

# WebAnno-MM: EXMARaLDA meets WebAnno

Steffen Remus\*   Hanna Hedeland   Anne Ferger   Kristin Bührig   Chris Biemann\*

**\*Language Technology Group**

Department of Informatics

Universität Hamburg, Germany

{lastname}@informatik.uni-hamburg.de

**Hamburg Centre for Language Corpora (HZSK)**

Universität Hamburg, Germany

{firstname.lastname}@uni-hamburg.de

## Abstract

In this paper, we present WebAnno-MM, an extension of the popular web-based annotation tool WebAnno, which is designed for the linguistic annotation of transcribed spoken data with time-aligned media files. Several new features have been implemented for our current use case: a novel teaching method based on pair-wise manual annotation of transcribed video data and systematic comparison of agreement between students. To enable the annotation of transcribed spoken language data, apart from technical and data model related challenges, WebAnno-MM offers an additional view to data: a (musical) score view for the inspection of parallel utterances, which is relevant for various methodological research questions regarding the analysis of interactions of spoken content.

## 1 Introduction

We present WebAnno-MM<sup>1</sup>, an extension of the popular web-based annotation tool WebAnno<sup>2</sup> (Yimam et al., 2013; Eckart de Castilho et al., 2014), which allows linguistic annotation of transcribed spoken data with time-aligned media files. Within a project aiming at developing innovative teaching methods, pair-wise manual annotation of transcribed video data and systematic comparison of agreement between annotators was chosen as a way to teach students to analyze and reflect *a)* on a authentic classroom communication, and *b)* on the linguistic transcription as a part of this process.

For the project, a set of video recordings were partly transcribed and compiled into a corpus with meta-data on communications and speakers using the EXMARaLDA system (Schmidt and Wörner, 2014), comprising a set of data models, XML transcription and metadata formats and software tools for the creation management and analysis of spoken corpora. The EXMARaLDA system could have been further used to implement the novel teaching method, since it allows for manual annotation of audio and video data and provides methods for visualizing the transcribed data (in HTML format) for further qualitative analysis. However, within the relevant context of university teaching, apart from such requirements, addressing the peculiarities of spoken data, several additional requirements regarding collaborative annotation, user management and data management becomes an increasingly important part of the list of

---

This work is licensed under a Creative Commons Attribution 4.0 International License.  
License details: <http://creativecommons.org/licenses/by/4.0/>

WebAnno-MM is licensed under Apache Version 2.0 License.  
License details: <https://www.apache.org/licenses/LICENSE-2.0>

<sup>1</sup>MM refers to Multi Modal. The application as well as the source can be found at <https://github.com/webanno/webanno-mm>

<sup>2</sup><https://webanno.github.io>

Steffen Remus, Hanna Hedeland, Anne Ferger, Kristin Bührig and Chris Biemann 2019. WebAnno-MM: EXMARaLDA meets WebAnno. *Selected papers from the CLARIN Annual Conference 2018*. Linköping Electronic Conference Proceedings 159: 166–176.

desirable features. The following list illustrates the main necessities for a successful implementation of the project:

- proper handling of spoken data (e.g. speaker and time information must be maintained),
- ease of playback and display of aligned media files must be properly endowed,
- support visualization of transcripts in the required layout, i.e the (musical) score view,
- support complex manual annotation of linguistic data,
- support collaborative (i.e. pair-wise) annotation,
- support the assessment of inter-annotator agreement scores,
- support reliable user management (for student grading).

Furthermore, a web-based user environment is preferred to avoid issues regarding installation, different versions of the software, and data distribution, particularly video content, to the users (students). Another important feature was to use a freely available tool, which allows others to use the teaching method, which was developed within the project, using the same technical set-up for other areas of analysis.

While WebAnno (in its default version) fulfills some of the requirements that are not met by the EXMARaLDA system or any other similar desktop application for transcription and annotation of spoken data, it is primarily designed for annotating sequential data (mostly occurring as written text). Thus, various extensions are required for the tool to interpret and display transcriptions and video data of spoken content, which is aligned by time and thus appears in parallel. Since there are several widely used tools for creating corpora of spoken language, we preferred to rely on an existing interoperable standardized exchange format in order to enable interoperability between the tools with advanced complementary features. For this, we chose the ISO/TEI format, which is the TEI-based ISO standard ‘Transcription of spoken language’ (ISO/TC 37/SC 4, 2016; Schmidt, 2011).

In Section 2, we will further describe the involved components and related work, in Section 3 we will outline WebAnno-MM in more detail. Section 4 describes the novel teaching method and the use of WebAnno-MM within the university teaching context. In Section 5, we report on another emerging use case for the WebAnno-MM based on the ISO/TEI format, and we present some ideas on how to implement further improvements of the extension in order to open and generalize it for additional, more general use case scenarios related to spoken and multimodal data annotation.

