Screen Concepts for Multi-Version Hypervideo Authoring

Britta Meixner[1,2]**, Stefan John**[1]**, Philipp Defner**[1]
[1]University of Passau, Innstrasse 43, 94032 Passau, Germany
[2]FX Palo Alto Laboratory, 3174 Porter Drive, Palo Alto, CA 94304, USA
[2]meixner@fxpal.com, [1]firstname.lastname@uni-passau.de

ABSTRACT
The creation of hypervideos usually requires a lot of planning and is time consuming with respect to media content creation. However, when structure and media are put together to author a hypervideo, it may only require minor changes to make the hypervideo available in other languages or for another user group (like beginners versus experts). However, to make the translation of media and all navigation elements of a hypervideo efficient and manageable, the authoring tool needs a GUI that provides a good overview of elements that can be translated and of missing translations. In this work, we propose screen concepts that help authors to provide different versions (for example language and/or experience level) of a hypervideo. We analyzed different variants of GUI elements and evaluated them in a survey. We draw guidelines from the results that can help with the creation of similar systems in the future.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

Author Keywords
Hypervideo; Navigation; Multilingual; Translation; Authoring

INTRODUCTION
Hypervideos consist of linked video scenes which are combined with other types of media to provide further information on certain topics. Linking media that way, non-linear structures are created that allow users to interactively explore facts fitted to one's personal level of knowledge and learning speed. This type of video is then presented in a player with additional presentation areas and navigational elements.

Hypervideo media content creation is time consuming and requires a lot of planning. Currently, if a new version of an already finished hypervideo is needed, for example in another language or for another target group (like beginners versus experts), a completely new hypervideo has to be created. This requires either to copy the existing video and replace contents in a database or (often XML- or JSON-based) transfer file,

or a new hypervideo has to be put together in the authoring tool from scratch. In case of changes in the hypervideo, both versions have to be altered and maintained, which can be error prone in larger projects. Translating a hypervideo means not only translating every text or medium, it also involves the table of contents, the terms for the keyword search, the titles of scenes, alternative texts for media, the captions and button labels of control panels, and all other elements that are shown in the player at some point. Providing facilitating functions for this in a hypervideo authoring tool can help keeping maintenance costs low while providing up to date hypervideos.

The aforementioned issues raise the question how a translation feature can be integrated into a desktop hypervideo authoring tool without overcrowding the GUI. However, possibilities need to be provided to add contents in another language for all hypervideo elements. Accordingly, the goal of this work was to understand:

- *How are the different versions of the hypervideo best indicated in the GUI?*

- *What is a suitable GUI approach to enable the translation of single elements in a hypervideo?*

- *What is a suitable GUI approach to give an overview of all elements and enable editing in the overview?*

We conducted a survey with 51 people mainly from Germany and Austria presenting them different screen concepts. The survey used conditional statements to present screen concepts always based on previous decisions to avoid confusion among the participants. We found out that users preferred a tidy user interface that can be controlled by keyboard and mouse, making it not overcrowded with too many buttons.

This paper is structured as follows: We first give an overview of related work, then we describe the procedure and participants of our study. The evaluation results lead to tentative recommendations for hypervideo authoring tools. The paper ends with a discussion and conclusion section.

RELATED WORK
Taking a look at hypervideo and multimedia authoring systems, it can be noted that neither early systems like Hyper-Hitchcock [16] nor current systems like the SIVA Producer [11, 9, 10] or Klynt [6] are capable of translating contents or providing different versions of contents. This is why we started looking at systems with similar underlying structures of the content, namely Content Management Systems (CMS)

for web-pages. According to [3], a CMS should "enable and disable as many languages" as required and provide a fallback locale. The fallback locale is used if content is not available in the language selected by the user.

