2017

Designing and Evaluating a Lightweight Video Player for Language Learning

Hu, Sathaporn

UNIVERSITY OF CALGARY

Designing and Evaluating a Lightweight Video Player for Language Learning

by

Sathaporn Hu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN COMPUTER SCIENCE

CALGARY, ALBERTA

DECEMBER, 2017

Abstract

Watching foreign language videos is a popular and convenient strategy used by many people for learning a new language. However, traditional video players, such as the YouTube player, are not designed to support language learning. We created two video players to explore and to address the issues of using traditional players as a language learning tool. Our players specifically target casual language learners. After evaluating the first player, we found that a traditional player makes it difficult for learners to (1) adjust the level of difficulty, (2) recover missed information, and (3) assess learning progress. We then created the second player to address these issues. The results of the evaluation of the second player demonstrate that people found the player to be helpful for language learning. We also found common usage patterns in the results and opportunities for future improvement.

Acknowledgements

First, I would like to thank Wesley Willett, for his generosity, patience, and kindness. Without his help and advice, I would not be able to complete this thesis. Where I use “we”, it refers to I and him. I also would like to thank Usman Alim and Parmit Chilana for their roles in the examination committee.

Secondly, I want to thank the members of my family for providing me supports throughout the years. Without their help and support, I will not be where I am today. They have sacrificed many things for me, including their dreams, so I could have my own.

Thirdly, I would like to thank my girlfriend and her support throughout the program. Without her, I would not be able to overcome some of the darkest moments in the program.

Fourthly, I would like to thank my colleagues at iLab for their advice, and friendship.

Fifthly, I would like to thank Juan David Hincapié-Ramos who has shown me a glimpse of research in the private sector, and my former colleagues at Lenovo, Beijing.

Dedication

To those who have or had confidence in me – even if I may not be worthy of it.

Table of Contents

List of Tables

List of Figures and Illustrations

Epigraph

"In my youth, I experienced overseas studies. The languages of the West, its literature, its political science, its customs, its mathematics, its geography, its physics and chemistry – all these I have had the chance to study."

Sun Yat-Sen

Chapter One: **Introduction**

Watching foreign-language televisions, films, and web videos is a convenient, affordable, and popular strategy for learning a foreign language [Vanderplank 2010]. Additional augmentations to the videos, such as subtitles, can also further improve learning experience [Cubbison 2005; Kovacs and Miller 2014]. Although traditional video players, such as the YouTube video players, may have these augmentations, their interfaces still make accessing these augmentations difficult. The main reason is that traditional video players do not expect their users to frequently interact with the interfaces. They also emphasize having sleek and clean interfaces. As such, the players hide many advanced options under layers of menus – making them difficult to access. However, prior research demonstrates that people learn better if they can have frequent and better interactions with videos [Schwan and Riempp 2004; Zhang et al. 2006]. We created and evaluated several video players to research improving casual language learning. Additionally, we also attempted to identify how people consume videos for language learning.

Vanderplank [2010] argues that videos are affordable and convenient media for language learning. He also argues that augmentations such as the subtitles can enhance learning experience. Prior research in the field of Computer-Assisted Language Learning (CALL) supports his argument [Cabello 2013; Grgurovic and Hegelheimer 2007]. Although subtitles are simplistic in nature, they are extremely powerful for enhancing viewing experience [Cubbison 2005], and language learning experience [Grgurovic and Hegelheimer 2007]. However, one should note that individual learning style [Homer et al. 2008], and video content [Guo et al. 2014] greatly influence the effectiveness of learning with video.

In the context of using a video player, being "interactive" means that viewers frequently use the controls provided in the player to perform actions such as, pausing, rewinding, fast-forwarding, and resuming. On the other hand, "passive" means that the viewer performs fewer actions. For instance, one can argue that a movie theater is a video player ultimately designed for passivity since it severely restricts viewers' interactivity through means such as having dim lighting, a loud speaker system, and a set of restrictive social etiquettes. Prior research demonstrates that interactivity can either facilitate or hinder learning. Schwan and Riempp [2004] argue that if a video player allows the learners to have meaningful interactions, they may be able to learn better. However, they also warn that if the interactions are not well thought-out, additional interactivity may instead hinder learning.

1.1 Traditional Video Players

We define a traditional player to be a video player with most of its controls and options available at the bottom edge of the player. Since it assumes that viewers do not frequently interact with its interface, some of its features are not easily accessible. The YouTube video player (Figure 1.1) is one of more prevalent examples of traditional video players. While traditional video players might have a simple and sleek appearance, they can still be powerful. For example, the YouTube player has an ability to provide automatically generated subtitles.

1.2 Casual Language Learning

Since there is no clear standard definition for casual language learning, we define it to be a type of language learning where learners try to adapt existing everyday activities into opportunities for learning. However, the learners also emphasize deriving entertainment from the activities over acquiring new knowledge. As such, they may not be very diligent. We can classify casual

