Kalgan: Video Player for Casual Language Learning

Sathaporn "Hubert" Hu[1]

University of Calgary

Calgary, AB T2N1N4, CA

sathaporn.hu@ucalgary.ca

Wesley J. Willett

University of Calgary

Calgary, AB T2N1N4, CA

wesley.willett@ucalgary.ca

Abstract

We present the initial design and evaluation of Kalgan, an interactive video player tailored to support casual language learning. Unlike traditional, more structured learning, casual language learning takes place alongside existing leisure activities such as watching videos. To better support learning during casual video viewing, we introduce several new features, including subtitle-aware rewind, interactive subtitle translation, and word lookup history that help people quickly access language content and recover missed information without disrupting their viewing experience. We evaluated this initial version of Kalgan using a mix of remote studies and in-person observation. The results of the evaluation highlight our participants' enthusiasm for subtitle-centric learning aids. They also suggest a variety of future research opportunities for casual language learning tools.

Author Keywords

Computer-assisted language learning; video player.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

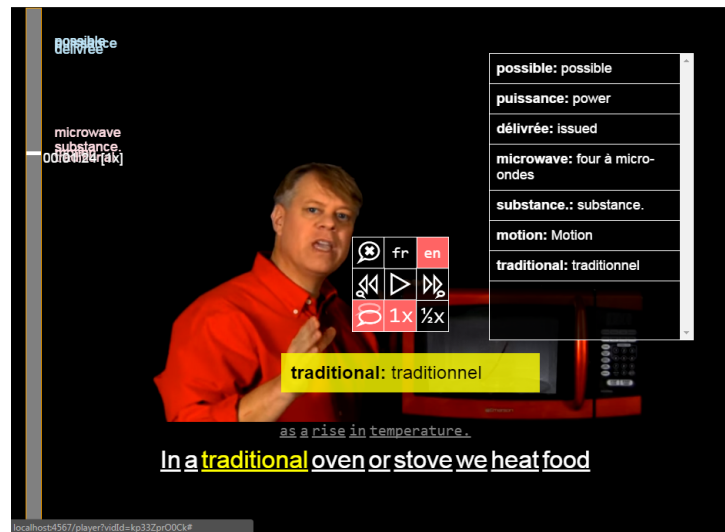1 The author is now with Dalhousie University, Halifax, NS B3H4R2, CA.

Figure 1. The general interface of Kalgan. More information available in the sidebars (Figure 2-4). Contains an image from https://www.youtube.com/watch?v=kp33ZprO0Ck, courtesy of Bill Hammack.

Introduction

Kalgan is a video player designed for casual language learning. It includes features such as touch-friendly interactive subtitles, subtitle-aware rewind, and a word history box. These features help viewers to more easily learn new words from videos — even when they aren't specifically focused on learning. We evaluated Kalgan using multiple methods, including techniques like an online remote study and a think-aloud protocol that have not frequently been deployed in the evaluation of video players for language learning.

Casual language learning is a type of learning in which learners do not typically prioritize acquiring new linguistic knowledge. Instead casual language learning occurs implicitly during casual everyday activities like reading, listening to podcasts, or playing games that involve content in a new language. Watching online video is already a common venue for causal language learning [11], yet few video players are tailored to support it. We designed Kalgan to explore how future video players might better facilitate these activities.

Although there are many video players designed for language learning [1,7,13], Kalgan provides a number of unique features. Kalgan also uses the YouTube APIs, providing access to a large volume of video content and subtitles in a wide range of languages.

We evaluated Kalgan with an online remote study, an interview, and a think-aloud protocol. Using multiple methods allowed us to obtain more holistic data. While our study is still preliminary, the results highlight opportunities for further improvement.

Figure 2. The 9-button Menu.

Top-Left: Turn off subtitles.

Top-Middle: Turn on the learning language subtitles.

Top-Right: Turn on the user's native language subtitles.

Centre-Left: Rewind to the previous subtitle/4 seconds.

Centre-Middle: Play/Pause

Centre-Right: Fast-forward to the next subtitle/4 seconds.

Bottom-Left: Turn on the previous subtitles.

Bottom-Centre: Play the video at full speed.

Bottom-Right: Play the video at half-speed.