Several CMS offer the option to translate the contents of a web-page. OctoberCMS [13] provides a default language and translations. New languages have to be added in the settings view, where already added languages can also be managed. Text can be translated via a table. The table provides two options, one shows all contents, the other allows to hide already translated contents giving a better overview. The Adobe Experience Manager (Adobe CQ5) [4] lets users translate whole web-pages. "The website will be displayed side by side in 2 parts with the source (reference page) and the target language page in order to compare, check difference, edit and translate as appropriate". ICanLocalize [7] is a translation service for web-pages by professional translators. The tool that is provided for the translators loads web-pages and shows a list of sentences which are extracted from the web-page. The translators can then translate sentence by sentence. The rubedo CMS [15] requires to first add a list of languages into which the currently available content is translated. The user selects between two strategies. The "only one" strategy displays only translated elements in the resulting web-page (nothing is displayed in case no translation is provided). The "with fallback" strategy displays the fallback language if content is not translated. In order to edit or translate contents of a language, it is selected in a combo-box. Then, only the contents in this language are shown, which allows no comparison with the default language. Images and other media files cannot be translated (or replaced with content in another language), only descriptions can be translated. Synapse [8] is a system for the management of translations of websites. Translation is done sentence by sentence (or parts of sentences). The translation interface shows the original text and the translated text alternating for the single parts separated by markups. Another way to deal with the translation issue is proposed by the Open Real Estate CMS which provides automatic translation of contents [12]. It first used the Google Translate API [5] and is now using the MyMemory API [18].

The systems described in this section offer different implementations of the translation functionality in the GUI. Ranging from translations of whole web-pages to translations of single text elements, most of the tools do not offer functions to replace media files.

STUDY/METHOD
To answer the questions stated in the introduction, we conducted an online survey comparing different interface designs.

Procedure/Data Collection
A brainstorming in a focus group [14] with four participants ($N = 4$) was used to find different ways of presenting the different versions of a hypervideo in the GUI of the SIVA Producer. Then, high fidelity paper prototypes were created and refined step by step that covered all alternative presentation approaches in a way with consistent representations of concepts. The high fidelity prototypes were then integrated
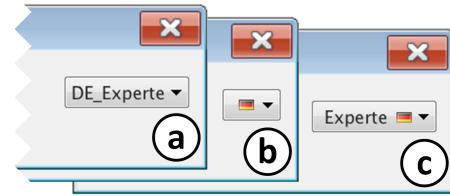

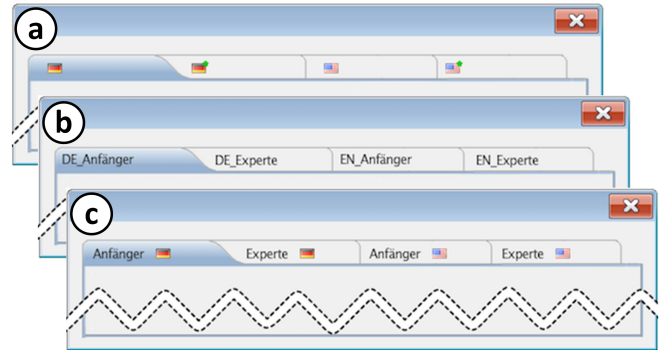Figure 1. Version representations for combo-boxes.

Figure 2. Version representations for tabs.

into an online survey that allowed conditional elements based on previous answers.

The survey as well as the high fidelity prototypes used a car repair scenario. The goal was to make the instructions on how to repair a car available for beginners and experts in both German and English. Accordingly, the original version of the hypervideo which was created for German novices had to be extended with a version for German experts and translated into English.

First, participants were asked about their demographic data (age, gender, level of education) and technical background (monitor size, monitor resolution, experience in video editing).

The second part of the survey showed three different representations of the versions used in a combo-box (see Figure 1) and on tabs (see Figure 2). The first representation (a) was purely text based combining language and experience level with a "_", the second representation (b) was purely image based showing a flag and a symbol, and the third representation (c) was a combination of text and image, where the experience level was written as text and the language was shown as a flag-symbol. In the remainder of the survey, the version representation option preferred by the participant in a ranking was utilized in the next screens.

