the most urgent CLCM authoring aid components and their evaluation, are envisaged.

6. Acknowledgements

The authors are extremely grateful to Dr. Galina Maneva, Dr. Petar Temnikov and Dr. Anke Büttner for their precious advices about the evaluation methodology and the help with the statistical processing of the results, as well as to Dr. Le An Ha for his ideas for setting the reading comprehension experiment and for the updates to the web interface. Special thanks go to Dr. Kevin Cohen who helped both with comments and by proofreading the initial abstract of this article.

7. References

- ASD, 2005. *ASD Simplified Technical English Specification ASD-STE10*. Aerospace and Defence Industries Association of Europe, Brussels, Belgium.
- S. C. M. de Sousa, W. Aziz, and L. Specia. 2011. Assessing the post-editing effort for automatic and semi-automatic translations of DVD subtitles. In *Proceedings of the International Conference Recent Advances in Natural Language Processing 2011*, pages 97–103, Hissar, Bulgaria. RANLP 2011 Organising Committee.
- Y. Canning. 2002. *Syntactic Simplification of Text*. Ph.d. thesis, University of Sunderland, UK.
- R. Chandrasekar, C. Doran, and B. Srinivas. 1996. Motivations and methods for text simplification. In *Proceedings of the 16th conference on Computational linguistics - Volume 2, COLING '96*, pages 1041–1044, Stroudsburg, PA, USA. Association for Computational Linguistics.
- M. Conway, N. Collier, and S. Doan. 2009. Using hedges to enhance a disease outbreak report text mining system. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing, BioNLP '09*, pages 142–143, Stroudsburg, PA, USA. Association for Computational Linguistics.
- W. J. Corvey, S. Vieweg, T. Rood, and M. Palmer. 2010. Twitter in mass emergency: what NLP techniques can contribute. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media, WSA '10*, pages 23–24, Stroudsburg, PA, USA. Association for Computational Linguistics.
- DuBay, W. H. 2004. The principles of readability. *Costa Mesa, CA: Impact Information*.
- C. Gasperin, E. Maziero, L. Specia, T. Pardo, and S. M. Aluisio. 2009. Natural language processing for social inclusion: a text simplification architecture for different literacy levels. In *Proceedings of SEMISH-XXXVI Seminário Integrado de Software e Hardware*, pages 387–401.
- S. Hale and S. Campbell. 2002. The interaction between text difficulty and translation accuracy. *Babel*, 48(1):14–33.
- T. A. Harley. 2008. *The Psychology of Language: From Data to Theory*. Taylor & Francis Group.
- L. Hirschman and I. Mani, 2001. *Evaluation*. Oxford University Press.
- J. House. 2001. Translation quality assessment: Linguistic description versus social evaluation. *Meta: Translators' Journal*, 46(2):243–257.
- N. Ireson. 2008. Local community situational awareness during an emergency. *3rd IEEE International Conference on Digital Ecosystems and Technologies*, pages 49–54.
- E. Johnson, M. Garner, S. Hick, and D. Matthews. 1993. *PoliceSpeak - Police Communications and Language and the Channel Tunnel - Research Report*. PoliceSpeak Publications, Cambridge.
- D. Jurafsky and J. H. Martin. 2008. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice Hall, second edition.
- D. Kiwan, A. Ahmed, and A. Pollitt. 1999. The effects of text comprehension and performance in examinations. In *Proceedings of BPS London Conference*.
- H.P. Krings. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Koby, G.S. (eds.) Kent, Ohio: The Kent State University.
- T. Kuhn. 2009. *Controlled English for Knowledge Representation*. Ph.d. thesis, University of Zurich, Switzerland.
- P. Lietz. 2006. A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies In Educational Evaluation*, 32(4):317–344.
- D. S. McNamara, M. M. Louwerse, P. M. McCarthy, and A. C. Graesser. 2010. Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes*, 47:292–330.
- S. O'Brien. 2005. Methodologies for measuring the correlations between post-editing effort and machine translatability. *Machine Translation*, 19(1):37–58.
- C. K. Ogden. 1930. *Basic English: a general introduction with rules and grammar*. London, Kegan Paul, Trench, Trubner.
- K. Papineni, S. Roukos, T. Ward, and W. Zhu. 2002. BLEU: a method for automatic evaluation of machine translation. pages 311–318.
- J. Renahy and et al. 2010. Development and Evaluation of a Controlled Language and of a computerized writing assistant LiSe to improve the quality and safety of medical protocols. In *International Forum on Quality and Safety of Health Care. Nice, France*.
- M. L. Sanz, L. De Acedo, M. T. Sanz, and B. De Acedo. 2007. Factors that affect decision making : gender and age differences. *International Journal of Psychology and Psychological Therapy*, 7(1993):381–391.
- A. Siddharthan. 2003. *Syntactic Simplification and Text Cohesion*. Ph.d. thesis, University of Cambridge, UK.
- M. Snover, B. Dorr, R. Schwartz, L. Micciulla, and J. Makhoul. 2006. A study of translation edit rate with targeted human annotation. In *Proceedings of Association for Machine Translation in the Americas*, pages 223–231.

- M. Snover, N. Madnani, B. J. Dorr, and R. Schwartz. 2009. Fluency, adequacy, or HTER? Exploring different human judgments with a tunable MT metric. In *Proceedings of the Fourth Workshop on Statistical Machine Translation, StatMT '09*, pages 259–268, Stroudsburg, PA, USA. Association for Computational Linguistics.
- I. Temnikova. 2010. Cognitive evaluation approach for a controlled language post-editing experiment. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta.
- I. Temnikova. 2011. Establishing implementation priorities in aiding writers of controlled crisis management texts. In *Proceedings of the International Conference Recent Advances in Natural Language Processing 2011*, pages 654–659, Hissar, Bulgaria, September.
- L. Winerman. 2009. Crisis communication. *Nature*, 457.
- G. K. Zipf. 1949. *Human Behavior and the Principle of Least Effort*. Addison-Wesley (Reading MA).