

## **PAT workbench: Annotation and Evaluation of Text and Pictures in Multimodal Instructions**

*Ielka van der Sluis* (i.f.van.der.sluis@rug.nl), *Lennart Kloppenburg* (lennartkloppen-  
burg@hotmail.com) and *Gisela Redeker* (g.redeker@rug.nl), CLCG, University of Groningen.

We present the PAT workbench, a tool to investigate the design of multimodal instructions (MIs), i.e., instructions that contain both text and pictures. The benefit of including pictures in information presentation has been established, but the characteristics of those pictures, their textual counterparts and the relation(s) between them have not been researched in a systematic manner. Our main goal is to systematically investigate how text and pictures are best combined in MIs in terms of effectiveness in their context of use. For now, we consider primarily MIs in health communication. Outcomes of our work will aid (semi-)automatic annotation, evaluation and generation of MIs as well as the formulation of authoring guidelines on how to combine text and pictures effectively according to judgments and performance of readers and users, dependent on, for instance, the function of the MI (e.g., to learn a task or to perform a task only once).

The PAT workbench [1] is a web application with a template-based design, written in PHP and facilitated through a MySQL database. The relational database structure allows an efficient design of connections between concepts such as users, user groups, group assignments, documents and annotations. The PAT workbench was built to facilitate the annotation, storage and retrieval of MIs based on a validated coding scheme [3] with currently 42 categories that describe instructions in terms of textual features, pictorial elements, and relations between text and pictures. Illocutionary properties of text and pictures are defined so that they can be aligned in terms of actions and control information. Correspondences between text and pictures are defined in terms of number of steps, indices and layout. Yearly about 190 annotated MIs are collected and added to the MI corpus by master students in Communication and Information Sciences at the University of Groningen. These MIs concern first-aid tasks like applying a band-aid, removing ticks, or reanimating a person.

The current version of the PAT workbench includes the following functionalities: a detailed MI search system with filtering options, a viewing panel to inspect MIs, a function to upload MIs to the workbench, an annotation panel to annotate MIs according to the PAT coding scheme, an assignment panel to create and manage collaborations with other annotators, a revision history for annotators, a function to add annotated MIs to the MI corpus, a MI browser to select MIs for viewing, a web-friendly manual for annotating MIs and documentation at the levels of installation/use, code, and database. Users of the PAT workbench can add (sets of) MIs in PDF format to their own corpus. The PAT workbench facilitates collaborative annotation and inter-annotator agreement calculation. Users can invite other users to annotate a subset of their owned MIs using the collaboration panel. Agreement between two annotations of each MI is calculated per subcategory to help the owner of the MI to double check and improve the annotation, which the PAT workbench administrator then adds to the MI corpus as the gold standard annotation of the MI. A PHP-based search engine is used to retrieve MIs in the gold standard annotated corpus based on a keyword search that can be augmented with meta-information.

Future functionalities we envisage include a solution to parsing problems where OCR fails due to complex layout with interwoven textual and pictorial elements. Fully parsed MIs will allow for automatic tagging of lexical and grammatical features and will considerably reduce the need for manual annotation of linguistic features. Subsequently, we plan to develop and implement an evaluation algorithm that scores features of MIs in terms of predicting readers' and users' ratings of the quality of MIs. These ratings will be based on crowdsourcing experiments in which readers are asked to rate MIs as well as on empirical studies in which users perform the actions instructed in MIs. Implementation of automatic methods for annotation and evaluation will allow us to annotate larger amounts of MIs and thus in the future extend and generalise the workbench to process other types of MIs (e.g., indoor navigation, cooking recipes, construction manuals) and possibly instruction videos. Other extensions include the annotation of individual pictures, features of pictures and textual units (to allow description of individual picture-text relations cf.[2]), a more extensive administrator panel (i.e., that allows to add new and disable existing annotation categories and their values) and an English interface.

## References

- [1] Van der Sluis, I., Kloppenburg, L. & Redeker, G. (2016). PAT Workbench: Annotation and Retrieval of Multimodal Instructions. To be published at the CLARIN-D workshop on Language Technology Resources and Tools for Digital Humanities Workshop (LT4DH'16) at COLING 2016, Osaka, Japan. CLARIN-D.
- [2] Van der Sluis, I., Leito, S., & Redeker, G. (2016). Text-Picture Relations in Cooking Instructions. In H. Bunt (Ed.), Proceedings of LREC 2016, Tenth International Conference on Language Resources and Evaluation: Proceedings of the Twelfth Joint ISO - ACL SIGSEM Workshop on Interoperable Semantic Annotation (ISA-12). (Vol. 16, pp. 22-27). Portoroz, Slovenia.
- [3] Van Dijk, J., Van der Sluis, I. & Redeker, G. (2016). Annotation of Text and Pictures in Health Care Instructions. Presented at TABU'2016, 3 June, 2016.