The third part of the survey showed different variants of the media repository and how the elements should be grouped (see Figure 3). Variant (x) grouped all images, videos, text-files, audios, and subtitles into separate folders not differentiating between the versions. Variant (y) adds a second layer for grouping the elements. Here the files are first grouped by medium and then by version. Variant (z) also has a second layer for grouping the elements. In this variant, the files are grouped by version first and then by medium. Again, the
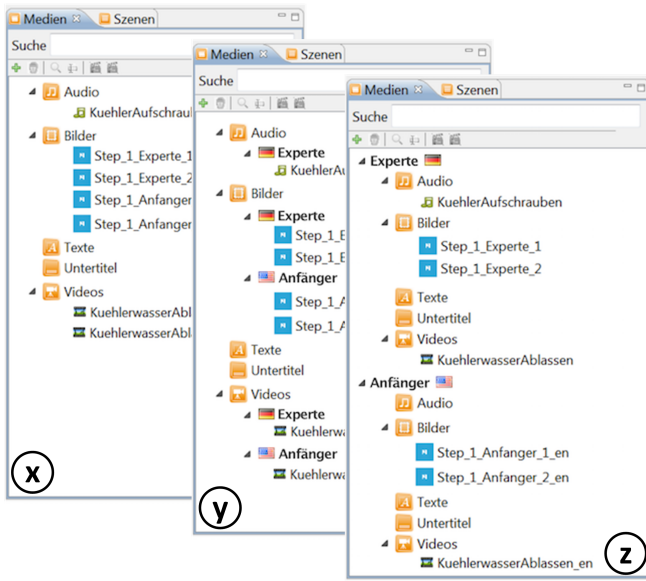
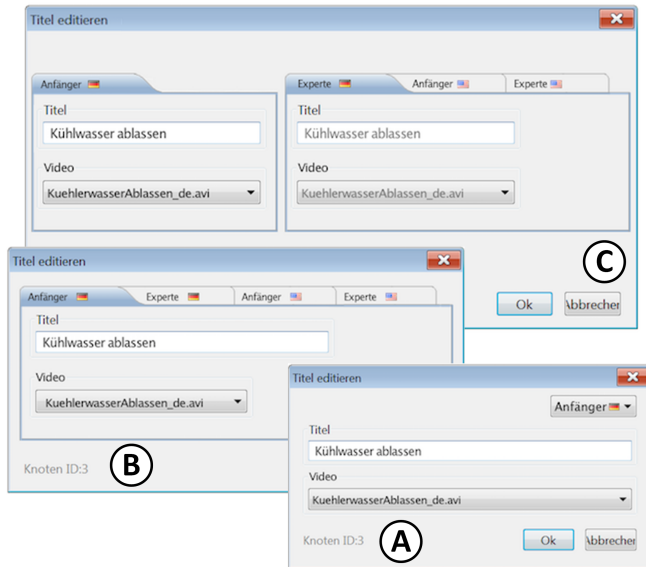**Figure 3. Different layers in media repositories.**

Figure 4. Variants for the translation of scene titles.

Figure 5. Variants for the translation of the conditional control scene graph element.

versions had to be ranked. Another question in this part of the survey explored various ways to import media into the media repository, which also had to be ranked. One way was first selecting the version and then the media which were then added to the right folders automatically. Another way was to first select the media and then the version. The third option imported the media into a "global" folder from where they had to be added to the according folders in the media repository by drag-and-drop. These solutions also had to be ranked. A text field was offered for suggestions of other solutions.