Related Work

The field of computer-assisted language learning (CALL) investigates the use of computers to facilitate language learning. In designing Kalgan, we relied on several important observations from previous CALL research. In particular, prior work has highlighted the importance of subtitles for supporting language learning with video [2]. Recent research also suggests that people learn better with an ability to immediately look up definitions of subtitle words [1,10]. In fact, many experimental video players designed for language learning (GliFlix [7], ViVo [13], Yabla [1], Smart Subtitles [5]) make use of these suggestions.

While they are not necessarily focused on language learning, Massively Open Online Courses (MOOC) are also very reliant on videos for content delivery [6]. We take inspiration from work by Kim et al. [4] which provides visual indicators of important content in videos in order to encourage re-visitation and learning.

System Description

We created Kalgan using HTML5 and the YouTube API, allowing it to work on a variety of different mobile and desktop devices. Since previous work [1,2,7,13] has demonstrated that subtitles are helpful for language learning, we prioritized subtitle-based features and interactions. Unlike most existing video players [7,13], we made our video player touch-friendly, using a numpad-style grid layout (Figure 2) suggested by Wang and Ren [12]. Kalgan works with any YouTube video that has subtitles in at least one language. This approach is different from GliFlix [7], which requires a custom set of video and subtitle files, as well as from ViVo [13] and Yabla [1] which can only play specific videos provided by the players themselves.

Features

In our initial design iterations, we observed that casual learners were often frustrated by videos or subtitles that moved too quickly, causing them to miss information like names, slang terms, and rapid dialogue. In traditional players, recovering these kinds of missed information is often difficult, requiring awkward rewinding or scrubbing that disrupts the casual viewing experience. We introduced features that help the user to recover missed information or to prevent information from being missed. These include:

- **Subtitle-aware rewind (Figure 2).** When the subtitles are on, the user can rewind in subtitle-by-subtitle increments, making it easier to quickly recover missed phrases without scrubbing or manually backtracking. Kalgan also has Subtitle-aware fast-forward which supports rapid resumption by fast-forward the video to the start of the next subtitle.
- **Half-speed playback (Figure 2).** This feature allows the video to be played at half-speed, so the user can better parse rapid speech.
- **Previous subtitles (Figure 3).** This feature shows the previous subtitle in gray just above the current one – giving the user more time to read and interact with the previous subtitle.
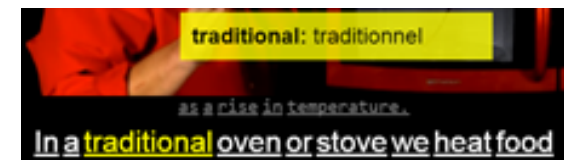

Figure 3. The previous subtitles (appearing as a gray text) and the current subtitles. The user can look up words in both. The yellow box contains the word's definition. Close up of Figure 1.
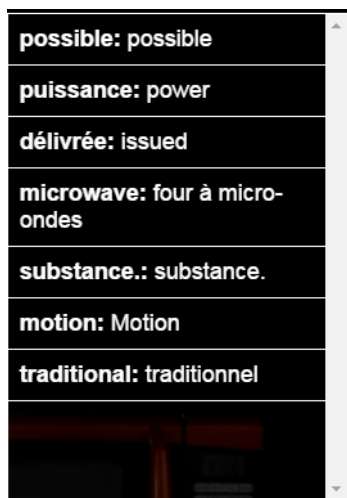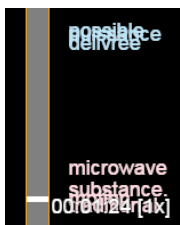
Figure 4. The word history box.

Figure 5. Annotated timeline. Each word that has been looked up appears along the timeline, color-coded by language.

We also added several features that simplify and support common learning tasks:

- **Interactive subtitle translation (Figure 3).** When the user taps/clicks on a word in a subtitle in one language, Kalgan instantly provides the word's definition in the other language. This makes word lookup more easier and quicker. We implemented this feature using the Microsoft Translate API.
- **Word history box (Figure 4).** Kalgan records words and definitions that the user has looked up using interactive subtitles. The feature can help the user generate flashcards for later review.
- **Annotated timeline (Figure 5).** Once the user looks up a word, Kalgan also appends the word on the player timeline to visualize learning activity.

The layout of the numpad-style menu also allows the user to quickly change between subtitle languages (Figure 2). There are three subtitle language settings: (1) no subtitles, (2) the language the user is learning, (3) the language the user is proficient in. Each setting represents a difficulty level. We envision that advanced learners would choose (1), intermediate users would choose (2) while novices would choose (3). Learners can also easily minimize the entire interface by clicking or tapping outside the menu. This allows learners to minimize distraction and simply enjoy the content.