## 2 Related work

### 2.1 The EXMARaLDA system

Within the context of our goals, the EXMARaLDA<sup>3</sup> (Schmidt and Wörner, 2014) transcription and annotation tool, including the corresponding data model and XML file format for transcription data, is the most relevant component. The tool was originally developed to support researchers in the field of discourse analysis and research on multilingualism, but has since then been used in various other contexts, e.g. for dialectology, language documentation and even with historical written data. Since spoken and multimodal data inherently displays parallel and overlapping structures, e.g. due to overlapping speech in a conversation or due to gestures accompanying speech, the tool has proven useful in various other contexts where this type of complexity in annotation structure needs to be modeled. EXMARaLDA provides support for common transcription conventions (e.g. GAT, HIAT, CHAT) and can visualize transcription data in various formats and layouts for qualitative analysis.

As shown in Figure 1, the score layout of the interface displays a stretch of speech corresponding to a couple of utterances or intonation phrases, which is well suited for transcription or annotations spanning at most an entire single utterance. Though scrolling through the transcription text is possible, a more comprehensive overview of larger spans of discourse is only available in the static visualizations generated from the transcription data. Another aspect relevant for the current use case is the rather simplistic

---

<sup>3</sup><http://exmaralda.org>

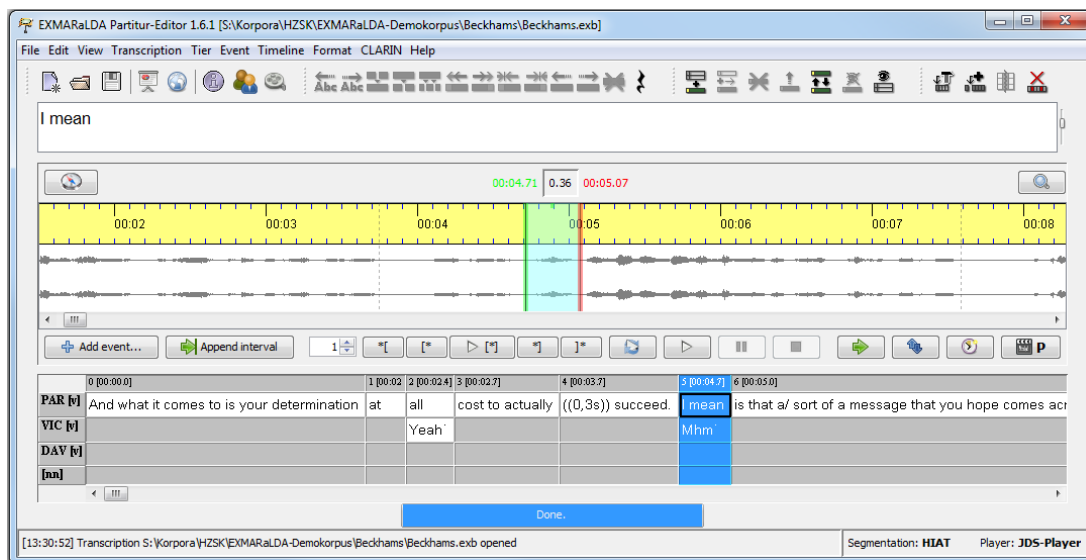


Figure 1: The musical score view of the EXMARaLDA transcription and annotation tool. The transcription is time-aligned with individual tiers for each speaker.

data model of the EXMARaLDA system, which only allows simple span annotations based directly on the transcribed text. On the one hand, this simplicity allows for efficient data processing, but on the other hand, it makes it impossible to model more complex tier dependencies or structured annotations. When annotating phenomena that occur repeatedly and interrelated over a larger span of the discourse, e.g. to analyze how two speakers discuss and arrive at a common understanding of a newly introduced concept, the narrow focus and the simple span annotations make it difficult to perform this task in an efficient and fully expressive manner.

## 2.2 WebAnno: a flexible, web-based annotation platform for CLARIN

In order to augment texts with linguistic annotations, automatic tools are required that support annotators and principals to collectively create, visualize, analyze, and compare annotations. WebAnno is an annotation platform that provides an interactive web interface accessible through standard web browsers while collecting, storing and processing of data takes place on a centralized server. This paradigm allows to perform shared annotation projects where annotators collectively annotate text documents with information that is immediately available on the server for further processing, e.g. for monitoring or curation purposes. Additionally, WebAnno is developed *by* the community *for* the community, and was first implemented in the CLARIN<sup>4</sup> context. Collaborative effort is made to increase the quality of the project and to address the needs of users who are using the tool. WebAnno offers standard means for linguistic analysis, such as span annotations, which are configurable to be locked to, or to be independent of, token or sentence annotations, relational annotations between spans, and chained relation annotations. Figure 2 shows a screenshot of WebAnno during a standard annotation task.

