

## University of Groningen

### The PAT annotation model for Multimodal Instructions

van der Sluis, Ielka; Redeker, Gisela

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2019

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

van der Sluis, I., & Redeker, G. (2019). *The PAT annotation model for Multimodal Instructions*. Abstract from 6th European and 9th Nordic Symposium on Multimodal Communication, Leuven, Belgium.  
[http://mmsym.org/?page\\_id=379](http://mmsym.org/?page_id=379)

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# The PAT Annotation Model for Multimodal Instructions

Ielka van der Sluis & Gisela Redeker  
Center for Language and Cognition Groningen  
{i.f.van.der.sluis, g.redeker}@rug.nl

**Keywords:** annotation, model, application, corpus, instructions

We present the development, implementation and application of the PAT annotation model, which describes instructional documents that consist of pictures and text (PAT). In document design research (Schriver, 1997) the combination of pictures and text has been noted but not so much investigated in terms of function and content. Although useful starting points have been provided (Bateman, 2014), we are unaware of a standard methodology to describe and evaluate picture-text relations. This leaves document designers without specific guidelines, while readers and users may experience difficulties in effectively processing multimodal content due to mismatches with their expectations and cognitive capacities.

## Development

As the possibilities to describe multimodal documents are infinite, we advocate conducting corpus studies and reader/user studies in tandem to determine the relevance of annotation categories and their values. Based on preliminary analyses of a corpus with first-aid instructions (currently 297), we conducted multiple reader and user studies to investigate the effectiveness of particular design features. With the results of these studies we developed and fine-tuned the specification of 51 types of functional and content relations between textual elements and between text and pictures.

## Implementation

Our corpus annotation is supported by the PAT Workbench, a custom-made online tool that provides a flexible environment to systematically investigate multimodal designs by facilitating storage, annotation, retrieval and evaluation of documents (See, Figures 1, 2 and 3 and <https://cosmo.service.rug.nl/patworkbench/login/>). The workbench includes 'smart' OCR for uploaded documents, user-defined specification of annotation categories, and a tool for creating gold standard annotations based on multiple annotations.

## Application

As a worked example, we will present the results of a comparative study that involved the application of the PAT annotation model to a subcorpus of 46 first-aid instructions from two editions of Het Oranje Kruis Boekje 2011 and 2016. Het Oranje Kruis is a Dutch organisation that provides learning materials for first-aid certification trainings. A comparison of multimodal instructions (117 pictures and 9416 words in total) for 23 tasks in both editions of Het Oranje Kruis Boekje allows us to conclude that the two editions are similar in terms of the visualised actions, but differ in terms of: text content (preambles, alternative actions, control information); the type of shot used in the pictures (close-up/medium shot versus long shot); and text-picture relations in terms of layout (alignment versus proximity).

## Future work

The PAT project (<http://www.rug.nl/let/pat>) will deliver theoretical results in terms of empirically validated models for effective multimodal presentations and authoring guidelines for multimodal documents. Future work will include more comprehensive textual analysis and finer-grained analysis of the pictorial materials, coverage of a greater number and a wider variety of instructions, (semi-)automatic annotation, more empirical evaluation, and (semi-)automatic generation of potentially effective text-picture combinations for multimodal instructions.

## Bibliography

- Bateman, J. (2014). Using multimodal corpora for empirical research. In *The Routledge handbook of multimodal analysis*, pp. 238–252. Routledge, London.
- Schriver, K. (1997). *Dynamics in document design: Creating text for readers*. Wiley, New York.

