

Establishing Implementation Priorities in Aiding Writers of Controlled Crisis Management Texts

Irina Temnikova

Research Institute in Information and Language Processing

University of Wolverhampton, UK

irina.temnikova@gmail.com

Abstract

As clarity in the Crisis Management domain is crucial, and there exists an enormous amount of Crisis Management documents, a specific language resource (the Controlled Language for Crisis Management, CLCM) for editing Crisis Management instructions in English has been previously developed. Based on a specially designed controlled language evaluation experiment, we have determined that manual simplification, far from being easy, is an extremely time-consuming process and thus automatization is essential in order to facilitate the writing of clear instructions. This article describes this experiment which also aims to determine which operations should be privileged and are more urgent to be implemented in order to address the most critical issues first.

1 Introduction

Attention paid to the Crisis Management domain has strongly increased in recent years (Schneid and Collins, 2001), due to the urgent need to guarantee safe and efficient management of emergency situations. There exists an enormous amount of already written crisis management documents and new ones are being created with exponentially growing speed. Efficient communication between crisis management teams and local populations is crucial in situations in which there is a very short reaction time (Ogrizek and Guillery, 1999; Winerman, 2009). It is also known that human comprehension under stress is different from the one in normal conditions (Kiwani et al., 1999). For this reason the clarity and conciseness of the information exchanged during emergency situations is crucial. A controlled language (CL) is a very good manual approach for ensuring clarity of crisis management texts. Unfortunately, it

has been previously shown (Goyvaerts, 1996; Huijsen, 1998) that manual editing of texts according to controlled language guidelines is a difficult and highly time-consuming process. Natural Language Processing (NLP) techniques are thus a good way to at least partially automatize and thus improve and speed up manual text simplification. This article introduces a controlled language for text simplification in the crisis management domain and its evaluation in terms of time and cognitive efforts for the human simplifiers and draws conclusions about which operations to implement first, in order to speed up and facilitate manual controlled language-based simplification. The article is structured as follows: Section 2 presents the related work in the crisis management domain and in Controlled Languages aids, Section 3 presents the Controlled Language for Crisis Management (CLCM), Section 4 describes the text simplification evaluation experiment, Section 5 discusses the results of the experiment and Section 6 provides some conclusions and future work.

2 Related Work

Although the large recent activity in crisis management, there are not many NLP works in this domain. Roman (2008) has worked on redundancy identification from personal web blogs on emergency topics. Ireson (2009) has worked in Information Extraction in detection and monitoring of emergency events from open discussion forums. During the project EPIC, managed by the University of Colorado at Boulder and the University of California, some work was done in the extraction of important information relevant to mass emergencies signaled in Twitter (Corvey et al., 2010). In medical crisis management, Chapman et al. (2005) have applied NLP approaches to syndromic surveillance in order to obtain free text classification of chief complaints. As the existing NLP approaches deal mainly with emergencies detec-

tion, this is the first known attempt to process text simplicity of instructions delivered to the general population during mass emergencies. In terms of controlled languages, CLCM is the first controlled language for English for this domain. The existing NLP tools for controlled language simplification are Controlled Language Editors (CLEs) and Controlled Language Checkers (CLCs). The CLEs are used in order to facilitate writing text according to controlled Language specifications, while the CLCs – to check whether an already created text is written in accordance with the controlled language rules. Compagnon LiSe is an example of a very simple CLE that facilitates writing sentences according to the controlled language rules but does not apply any NLP techniques. It has been developed for the French version of CLCM (Renahy et al., 2010). Two other CLEs have been developed for the machine-oriented controlled languages PENG and ACE (Schwitter, 2008; Kuhn, 2009). Mitamura and Nyberg (2001) have developed the KANT controlled language re-writing system which checks compliance with a machine translation-oriented controlled language. This article aims to assist with the initial design of an NLP-based CLCM Editing Aid, in order to facilitate text simplification in the crisis management domain.