WebAnno is able to read various pre-defined input formats. The UIMA<sup>5</sup> (Ferrucci and Lally, 2004) framework is the foundation of WebAnno's backend. Hence, the input data, that is specified in a particular format and that might already contain prior annotations, are converted into UIMA's internal representation. UIMA stores text information, specifically the text itself and its annotations, in a stand-off fashion in so-called CASs (Common Analysis Structures). The basic elements needed for utilizing the underlying BRAT<sup>6</sup> visualization (Stenetorp et al., 2011) are then *Sentence* and *Token* annotations, which are ideally specified in the input data, and heuristically created otherwise. While *Sentence* annotations directly influence the visual segments (c.f. Figure 3), *Token* annotations are used for maintaining

<sup>4</sup><https://www.clarin.eu/>

<sup>5</sup>Unstructured Information Management Architecture: <https://uima.apache.org/>

<sup>6</sup><http://brat.nlplab.org>

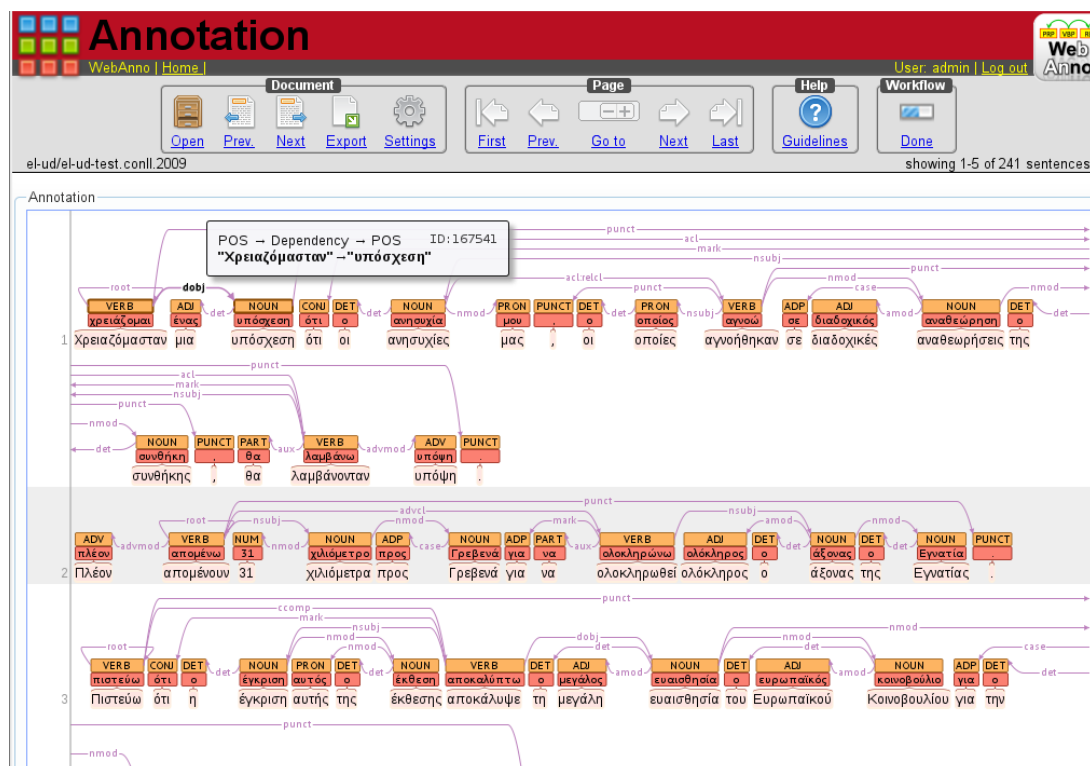


Figure 2: The default WebAnno user view during a standard annotation task. Span annotations are rendered above the annotated span showing their value, and span annotations are connected via relational annotations.

11	Aber diesen Scheiß der da immer gelabert wird!
12	Da sollten sich alle mal wirklich mal Gedanken machen, ob das / ob wir in der Zukunft so weitermachen können!

Figure 3: Sentence annotations define the visual segments for further annotations.

offsets between the visualization and the backend, and are further used for locking other, higher level annotations to specific token offsets. By using sentences and tokens as basic units, any input text data that is used for annotation is defined to be sequential.

For further analysis of and management of the collected annotations, WebAnno has been equipped with a set of assistive instruments, some examples of which include:

- web-based user- and project management,
- curation of annotations made by multiple users,
- built-in inter-annotator agreement measures such as Krippendorff's  $\alpha$ , Cohen's  $\kappa$  and Fleiss'  $\kappa$ , and
- flexible and configurable annotations including extensible tagsets.

All this is available via easy web access for users (annotators), which makes it particularly suitable for research organizations and a perfect fit for the targeted use case in this work. Several extensions have been introduced, one such component is the adaptive learning component introduced by Yimam et al. (2014), where annotations and annotation values are learned by the system during the annotation tasks progress. In this setting, annotators are presented with system generated suggestions that improve by usage.

