

# Smart Subtitles for Vocabulary Learning

Geza Kovacs  
Stanford University  
Palo Alto, California  
gkovacs@stanford.edu

Robert C. Miller  
MIT CSAIL  
Cambridge, MA  
rcm@mit.edu

## ABSTRACT

Language learners often use subtitled videos to help them learn the language. However, standard subtitles are suboptimal for vocabulary learning, as translations are nonliteral and made at the phrase level, making it hard to find connections between the subtitle text and the words in the video. This paper presents Smart Subtitles, which are interactive subtitles tailored towards vocabulary learning. Smart Subtitles can be automatically generated from common video sources such as subtitled DVDs. They provide features such as vocabulary definitions on hover, and dialog-based video navigation. Our user study shows that students studying Chinese learn over twice as much vocabulary with Smart Subtitles than with dual Chinese-English subtitles. Learners' self-assessed enjoyment of the viewing experience, as well as their comprehension of the video, both self-assessed and as indicated by independent evaluations of their summaries, remain unchanged.

## Author Keywords

subtitles; interactive videos; language learning

## ACM Classification Keywords

H.5.2. Information Interfaces and Presentation: Graphical User Interfaces

## INTRODUCTION

Students studying foreign languages often wish to enjoy authentic foreign-language video content. For example, many students cite a desire to be able to watch anime in its original form as their motivation for starting to study Japanese. However, the standard presentations of videos are not accommodating towards language learners. For example, if a learner were watching anime, and did not recognize a word in the dialog, the learners would normally have to listen carefully to the word, and look it up in a dictionary. This is a time-consuming process which detracts from the enjoyability of watching the content. Alternatively, the learner could simply watch a version with subtitles in their native language to enjoy the content. However, they would not learn the foreign language this way.

We aim to build a foreign-language video viewing tool that maximizes vocabulary learning, while ensuring that the learner fully understands the video and enjoys watching it.

## BACKGROUND

Videos in foreign languages have been adapted for foreign viewers and languages learners in many ways. These are summarized in Figure 1.

### Presenting Videos to Foreign Viewers

One approach used to adapt videos to viewers who do not understand the original language is *dubbing*. Here, the original foreign-language voice track is removed, and is replaced with a voice track in the viewers native language. Because the foreign language is no longer present in the dubbed version, this medium is ineffective for foreign language learning.

Another approach is to provide *subtitles* with the video. Here, the foreign-language audio is retained as-is, and the native-language translation is provided in textual format, generally as a line presented at the bottom of the screen. Thus, the learner will hear the foreign language, but will not see its written form - therefore, they will need to pay attention to the audio. Subtitles have had mixed reactions in the context of language learning - some studies have found them to be beneficial for vocabulary acquisition, compared to watching videos without them [1]. That said, other studies have found them to provide little benefit to language learners in learning vocabulary [4]. Additionally, the presence of subtitles are considered to detract attention from the foreign-language audio and pronunciation [2]. The mixed results that studies have found on the effects of subtitles on language learning suggests that their effectiveness depends on factors such as the experience level of the learners [6].

### Presenting Videos to Language Learners

Whether or not videos should at all be presented to language learners with subtitles or similar aids is itself a matter of debate. Subtitles, in particular, are frowned upon, because they do nothing to deter learners from simply reading in their native language. Some language educators are of the opinion that since students will not have subtitles or similar aids when they visit the foreign country, then they should not be provided with any when viewing videos either. Nevertheless, various video comprehension aids have been experimented with in the context of language learning:

With a *transcript* (also referred to as a *caption*), the video is shown along with the text in the language of the audio (in this case, the foreign language). Transcripts are generally used to assist hearing-impaired viewers; however, they

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single spaced.

can also be beneficial to language learners for comprehension, particularly if they have better reading ability in the foreign language than listening comprehension ability [1]. However, transcripts are only beneficial to more advanced learners whose language competence is already near the level of the video [6].