3 The Controlled Language for Crisis Management

The Controlled Language for Crisis Management has been adapted to English on the basis of a controlled language for French (Renahy, 2009) in the context of MESSAGE Project¹. As a result of MESSAGE, four controlled languages for different European languages have been developed (French, Spanish, Polish and English), together with two prototypes for Modern Greek and Bulgarian (Temnikova and Margova, 2009). CLCM has been developed on the basis of a collected corpus of crisis management documents, amounting to over 2.5 million words and collected from the web.

The existing version of CLCM applies only to instructions for the general public (GP), as these are considered to be the documents which most need simplification, as their audience members are not trained specialists. Although covering differ-

¹<http://message-project.univ-fcomte.fr/> Accessed 12 May 2011.

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.

ent topics, these instructions have high document structure and language similarity, which allowed the development of a specific CL tailored to them. The role of CLCM is two-fold: on one hand to provide rules for the efficient simplification of existing crisis management documents, and on the other hand - to provide rules for writing new crisis management documents. CLCM is easily transferable to other domains' documents, containing instructions.

The CLCM features thirty pages of over eighty simplification rules, which address different text aspects, starting from general text structure and ending with punctuation, as well as different document elements (titles, conditions, instructions, lists). Below are provided examples of some of the existing rule types and in Figure 1 - a screenshot of a rule taken from the CLCM guidelines:

- **General:** If there are distinguished situations:
 - Identify the specific situations.
 - Divide the blocks of instructions regarding the specific situations into subsections.
 - Write first the most specific situation.
 - Write the next more general situation.
 - End with the most general situation.
- **Formatting:** Separate with a new line each block of instructions.
- **Lexical:** Use only words defined in the dictionary.
- **Syntactic:** Avoid passive voice
- **Punctuation:** Avoid any punctuation signs at the end of the titles.

As can be seen from Figure 1, each rule has a reference number which is formed by: the type of document (“In” = “instructions”), a number of rules are document type-specific; the type of rule (“L” = “lexical”); and a standard number. Also below each rule is shown an example of how text should not look according to this rule (stroke

through) and how it should look instead. Sometimes below these illustrative examples less important information is provided, as for example an explanation of why the rule is necessary. Previously two experiments have been conducted in order to evaluate CLCM. One experiment evaluated the impact of the controlled language on human translation (Temnikova and Orasan, 2009), while the second experiment evaluated the impact of the controlled language simplification on machine translation (Temnikova and Orasan, 2009; Temnikova, 2010). These experiments have shown that although CLCM was written for human readers, it had a significant impact on and improved the results of both human and machine translation.

4 Description of the Evaluation Experiment

The aim of the experiment carried out was to evaluate the quality of the CLCM guidelines. The experiment consisted in asking six linguists - English advanced and native speakers with a computational linguistics background, to read carefully and familiarise themselves with the CLCM simplification guidelines and to simplify manually four texts of a total of two thousand words according to the simplification rules in these guidelines. In order to direct the participants and simplify their task in remembering over eighty rules, an assisting leaflet was provided. The leaflet contained the thirty most important rules to be consulted during simplification. The rules in the leaflet were classified into three natural language generation-like groups:

1. Rules for discourse structure organisation at text level;
2. Rules for discourse structure organisation at paragraph level;
3. Concrete linguistic realization rules.

The experiment was performed in two stages distributed over two days to avoid the impact of the factor of tiredness. The four texts were taken from the previously collected Crisis Management Corpus and represent instructions for the general population in different emergency situations: precautions to be taken after a flood, instructions how to clean chemicals from clothing, actions to be taken after volcanic eruptions. Time is measured during the first time reading guidelines and during the

manual simplification of each text. Table 1 shows the text lengths per text for each day of the experiment calculated in words.

Day	Text	Words	Chars
Day 1	Text 1	166	900
	Text 2	833	5018
Total	Day 1	999	5918
Day 2	Text 3	271	1562
	Text 4	728	4486
Total	Day 2	999	6048
Total	Day 1 and 2	1998	11966

Table 1: Lengths of texts used for the CLCM guidelines evaluation.

As can be seen from the Table 1, the first two columns show which texts were presented to the participants each day while the last two columns provide the text lengths in words and in characters. A text complexity analysis of the four original texts was run, by examining the main text complexity features according to literature. The results of this analysis are provided in Table 2.