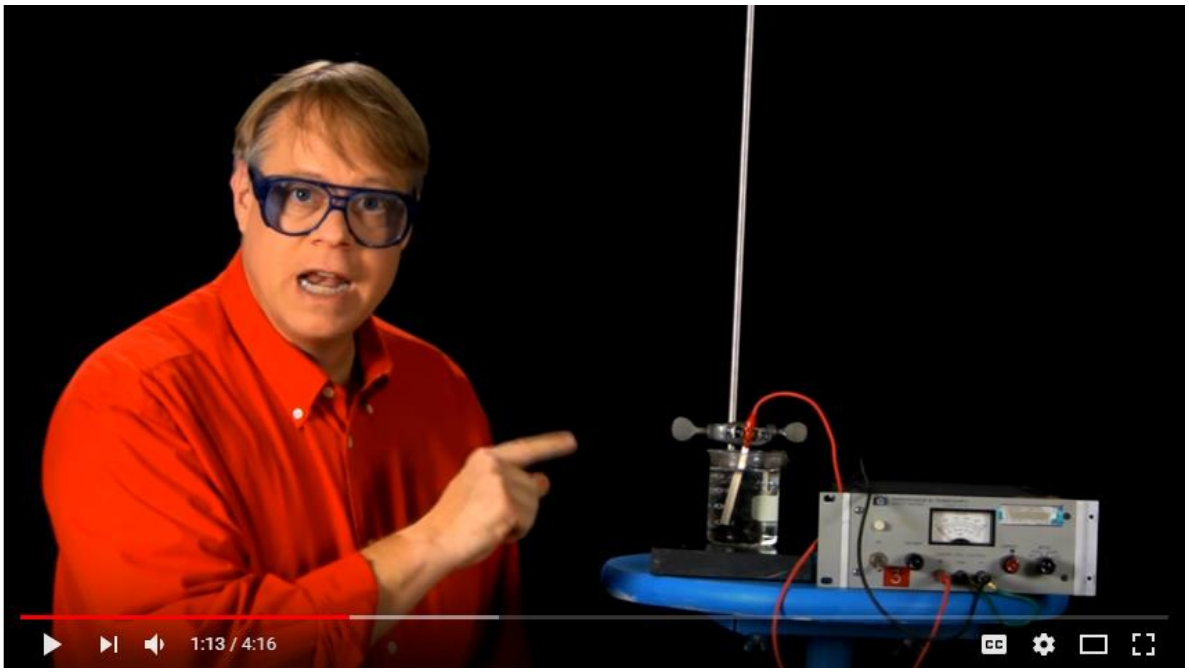


Figure 1.1. A screenshot of the YouTube video player with a clip from [Hammack 2012a]**.**

language learning as a subtype of informal learning which occurs outside a classroom and does not follow any curriculum.

1.3 Primary Research Questions

The primary research question is how do we design video players for casual language learning? Throughout this work, we provide several designs of video players that improve the experience of casual language learning. The secondary question is how do people use video players for language learning? This question is important and supplementary to the first one. During the evaluations of our video players, we recorded how our participants used the players as well as their feedback. We improved later the designs of our players based on their usage patterns and their suggestions.

Figure 1.2. Kalgan I's interface with a clip from [Hammack 2017]**.**

1.4 Kalgan: The Video Player for Our Research

To address the primary research questions, we created two video players. We named these video players Kalgan [1] I and Kalgan II.

We designed Kalgan I to resemble a traditional video player, so new users can easily learn to master its interface. Kalgan I includes some touch gesture interactions; for example, the learner can swipe up on the screen in order to change the subtitles. It also has some subtitle-based features such as interactive subtitles that allow learners to quickly look up a meaning of a subtitle word by tapping/clicking on it. We evaluated Kalgan I inside a laboratory where the participants had to watch some short videos with Kalgan I and a control video player.

[1] Kalgan is the former name of Zhangjiakou, China. It is based on a Mongolian word meaning “the Gate” or the “Frontier” [The Editors of Encyclopædia Brittanica 1998].

Figure 1.3. Kalgan II's interface with a clip from [Hammack 2012b]**.**

Kalgan II has a design that is different than that of a traditional video player, because we adopted some design suggestions from some prior work. For example, we adopted a numpad grid arrangement for Kalgan II since there is some evidence that the arrangement is more user-friendly [Wang and Ren 2009]. We also altered the evaluation process. Instead of conducting the evaluation solely in a laboratory, we used a combination of an online study, interviews, and a think-aloud protocol. We decided to change the evaluation method because we wished to have better triangulation of data. The results demonstrate that the interface of Kalgan II is generally helpful for language learning. However, they also highlight opportunities for further improvement. We also found that Cognitive Load Theory (CLT), a learning model from the field of psychology proposed by Sweller [1994], may also be able to explain some of the results.

Although we designed Kalgan I and Kalgan II mostly independently from other existing video players, their sets of features still some converge upon those that exist in video players

such as Yabla [Cabello 2013], GliFlix [Sakunkoo and Sakunkoo 2009], ViVo [Zhu et al. 2017], and a video player described by Kim et al. [2014]. The major difference of our players and these players is in design philosophy. We prioritize identifying and simplifying existing tasks that a learner would perform on traditional video players, instead of adding new learning mechanisms and activities like most of the aforementioned players.

1.5 Contributions

The thesis makes several contributions. First, we provide several video player designs specifically for casual language learning. These video players contain features that address some of the problems using traditional players for language learning. For example, they allow the learner to more precisely rewind so the learner can better recover the information that they have missed. We also identified how people may interact with a video player to learn a language. For example, we found that people usually turned on the subtitles as soon as possible. Furthermore, we also identified several types of intentions or reasons for learning a language.

1.6 Thesis Overview

- **Chapter 1 (This Chapter).** We provide an outline to what we plan to accomplish in this thesis as well as providing some basic definitions.
- **Chapter 2.** We present the related work that is both relevant to the creations and the evaluations of Kalgan I and II.
- **Chapter 3.** We introduce the design of Kalgan I and its evaluation. We also discuss the consequence of designing and evaluating Kalgan I.
- **Chapter 4.** We showcase the design of Kalgan II.
- **Chapter 5.** We present the evaluation of Kalgan II and its results.

- **Chapter 6.** We provide a broader discussion for the whole thesis. We also present the prototype of Kalgan III. Kalgan III is a hypothetical video player that we would like to develop if we have more resource.
- **Chapter 7.** This chapter contains the conclusion of the thesis.

Chapter Two: **Related Work**

Since analyzing and designing a video player is multidisciplinary in nature, we explored work in many fields such as education, and computer science. Here, we introduce concepts and research in learning, Computer-Assisted Language Learning (CALL), and interaction design.

2.1 Learning

There are many relevant concepts in learning. First, there is the concept of classifying learning into multiple types. This is important because each type of learning has a different set of stakeholders. For example, when designing a tool for informal learning, an instructor may not be present as one of the stakeholders as informal learning does not involve any instructor. Next, there is the concept of providing and receiving learning feedback which prior research argues to be important for acquiring knowledge. Then, there is the concept of using videos as a learning tool. Existing work suggests that while videos have their own advantages, they also have their own pitfalls. For example, while viewing a video is simpler than attending a class, individual learning style may prevent learners from effectively learning from the video. Lastly, learning platforms, such as Massively Open Online Courses (MOOCs) and specially designed video players, are also quite important, because they might also have additional augmentation that enhances learning experience.