The fourth part of the survey showed three variants of exemplary GUI elements needed for the creation and translation of the hypervideo itself. The shown GUI elements were part of the scene editor (simple text field, Figure 4), the editor for
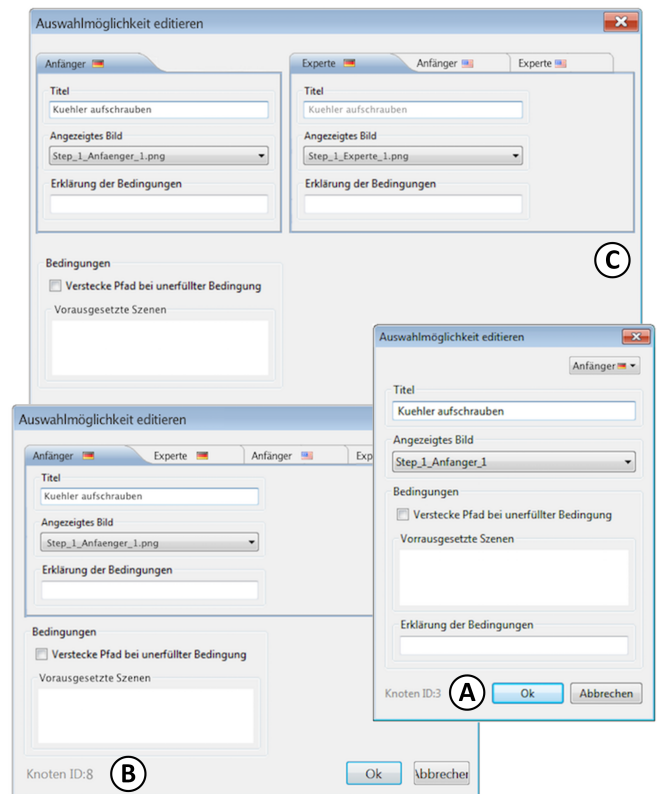
the path selectors at a fork in the scene graph (more complex set of text fields, combo- and check-boxes, Figure 5), the annotation editor for image galleries (text fields and drop areas for images, Figure 6[1]), and the editor for the table of contents (tree structure with entries and list of scenes, Figure 7[1]). Each GUI element had a variant that showed the current implementation in the SIVA Producer, but extended with a combo-box for version selection (variant (A), see Figure 4 and Figure 5). Variant (B) showed a very similar view compared to the current implementation but arranged on tabs (one tab for each version) (see Figure 4 and Figure 5). Variant (C) extended the screen. It showed both the original version of the element and three tabs for the other experience levels and translations (see Figures 4, 5, 6, and 7). The variants for each GUI element were ranked by the participants.

The last part of the survey showed different versions of an overview table that allows the author to detect missing translations. One version of the table had equally high rows for all elements (see Figure 8 (Y)), the other had collapsible rows that showed previews and full texts if they were expanded and only basic information (like the beginning of a text or the name of a media file) when they were collapsed (see Figure 8 (Z)). The participants were asked several questions about each version as well as to select the preferred variant. Other questions were

---

[1]Variants with combo-box and tabs are not shown here to avoid repetitive images and space issues.
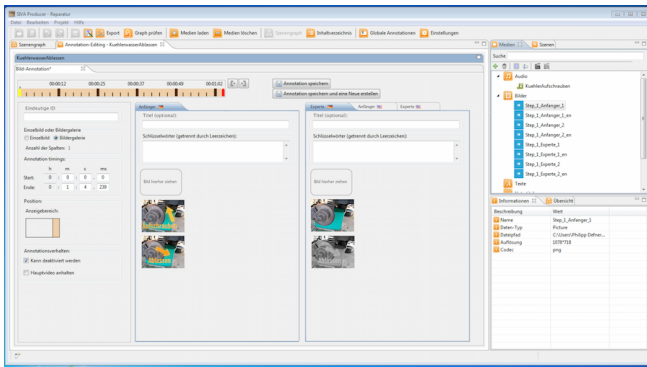
Figure 6. Default version and translation tabs for the image gallery view.
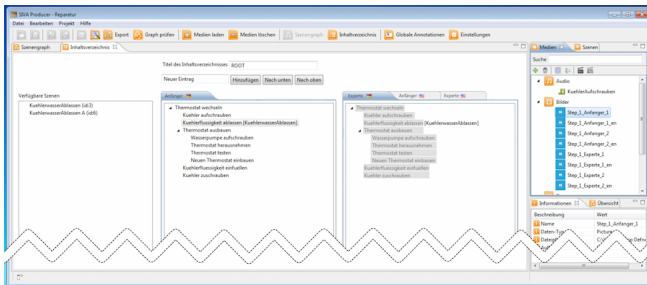