Text	SL	WL	SM	LD	WS
Text 1	12.69	4.24	3.64	0.50	11.48
Text 2	16.76	4.94	4.67	0.42	8.08
Text 3	14.83	4.68	5.62	0.39	7.95
Text 4	14.74	5.03	4.87	0.44	8.86

Table 2: Text Complexity analysis of the four texts.

In Table 2, the first column contains the text reference number, while the following five columns - the text complexity features they have been analysed for, namely:

- SL - average sentence length, measured in number of words;
- WL - average word length, measured in number of letters;
- SM - percentage of subordinating markers;
- LD - lexical diversity, measured as types/tokens ratio;
- WS - average number of word senses per word

In order to do this text complexity analysis, the texts were pre-processed using Connexor parser ²

²www.connexor.eu Accessed 12 May 2011.

and WordNet (Fellbaum, 1998) was used for calculating the average number of senses per word.

At the end of the second day, the participants were asked to fill in a questionnaire asking details regarding the work they had done in the previous two days. The questionnaire collected data in three parts - Part 1 was asking for personal comments in the form of free text, Part 2 was providing a list of rules to be evaluated in terms of how difficult they are to be applied, while Part 3 was suggesting a list of implementations to be rated. The personal comments in the first part of the questionnaire were requested by the following question: “Could you think of what was most difficult for you while simplifying?”. The rules, provided in the second part were those thirty most important rules, contained in the assisting leaflet. The participants were asked to write next to each rule whether it was easy or difficult to apply and mark those ones which, if automated, would speed up their work. The suggested implementations in Part 3 were offering preliminary easy-to-implement operations which would result in highlighting different text elements. The text elements to be highlighted ranged from single words to whole paragraphs and were CLCM-specific. The participants were asked to give scores to these operations, according to the following ranking: 1 - Implementing this operation will not help at all; 2 - Implementing this operation will help to a certain extent; 3 - Implementing this operation will help very much.

Additionally, most of the participants provided useful feedback on the design of the experiment and the future implementation, thanks to their NLP background.

5 Results of the Experiment

The analysis of the results of the experiment provided useful information about the internal process of manual simplification of texts according to the CLCM rules and shows that text simplification is not a trivial task, even when precise guidelines are provided. The analysis of the times and speed for reading the guidelines show that the average time for reading the guidelines for the first time was between 30 to 45 minutes. The speeds for simplifying manually the texts are given in Table 3.

The first column of Table 3 indicates the participant, while the next four columns - the four texts. The values in the table are given in charac-

Subject	Text 1	Text 2	Text 3	Text 4
Sub1	21.9	69.7	71.0	203.9
Sub2	75.0	358.4	173.6	263.9
Sub3	30.0	47.8	78.1	149.5
Sub4	30.0	83.6	97.6	121.2
Sub5	18.7	52.3	33.9	48.2
Sub6	33.3	72.7	67.9	115.0
mean 6	34.8	114.1	86.5	150.3
st.dev. 6	18.7	110.0	43.0	68.7
mean 5	26.8	65.2	69.7	127.6
st.dev. 5	5.5	13.3	20.7	50.6

Table 3: Manual simplifying speed per subject and per text, measured in characters per minute.

ters/minute, in order to take into consideration the different length of texts. Obviously, although the value of Subject 2 is the outlier of the sample. can be clearly seen that the speed ranges between 18.7 to 358.4 characters per minute, depending on the text complexity and the subject. I.e. the speed difference is 20 times. Even if a clear learning effect is visible from the data, it is still clear that simplifying text takes a large amount of time. The table also provides the mean and standard deviation values with and without the outlier. Row “mean 6” provides the mean values of all six participants, together with the outlier, while row “mean 5” - only of the five participants, excluding the outlier. In a similar way, row “st.dev. 6” provides the standard deviation values of all six participants, while row “st.dev. 5” the standard deviations without the outlier. The standard deviation values excluding Subject 2 decreases significantly, for example, for Text 2 from 120.4 to 14.9.