### 2.3 The ISO/TEI Standard for Transcription of Spoken Language

The ISO standard ISO 24624:2016 Transcription of spoken language is based on Chapter 8, Transcriptions of Speech, of the TEI Guidelines<sup>7</sup> as an effort to create a standardized solution for transcription data (Schmidt, 2011). Previously, TEI has rarely been used to model spoken language, since its flexibility makes various compliant versions of the format equally possible to model even the most basic elements of transcriptions, i.e. speakers' contributions, including information on their type and relative order and the alignment with media files. A number of well-established formats for transcription data exist and are only to a certain extent interoperable, due to their varying degree of complexity. These are formats of widely used transcription tools, which are usually time-based and group information in different tiers for speakers and various annotation layers. As outlined in Schmidt et al. (2017), most common transcription tool formats, including ELAN (Sloetjes, 2014) and Transcriber (Barras et al., 2000), were taken into account during the standardization process and can be modelled and converted to ISO/TEI. By using the standard, common concepts and structural information, such as speaker or time information, are modeled in a uniform way regardless of the tool used to create the data. It also becomes possible to achieve basic interoperability across transcription convention specific variants, since the standard allows for transcription convention specific units (e.g. utterances vs. intonational phrases) and labels, while still using the same elements for shared concepts.

## 3 WebAnno-MM: Adapting WebAnno for multimodal content and spoken data

### 3.1 Transcription, theory and user interfaces

A fundamental difference between linguistic analysis of written and spoken language is that the latter usually requires a preparatory step; the transcription. Most annotations are based not on the conversation or even the recorded signal itself but on its written representation. That the creation of such a representation is not an objective task, but rather highly interpretative and selective, and the analysis thus highly influenced by decisions regarding layout and symbol conventions during the transcription process, was addressed already by Ochs (1979). In particular, different arrangements of the speakers' contributions stress different aspects of conversation. If all contributions are sorted and placed underneath each other, so-called line notation, the conversation might appear more ordered with focus on the transitions, i.e. the turn-taking, than when using a musical score layout. Since the exact onset of a speaker's contributions in relation to other speakers' contributions is visualized accordingly in the score layout, this arrangement of the contributions rather stresses the simultaneity and the overlapping nature of speech.

It is therefore crucial that tools for manual annotation of transcription data respect these theory-laden decisions comprising the various transcription systems in use within various research fields and disciplines. Apart from this requirement on the GUI, the tool also has to handle the increased complexity of "context" inherent to spoken language: While a written text can mostly be considered a single stream of tokens, spoken language features parallel structures through simultaneous speaker contributions or additional non-verbal information or other relevant events in the transcription. In addition to the written representation of spoken language, playback of the aligned original media file is another imperative requirement.

### 3.2 From EXMARaLDA to ISO/TEI

To allow for transcription system specific characteristics, e.g. units of the transcription such as utterances or intonation phrases, the existing conversion from the EXMARaLDA format to the tool-independent ISO/TEI standard is specific to the conventions used for transcription. For our use case, this was the HIAT transcription system as defined for EXMARaLDA by Rehbein et al. (2004). Apart from the generic transcription tier holding verbal information, and non-verbal or non-linguistic information encoded as incidents in ISO/TEI, the HIAT conventions also define the following optional but standardized annotation layers: *a)* akz for accentuation/stress *b)* k for free comments *c)* sup for suprasegmental information *d)* pho for phonetic transcription. Though some common features can be represented in a generic

---

<sup>7</sup><http://www.tei-c.org/Guidelines/P5/>

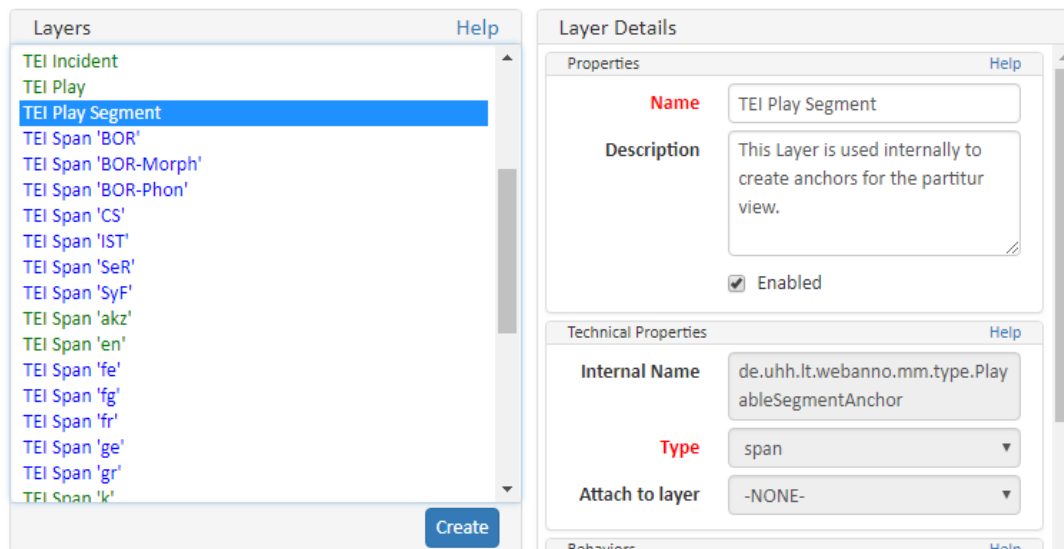


Figure 4: Definition of predefined WebAnno-MM annotation layers (green), and custom annotation layers (blue), which are both resolved and matched during TEI import.

way by the ISO/TEI standard, for reasons described above, several aspects of the representation of conversations must remain transcription convention specific, e.g. the kind of linguistic units defined below the level of speaker contributions.