Figure 7. Default version and translation tabs for the table of contents.

about editing and deleting of elements and contents in the table and the presentation of editing options.

Participants

Survey participants were recruited through emails and postings on social networks[2]. The participants had a mean age of 23.23 years ($SD = 3.47$). 39 of the participants were male and 11 were female, one person did not want to tell. All participants had at least some school degree; 6 had finished a training on the job, 21 finished university. About half of the participants used screen resolutions of 1920x1080 and monitor sizes of 20 inch and larger. Most of the participants did not have a lot of experience with video editing software, only two categorized themselves as experts for Adobe Premiere [2], some had knowledge of Adobe Flash [1], Blender [17], and Adobe After Effects [2].

EVALUATION

The question regarding the representation of a version clearly showed that a combination of text and symbol is preferred (42 participants) (see Figures 1 and 2 variant (c)), 6 participants chose the symbol and 3 participants chose the plain text variant.

For the question about media grouping in the media repository, no clear statement regarding a preferred version can be made. However, it can be stated that a grouping with only one layer is not preferred, as variant (x) was selected by only 8 participants. Variant (y) was picked by 20 and variant (z) by 23 participants (see Figure 3). The steps for importing media into the desired
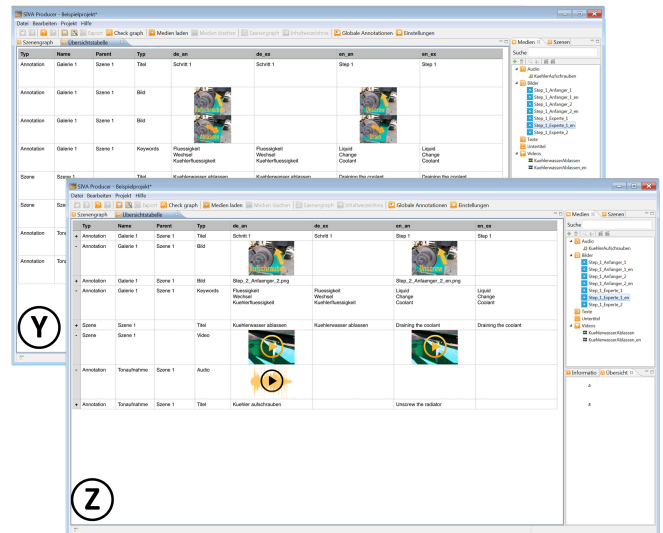
Figure 8. Versions for the translation table with all elements.

media repository folders also revealed no clear answer with an even distribution of the answers for each option.

Taking a look at the variants of the four different GUI elements (scene settings, path settings, image gallery editor, and table of contents) it can be noted that the tab-based variant (B) was considered as most appealing (average over all elements: 25 participants) and clearly arranged (average over all elements: 23.75 participants). The representation of the different versions was considered most obvious (average over all elements: 28 participants) and the distinction of the versions from each other was most clearly visible (average over all elements: 26.5 participants) for variant (B). Asking the participants which variant they would want to use, on average 24.75 chose variant (B), 15 chose variant (C), and 11.25 chose variant (A). Variant (A) with the combo-box seems to be considered better for the simpler GUI elements (scene settings and path settings), compared to the more complex ones. However, variant (C) achieved the best results regarding the question which variant makes the translation of contents the most pleasant (average over all elements: 26 participants).

