

# Multimodal Summarization for people with cognitive disabilities in reading, linguistic and verbal comprehension

**Naushad UzZaman**

Computer Science Department  
University of Rochester  
naushad@cs.rochester.edu

**Jeffrey P. Bigham**

Computer Science Department  
University of Rochester  
jbigham@cs.rochester.edu

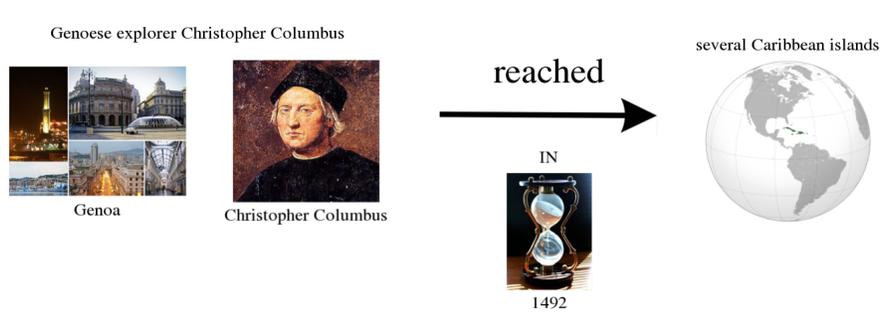
**James F. Allen**

Computer Science Department  
University of Rochester  
james@cs.rochester.edu

## ABSTRACT

Cognitive disabilities are the least understood and the least discussed type of disability among web developers [1]. Some work has explained how developers should create web contents that could be accessible to people with cognitive disabilities [1-4]. While these guidelines are definitely a valuable first step toward making the web more accessible to people with disabilities, the task of creating accessible content is increasingly falling to content creators who are not programmers and are unaware of accessibility guidelines.

Our goal is to use automatic approaches to display text in a way that is easier for people with cognitive disabilities to read. In this paper we introduce the idea of automatically illustrating *complex sentences* as multimodal summaries (MMS) combining pictures, structure and simplified text. By including text and structure in addition to pictures, multimodal summaries provide additional clues of what happened, who did it, to whom and how - to people who may have difficulty reading or looking to skim quickly.



**Figure 1: Multimodal summary of the sentence, “In 1492, Genoese explorer Christopher Columbus, under contract to the Spanish crown, reached several Caribbean islands, making first contact with the indigenous people.”**

Some researchers [5-9] approached the problem of automatic illustration to assist children and their work is mainly in children’s storybook domain. This prior work was not explored in the context of helping people with cognitive disabilities or applicable to complex sentences. We include both pictures and text in our diagrams. In this way, we can handle cases in which we lack a good picture and address cases that are hard to illustrate. Presenting pictures and text together can also improve both the understanding and memorability of concepts. According to dual code theory [10], text and pictures result in two different kinds of conceptual representations. These representations may allow independent access to information and hence benefit retention. Picture and text repeat important information, and may have similar beneficial effects on memory as explicit repetitions [11, 12]. Processing the information twice, once as text and once as a picture, may facilitate comprehension and memory. So our decision for inclusion of text with pictures is backed by theories that supports that it helps people for better understanding and memorizing.

We conducted initial experiments to compare our MMS diagram against diagrams that include only illustration. We showed these diagrams to users in Amazon Mechanical Turk and asked them to explain these diagrams in text. As expected, our experiment showed people understand more from MMS diagram than diagrams with only illustrations. We focus on main

event of the sentence and its related entities, hence summarizing the content of the sentence and our final structure has the structure similar to subject verb (event) object (SVO), which basically simplifies a complex sentence.

Bohman and Anderson [3] explain different functional cognitive disabilities, among which one of them is *reading, linguistic and verbal comprehension*. Our MMS representation can be useful for this particular functional disability, but could also be used for other disabilities, since it captures most of the principles of cognitive disability accessibilities mentioned in [2, 3] e.g. simplicity – with our simple sentence structure; consistent – by using the same structure; clear – with subject, verb (event), object, preposition representation and presenting only one event and its related entities; multi-modal – with illustration and text; and attention focusing – with illustration and structure.

Our immediate goal is to better adapt the MMS diagram representation to be appropriate for those with cognitive impairments, and eventually to conduct user evaluations that explore how automatically simplified text can improve reading for this group.

#### REFERENCE:

- [1] P. Bohman. (2004, Retrieved: September 2010). Cognitive disabilities part 1: we still know too little and we do even less. . Available: [http://www.webaim.org/techniques/articles/cognitive\\_too\\_little/](http://www.webaim.org/techniques/articles/cognitive_too_little/)
- [2] C. Rowland. (2004, Retrieved: September 2010). Cognitive disabilities part 2: conceptualizing design considerations. . Available: <http://www.webaim.org/techniques/articles/conceptualize/>
- [3] P. R. Bohman and S. Anderson, "A conceptual framework for accessibility tools to benefit users with cognitive disabilities," presented at the Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A), Chiba, Japan, 2005.
- [4] B. Caldwell, *et al.* (2008, Retrieved: September 2010). Web Content Accessibility Guidelines (WCAG) 2.0. Available: <http://www.w3.org/TR/WCAG20/>
- [5] K. Barnard, *et al.*, "Matching words and pictures.," *Machine Learning Research*, vol. 3, pp. 1107–1135, 2003.
- [6] A. B. Goldberg, *et al.*, "Easy as ABC? Facilitating pictorial communication via semantically enhanced layout.," in *Twelfth Conference on Computational Natural Language Learning (CoNLL 2008)*, 2008.
- [7] D. Joshi, *et al.*, "The story picturing engine—a system for automatic text illustration.," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 2(1), 2006.
- [8] R. Mihalcea and B. Leong, "Toward communicating simple sentences using pictorial representations," presented at the Association of Machine Translation in the Americas., 2006.
- [9] J. Zhu, *et al.*, "A text-to-picture synthesis system for augmenting communication.," in *The Integrated Intelligence Track of the Twenty-Second AAAI Conference on Artificial Intelligence (AAAI 2007)*, 2007.
- [10] A. Paivio, "Mental representations: A dual coding approach," *New York: Oxford University Press.*, 1986.
- [11] A. M. Glenberg, "Component-levels theory of the effects of spacing of repetitions on recall and recognition.," *Memory and Cognition*, vol. 7, pp. 95-112, 1979.
- [12] R. G. Greene, "Spacing effects in memory: Evidence for a two-process account.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 15, pp. 371-377, 1989.