PAT Workbench: Annotation and Evaluation of Text and Pictures in Multimodal Instructions

Ielka van der Sluis
Center for Language
and Cognition
University of Groningen
i.f.van.der.sluis
@rug.nl

Lennart Kloppenburg
Center for Language
and Cognition
University of Groningen
l.kloppenburg
@rug.nl

Gisela Redeker
Center for Language and
Cognition
University of Groningen
g.redeker
@rug.nl

Abstract

This paper presents 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 and of their textual counterparts and the relation(s) between them have not been researched in a systematic manner. We present the PAT Workbench, a tool to store, annotate and retrieve MIs based on a validated coding scheme with currently 42 categories that describe instructions in terms of textual features, pictorial elements, and relations between text and pictures. We describe how the PAT Workbench facilitates collaborative annotation and inter-annotator agreement calculation. Future work on the tool includes expanding its functionality and usability by (i) making the MI annotation scheme dynamic for adding relevant features based on empirical evaluations of the MIs, (ii) implementing algorithms for automatic tagging of MI features, and (iii) implementing automatic MI evaluation algorithms based on results obtained via e.g. crowdsourced assessments of MIs.

1 Introduction

This paper presents a tool to facilitate a rigorous empirically oriented study on the design and use of multimodal instructions (MIs), i.e., instructions that contain both text and pictures. MIs are ubiquitous in all walks of modern life (medicine, electronics, flatpack furniture, recipes, etc.). In general, it has been established that including pictures in information presentation is beneficial for readers and users (e.g., Glenberg and Roberts, 1999; Kjelldahl, 1992; Mayer, 2009; Schriver, 1997). But what are the characteristics of the pictures, texts and the relation(s) between them in these presentations? To illustrate the abundant variety of presentational aspects in MIs and the apparent lack of authoring guidelines, Figures 1 and 2 present two MIs for operating an automated external defibrillator (AED). Already at first sight the two instructions differ in numerous ways, e.g. the number of steps, the number of actions to carry out per step, the type of pictures, layout in terms of columns and rows, the amount of text (per step), the occurrence and type of arrows used in pictures, the use of indices, labels and references.

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. Our corpus-based studies will allow investigation of the breadth of instructional design, while existing studies on MIs (Houts et al., 2006; Katz et al., 2006) generally focus on human processing of particular instructions. 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 and dependent on, for instance, the function of the MI (e.g., to learn a task or to perform a task only once).

Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities (LT4DH), pages 131–139, Osaka, Japan, December 11-17 2016.

Various tools exist for picture annotation (e.g., Cusano et al., 2003, Russel et al., 2008), text annotation (e.g., Erdman et al., 2000; Ogren, 2006; Stenetorp et al., 2012) and video annotation (e.g., Brugman & Russel, 2004; Do et al., 2016; Kipp, 2001). However, to our knowledge only the UAM tool (O'Donnell 2008) supports the annotation of both text and pictures. UAM supports semi-automatic tagging of text and allows parts of pictures to be selected and labelled, but not to our knowledge relational annotations.

In this paper we present the PAT Workbench for annotation, storage and retrieval of MIs. The workbench supports manual annotation based on a coding scheme that describes MIs in terms of the factors that potentially influence the effectiveness of MIs. The 42 categories of the coding scheme were inspired by studies on human processing of multimodal presentations (e.g., Arts et al., 2011; Dupont & Bestgen, 2006; Florax & Ploetzner, 2010; Heiser & Tversky 2006; Van Hooijdonk & Krahmer, 2008; Maes & Noordman, 2004; Morrow et al., 2005). Like UAM (O'Donnell, 2008) and Anvil (Kipp, 2001), the workbench includes inter-annotator agreement calculation and a method to resolve any differences between annotations and obtain a gold standard annotation.

In Section 2 we present the PAT Workbench, In Section 3, we present an overview of our current MI corpus with 194 annotated first-aid instructions. In Section 4, we conclude with a discussion of future work on the PAT Workbench.