2.1.1 Formal, Informal, and Casual Learning

The most important classifications of learning are: formal and informal learning. Before discussing casual learning, it is useful to first discuss the difference between formal and informal learning. The most important difference between both of them is that formal learning occurs within a class setting and that there is an instructor and a curriculum [Sockett 2014]. Sockett

[2014] argues that defining informal learning is not so straightforward. He points out while the consensus is that informal learning happens outside of the classroom with no structured curriculum, there is no clear and precise definition of informal learning that everyone agrees on. Sockett also argues that there are multiple types of informal learning. He states that Shugurensky thinks there are two main variables for informal learning: intention and awareness. Intention indicates whether the learner is the one initiating the learning process while awareness indicates if the learner is aware that they are learning. Shugurensky also states that there are three subtypes of informal learning: (1) self-directed study, where the learner has an intention to learn and is aware, (2) incidental learning, where the learner does not start the process but is aware, and (3) tacit learning/socialization where the learner does not seek to learn and is unaware. Sockett argues that in an informal learning, a learner may have multiple intentions that conflict with each other. For example, a learner may want to prioritize leisure over learning a language when watching a foreign film.

While casual learning is a type of informal learning, there is no precise definition for this term as each work defines this term differently. For example, Pivec [2007] defines casual learning as one that occurs when someone plays a video game. On the other hand, Haefner [1932] defines casual learning to be like incidental learning. We choose to define casual learning as a type of incidental intentional learning that happens by performing everyday tasks, and where the learner decides to prioritize leisure over acquiring new knowledge.

2.1.2 Feedback

Feedback is important for language learning as Liddicoat and Scarino [2013] argue that learners can reuse it to improve future learning. Łuczak [2017] specifically argues that if a learner acquires a list of new vocabulary, they can turn it into flashcards – a common language learning

tool. Furthermore, Qassemzadeh and Soleimani [2016] found that feedback from a piece of autocorrection software can better assist Iranian learners in learning English passive structure in certain cases.

2.1.3 Interactive Learning with Videos

Although interactive learning with videos does not need to involve language learning, work in this area is still quite useful. The work can still provide insight and guidance for creating a video player for language learning. Schwann and Riempp [2004] present a study where they asked the participants to learn about tying nautical knots. There were two groups of participants. The first group could not interact with the video, except for re-watching the video from the beginning. The second group could interact more with the video; they could rewind, pause, and perform other actions. Schwann and Riempp found that although both groups spent approximately the same amount of time learning, the second group developed a deeper understanding. Another study by Zhang et al. [2006] suggests that using a video player that encourages interaction may feel more satisfying than using a video player with less interactivity.

Although there are benefits with interactive learning, there are also disadvantages. First, Schwan and Riempp [2004] warn that poorly designed interaction can hinder learning instead of enhancing it. Secondly, Homer et al. [2008] conducted a study which involved learning with a video recording of a university-level lecture. They found that while some participants liked learning with the video recording, they also found some participants who disliked learning with the video. These participants stated that they preferred other learning methods such as listening, and reading. Homer et al. concluded that an individual's learning style affects how well they can learn from a video.

2.1.4 Evaluation

Evaluating a Piece of CALL Software

Evaluation within the field of CALL can use techniques in HCI such as quantitative studies, qualitative studies, and think-aloud protocols [Levy 2015]. However, since CALL also involves language learning, it may also use other evaluation methods found in Second Language Acquisition. For example, Qassemzadeh and Soleimani [2016] conducted a study which used Grammarly, a piece of autocorrection software as a language learning tool. Instead of focusing on the HCI aspects of Grammarly (such as usability of the software), they focused on improvement in language performance.

Evaluating Language Proficiency of a User

Evaluating someone's language proficiency can provide some insight on how a piece of CALL software influence someone's learning process. This can be either very easy or very difficult – depending on the research. While Maki et al. [2003] argue that it is possible to evaluate someone's proficiency in Japanese using simple quizzes, Başaran and Köse [2013] found that they could not find whether subtitles are helpful for language learning through a quiz. On the other hand, Winke et al. [2010] found that languages can actually affect how subtitles benefit learners. It seems that there are many variables that can influence whether the evaluation will be effective or not.

2.1.5 Learning Platforms

Massively Open Online Courses (MOOCs)

MOOCs are a type of online course which is usually free to participate and is open to everyone. Therefore, a MOOC can have a massive number of students. Due to having a large student body, a MOOC instructor is often unable to assist all students, and the study becomes self-directed and

self-paced [McAuley et al. 2010]. A MOOC may also use videos as a mean of delivering its class content [Guàrdia et al. 2013]. Because of this, MOOC research often contain work relevant to designing a video player for learning. For example, a MOOC paper by Li et al. [2015] suggests if students frequently interact with a video, then that video may be difficult for the students to understand. In another paper, Guo et al. [2014] argue that a MOOC's videos can affect how students learn. They noted that certain MOOCs, such as those provided by Khan Academy, are successful, because their videos have certain styles that students find engaging. Finally, Krzysztof et al. [2015] directly analyzed and researched interactions with video players in hope that they can create a better video player interface for MOOCs.

Video Players Designed for General Learning

There are video players designed specifically for learning, rather than for entertainment. For example, Rich and Hannafin [2008] present various video players designed to allow learners to communicate with their instructors through annotations. Additionally, Ullrich et al. [2010] created a video player that students of a university class had to use to participate the class's lectures. Furthermore, Kim et al. [2014] designed a video player specifically for MOOC. The video player includes data visualization tools that allow the learner to better analyze a video's content. Although the video players that Schoeffmann [2010] presented may not be strictly for learning, some of the players that he presented still have features similar to those present in Kim et al.'s player.

2.2 Computer-Assisted Language Learning (CALL)

Language learning is a specific kind of learning where the learner attempts to become more proficient at a language. Some learners may use technologies such as the Internet or language learning software. CALL is a multidisciplinary field that deals with creating software for

language learning [Gündüz 2005]. Examples of CALL software include Duolingo (a language learning website) [Garcia 2013], and Tactical Iraqi (a 3D game designed to train American soldiers stationed in Iraq) [Peterson 2010]. Work in CALL suggests that (1) a learner may perform better if they can obtain instantaneous translations of the words that they are trying to learn [Taylor 2013] and (2) that subtitles are very helpful for language learning [Grgurovic and Hegelheimer 2007]. CALL also has a subfield called Mobile-Assisted Language Learning (MALL) which specifically deals with language learning on mobile devices [Park 2014].

2.2.1 Subtitles

Subtitles is a method of translating foreign videos that preserves the original audio. They are relatively inexpensive and are easy to produce when compared to other translation methods such as voice over [Cubbison 2005]. Grguvoric and Hegelheimer [2007] argue that they can also facilitate language learning.