With *reverse subtitles*, the video has an audio track and a single subtitle, just as with regular subtitles. However, the key distinction from conventional subtitles in the viewer's native language is that here, the audio is in the native language, and the subtitle shows the foreign language. This takes advantage of the fact that subtitle reading is a semi-automatic behavior [5], meaning that the presence of text on the screen tends to attract people's eyes to it, causing them to read it. Therefore, this should attract attention to the foreign-language text. The presentation of the foreign language in written form may also be helpful with certain learners whose reading comprehension ability is stronger than their listening comprehension. That said, because the foreign language is presented only in written form, the learner may not end up learning the pronunciation, particularly with a language with a non-phonetic writing system, such as Chinese.

With *dual subtitles*, the audio track for the video is kept as the original, foreign language. However, in addition to the subtitle displaying the foreign-language, they also display the viewers native language as well. In this way, a learner can both read the written representation, as well as hear the spoken representation of the dialog, and will still have the translation available. Thus, of these options, dual subtitles provide the most information to the learner. Indeed, dual subtitles have been found to be at least as effective for vocabulary acquisition as either captions or subtitles alone [3]

*GliFlix* is a variant on the conventional, native-language subtitle, which inserts translations to the foreign language for the most common words that appear in the dialog. They attain larger rates of vocabulary acquisition compared to regular subtitles (though not dual subtitles) in user studies. However, compared to dual subtitles, *GliFlix* has the disadvantage that because it shows only the most common words in a dialog, then learners will not learn all the vocabulary in the video. Additionally, because *GliFlix* presents the foreign vocabulary in the order of the viewers native language, this approach is likely ineffective for other language-learning tasks such as learning pronunciation and grammar.

In addition to dual subtitles, there have been other attempts to boost vocabulary acquisition rates via unconventional, non-interactive subtitles.

Much work has been done studying subtitles' effects on language learning, which have generally been shown to help language learners attain vocabulary. There have been some variants on subtitling to help learners – dual subtitles .

On each page your material (not including the page number) should fit within a rectangle of 18 x 23.5 cm (7 x 9.25 in.), centered on a US letter page, beginning 1.9 cm (.75 in.) from the top of the page, with a .85 cm (.33 in.) space between two 8.4 cm (3.3 in.) columns. Right margins should be justified,

not ragged. Beware, especially when using this template on a Macintosh, Word can change these dimensions in unexpected ways. Please be sure that your PDF is US letter and not A4. If your PDF or paper are formatted for A4, the submission will be returned to you to fix.

## TYPESET TEXT

Prepare your submissions on a word processor or typesetter. Please note that page layout may change slightly depending upon the printer you have specified.  $\LaTeX$  sometimes will create overfull lines that extend into columns. To attempt to combat this, the .cls file has a command, `\sloppy`, that essentially asks  $\LaTeX$  to prefer underfull lines with extra white-space. For more details on this, and info on how to control it more finely, check out <http://www.economics.utoronto.ca/osborne/latex/PMAKEUP.HTM>.

## Title and Authors

Your paper's title, authors and affiliations should run across the full width of the page in a single column 17.8 cm (7 in.) wide. The title should be in Helvetica 18-point bold; use Arial if Helvetica is not available. Authors' names should be in Times Roman 12-point bold, and affiliations in Times Roman 12-point. For more than three authors, you may have to place some address information in a footnote, or in a named section at the end of your paper. Please use full international addresses and telephone dialing prefixes. Leave one 10-pt line of white space below the last line of affiliations.

## Abstract and Keywords

Every submission should begin with an abstract of about 150 words, followed by a set of keywords. The abstract and keywords should be placed in the left column of the first page under the left half of the title. The abstract should be a concise statement of the problem, approach and conclusions of the work described. It should clearly state the paper's contribution to the field of HCI.

The first set of keywords will be used to index the paper in the proceedings. The second set are used to catalogue the paper in the ACM Digital Library. The latter are entries from the ACM Classification System [?]. In general, it should only be necessary to pick one or more of the H5 subcategories, see <http://www.acm.org/class/1998/ccs98.html>

## Normal or Body Text

Please use a 10-point Times Roman font or, if this is unavailable, another proportional font with serifs, as close as possible in appearance to Times Roman 10-point. The Press 10-point font available to users of Script is a good substitute for Times Roman. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times and not Times New Roman. Please use sans-serif or non-proportional fonts only for special purposes, such as headings or source code text.