Another demonstration of the fact that manual simplification is not a trivial task is the results provided by the questionnaire the language experts were asked to fill in at the end of the second day. The aim of this questionnaire was to determine which rules were most difficult to apply and which simplification operations would need to be automatized. As mentioned before, the questionnaire was composed of three parts: Part 1 containing personal comments in free text, Part 2 containing the list of the thirty main rules to be evaluated as “easy”/“difficult” to be manually applied and whether to be implemented or not and Part 3 containing a list of suggested implementations to be ranked as “will not help at all”/“will help to a certain extent”/“will help very much”. Surprisingly, most of the subjects have very similar personal comments in answering the question, posed in Part 1 “Could you think of what was most dif-

difficult for you while simplifying?”. The answers were put in a common table and were given a mark “1” if a Subject has mentioned it in the free comments and “0” if not. The marks were added and averages were obtained. In this way, the top four results were, ordered from the one with the highest score to the one with the lowest score:

- Avoiding negatives/Re-phrasing negative phrases.
- Remembering to remove pronouns/Avoiding pronouns.
- Being mindful of word difficulty/Replacing technical terms.
- Re-organizing and re-grouping the content of the original.

In Part 2, the marks were given different weights in the following way:

- “no answer” or “easy” = 0
- “simplify” = 1
- “moderate” = 1.5
- “difficult” = 2
- “difficult” and “simplify” = 3
- “very difficult” = 4
- “very difficult” and “simplify” = 5

As the participants have used different combinations of marks, the conclusive marks have been given weights in correspondence with the apparent ranks of the different marks. The results of Questionnaire Part 2 were added and averages were obtained. The top twelve results are shown in the Table 4.

Table 4 has two columns, the first one indicating the CLCM rule, while the second one - the average score, obtained after adding the scores, provided by the participants. “N” stands for “noun”, while “V” stands for “verb”. The top ten results of Part 3 are given in Table 5. The results were again added and averages were ordered from the highest to the lowest one.

Table 5 is composed, like Table 4, by two columns. The first column contains the suggested implementations, while the second column – the average scores, obtained by adding the participant

Rule	Score
Try to avoid negative forms	3
Replace passive with active voice	2.17
Avoid any pronouns (person., poss., demonstr.)	2
Avoid ambiguous words	1.83
Replace idiomatic expressions with literal ones	1.83
Replace techn. terms with common synonyms	1.83
Order instructions in logic. and chronol. order	1.67
If 2+ complem. determ. the same N, repeat the N	1.67
Write only one action per line	1.67
If a prep./adj. refers to 2+ N, repeat the prep/adj.	1.67
Use standard word order	1.67
Place conditions before instructions	1.5

Table 4: Questionnaire Part 2 results.

Rule	Score
Highlighting the ambiguous lexical terms	2.67
Highlighting the phrasal verbs	2.5
Highlighting the separate thematic situations	2.5
Highlighting the negative phrases	2.33
Highlighting the ambiguous syntact. expressions	2.33
Highlighting the technical terms	2.33
Highlighting the beginning of instructions	2.17
Highlighting the beginning of conditions	2.17
Highlighting the beginning of explanations	2.17
Highlighting the acronyms and abbreviations	2.17

Table 5: Questionnaire Part 3 results.

scores and dividing them per number of participants. The top results confirm that the rules found more difficult to apply manually by participants are those which are tackling cognitively hard to process linguistic phenomena (negation, passive, ambiguity). This makes an NLP application a good solution to the aforementioned problems.

6 Conclusions and Future Work

The analysis of the results of the experiment shows that human simplifiers employ too much time in simplifying even short texts and thus simplifying is not a trivial task. More particularly, the results collected from the questionnaire show that the simplifiers mostly agreed on the set of difficult rules and on the set of suggested implementations. Future work will include cognitively analysing rules’ formulations before proceeding with any NLP implementation. For example the rule “Avoid negative forms” may be difficult to apply, as which negative forms to avoid are not concretely defined in the guidelines. Otherwise NLP techniques may be applied in a way to perform negative forms recognition and suggestion of alternative positive forms. While some of the suggested implementations could be solved by an