The questions regarding the overview table revealed somewhat controversial results. Half of the participants found the GUI of both versions appealing and clearly arranged. They found that the contents can be recognized well (35 participants) and it is easy to find out which elements do not have an explanation or translation in a certain variant (25 participants) in the variant showing all the available information (i.e. variant (Y). However, 30 participants preferred the collapsible variant. Regarding the editing of contents in the table, 38 participants preferred "drag & drop from the media repository or input of text directly in the cell of the table", 8 participants wanted an editor window that only shows the content of the cell as well as editing tools, and 5 participants wanted to open the standard annotation editor. Regarding the question how to delete contents, 26 participants preferred marking a cell and pressing the delete button, 12 participants wanted a right-click-menu with a delete option, 9 participants wanted to drag the

element and drop it on a trash-bin symbol, and only 4 participants preferred an additional delete button in each cell. This is also supported by the results of another question, where 42 participants wanted a tidy user interface instead of providing additional buttons in each cell for editing contents. Regarding the question how missing versions should be presented, 22 participants preferred empty cells and 29 preferred a grayed-out miniature view of the standard version.

RECOMMENDATIONS

The evaluation results of the versioning interfaces for the desktop hypervideo authoring tool allow us to draw tentative guidelines and recommendations for similar systems where contents of hypervideo have to be translated or edited for another version:

- If suitable, provide a combination of text and symbol for representing a version in the GUI.

- The media repository should provide more layers than just sorting by medium.

- A view with tabs is preferred to switching between versions with a combo-box.

- Showing the default and translated versions next to each other makes the translation task more pleasant.

- In an overview, elements should be editable as easily as possible (direct entry, drag & drop, keyboard shortcuts) and only necessary information should be visible without overcrowding the GUI with buttons for editing the contents.

- Missing information should be either visualized by empty fields or by grayed-out versions of the originals.

DISCUSSION AND CONCLUSIONS

In this paper we presented a GUI concepts study for providing hypervideo contents in different versions, like for beginners and experts and/or in different languages. We created high fidelity prototypes of GUI concepts and integrated them in a survey which was answered by 51 participants. However, most of the participants were less than 30 years old, from Germany/Austria, and had very limited to no experience in video editing. A more diverse user group with more participants would make the results more reliable. Rough guidelines could nonetheless be drawn from the results where most of the users agreed. Our prototypes have a somewhat complex scenario showing different experience levels and languages making it harder for the participants to understand the scenario. Using a simpler scenario would have made understanding the scenario easier, but would also have over-simplified the prototypes. This may have resulted in results that are not applicable to more difficult scenarios.

A combination of text and symbol are preferred for the description and differentiation of the different versions. Media organization in the media repository should be enabled on two levels, either medium-version or version-medium; an organization structure with only one level is not desired. Comparing the different GUI variants for editing hypervideo elements, variant (B) with a list of tabs is preferred to variant (C) with

the view of the standard version and the tabs of the alternative versions, and to variant (A) with the combo-box. No clear statement can be made regarding the design of the overview table where participants evaluated the version showing more information at one time better, but would prefer the more compact version when they were asked directly.

Future work should include follow-up studies to confirm and refine the tentative guidelines and recommendations formulated in this work. Further work in this area should also include hypervideo authoring on tablets and other devices.

REFERENCES

1. Adobe Systems Incorporated. 2016a. Adobe Flash Player. Website `http://www.adobe.com/products/flashplayer.html` (accessed May 26, 2016). (2016).

2. Adobe Systems Incorporated. 2016b. Adobe Premiere Pro CC. Website `http://www.adobe.com/products/premiere.html` (accessed May 26, 2016). (2016).