Preliminary Evaluation

We performed an initial evaluation of Kalgan, primarily using an online remote study. This approach is different from the evaluation protocols for other video-based language learning tools. For example, GliFlix and ViVo were evaluated via controlled studies conducted inside laboratories [7,13] and Yabla was never evaluated [1].

We recruited 10 participants from the University of Calgary. Each participant used Kalgan for at least a week. While we provided a list of suggested videos in English and French, participants could watch any video at anytime, anywhere, and in any language. After this period, we interviewed participants in person. During the interview, we asked questions about: (1) participants' own linguistic background, (2) bugs that they encountered, (3) preferences and dislikes toward the features of Kalgan, (4) comparison between Kalgan and an ordinary player, and (5) devices that they used during the remote study. Finally, we conducted a think-aloud session during which each participant used Kalgan for a short period in our presence. We compensated each participant with a $20 payment.

Findings and Discussion

We analyzed participants' overall usage of the tool using a combination of log data, questionnaire responses, and qualitative observation. We refer to our participants as P1 to P10. Since our quantitative data are non-normal, we report their medians instead of their means. For each median, we also generated a 95% confidence interval using bootstrap with 10,000 iterations. Our results demonstrate the potential of some of Kalgan's features and also highlights opportunities for future casual learning tools.

General Statistics

We computed the total time each participant spent with the tool by summing the lengths of the videos that the they had watched, excluding re-watches (Figure 7-A). The median total viewing time per participant was 75 minutes (CI: [47.8, 137.1]). We also computed the number of unique videos that a participant watched (Figure 7-B). We excluded P8 from the calculation,

Figure 6. The score and rank distributions for the features in Kalgan. Sub 1 = Subtitles in the Language the User Is Already Proficient In. Sub 2= Subtitles in the Language the User is Learning.
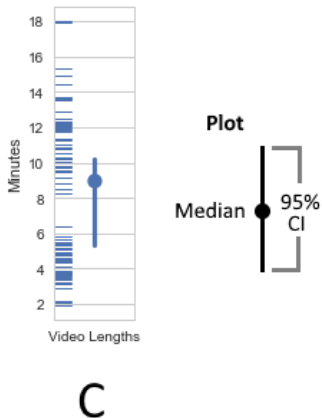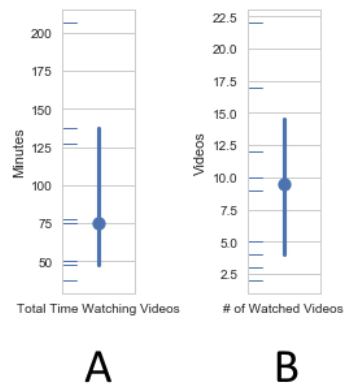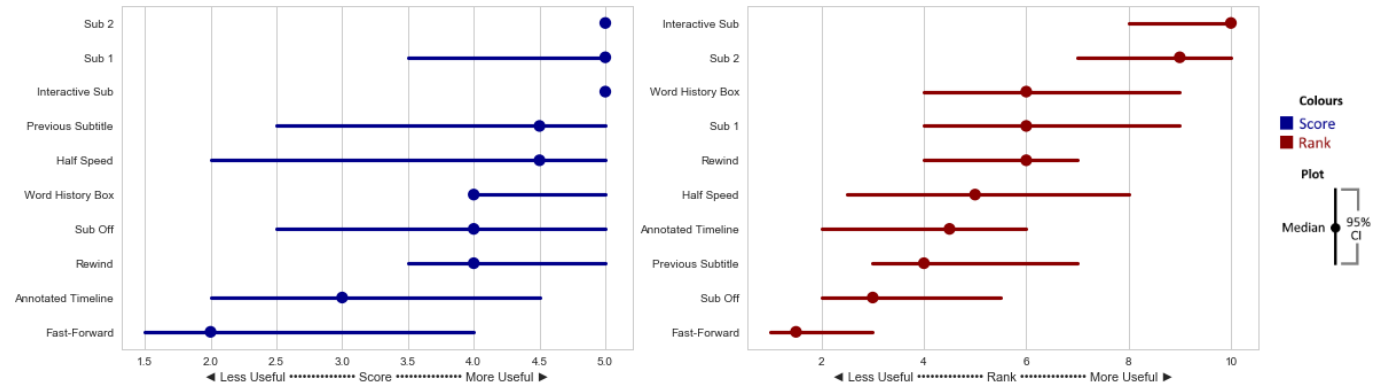

Figure 7. A: Total time spent watching video per participant (Each tick represents a participant). B: Total number of videos watched per participant (Each tick represents a participant). C: Video duration (Each tick represents a video).

because some of the videos that he watched were no longer available on YouTube. The median number of videos per participant was 9.5 (CI: [4, 14.5]), while the median video length (Figure 7-C) was just over 9 minutes (CI: [5.3, 10.2]). Interestingly, the distribution of the video lengths seems to be bimodal.