Furthermore, metadata is stored in different ways for various transcription formats. Within the EXMARaLDA system, metadata on sessions and speakers is stored and managed separated from the transcriptions in the EXMARaLDA Corpus Manager XML format to enhance consistency. The ISO/TEI standard on the other hand, as any TEI variant, can make use of the TEI Header to allow transcription and annotation data and various kinds of metadata to be exported and further processed in one single file, independent of the original format. This approach has also been implemented for the WebAnno-MM extension to be able to retrieve e.g. basic metadata on speakers, such as their age or linguistic knowledge, while annotating.

### 3.3 Parsing ISO/TEI to UIMA CAS

A major challenge is the presentation of time-aligned parallel transcriptions (and their annotations) of multiple speakers in a sequence without disrupting the perception of a conversation, while still keeping the individual segmented utterances of speakers as a whole, in order to allow continuous annotations. For our use case, we restrict WebAnno-MM to the ISO/TEI standard with HIAT conventions as described in Section 3.2. Annotations in the standardized HIAT annotation format (cf. Rehbein et al., 2004) are recognized as pre-defined annotations, where each TEI (HIAT) type is linked with a matching WebAnno-MM annotation layer on import. By default, unrecognized TEI types (non-HIAT) are merged into a single generic WebAnno-MM annotation layer where the TEI type information is preserved as an annotation feature. Since the TEI standard does generally not restrict annotation types, and in order to still allow the import of HIAT unrelated TEI annotations as new WebAnno-MM layers, which might be required for future projects or research purposes, we adapt the importer, such that any TEI annotation type<sup>8</sup> which occurs in the transcription file and for which a matching WebAnno-MM layer (including certain annotation features) must have been created manually before the actual import, is linked to the particular matching annotation layer. Figure 4 shows the generated ISO/TEI layers in addition to WebAnno's existing annotation layers (e.g. 'Dependency' or 'Lemma'). While 'TEI Incident' or 'TEI Play (Segment)' are necessary layers corresponding to generic components of the ISO/TEI standard, and will thus be generated for all ISO/TEI data, other layers have been predefined according to the HIAT transcription

<sup>8</sup>Note that currently only time-based span annotations are supported. Word-based span annotations will be implemented in the future.



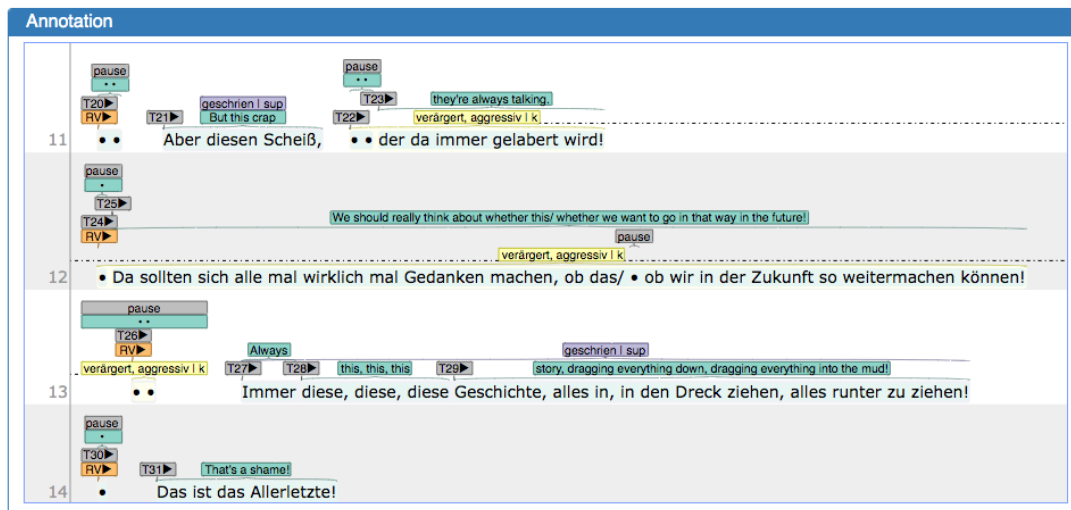


Figure 5: Screenshot of WebAnno-MM’s annotation view. Note the additional time marker annotations (with ▶ as suffix), which are used for synchronizing WebAnno-MM’s score view.

conventions used to create the transcripts (akz, en, k and sup, cf. Rehbein et al., 2004). Additionally, it is also possible to add custom annotation layers describing project specific annotation, which might be independent of the HIAT standard and can be freely defined in the ISO/TEI. The creation of these new layers follow standard naming policies for automatic matching of TEI span types to the actual annotation layer. The definition must be finished prior the import of data, otherwise the system ignores unknown TEI span types. During the import, the ISO/TEI XML content is parsed and utterances of individual speakers are stored in different ‘views’ (sub-spaces of CASes with different textual content). Time alignments are kept as metadata within a specific CAS which we call the *speakercas*. Segments are considered as the basic unit of the transcription in TEI format and can be considered to be roughly equivalent to sentences. As such, segments are directly converted to sentence annotations, which are sequentially aligned in the order of their occurrence in the so-called *annotation view*.