Knowing the time required for a person to read a subtitle can be helpful when designing a video player that has subtitle-based features. Jensema [2000] conducted a study which tried to find the time needed for people with hearing disability to read a subtitle. Similarly, Bisson et al. [2014] conducted a study to find the time required for a person to read a subtitle using an eye tracker. They found that the time required was approximately 2-3 seconds.

2.2.2 Subtitles-Based Video Players for Language Learning

There are also video players that specifically target language learning. Many of these video players have features that use subtitles in some way or another. Yabla is an example that is also commercially available. It allows learners to quickly look up the meaning of a subtitle word by clicking/tapping on it [Cabello 2013]. Although the creator of Yabla does not use any research to justify its effectiveness, Cabello [2013] argues that Yabla still applies some of the best practices

in CALL. In addition to Yabla, there are also experimental video players that use subtitles to improve language learning experience. GliFlix [Sakunkoo and Sakunkoo 2009] is an experimental video player that has word definitions embedded in the subtitles. There is also ViVo [Zhu et al. 2017], a video player that can highlight important words in the subtitles and has a word look up system. Finally, Kovacs and Miller [2014] created an augmented subtitles system that helps learners with acquiring Mandarin vocabulary.

2.3 Interaction Design

Since CALL also involves software design, interaction design is also quite important to CALL. There are several modes of interactions such as mouse or touch interaction. This section discusses designing a menu for multimodal interaction. It also talks about video players that are designed for less common interactions such as touch interaction with stylus.

2.3.1 Menu

In a traditional video player, the controls are mostly located at the bottom edge of the player. Their arrangement greatly resembles that of a horizontal linear menu design which the menu options are located side by side [Backs et al. 1987]. Backs et al. [1987] demonstrated that this layout requires more time for a user to reach an option than the vertical one (where options are stacked on top of each other). Tang [2001] further argues that linear menus also makes it difficult for someone discover sub-menus. On the other hand, Callahan et al. [1988] suggest that a better alternative is a pie menu (Example in Figure 2.4). The menu options for a pie menu are located radially and it takes the same amount of time for the user to reach each of the options. This means that it is very efficient for mouse interaction. For touch interaction, Wang and Ren [2009] provide some examples of menu layouts which are comfortable to use.

Figure 2.4. An example of a pie menu with arbitrary menu options.

2.3.2 Video Players with Novel Interactions

There are also innovative players that are designed for mobile and touch devices. The player created by Ullrich et al. [2010] is a mobile video player for Nokia Symbian OS devices. Similarly, Buchinger et al. [2010] created two video players for newer smart phones. The first one could only be controlled through tilting or rotating the device. The second one could only be controlled through touch gestures. In addition, Ramos and Balakrishnan [2003] mentioned an experimental video player created for touch devices and to be used with a stylus. Similarly, Hürst and Meier [2008] created a mobile video player that viewers could control by swiping a stylus across the screen.

Chapter Three: **Design and Evaluation of Kalgan I**

3.1 Introduction

Kalgan I is the first of the video players that we created for our research. It greatly resembles a traditional player so that learners can quickly become familiar with its interface. Since we implemented Kalgan I using HTML5, it is also multiplatform and should work on many devices. While we designed Kalgan I primarily based on our own experiences as language learners, our design also happens to share features with video players found in prior research [Cabello 2013; Taylor 2013; Grgurovic and Hegelheimer 2007; Kim et al. 2014; Sakunkoo and Sakunkoo 2009].

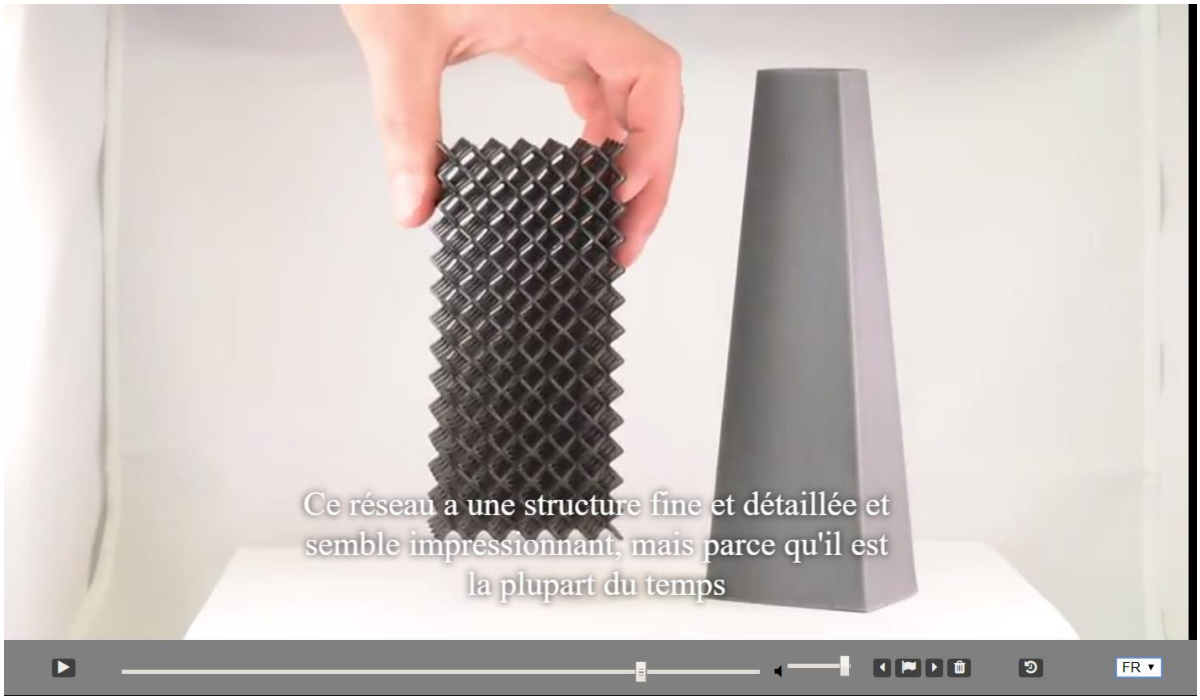


Figure 3.5. Kalgan I with French subtitles and a clip from [Hammack 2017]**.**

Figure 3.6. The interactive subtitles in action. Clip from [Hammack 2017]**.**

3.2 Features for Language Learning

3.2.1 Interactive subtitles and word history box

The learner can look up the definition of a word in the subtitle by tapping or clicking on it. Once tapped or clicked, Kalgan I uses the Yandex Translator API to translate the word and shows its definition in a blue band above the subtitle. Kalgan I also pauses the video so the learner can record the word. Kalgan I also adds the word and its definition into the word history box. Yabla [Cabello 2013], and ViVo [Zhu et al. 2017] have features that are similar to this one.