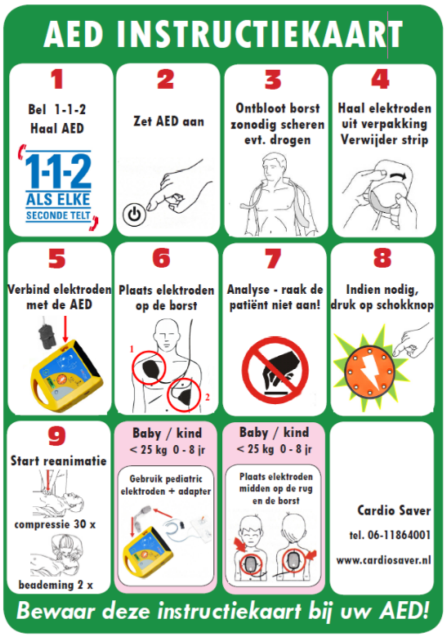
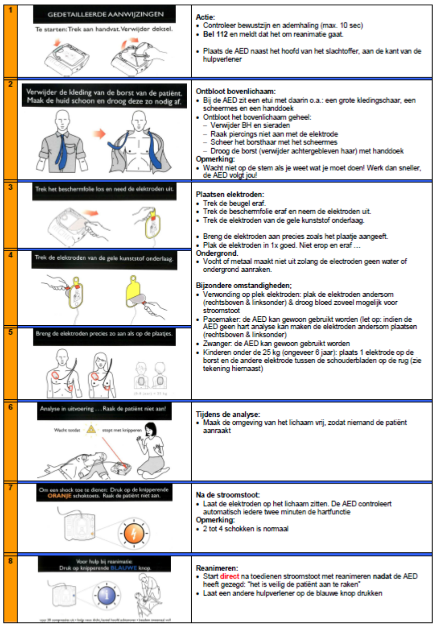


Figure 1: AED MIT by cardiosaver.nl.


1
GEDETAILLEERDE AANWIJZINGEN
Te starten: Trek aan handvat. Verwijder deksel.
2
Verwijder de kleding van de borst van de patiënt. Maak de huid schoon en droog deze zo nodig af.
3
Trek het beschermfolie los en need de elektroden uit.
4
Trek de elektroden van de gele kunststof onderlaag.
5
Breng de elektroden precies zo aan als op de plaatjes.
6
Analyse in uitvoering … Raak de patiënt niet aan!
Wacht totdat
stopt met knipperen
7
Om een shock toe te dienen: Druk op de knipperende ORANJE schoktoets. Raak de patiënt niet aan.
8
Voor hulp bij reanimatie:
Druk op knipperende BLAUWE knop.

Actie:
- Controleer bewustzijn en ademhaling (max. 10 sec)
- **Bel 112** en meldt dat het om reanimatie gaat.
- Plaats de AED naast het hoofd van het slachtoffer, aan de kant van de hulpverlener

Ontbloot bovenlichaam:
- Bij de AED zit een etui met daarin o.a.: een grote kledingschaar, een scheermes en een handdoek
- Ontbloot het bovenlichaam geheel:
  - Verwijder BH en sieraden
  - Raak piercings niet aan met de elektrode
  - Scheer het borsthaar met het scheermes
  - Droog de borst (verwijder achtergebleven haar) met handdoek

Opmerking:
- Wacht niet op de stem als je weet wat je moet doen! Werk dan sneller, de AED volgt jou!

Plaatsen elektroden:
- Trek de beugel eraf.
- Trek de beschermfolie eraf en neem de elektroden uit.
- Trek de elektroden van de gele kunststof onderlaag.
- Breng de elektroden aan precies zoals het plaatje aangeeft.
- Plak de elektroden in 1x goed. Niet erop en eraf …

Ondergrond.
- Vocht of metaal maakt niet uit zolang de electroden geen water of ondergrond aanraken.

Bijzondere omstandigheden;
- Verwonding op plek elektroden: plak de elektroden andersom (rechtsboven & linksonder) & droog bloed zoveel mogelijk voor stroomstoot
- Pacemaker: de AED kan gewoon gebruikt worden (let op: indien de AED geen hart analyse kan maken de elektroden andersom plaatsen (rechtsboven & linksonder)
- Zwanger: de AED kan gewoon gebruikt worden
- Kinderen onder de 25 kg (ongeveer 6 jaar): plaats 1 elektrode op de borst en de andere elektrode tussen de schouderbladen op de rug (zie tekening hiernaast)

Tijdens de analyse:
- Maak de omgeving van het lichaam vrij, zodat niemand de patiënt aanraakt

Na de stroomstoot:
- Laat de elektroden op het lichaam zitten. De AED controleert automatisch iedere twee minuten de hartfunctie

Opmerking:
- 2 tot 4 schokken is normaal

Reanimeren:
- Start direct na toedienen stroomstoot met reanimeren **nadat** de AED heeft gezegd: "het is veilig de patiënt aan te raken"
- Laat een andere hulpverlener op de blauwe knop drukken

Figure 2: AED MIT by ProCardio.nl.