Note, that a segment within an `annotationBlock` XML element is considered non-disruptive. It can safely assumed that ISO/TEI span annotations are within the time limits of the one utterance that is bound to the annotation block. Hence we can map easily cross-segment annotations, but not cross-utterance annotations. This mainly happens for incident annotations that are speaker independent. We neglect such cases in the annotation view and remark that those occurrences are correctly presented in WebAnno-MM’s (musical) score view (explained in detail in the next section).

### 3.4 New GUI features

In order to show utterances and annotations in an established and widely accepted environment for parallel visualization of transcribed audio content, e.g. similar to EXMARaLDA’s score layout of the partitur editor, we adapt the existing html show case demos<sup>9</sup>, and call this view the *score view* of WebAnno-MM henceforth. From a user perspective, a new browser window will open on click of a time marker in the annotation view – time markers are implemented as zero width span annotations starting with the respective speaker abbreviation or time marker id and ending with a play button character (▶, see Figure 5). All such markers are clickable and trigger the focus change in the score view and start or pause the media replay. The design choice of using two parallel and synchronized browser windows was chosen because of maximum flexibility and the assumption that users have ideally two screens available, where each view sides on a different physical screen. Figure 6 schematically illustrates this operation mode.

Figure 7 shows a screenshot of the adjustable (musical) score view. Metadata (if provided in the TEI source file), such as speaker information, can be accessed via clicking on a specific speaker row, and

<sup>9</sup>EXMARaLDA show case demos are available at <http://hdl.handle.net/11022/0000-0000-4F70-A>



Figure 6: Schematic overview of the synchronization feature and the intended application operation mode: Two monitors, where one holds the browser window of the annotation view and the other holds the browser window of the score view, synchronized by clicking on the respective time markers.

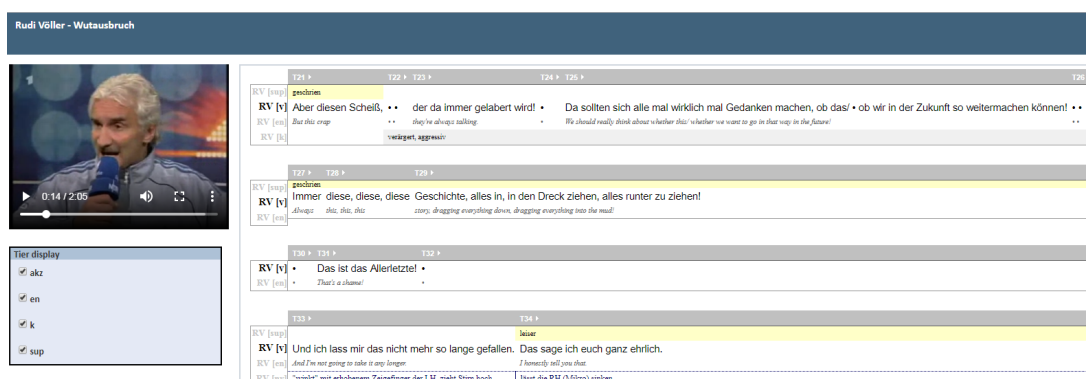


Figure 7: Screenshot of WebAnno-MM's score view, which resembles EXMARaLDA's score view in the partitur editor. WebAnno-MM's annotation view and the score view are synchronized by clicking the respective time markers. The media playback (audio or video) is shown on the top left and reactive to click events on time markers. The parallel tiers are displayed on right which can be activated and deactivated via the checkboxes in the lower left.

further metadata, i.e. relevant to the recording itself can be accessed via clicking the title of the document (cf. Figure 8).

The score view itself is reverse synchronized by clicking on the time marker. The selection of the focus for the annotation view is heuristically determined by using the first marker annotation (independent of the speaker utterance). Upon a synchronizing action, the focus will immediately switch to the other window. Additionally, users are able to select from multiple media formats (if provided during project setup), such as plain audio media, or additionally video media. Also, the score view offers multiple media formats for selection, viewing speaker- or recording related details and a selectable width of the transcribed speaker tracks (rows).

Each segment starts with a marker showing the respective speaker. All markers are clickable and trigger the focus change in the score view and start or pause the media media replay.

For the purpose of managing media per document, a media pane was added to the project settings (cf. Figure 9). Here, media files can be uploaded or web URLs can be specified and linked to uploaded documents. Uploaded media files are hosted within the WebAnno-MM server infrastructure, benefitting from access restrictions through its user management. Additionally, we added support for streaming media files that are accessible in the web by providing a URL instead of a file. Multiple media files can be mapped to multiple documents, which allows user preferred reuse of different media locations of multiple document recordings.