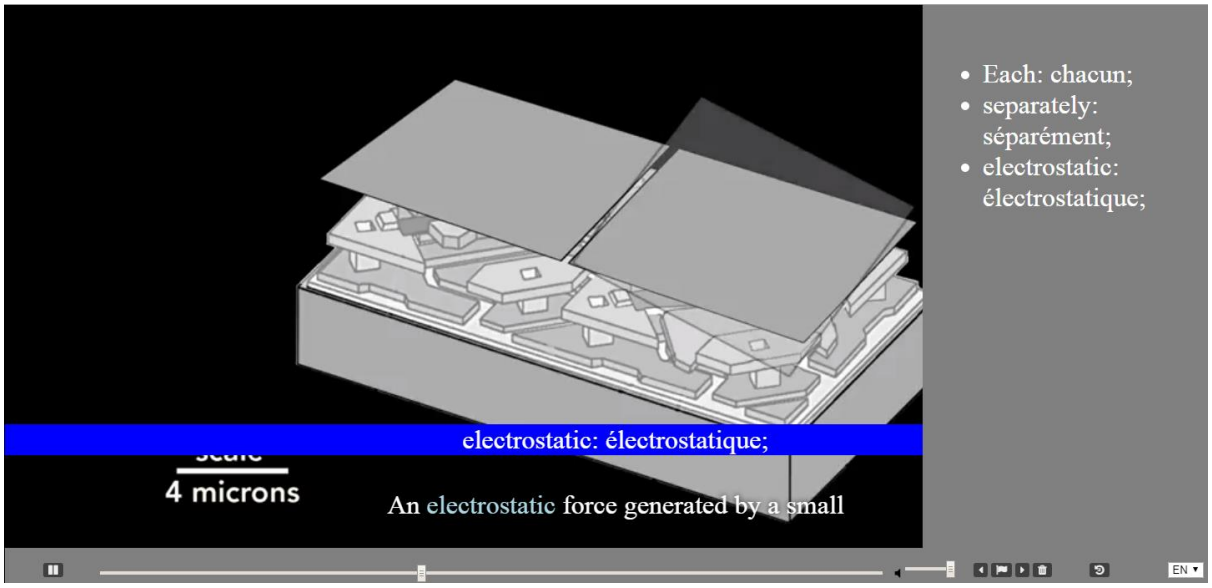


Figure 3.7. The word box history turned on. Clip from [Hammack 2017]**.**

3.2.2 Easy subtitle change

Kalgan I allows the learner to quickly change the subtitle language or turn off the subtitles by swiping up on the screen. Kalgan I changes the subtitles setting based on the current setting. The list below summarizes what the player will do when the learner swipes up.

- If the subtitles are off → Turn on subtitles in the learner's proficient language.
- If the subtitles are in the learner's fluent language → Change to the learning language.
- If the subtitles in the learning language → Turn off the subtitles.

3.2.3 Flagging

When the learner feels that the video is progressing too quickly for them to interact in time, they can either tap/click the flag icon in the control bar (Figure 3.8) or swipe down the screen. This will leave a small flag on the video timeline as a marker. Tapping/clicking on the left arrow in Figure 3.8 rewinds the video to the closest flag left to the current position. On the other hand, tapping/clicking on the right arrow will fast-forward the video the closest right flag.

3.2.4 Quick Rewind/Fast-forward

When the learner wants to rewind for a short amount of time (5 seconds), they can swipe left on the screen. If they swipe right, they will fast-forward for 5 seconds. The ability to quickly rewind allows the learner to quickly recover information from the sections of the video that they have missed. For example, if the learner missed a phrase uttered by a character in the video, they can swipe left to recover this information.

Figure 3.8. The flag controls.

3.2.5 On-screen Graphics

When the learner performs one of the following actions: quick rewind/fast-forward, changing subtitles setting, or planting a flag, an icon appears on the top-left corner of the screen to indicate that the action has been performed. Since swiping, by default, is not very visible, the icon provides a visual clue to what the learner has done. Figure 3.9 contains the on-screen graphics and how they can appear on the screen.

3.2.6 Summary of Swipe Interactions

There are four swiping directions. Swiping in each of the four direction causes the player to perform a certain action. The list below provides a summary of the swipe directions and how they affect the player.

- **Up.** Change or turn off the subtitles.
- **Down.** Insert a flag on the timeline.
- **Left.** Quick rewind.
- **Right.** Quick fast-forward.

Figure 3.9. LEFT: On-screen graphic in action, clip from [Hammack 2017]**. RIGHT: All on-screen graphics.**

3.3 Evaluation

The goal of the evaluation is two-fold: (1) to see if the first prototype is superior to a normal video player, and (2) to see if a learner can improve their language learning performance using Kalgan I. We conducted the study in a laboratory where we asked the participants to use Kalgan I and a control video player. We removed the following features: swiping, word history box, and interactive subtitles. After the participants watched a video, they completed a two-minute quiz. After they watched three or six videos, they completed a questionnaire.

3.3.1 Participants

We recruited participants with knowledge in French using posters attached to many notice boards at the University of Calgary. Since French is the second official language of Canada, it was easy for us to find participants with some French knowledge. We collected results from 12 participants, and we paid each participant $20 for completing the study.

3.3.2 Apparatus

To encourage the participants to use the swiping interactions, we conducted the study using a Microsoft Surface Pro 4 which is a touch tablet device. We also created a control video player by removing some features of Kalgan I. The features that we removed were: swiping, interactive subtitles, word history box, and flagging. Without these features, the control player behaved like a basic video player.