2 The PAT Workbench

2.1 System description

The PAT Workbench is an online tool that was built to facilitate the annotation, storage and retrieval of MIs collected by master students in Communication and Information Sciences at the University of Groningen. Each year, about 200 new annotated MIs are added to the corpus. These MIs concern first-aid tasks like applying a band-aid, removing ticks, or reanimating a person. The PAT Workbench is a web application written in PHP using the CodeIgniter Framework[1]. The design of the website features a Bootstrap[2] template called Dark Admin[3]. Data creation and manipulation is facilitated through a MySQL database. This relational database structure allows an efficient design of connections between concepts such as users, user groups, group assignments, documents and annotations. Documents are

[1] https://www.codeigniter.com/

[2] http://getbootstrap.com/

[3] http://www.prepbootstrap.com/bootstrap-theme/dark-admin

not stored in this database, but in a separate directory structure where each document is linked by its identification code to its entry in the database (which contains the metadata).

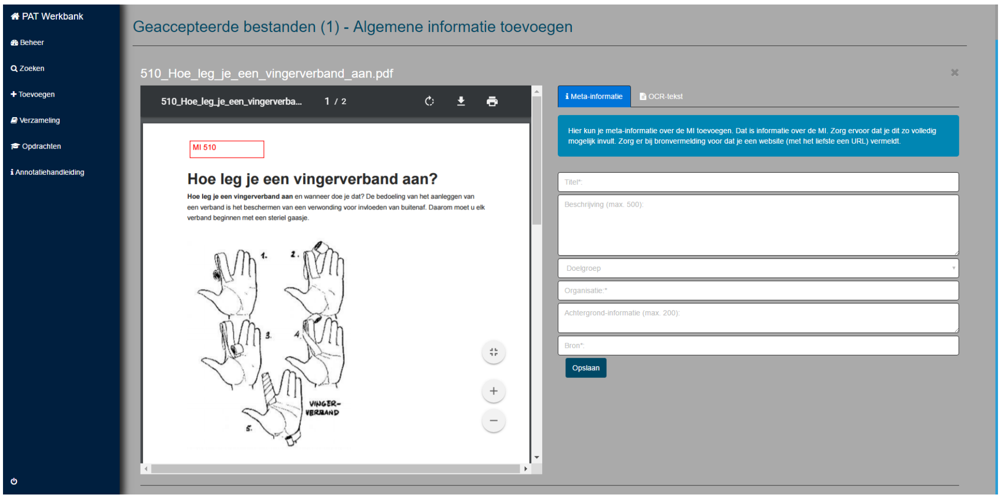
The current version of the PAT Workbench has a menu structure with five main topics: Search, Add, Assignment, Manual, Collection and Manual, and includes the following functionalities:

- Detailed MI search system with filtering options
- Viewing panel to inspect MIs
- Function to upload MIs to the workbench
- Annotation panel to annotate MIs according to the PAT coding scheme
- Assignment panel to create and manage collaborations with other annotators
- Revision history for annotators
- Function to add annotated MIs to the MI corpus
- MI browser to select MIs for viewing
- Web-friendly manual for annotating MIs
- Documentation at the levels of installation/use, code, and database.

The interface of the current version of PAT is in Dutch; an English version is in preparation.

2.2 Adding MIs

Users of the PAT Workbench can add (sets of) MIs in PDF format to their own corpus via the menu item ‘Add’. A user becomes the owner of MIs that he/she adds from the collection of free MIs in the MI corpus and of the new MIs that he/she uploads to the workbench. When uploading MIs, the user is prompted to select a set of documents. For each document the user needs to specify some metadata about the MI it contains (e.g., title, description, target audience, background information, source), aided by a simultaneously offered view of the MI (see Figure 3). Based on the metadata, the system checks if the MI is not uploaded already. To make sure that no duplicates exist, the user can also manually check the uploaded MIs based on a keyword search that retrieves MIs using the metadata; the retrieved MIs can be browsed and inspected with a simple document viewer. In addition, uploaded PDF documents are converted to plain text using OCR (ABBYY[4]).



Geaccepteerde bestanden (1) - Algemene informatie toevoegen

510_Hoe_leg_je_een_vingerverband_aan.pdf

MI 510

Hoe leg je een vingerverband aan?

Hoe leg je een vingerverband aan en wanneer doe je dat? De bedoeling van het aanleggen van een verband is het beschermen van een verwonding voor invloeden van buitenaf. Daarom moet u elk verband beginnen met een steriel gaasje.