3. Globalization Partners International. 2013a. Selecting the Right Multilingual CMS to Manage International Websites. Website `http://blog.globalizationpartners.com/cms-international-websites.aspx` (accessed May 06, 2016). (2013).

4. Globalization Partners International. 2013b. Website Translation in Adobe Experience Manager - Adobe CQ5. Website `http://blog.globalizationpartners.com/multilanguage-website-adobe-experience-manager.aspx` (accessed May 06, 2016). (2013).

5. Google Inc. 2015. Google Translate API - Dynamically translate between thousands of available language pairs. Website `https://cloud.google.com/translate/` (accessed May 06, 2016). (2015).

6. Honkytonk Films. 2015. Klynt. Website, `http://www.klynt.net/` (accessed April 27, 2015). (2015).

7. ICanLocalize. 2016. ICanLocalize Documentation - ICanLocalize documentation center. Website `http://docs.icanlocalize.com/information-for-clients/tutorials/how-to-translate-cms-websites-with-icanlocalize/` (accessed May 06, 2016). (2016).

8. iData Technologies. 2016. Multilingual Content Management with the Synapse Publisher CMS. Website `http://www.idatatechnologies.com/en-US/Multilingual-Web-Content-Management.aspx` (accessed May 06, 2016). (2016).

9. Britta Meixner, Stefan John, and Christian Handschigl. 2015. SIVA Suite: Framework for Hypervideo Creation, Playback and Management. In *Proceedings of the 23rd ACM International Conference on Multimedia (MM '15)*. ACM, New York, NY, USA, 713–716. DOI: `http://dx.doi.org/10.1145/2733373.2807413`

10. Britta Meixner, Katarzyna Matusik, Christoph Grill, and Harald Kosch. 2012. Towards an easy to use authoring tool for interactive non-linear video. *Multimedia Tools*

and Applications 70, 2 (2012), 1251–1276. `DOI:`
`http://dx.doi.org/10.1007/s11042-012-1218-6`

11. Britta Meixner, Beate Siegel, Günther Hölbling, Franz Lehner, and Harald Kosch. 2010. SIVA Suite: Authoring System and Player for Interactive Non-linear Videos. In *Proceedings of the 18th ACM International Conference on Multimedia (MM '10)*. ACM, New York, NY, USA, 1563–1566. `DOI:`
`http://dx.doi.org/10.1145/1873951.1874287`

12. MonoRay. 2016. Machine (automated) translation in Open Real Estate CMS. Website `https://open-real-estate.info/en/blog/` `mashinnyj-perevod-v-open-real-estate-cms` (accessed May 06, 2016). (2016).

13. octobercms.com. 2014. OctoberCMS Plugins - Translate - Enables multi-lingual websites and translate website contents. Website `https://octobercms.com/plugin/rainlab-translate` (accessed May 06, 2016). (2014).

14. D. Remenyi. 2012. *Field Methods for Academic Research: Interviews, Focus Groups and Questionnaires 3rd Edition:*. Academic Conferences and Publishing Limited.

15. rubedo. 2013. Multilingualism tutorial : rubedo cms. Website `http://www.slideshare.net/Rubedo-project/` `multilingualism-tutorial-rubedo-cms` (accessed May 06, 2016). (2013).

16. Frank Shipman, Andreas Girgensohn, and Lynn Wilcox. 2008. Authoring, Viewing, and Generating Hypervideo: An Overview of Hyper-Hitchcock. *ACM Trans. Multimedia Comput. Commun. Appl.* 5, 2, Article 15 (Nov. 2008), 19 pages. `DOI:`
`http://dx.doi.org/10.1145/1413862.1413868`

17. Stichting Blender Foundation. 2016. Blender. Website `https://www.blender.org/` (accessed May 26, 2016). (2016).

18. Translated.net. 2015. MyMemory: API technical specifications. Website `http://mymemory.translated.net/doc/spec.php` (accessed May 06, 2016). (2015).