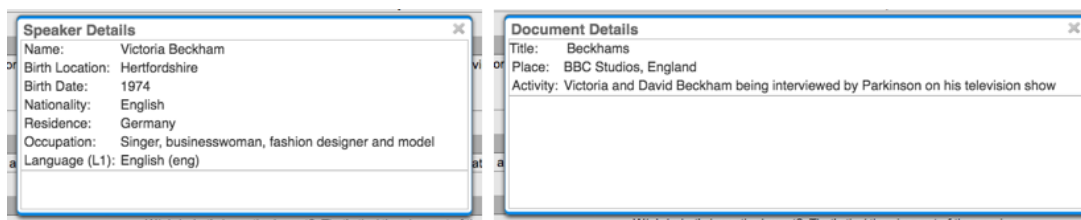


Figure 8: Speaker details (left) and recording details (right), provided by the score view and initially specified in the ISO/TEI source file.

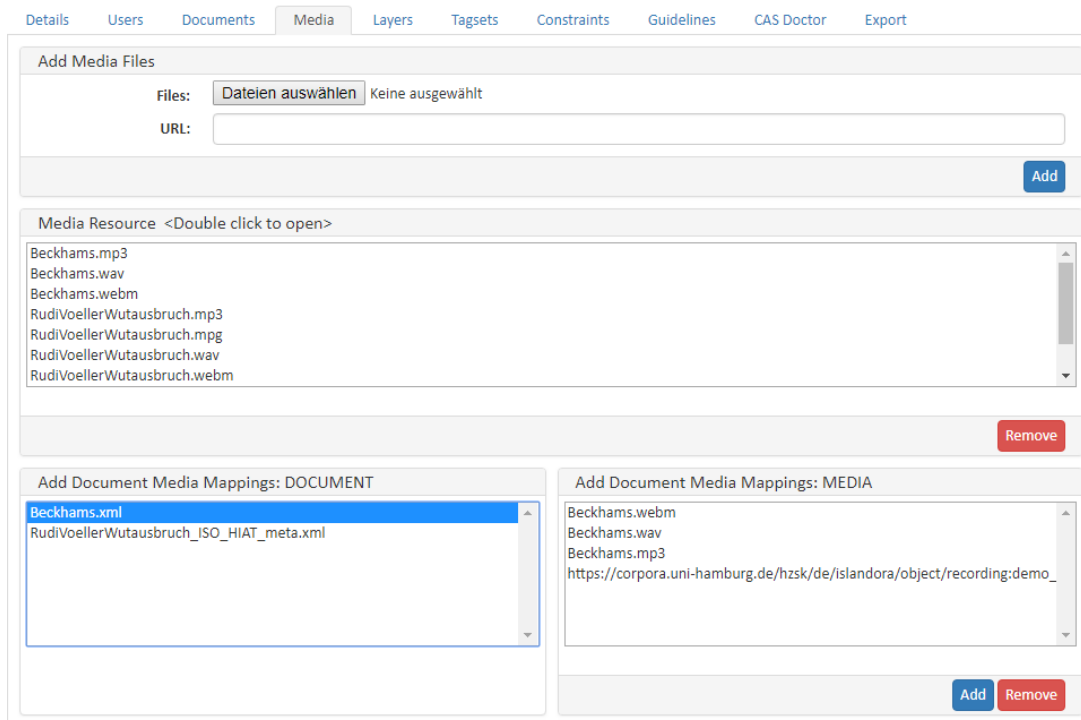


Figure 9: Screenshot of the WebAnno-MM media management pane.

#### 4 WebAnno-MM in practice: An innovative teaching study

As part of a so-called “teaching lab”, WebAnno-MM was used by teams of students participating in a university seminar to collaboratively annotate videotaped authentic classroom discourse. Thematically, the seminar covered the linguistic analysis of comprehension processes displayed in classroom discourse. The seminar was addressed to students in pre-service teacher training and students of linguistics. Students of both programs were supposed to cooperate on interdisciplinary teams in order to gain the most from their pedagogic as well as their linguistic expertise. The students had to choose their material according to their own interest from a set of extracts of classroom discourses from various subject matter classes. Benefiting from the innovative ways to decide on units of analysis such as spans, chains, etc., different stages of the process of comprehension were to be identified and then to be described along various dimensions relevant to comprehension. This approach made single steps of analysis transparent for the students, and thus allowed for their precise and explicit discussion in close alignment with existing academic literature. Compared to past seminars with a similar focus, but lacking the technological support, these discussions appeared more thoughtful and more in-depth. The students easily developed independent ideas for their research projects. Students remarked on this very positively in the evaluation of the seminar.

## 5 Outlook

By implementing an extension of WebAnno, we showed that it is possible to repurpose a linguistic annotation tool for multimodal data. According to the intended use case, the first release focused only on data transcribed according to the HIAT conventions using the EXMARaLDA transcription and annotation tool, thus at first restricting the possible annotation layers. Later releases allow for the creation of custom annotation layers for existing project or research question specific annotation tiers in the transcription data.