Hier kun je meta-informatie over de MI toevoegen. Dat is informatie over de MI. Zorg ervoor dat je dit zo volledig mogelijk invult. Zorg er bij bronvermelding voor dat je een website (met het liefste een URL) vermeldt.

Titel*:

Beschrijving (max. 500):

Doelgroep

Organisatie:*

Achtergrond-informatie (max. 200):

Bron*:

Figure 3: Panel for uploading MIs and adding metadata.

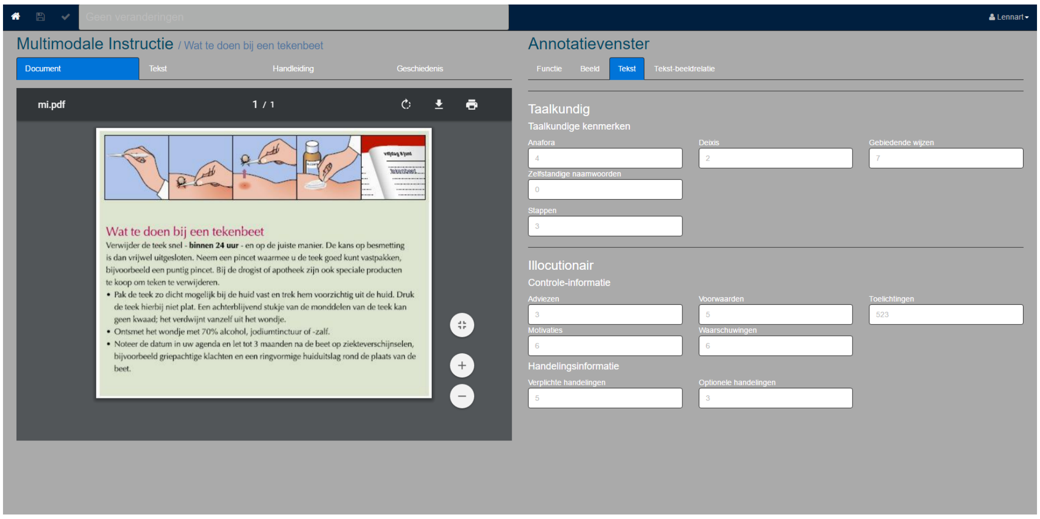
2.3 Annotation panel

Figure 4 presents the annotation panel. The grey part of the screen offers a two-column view of an uploaded MI. The left side offers either a view of the document as it was uploaded to the PAT Workbench, a view of OCR output, the annotation manual, or the annotation history of the document.

[4] https://www.abbyy.com/en-eu/en/cloud-ocr-sdk/

The right side presents a tabulated view of the main annotation categories: (1) function, (2) text, (3) pictures and (4) text-picture relations.

The annotation scheme used in the PAT Workbench is the improved version of a scheme that was used by 13 annotators who annotated a corpus of 227 health care instructions (Van Dijk et al. 2016). The corpus described in Section 3 was annotated with the improved scheme. In Tables 1 to 4 below the scheme is presented in terms of the four main categories, including their subcategories with values. Note that the 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. Future work will also consider relations in terms of implements, agents and actions. Other extensions will include annotations of MIs per picture and per textual unit. Currently all annotations concern the MI as a whole.



Multimodale Instructie / Wat te doen bij een tekenbeet

Wat te doen bij een tekenbeet

Verwijder de teek snel - **binnen 24 uur** - en op de juiste manier. De kans op besmetting is dan vrijwel uitgesloten. Neem een pincet waarmee u de teek goed kunt vastpakken, bijvoorbeeld een puntig pincet. Bij de drogist of apotheek zijn ook speciale producten te koop om teken te verwijderen.

- Pak de teek zo dicht mogelijk bij de huid vast en trek hem voorzichtig uit de huid. Druk de teek hierbij niet plat. Een achterblijvend stukje van de monddelen van de teek kan geen kwaad; het verdwijnt vanzelf uit het wondje.
- Ontsmet het wondje met 70% alcohol, jodiumtinctuur of -zalf.
- Noteer de datum in uw agenda en let tot 3 maanden na de beet op ziekteverschijnselen, bijvoorbeeld griepachtige klachten en een ringvormige huiduitslag rond de plaats van de beet.

Annotatievenster

Taalkundig

Taalkundige kenmerken

Anafora: 4
Deixis: 2
Gebiedende wijzen: 7
Zelfstandige naamwoorden: 0
Stappen: 3

Illocutionair

Controle-informatie

Adviezen: 3
Voorwaarden: 5
Toelichtingen: 523
Motivaties: 6
Waarschuwingen: 6

Handelingsinformatie

Verplichte handelingen: 5
Optionele handelingen: 3

Figure 4: Panel for annotating MIs.