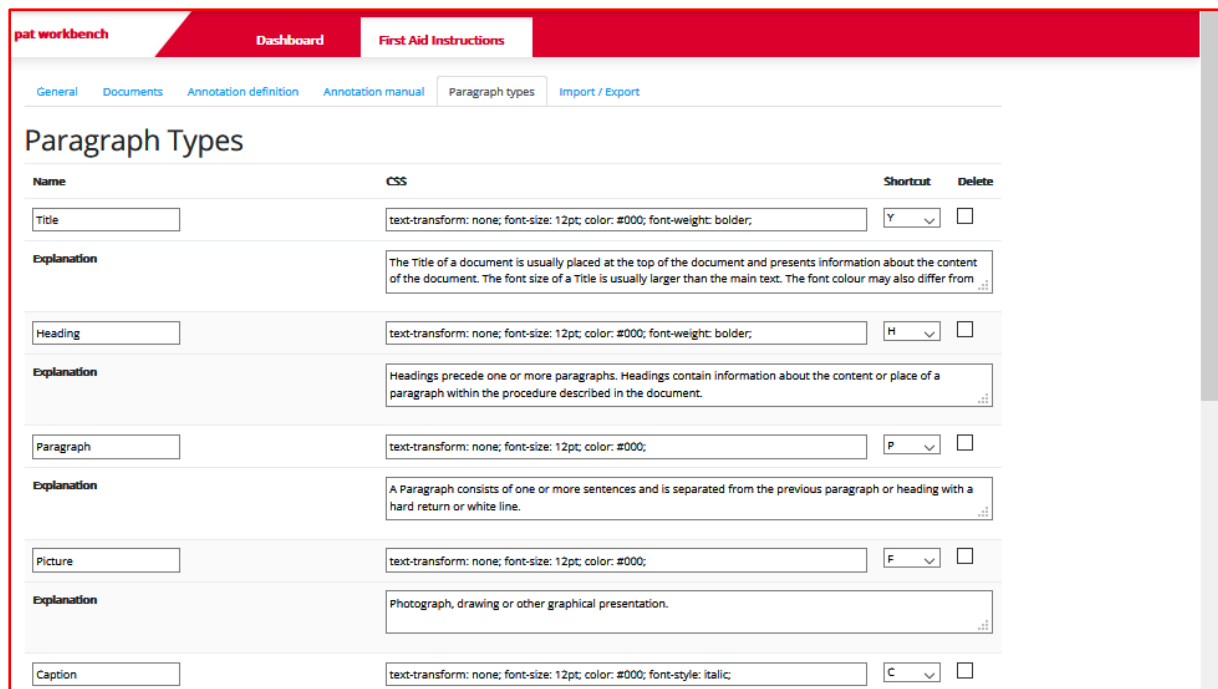


Figure 1: Screenshot of description of Paragraph Types in the PAT workbench.

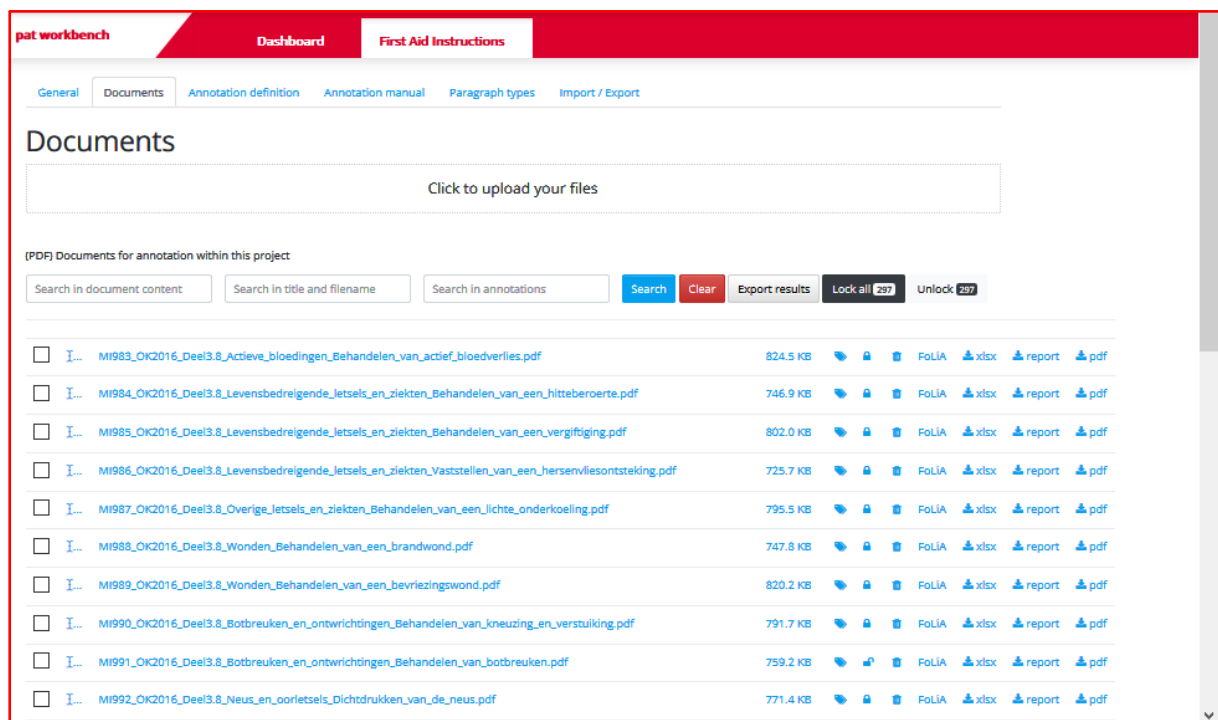


Figure 2: Screenshot of corpus documents in the PAT workbench.