appropriate training, others, such as “highlight in the text the phrasal verbs in case the main verb and the preposition are split up” cannot and would help to be implemented. On the basis of the conclusions drawn from this very useful experiment, future work will be to apply some of the suggested implementations and to find automatic solutions to the highest ranked manual rules as first steps towards a high-level NLP-based Controlled Language Editing Aid. As “avoiding negatives” was listed as first choice in Part 1 and Part 2 and also had one of the highest scores in Part 3, we choose it as the most urgent issue to be solved and possibly implemented. Negation implementation would include constructing patterns for recognizing negation to avoid in emergency instructions, based on the collected corpus and building a grammar to help supplying the user with positive alternatives to negated phrases. Another candidate for implementation is, of course, “Highlighting the ambiguous lexical terms”, which has emerged as the suggested implementation with the highest ranking score (2.67). Future work would also include testing whether more appropriate training of human simplifiers would change the rules considered difficult to apply.

References

- Chapman, W.W., et al. (2005) *Classifying free-text triage chief complaints into syndromic categories with natural language processing*. *Artif Intell Med*. 2005 Jan; 33(1):31-40.
- Corvey, W. J., Vieweg, S., Rood, T. and Palmer, M. (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 (Los Angeles, California, June 2010), 2324.
- Fellbaum, C. (1998) *WordNet: an Electronic Lexical Database* MIT Press.
- Goyvaerts P. (1996) *Controlled English, curse or blessing? A users perspective*. Proceedings of CLAW 1996.
- Huijsen, W.O. (1998) *Controlled Language An Introduction*. Proceedings of the Second International Workshop on Controlled Language Applications. Pittsburgh, Pennsylvania.
- Ireson, N. (2009) *Local Community Situation Awareness During an Emergency*. Proceedings of the IEEE International Conference on Digital Ecosystems and Technologies (IEEE-DEST 2009).
- Kiwan, D., Ahmed, A. and Pollitt, A. (1999) *The effects of text comprehension and performance in examinations*. Proceedings of BPS London Conference, December, 1999.
- Kuhn, T. (2009) *Controlled English for Knowledge Representation*. Ph.D. Thesis. University of Zurich, Switzerland.
- Mitamura, T. and Nyberg, E. (2001) *Automatic rewriting for controlled language translation*. Proceedings of the NLPRS-2001 Workshop on Automatic Paraphrasing: Theories and Applications. Tokyo, Japan. Pages 1-12.
- Ogrizek, M. and Guillery, J-M. (1999) *Communicating in crisis*. Transaction Publishers.
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J. (1985) *A comprehensive grammar of the English language*. Harlow: Longman. Pp. 1779.
- Renahy J. (2009) *Controlled Languages: a Scientific Popularization through the Example of the Controlled Language “LiSe”*. ISMTCL Proceedings, International Review Bulag, pp 215-222.
- Renahy, J. 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*. International Forum on Quality and Safety of Health Care. 20-23 April 2010, The Nice Acropolis, Nice, France.
- Roman, J.H., et al. (2008) *Reducing information overload in emergencies by detecting themes in Web content*. Proceedings of the 5th International ISCRAM Conference 2008;101-6.
- Schneid, T.D. and Collins, L. (2001) *Disaster management and preparedness*. Lewis Publishers.
- Schwitler, R. (2008) *A Controlled Natural Language for the Semantic Web*. *Journal of Intelligent Systems*, 17, pp.125-141.
- Temnikova, I. (2010) *A Cognitive Evaluation Approach for a Controlled Language Post-Editing Experiment*. Proceedings of the International Conference “Language Resources and Evaluation” (LREC2010), Valletta, Malta.
- Temnikova, I. and Margova, R. (2009) *Towards a Controlled Language in Crisis Management: The Case of Bulgarian*. Proceedings of the International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages (ISMTCL), Besancon, France, July 1-3, 2009.
- Temnikova, I. and Orasan, C. (2009) *Post-editing Experiments with MT for a Controlled Language*. Proceedings of the International Symposium on Data and Sense Mining, Machine Translation and Controlled Languages (ISMTCL), Besancon, France, July 1-3, 2009.
- Winerman, L. (2009) *Crisis Communication*. *Nature*, vol. 457, p 376.