Main category	**Subcategory**	**Value**
1. Identification MI		
	A) Title	Text
	B) Description	Text
	C) Target group	*Adults, children, medics*
	D) Organisation	Text
	E) Background Knowledge	Text
	F) Source	Text
2. Function MI		
	A) Reading-to-do	0 or 1
	B) Reading-to-learn	0 or 1
	C) Reading-to-decide during instruction	0 or 1
	D) Reading-to-decide after instruction	0 or 1

Table 1: Functional aspects of MIs.

Main category	**Subcategory**	**Value**
3. Text: Length		
	A) Number of steps	*Count*
	B) Number of sentences	Count
	C) Word count	Count
4. Text: Linguistic properties		
	A) Nouns	Count

	B) Anaphora	*Count*
	C) Deixis	*Count*
	D) Imperatives	*Count*
	E) Form of address	Formal, informal, avoiding
5. Text: Illocutionary properties		
	A) Compulsory actions	*Count*
	B) Optional actions	*Count*
	C) Control information, warnings, conditionals, motivations, advisements, explanations, other	*Count*
	D) Extra information	
	Notes	Text

Table 2: Textual aspects of MIs.

Main category	**Subcategory**	**Value**
6. Picture: Visual properties		
	A) Number of pictures	*Count*
	B) Picture type	*Drawing, photo, other*
	C) Human appearance	*Count*
	D) Text in picture	*Count*
	E) Pictograms	*Count*
	F) Arrows	Count
	G) Indication	*Numbers, letters, none*
	H) Clock-time indications	*Count*
7. Picture: Illocutionary properties		
	A) Compulsory actions	*Count*
	B) Optional actions	*Count*
	C) Result of action	*Final, partial result*
	D) Function	*Localisation, identification*
	E) Control information, warning, explanation, other	*Count*
	F) Extra information	0 or 1
	Notes	

Table 3: Pictorial aspects of MIs.

Main category	**Subcategory**	**Value**
9. Text-picture relation		
	A) Correspondence in steps	*0 or 1*
	B) Text-picture layout	*Numbers, lines, blocks, titles, other*
	C) Textual reference to pictures	*Count*
	Notes	Text

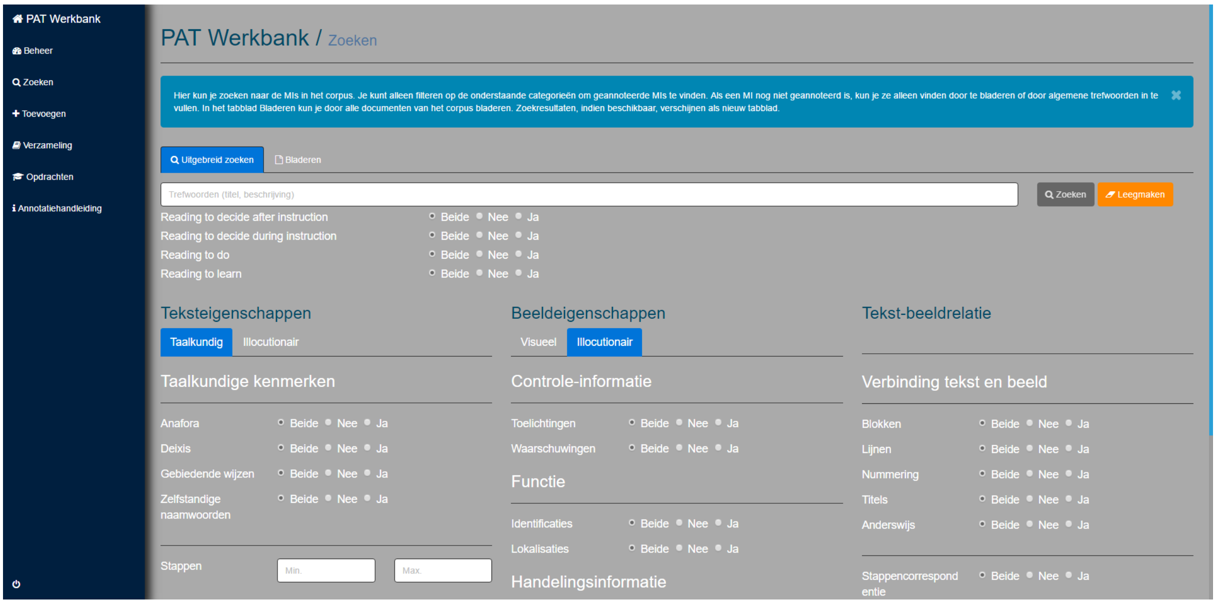
Table 4: Text-picture relations in MIs.