## First Page Copyright Notice

Leave 3 cm (1.25 in.) of blank space for the copyright notice at the bottom of the left column of the first page. In this template a floating text box will automatically generate the

Objects	Caption — pre-2002	Caption — 2003 and afterwards
Tables	Above	Below
Figures	Below	Below

**Table 1.** Table captions should be placed below the table.

required space. Note however that the text box is anchored to the **ABSTRACT** heading, so if that heading is deleted the text box will disappear as well. You can replace the default copyright notice by uncommenting the \toappear block at the beginning of the document and inserting your own text, for example, for versions under review.

### Subsequent Pages

On pages beyond the first, start at the top of the page and continue in double-column format. The two columns on the last page should be of equal length.



**Figure 1.** With Caption Below, be sure to have a good resolution image (see item D within the preparation instructions).

### References and Citations

Use a numbered list of references at the end of the article, ordered alphabetically by first author, and referenced by numbers in brackets [?, ?, ?, ?]. For papers from conference proceedings, include the title of the paper and an abbreviated name of the conference (e.g., for Interact 2003 proceedings, use *Proc. Interact 2003*). Do not include the location of the conference or the exact date; do include the page numbers if available. See the examples of citations at the end of this document. Within this template file, use the References style for the text of your citation.

Your references should be published materials accessible to the public. Internal technical reports may be cited only if they are easily accessible (i.e., you provide the address for obtaining the report within your citation) and may be obtained by any reader for a nominal fee. Proprietary information may not be cited. Private communications should be acknowledged in the main text, not referenced (e.g., “[Robertson, personal communication]”).

### SECTIONS

The heading of a section should be in Helvetica 9-point bold, all in capitals. Use Arial if Helvetica is not available. Sections should not be numbered.

### Subsections

Headings of subsections should be in Helvetica 9-point bold with initial letters capitalized. For sub-sections and sub-subsections, a word like *the* or *of* is not capitalized unless it is the first word of the heading.)

#### Sub-subsections

Headings for sub-subsections should be in Helvetica 9-point italic with initial letters capitalized. Standard \section, \subsection, and \subsubsection commands will work fine.

### FIGURES/CAPTIONS

Place figures and tables at the top or bottom of the appropriate column or columns, on the same page as the relevant text (see Figure 1). A figure or table may extend across both columns to a maximum width of 17.78 cm (7 in.).

Captions should be Times New Roman 9-point bold. They should be numbered (e.g., “Table 1” or “Figure ??”), centered and placed beneath the figure or table. Please note that the words “Figure” and “Table” should be spelled out (e.g., “Figure” rather than “Fig.”) wherever they occur.

Papers and notes may use color figures, which are included in the page limit; the figures must be usable when printed in black and white in the proceedings. The paper may be accompanied by a short video figure up to five minutes in length. However, the paper should stand on its own without the video figure, as the video may not be available to everyone who reads the paper.

### LANGUAGE, STYLE AND CONTENT

The written and spoken language of SIGCHI is English. Spelling and punctuation may use any dialect of English (e.g., British, Canadian, US, etc.) provided this is done consistently. Hyphenation is optional. To ensure suitability for an international audience, please pay attention to the following:

- Write in a straightforward style.
- Try to avoid long or complex sentence structures.
- Briefly define or explain all technical terms that may be unfamiliar to readers.
- Explain all acronyms the first time they are used in your text—e.g., “Digital Signal Processing (DSP)”.
- Explain local references (e.g., not everyone knows all city names in a particular country).
- Explain “insider” comments. Ensure that your whole audience understands any reference whose meaning you do not describe (e.g., do not assume that everyone has used a Macintosh or a particular application).
- Explain colloquial language and puns. Understanding phrases like “red herring” may require a local knowledge of English. Humor and irony are difficult to translate.