We also prepared six short video clips (around 5-10 minutes) from three French video sources: TED x Talks at Cannes [Dall'Aglio 2012; Meunier 2014], an animated film named "La Ballade des Dalton" [Goscinny, René et al. 1987], and a film named "La Classe américaine" [Hazanavicius and Mézerette 1993]. For the sake of convenience, we assigned the following names for these clips:

- **TED 1.** A clip from “happiness is a matter of choice” by Menuier [Meunier 2014].
- **TED 2.** A clip from “Comment sauver l'amour?” by Dall’Aglio [Dall’Aglio 2012].
- **Dalton 1.** The first clip from “La Ballade des Dalton.”
- **Dalton 2.** The second clip from “La Ballade des Dalton.”
- **Classe 1.** The first clip from “La Classe américaine.”
- **Classe 2.** The second clip from “La Classe américaine.”

For each clip, we also created a 2-minute quiz. Each quiz had two components: (1) four questions that asked the participants to translate French words into English, and (2) four short-answer questions based on the content of the clip. The quizzes are available in Appendix A.

In addition to the quizzes, there were also questionnaires designed for the participants to rate features of Kalgan I and the control player. The participants also had opportunities to provide their opinions of the players on the questionnaires. The questionnaire for the control player consisted of Likert-scale questions asking them about how they felt using the player for the videos. The questionnaire for Kalgan I contained questions similar to those of the control player questionnaire plus questions about the features specific to Kalgan I. The questionnaires are available in Appendix A.

3.3.3 Procedure

The two groups of participants performed largely similar tasks. The only difference was the order of the video players that they had to use. The first group used the control video player before Kalgan I while the opposite was true for the second group. The table below summarizes the tasks that the participants had to perform.

Table 3.1. The Summary of the Procedure for Kalgan I Evaluation.

Tasks	**Player for Group I**	**Player for Group II**
Watched **TED 1**.	**Control**	**Kalgan I**
Completed a 2-minute quiz based on **TED 1**.	N/A	N/A
Watched **Dalton 1**.	**Control**	**Kalgan I**
Completed a 2-minute quiz based on **Dalton 1**.	N/A	N/A
Watched **Classe 1**.	**Control**	**Kalgan I**
Completed a 2-minute quiz based on **Classe 1**.	N/A	N/A
Completed a questionnaire on the video player used to view the last 3 videos.	N/A	N/A
Watched **TED 2**.	**Kalgan I**	**Control**
Completed a 2-minute quiz based on **TED 2**.	N/A	N/A
Watched **Dalton 2**.	**Kalgan I**	**Control**
Completed a 2-minute quiz based on **Dalton 2**.	N/A	N/A
Watched **Classe 2**.	**Kalgan I**	**Control**
Completed a 2-minute quiz based on **Classe 2**.	N/A	N/A
Completed a questionnaire on the video player used to view the last 3 videos.	N/A	N/A

3.4 Results and Discussion

In this section, we summarize the results of the study, and discuss them. Since the sample size is small and not normally distributed, we use non-parametric statistics. We report the medians and their confidence intervals which are 95% bootstrap confidence intervals. We use 10,000 iterations for each bootstrap.

3.4.1 Numbers of Interactions

The median number of interactions performed by Group I when using the control player was 15.0 (95% CI: [9, 164.5]). The median number interactions performed by of Group I when using Kalgan I was 58.5 (95% CI: [46.5, 480]). The median number performed Group II when using Kalgan I was 69.5 (95% CI: [23, 231]). The median number performed Group II when using the control player was 195 (95% CI: [66.5, 371]). More information is available in Figure 3.10. The

data hinted that the Group I may have been less interactive overall than Group II. This means that the order of introduction for video players may introduce an effect that influence how people interaction with the players later.

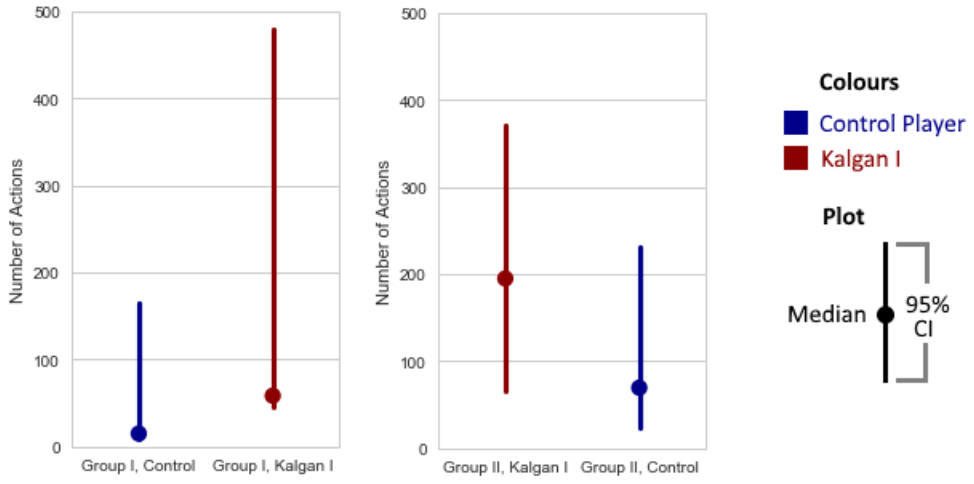


Figure 3.10. Frequencies of interactions by group and video player.

3.4.2 Improvement in French

We could not detect any improvement in French language skill from any participants. Instead, the score seemed to reflect the difficulty levels of the videos. For example, we found that the participants did not perform well on the quiz based on **Classe 2**. **Classe 2** was a very difficult video since it used slang, and it progressed very quickly. Overall, the median scores for Group I were: Q1 = 5 (95% CI: [2, 7]), Q2 = 3.5 (95% CI: [3, 5.5]), Q3 = 4.5 (95% CI: [3, 5.5]), Q4 = 4 (95% CI: [2.5, 7]), Q5 = 3 (95% CI: [2, 6]), Q6 = 3.5 (95% CI: [1, 5.5]). The media scores for Group II were: Q1 = 5 (95% CI: [3, 5.5]), Q2 = 3.5 (95% CI: [3, 5]), Q3 = 4 (95% CI: [3, 5.5]), Q4 = 4.5 (95% CI: [1.5, 6.5]), Q5 = 2.5 (95% CI: [1.5, 4]), Q6 = 2.5 (95% CI: [1.5, 4]). More information is available in Figure 3.11.