2.4 Collaborative annotation

The PAT Workbench administrator can formulate a collaboration assignment for a particular group of users, and the members of the group can then invite other members to annotate a subset of the MIs they own using the collaboration panel. The collaboration panel includes a progress indicator for the annotations. 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 administrator then adds to the MI corpus as the gold standard annotation of the MI.

2.5 MI Retrieval

MIs of which the annotation is agreed on by at least two annotators and which have thus been added to the gold standard annotated corpus, can be retrieved via the main menu item ‘Search’. The link points to a panel with two tabs (Figure 5), one to perform an extensive search and one to browse the database. The latter option offers a tiled view of all MIs in the corpus that includes a picture of the MIs and the metadata (i.e. title, description, uploader, owner, target audience, source, web address and date). The extensive search offers a keyword search, which can be augmented with the values of the subcategories presented in italics in Tables 1 to 4. The output of the PhP-based search engine is displayed with the simple MI browser that is also used to browse the corpus, which allows selection and display of individual MIs.



Hier kun je zoeken naar de MIs in het corpus. Je kunt alleen filteren op de onderstaande categorieën om geannoteerde MIs te vinden. Als een MI nog niet geannoteerd is, kun je ze alleen vinden door te bladeren of door algemene trefwoorden in te vullen. In het tabblad Bladeren kun je door alle documenten van het corpus bladeren. Zoekresultaten, indien beschikbaar, verschijnen als nieuw tabblad.

PAT Werkbank
Beheer
Zoeken
Toevoegen
Verzameling
Opdrachten
Annotatiehandleiding
PAT Werkbank / Zoeken
Uitgebreid zoeken
Bladeren
Trefwoorden (titel, beschrijving)
Zoeken
Leegmaken
Reading to decide after instruction ○ Beide ● Nee ● Ja
Reading to decide during instruction ○ Beide ● Nee ● Ja
Reading to do ○ Beide ● Nee ● Ja
Reading to learn ○ Beide ● Nee ● Ja
Teksteigenschappen
Taalkundig
Illocutionair
Taalkundige kenmerken
Anafora ○ Beide ● Nee ● Ja
Deixis ○ Beide ● Nee ● Ja
Gebiedende wijzen ○ Beide ● Nee ● Ja
Zelfstandige naamwoorden ○ Beide ● Nee ● Ja
Stappen
Min.
Max.
Beeldeigenschappen
Visueel
Illocutionair
Controle-informatie
Toelichtingen ○ Beide ● Nee ● Ja
Waarschuwingen ○ Beide ● Nee ● Ja
Functie
Identificaties ○ Beide ● Nee ● Ja
Lokalisaties ○ Beide ● Nee ● Ja
Handelingsinformatie
Tekst-beeldrelatie
Verbinding tekst en beeld
Blokken ○ Beide ● Nee ● Ja
Lijnen ○ Beide ● Nee ● Ja
Nummering ○ Beide ● Nee ● Ja
Titels ○ Beide ● Nee ● Ja
Anderswijs ○ Beide ● Nee ● Ja
Stappencorrespondentie ○ Beide ● Nee ● Ja

Figure 5: Search panel of the PAT Workbench.

2.6 Manuals and documentation

The PAT Workbench includes an illustrated annotation manual (Van Dijk et al. 2016) which is available in PDF format via the main menu as well as via the annotation panel. In addition, the system itself is documented at the installation/use level, code level, and database level.

2.7 Administrator Interface to manage MIs and Users

The administrator interface includes three different administration panels:

- On the User panel, the administrator may add or remove users of three types (student, researcher, administrator) or edit details of users (i.e. name, student number, email address and type).
- On the Group panel, the administrator can create a user group for a particular time span. For each group the administrator can add and remove users and formulate collaboration assignments.
- On the Corpus panel, the administrator can edit and remove MIs. The administrator may also unlock MIs from their owners to allow other users to change or extend the annotations. In all cases ‘remove’ means ‘flag as inactive or invisible’, as no user or MI is actually removed from the database to allow restore and repair.

2.8 Evaluation and Usability

Currently, an extensive evaluation of the system is in progress. A preliminary task-based user study on the administrator interface included an observation of the administrator (the first author of this paper) performing tasks while thinking aloud. She had not seen the administrator interface before, but was closely involved in the design and implementation of the underlying functionality. After she had performed the tasks assigned to her, she filled out two standard usability questionnaires, SUS (Brook, 1986) and CSUQ (Lewis, 1995), and was interviewed by the experimenter who asked her to reflect on her task performance and the way she filled out the questionnaires.