Instead of relying on one of the widely used transcription tool formats, we used the ISO/TEI standard, which can model transcription data produced by various tools and according to different transcription conventions, as an exchange format. Obvious next steps would therefore be to extend the interoperability to include full support for the ISO/TEI format. This would require adapted import functionality and transcript visualization functionality for further transcription systems, as well as a generic fallback option. With a more general interoperability, annotation projects could be based on data from several contexts (cf. Ligeois et al. (2015)). Another relevant aspect is supporting the standard-compliant modelling of segment-based annotations, in contrast to annotations that are time-based, i.e. aligned with the base layer and sharing features such as the speaker identity. Segment-based annotations also allow for annotation layers with a more complex structure, which is in some cases required to explicitly model grammatical information (Arkhangelskiy et al., 2019). The word or segment-based annotations are closer to the text-oriented data models used by WebAnno and will be supported in a future release of WebAnno-MM. Other important tasks to take on are extensions of the ISO/TEI standard to model both metadata in the TEI Header and the complex annotations generated in WebAnno in a standardized way.

## Acknowledgments

This work was partially funded by the project ‘Interaktives Annotieren von Unterrichtskommunikation’ in the program Lehrlabor Lehrerprofessionalisierung (L3Prof), Universität Hamburg.<sup>10</sup>

## References

- Timofey Arkhangelskiy, Anne Ferger, and Hanna Hedeland. 2019. Uralic multimedia corpora: ISO/TEI corpus data in the project INEL. In *Proceedings of the Fifth International Workshop on Computational Linguistics for Uralic Languages*, pages 115–124, Tartu, Estonia.
- Claude Barras, Edouard Geoffrois, Zhibiao Wu, and Mark Liberman. 2000. Transcriber: development and use of a tool for assisting speech corpora production. *Speech Communication – Special issue on Speech Annotation and Corpus Tools*, 33(1–2).
- David Ferrucci and Adam Lally. 2004. UIMA: An Architectural Approach to Unstructured Information Processing in the Corporate Research Environment. *Natural Language Engineering*, 10(3–4):327–348.
- ISO/TC 37/SC 4. 2016. Language resource management – Transcription of spoken language. Standard ISO 2462:2016, International Organization for Standardization, Geneva, Switzerland.
- Loc Ligeois, Carole Etienne, Christophe Parisse, Christophe Benzitoun, and Christian Chanard. 2015. Using the TEI as a pivot format for oral and multimodal language corpora. Paper presented at Text Encoding Initiative Conference, Lyon, 28–31, 2015.
- Richard Eckart de Castilho, Chris Biemann, Iryna Gurevych, and Seid Muhie Yimam. 2014. WebAnno: a flexible, web-based annotation tool for CLARIN. In *Proceedings of the CLARIN Annual Conference 2014*, pages 1–3, Soesterberg, Netherlands.
- Elinor Ochs. 1979. Transcription as theory. In E. Ochs and B.B. Schieffelin, editors, *Developmental pragmatics*, pages 43–72. Academic Press, New York.
- Jochen Rehbein, Thomas Schmidt, Bernd Meyer, Franziska Watzke, and Annette Herkenrath. 2004. Handbuch für das computergestützte Transkribieren nach HIAT. *Arbeiten zur Mehrsprachigkeit, Folge B*, 56:1 ff. (in German).

<sup>10</sup><https://www.zlh-hamburg.de/entwicklungsvorhaben/lehrlabore/lehrlabor-lehrerbildung>

- Thomas Schmidt and Kai Wörner. 2014. EXMARaLDA. In Jacques Durand, Ulrike Gut, and Gjert Kristoffersen, editors, *Handbook on Corpus Phonology*, pages 402–419. Oxford University Press.
- Thomas Schmidt, Hanna Hedeland, and Daniel Jettka. 2017. Conversion and annotation web services for spoken language data in clarin. In *Selected papers from the CLARIN Annual Conference*, number 136, pages 113–130, Aix-en-Provence, France. Linköping University Electronic Press, Linköpings Universitet.
- Thomas Schmidt. 2011. A TEI-based Approach to Standardising Spoken Language Transcription. *Journal of the Text Encoding Initiative*, 1:1–28.
- Han Sloetjes. 2014. ELAN: Multimedia annotation application. In Jacques Durand, Ulrike Gut, and Gjert Kristoffersen, editors, *Handbook on Corpus Phonology*, pages 305–320. Oxford University Press.
- Pontus Stenetorp, Goran Topić, Sampo Pyysalo, Tomoko Ohta, Jin-Dong Kim, and Jun’ichi Tsujii. 2011. Bionlp shared task 2011: Supporting resources. In *Proceedings of BioNLP Shared Task 2011 Workshop*, pages 112–120, Portland, OR, USA.
- Seid Muhie Yimam, Iryna Gurevych, Richard Eckart de Castilho, and Chris Biemann. 2013. WebAnno: A flexible, web-based and visually supported system for distributed annotations. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 1–6, Sofia, Bulgaria.
- Seid Muhie Yimam, Chris Biemann, Richard Eckart de Castilho, and Iryna Gurevych. 2014. Automatic annotation suggestions and custom annotation layers in WebAnno. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 91–96, Baltimore, MD, USA.