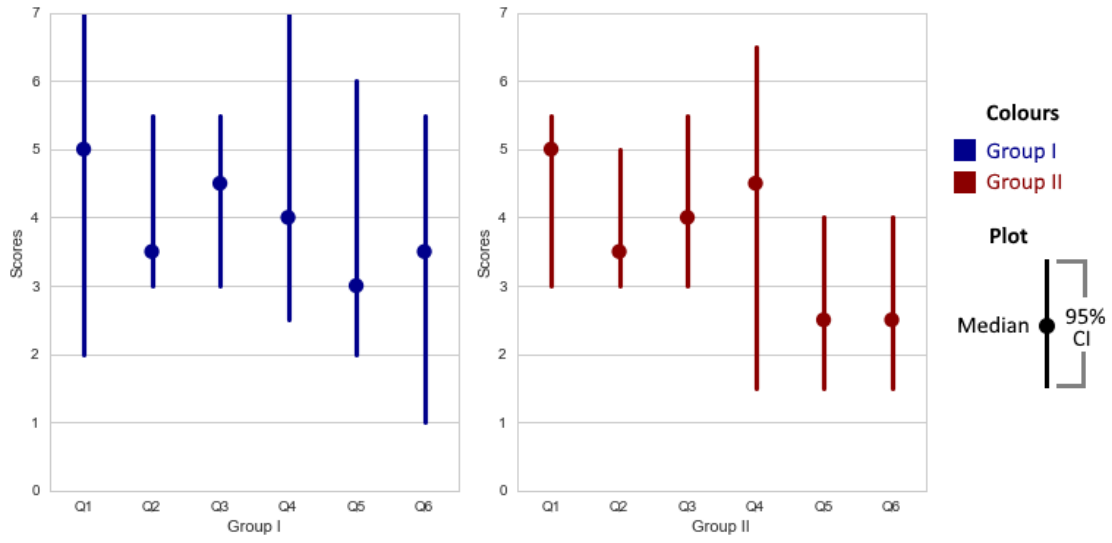


Figure 3.11. The distributions of quiz scores.

3.4.3 Fatigue

Each questionnaire contained a question which asked a participant to rank their fatigue level from one to seven. We found that the median fatigue scores for Group I were: 2 (95% CI: [1, 4.5]) for Questionnaire #1 and 3.5 (95% CI: [1, 5]) for Questionnaire #2. The median scores for Group II were: 2.5 (95% CI: [1, 5.5]) for Questionnaire #1 and 5 (95% CI: [2, 6]) for Questionnaire #5.

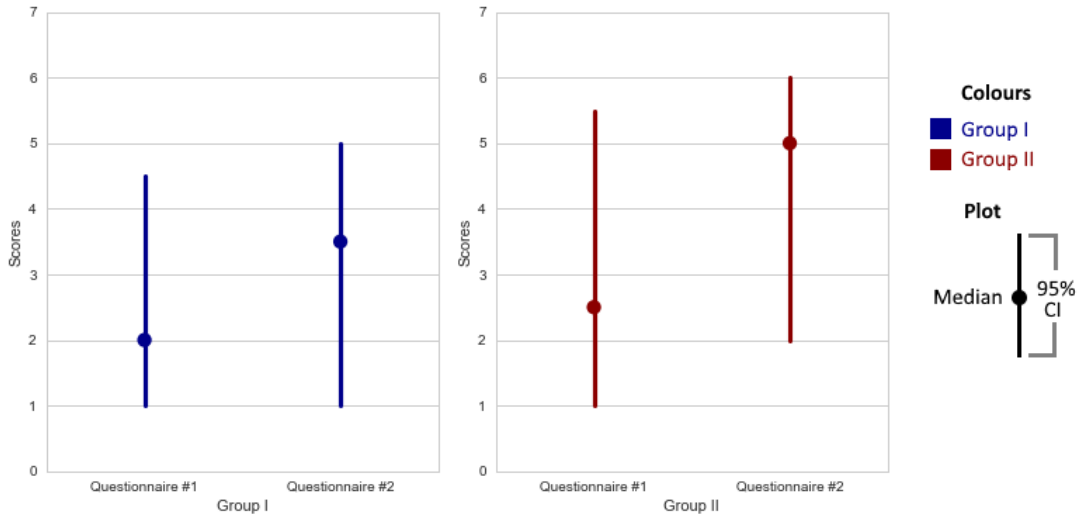


Figure 3.12. The distributions of fatigue scores by group and questionnaire.

3.4.4 Features of Kalgan I

On the questionnaire, we asked our participants to give a score between one to seven for each feature. The median scores for Group I were: Flagging = 6.5 (95% CI: [2.5, 7]), Interactive Sub = 6 (95% CI: [4.5, 6]), Pause after Using Interactive Subtitles = 5.5 (95% CI: [3, 6.5]), Quick Rewind/Fast = 6.5 (95% CI: [5.5, 7]), Swiping = 5.5 (95% CI: [4.5, 7]), Word History = 5.5 (95% CI: [3.5, 6.5]). The median scores for Group II were: Flagging = 3 (95% CI: [3, 6]), Interactive Subtitles = 6 (95% CI: [5, 7]), Pause after Using Interactive Subtitles = 5 (95% CI: [3, 6.5]), Quick Rewind/Fast = 6.5 (95% CI: [5, 6.5]), Swiping = 6 (95% CI: [2.5, 7]), Word History = 5.5 (95% CI: [3.5, 6]). Figure 3.13 provides a visualization of the statistics that we have just described. Since we found that quick rewind/fast-forward, interactive subtitles, and word history box have high medians, we decided to keep these features for Kalgan II. Since flagging had a low average, we decided to eliminate it from the future versions of Kalgan II. We also eliminated the pausing behaviour for the interactive subtitles for Kalgan II.