Observations of the task-based performance brought to light various flaws that are currently being remedied. Most prominent were issues with system feedback (e.g., system confirmations, visual cues, mouse-over information) and uniformity (e.g., potential user actions are indicated with buttons as well as icons). Overall this (admittedly biased) evaluation was very positive, but also identified a number of minor points where the interface could be improved.

After the implementation of those improvements, an analysis of the PAT Workbench is currently being conducted by an independent expert who will perform a functional analysis, a heuristic inspection of the whole application, identification of the main tasks and a cognitive walkthrough of these main tasks. A more extensive user study with potential users of the PAT Workbench will be conducted inspecting the usability of the PAT Workbench in terms of its main tasks using a think-aloud protocol, usability questionnaires and interviews.

3 MI corpus

The annotated MI corpus currently contains 192 MIs, 166 of which are designed for adult users, 9 for children and 17 for medics. The MI functions are distributed as follows: reading-to-do (120), reading-to-learn (85), reading-to-decide-during-instruction (17), and reading-to-decide-after-instruction (6); some MIs have multiple functions. Thirteen annotators familiar with the annotation scheme annotated, double-checked and agreed on the annotations of 106 MIs, subsequently the annotators coded 86 new MIs.

The annotation of linguistic properties shows that the mean number of steps in the MIs is 5.78, varying from 1 to 14. The mean number of sentences in the MIs is 10.37, varying from 1 to 41. The mean number of imperatives in the MIs is 6.93, ranging from 0 (24 MIs) to 25. The mean number of anaphora used is 1.01, where 90 MIs do not include any anaphora and the maximum number of anaphora used in an MI is 11.

Table 5 presents counts of illocutionary properties of text and pictures in the 192 MIs. By far the most text segments and pictures refer to compulsory actions. Conditionals, motivations, and advisements cannot be reliably identified in pictures and were thus only annotated for the texts.

	Text	**Picture**
Compulsory actions	1266	767
Optional actions	190	45
Control information	110	26
Warnings	171	4
Conditionals	113	NA
Motivations	98	NA
Advisements	244	NA
Explanations	141	166

Table 5: Total number of illocutionary properties of text and pictures (NA = ‘Not annotated’).

The annotation of visual properties shows that the mean number of pictures included in the MIs is 4.25 ranging from 1 to 23. In 66 MIs these pictures are photographs, 125 MIs use drawings and one MI uses a combination of photographs and drawings. The total of 816 pictures include 95 instances of text in picture and 8 instances of pictograms, 258 arrows and 12 time indications. Human body parts are included in 764 pictures. Pictures are used for identification purposes in 535 cases and for localisation purposes in 28 cases. While actions are usually referred to by action verbs in the text, pictures often present not the process, but the results (439 pictures) or partial results of actions (302 pictures).

Annotation of text-picture relations reveals a correspondence in 71 MIs in terms of the number of steps in the text and the number of pictures. In 27 MIs text and pictures are connected through numbering, 28 MIs use lines, 32 MIs use blocks and 13 MIs use titles. In 21 MIs textual references to pictures are included ranging from 1 (8 MIs) to 14 (1 MI).

4 Future Work

The current version of the PAT Workbench will be improved based on results from the planned expert and user evaluations. Additional functionalities we envisage include:

- Recognition of the document structure where OCR fails due to e.g., columns and pictures in the MIs. Fully parsed MIs will allow for automatic tagging of lexical and grammatical features and will considerably reduce the need for manual annotation of linguistic text features as presented in Table 2. Currently, the Alpino parser (Van Noord, 2006) is used for simple NLP tasks like counting words and sentences.
- Development and implementation of an algorithm to annotate illocutionary aspects of text and pictures in MIs.
- Annotation of individual pictures and textual units, which allow description of individual picture-text relations. These relations will include actions or illocutionary aspects that are described and visualised. As implemented in the UAM tool for pictures (O'Donnell, 2008), identification and annotation of features of pictures is envisioned.
- Development and implementation of 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.
- A more extensive administrator panel will be implemented that allows a dynamic annotation scheme, i.e. to add new and disable existing annotation categories and their values. Obviously, these changes need to be made in tandem with the search options and the method used to calculate inter-annotator agreement.
- An English interface will be provided to improve the system's accessibility.