- Use unambiguous forms for culturally localized concepts, such as times, dates, currencies and numbers (e.g., “1-5-97” or “5/1/97” may mean 5 January or 1 May, and “seven o’clock” may mean 7:00 am or 19:00). For currencies, indicate equivalences—e.g., “Participants were paid 10,000 lire, or roughly \$5.”
- Be careful with the use of gender-specific pronouns (he, she) and other gendered words (chairman, manpower, man-months). Use inclusive language that is gender-neutral (e.g., she or he, they, s/he, chair, staff, staff-hours, person-years). See [?] for further advice and examples regarding gender and other personal attributes.
- If possible, use the full (extended) alphabetic character set for names of persons, institutions, and places (e.g., Grønbæk, Lafrenière, Sánchez, Universität, Weißenbach, Züllighoven, Århus, etc.). These characters are already included in most versions of Times, Helvetica, and Arial fonts.

### ACCESSIBILITY

The Executive Council of SIGCHI has committed to making SIGCHI conferences more inclusive for researchers, practitioners, and educators with disabilities. As a part of this goal, the all authors are asked to work on improving the accessibility of their submissions. Specifically, we encourage authors to carry out the following five steps:

1. Add alternative text to all figures
2. Mark table headings
3. Add tags to the PDF
4. Verify the default language
5. Set the tab order to “Use Document Structure”

Unfortunately good tools do not yet exist to create tagged PDF files from Latex. LaTeX users will need to carry out all of the above steps in the PDF directly using Adobe Acrobat, after the PDF has been generated.

For more information and links to instructions and resources, please see: <http://chi2014.acm.org/authors/guide-to-an-accessible-submission>.

### PAGE NUMBERING, HEADERS AND FOOTERS

Please submit your anonymous version for reviewing with page numbers centered in the footer. These must be removed in the final version of accepted papers, as page numbers, headers, and footers will be added by the conference printers. Comment out the \pagenumbering command at the top of the document to remove page numbers.

### PRODUCING AND TESTING PDF FILES

We recommend that you produce a PDF version of your submission well before the final deadline. Your PDF file must be ACM DL Compliant. The requirements for an ACM Compliant PDF are available at: <http://www.sheridanprinting.com/typedept/ACM-distilling-settings.htm>.

Test your PDF file by viewing or printing it with the same software we will use when we receive it, Adobe Acrobat Reader Version 7. This is widely available at no cost from [?]. Note that most reviewers will use a North American/European version of Acrobat reader, which cannot handle documents containing non-North American or non-European fonts (e.g. Asian fonts). Please therefore do not use Asian fonts, and verify this by testing with a North American/European Acrobat reader (obtainable as above). Something as minor as including a space or punctuation character in a two-byte font can render a file unreadable.

### BLIND REVIEW

For archival submissions, CHI requires a “blind review.” To prepare your submission for blind review, remove author and institutional identities in the title and header areas of the paper. You may also need to remove part or all of the Acknowledgments text. Further suppression of identity in the body of the paper and references is left to the authors’ discretion. For more details, see the submission guidelines and checklist for your submission category.

### CONCLUSION

It is important that you write for the SIGCHI audience. Please read previous years’ Proceedings to understand the writing style and conventions that successful authors have used. It is particularly important that you state clearly what you have done, not merely what you plan to do, and explain how your work is different from previously published work, i.e., what is the unique contribution that your work makes to the field? Please consider what the reader will learn from your submission, and how they will find your work useful. If you write with these questions in mind, your work is more likely to be successful, both in being accepted into the Conference, and in influencing the work of our field.

### ACKNOWLEDGMENTS

We thank CHI, PDC and CSCW volunteers, and all publications support and staff, who wrote and provided helpful comments on previous versions of this document. Some of the references cited in this paper are included for illustrative purposes only. **Don’t forget to acknowledge funding sources as well**, so you don’t wind up having to correct it later.

### REFERENCES FORMAT

References must be the same font size as other body text.