Preferences and Dislikes toward Features
To collect data about participants' preferences and dislikes, we asked each participant to rate ten different features of the player (Figure 6, left) using a 5-point Likert scale (where 1 = least useful, 5 = most useful). We then asked participants to rank-order the features in terms of their usefulness (Figure 6, right). We found that the most popular features in Kalgan were subtitle-based, with interactive subtitle translation being the most acclaimed. Participants were also moderately enthusiastic about subtitle-aware rewind, the word history box, and the annotated timeline. On the other hand, participants identified subtitle-aware fast-forward

as the least useful feature. This is not surprising considering that the participants focused much more on rewinding to catch missed content than going forward.

Future Work
Our initial explorations with Kalgan highlight a number of opportunities for future video tools for casual language learning. In particular, our observations confirm that subtitles provide a useful double-encoding of foreign-language content that is particularly helpful for casual learners. Subtitle-based interactions that allowed viewers to clarify confusion about the meanings of words and recover missed content without disrupting the flow of the video content were also very popular. As a result, we believe that subtitle-centric interfaces and interactions may also benefit other types of casual language learning beyond video. For example, a podcast application tailored for casual language learning could provide onscreen subtitles and instant translations synchronized with the podcast audio.

Our participants also suggested that better control over word lists generated during their viewing sessions could be beneficial to the learning experience. Similarly, we believe that gamifying video interfaces to encourage viewers to watch and translate more challenging foreign language content could create further opportunities for learning. Although video players like Yabla [1] already include some degree of gamification, this topic is underexplored and warrants further work. Future evaluations of casual learning tools may also benefit from the application of Cognitive Load Theory (CLT) [9]. In particular, CLT may be able to better explain some of our results, including participants' strong preference for features like interactive translation which help reduce the load associated with simultaneously processing both onscreen action and unfamiliar speech in the video [3,8].

Video language learning tools are fast becoming a popular and affordable alternative to conventional language courses [11]. Our initial work on Kalgan further highlights their potential and points the way to an even wider set of casual learning tools to come.

References

1. Adolfo Alfredo Carrillo Cabello. 2013. Website Review: LomásTV. *CALICO Journal* 30, 1: 142–153.

2. Maja Grgurovic and Volker Hegelheimer. 2007. Help options and multimedia listening: Students' use of subtitles and the transcript. *Language Learning & Technology* 11, 1: 45–66.

3. Bruce D. Homer, Jan L. Plass, and Linda Blake. 2008. The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior* 24, 3: 786–797.

4. Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Data-driven interaction techniques for improving navigation of educational videos. *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*: 563–572.

5. Geza Kovacs and Robert C. Miller. 2014. Smart subtitles for vocabulary learning. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, ACM Press, 853–862.

6. A McAuley, B Stewart, G Siemens, D Cormier, and Creative Commons. 2010. The MOOC model for digital practice. *Massive Open Online Courses: digital ways of knowing and learning*: 1–64.

7. N. Sakunkoo and P. Sakunkoo. 2009. GliFlix: Using Movie Subtitles for Language Learning. *UIST*: 7–8.

8. Stephan Schwan and Roland Riempp. 2004. The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction* 14, 3: 293–305.

9. John Sweller. 1994. Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction* 4, 4: 295–312.

10. Alan M. Taylor. 2013. CALL versus Paper: In Which Context Are L1 Glosses More Effective? *CALICO Journal* 30, 1: 63–81.

11. Robert Vanderplank. 2010. Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching* 43, 1: 1.

12. Feng Wang and Xiangshi Ren. 2009. Empirical evaluation for finger input properties in multi-touch interaction. *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, ACM Press, 1063.

13. Yeshuang Zhu, Yuntao Wang, Chun Yu, et al. 2017. ViVo: Video-Augmented Dictionary for Vocabulary Learning. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*: 5568–5579.