Implementation of automatic methods for annotation and evaluation will allow us to annotate larger amounts of MIs and thus to extend and generalise the workbench to process other types of MIs (e.g., indoor navigation, cooking recipes, construction manuals) and possibly instruction videos in the future.

References

Arts, A., Maes, A., Noordman, L. & Jansen C. (2011). Overspecification in written instruction. *Linguistics*, *49* (3), 555-574.

Brooke, J. (1996). SUS: A quick and dirty usability scale. In Jordan P. W., Thomas B., Weerdmeester B. A., McClelland I. L. (Eds.), *Usability Evaluation in Industry* (pp. 189-194). London: Taylor & Francis. (Also see http://www.cee.hw.ac.uk/~ph/sus.html).

Brugman, H. & Russel, A. (2004): Annotating multi-media / multi-modal resources with ELAN. In *Proceedings of the fourth International Conference on Language Resources and Evaluation (LREC)*. Lisbon: Portugal, 2065–2068.

Cavicchio, F., & Poesio, M. (2009). Multimodal corpora annotation: Validation methods to assess coding scheme reliability. In M. Kipp, J.-C. Martin, P. Paggio, and D. Heylen (Eds.), *Multimodal Corpora* LNAI 5509, Berlin: Springer, 109-121.

Cusano, C., Ciocca, G., & Schettini, R. (2003). Image annotation using SVM. In *Electronic Imaging 2004* (pp. 330-338). International Society for Optics and Photonics.

Van Dijk, J., Van der Sluis, I. and Redeker, G. (2016). Annotation of Text and Pictures in Health Care Instructions. Presented at TABU'2016, 3 June, 2016.

Tuan Do, Nikhil Krishnaswamy, and James Pustejovsky. (2016). "ECAT: Event Capture Annotation Tool." Interoperability for Semantic Annotation (ISA) 2016, at the 10th Edition of the Language Resources and Evaluation Conference, Portoroz, Slovenia, May 28, 2016.

Dupont, V., & Bestgen, Y. (2006). Learning from technical documents: The role of intermodal referring expressions. *Human Factors, 48* (2), 257-264.

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity. *Learning and Instruction, 20* (3), 216-224.

Glenberg, A., & Robertson, D. (1999). Indexical understanding of instructions. *Discourse Processes*, *28* (1), 1-26.

Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, *30* (3), 581-592.

Van Hooijdonk, C. & Krahmer, E. (2008). Information modalities for procedural instructions: The influence of text, static and dynamic visuals on learning and executing RSI exercises. *IEEE Transactions on Professional Communication, 51* (1), 50-62.

Houts, P., Doak, C., Doak, L., & Loscalzo, M. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling, 61* (2),173-190.

Katz, M., Kripalani, S., & Weiss, B. (2006). Use of pictorial aids in medication instructions: a review of the literature. *American Journal of Health-System Pharmacy*, *63* (23), 2391-2398.

Kipp, M. (2001) Anvil - A Generic Annotation Tool for Multimodal Dialogue. *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)*, pp. 1367-1370.

Kjelldahl, L. (1992) *Multimedia: Systems, Interaction and Applications.* Berlin: Springer.

Lewis, J. R. (1995) IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction, 7* (1), 57-78.

Maes, A., Arts, A. & Noordman, L. (2004). Reference management in instructive discourse. *Discourse Processes, 37*(2), 117-144.

Mayer, R. (2009). *Multimedia learning*. Cambridge University Press.

Morrow, D., Weiner, M., Young, J., Steinley, D., Deer, M., & Murray, M. (2005). Improving medication knowledge among older adults with heart failure: a patient-centered approach to instruction design. *The Gerontologist*, *45* (4), 545-552.

Van Noord, G. (2006, April). At last parsing is now operational. In *TALN06. Verbum Ex Machina. Actes de la 13e conference sur le traitement automatique des langues naturelles* (pp. 20-42).

Ogren, P. V. (2006, June). Knowtator: a protégé plug-in for annotated corpus construction. In *Proceedings of the 2006 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: Companion volume: Demonstrations* (pp. 273-275).

O'Donnell, M. (2008). Demonstration of the UAM CorpusTool for text and image annotation. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Demo session* (pp. 13-16).

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. (2008). LabelMe: a database and web-based tool for image annotation. *International Journal of Computer Vision*, *77*(1-3), 157-173.

Schriver K. (1997). *Dynamics in Document Design: Creating Texts for Readers*. John Wiley & Sons: New York, NY.

Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., & Tsujii, J. I. (2012). BRAT: a web-based tool for NLP-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 102-107).