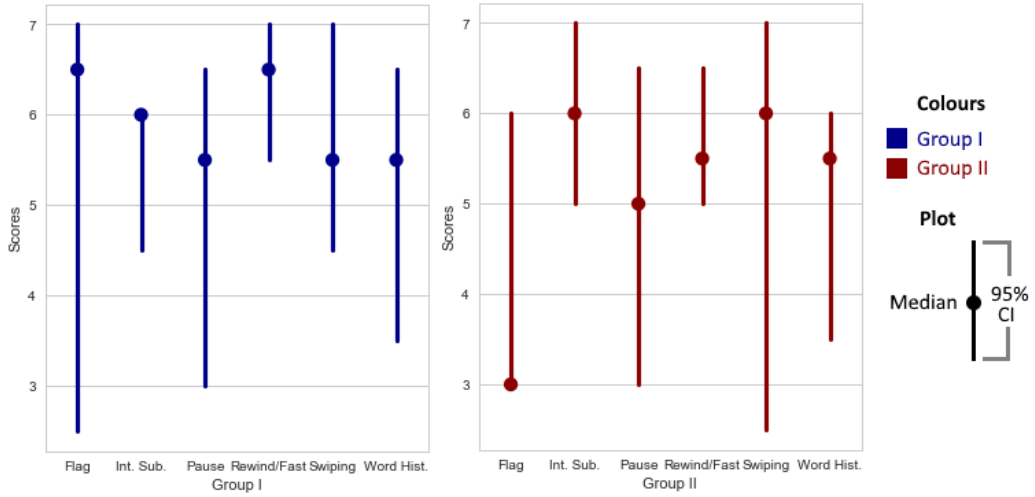


Figure 3.13. The distributions of scores given to Kalgan I features.

3.4.5 Qualitative Analysis

In the open-ended questions, we found that 9 out of 12 participants complained about the interactive subtitles not working well for them. However, the feature still received high scores. We also found that certain participants complained about the inability to make precise interactions since the videos were progressing too quickly. For instance, one participant complained that they could not change the subtitle quickly enough. Many participants also requested the ability to simultaneously show both English and French subtitles. Some participants found that pausing after tapping on a subtitle word to be inconvenient.

3.4.6 Issues of Using a Traditional Video Player

From prior research in CALL [Vanderplank 2010; Schwan and Riempp 2004] and from our own observations when designing and evaluating Kalgan I, we identify three issues that arise when someone attempts to use a traditional video player for casual language learning.

Issue #1: Difficulty Adjusting the Difficulty Level

A video may be difficult due to many reasons, such as the accents of the characters, the speed of conversation, or the use of slang. The learner's own background also influences how well they comprehend the video. Furthermore, the level of difficulty may fluctuate throughout the video [Canning-Wilson 2000]. Many traditional video players have mechanisms that make foreign language videos easier to comprehend such as subtitles and the ability to slow down the video. However, locating and adjusting these mechanisms can be difficult. Figure 3.14 demonstrates that a person may need to perform up to three steps to turn on the English subtitles. Prior research suggests that extra time performing an unnecessary task can hinder learning [DeLeeuw and Mayer 2008].

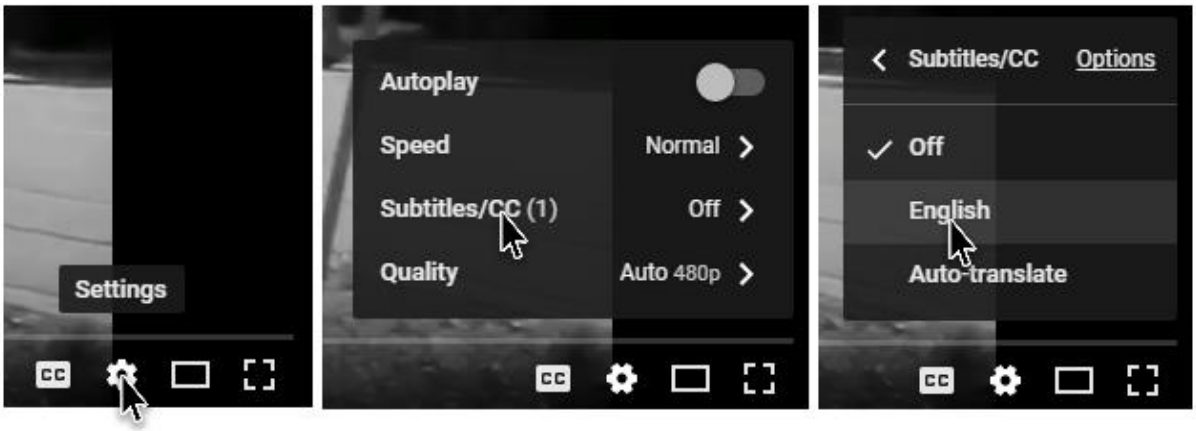


Figure 3.14. The steps required to change the subtitle language in YouTube.

Issue #2: Recovering Missed Information

Learners tend to miss information (such as words, phrases, and sentences) in the video, even when subtitles are present. As a result, video players designed for casual language learning need mechanisms for helping people to recover the information that they have missed. An alternative strategy for managing missed information is to prevent information from being missed to begin with. For example, some traditional video players allow learners to reduce the playback rate, so the learners can have more time to interact.

Issue #3: Tracking Learning Progress

Feedback is important in language learning [Qassemzadeh and Soleimani 2016; Nicholas et al. 2001; Liddicoat and Scarino 2013]. It can provide information on what people have learned and what they can do to improve. Some types of feedback can also be reused afterward in follow-on learning exercises. For example, when the learner generates a list of foreign words during a language lesson, they can later reuse the list by turning the words into flashcards [Łuczak 2017]. Traditional video players do not provide any feedback which means that the learner does not have any way of tracking their progress.

Chapter Four: **Design of Kalgan II**

Building on the findings from the evaluation of Kalgan I, we created Kalgan II. Kalgan II has many features that are available in Kalgan I. However, we also made some changes based on the findings. For example, we replaced quick rewind/fast-forward with the more precise subtitle-aware rewind/fast-forward. We also removed some features such as flagging since the participants of the Kalgan I study did not find it useful. We also replaced the Yandex translator API with the Microsoft Translator API, since the latter is more reliable. While Kalgan II has a very differently layout from a traditional video player, we designed it to be simple to use.

Figure 4.15. The general interface of Kalgan with a clip from [Hammack 2012b]**.**

4.1 Features and Layout of Kalgan II

Here is a brief introduction of the features in Kalgan II.

- **Figure 4.15-A.** Annotated timeline; when a learner interacts with the subtitle, a word appears beside the timeline.
- **Figure 4.15-B.** The 9-button menu which hosts most of the features.
- **Figure 4.15-C.** The current subtitle; when a learner taps/clicks a word, it displays the word's definition in the yellow box.
- **Figure 4.15-D.** The previous subtitle that appeared before the current one.
- **Figure 4.15-E.** The word history box with the list of words and their meanings that the learner has looked up.

Before playing a video, the learner selects two subtitle languages. Once the languages are chosen, the video will play right away. Most of the interface is initially hidden, but the learner can display it by tapping or clicking on the video. The learner can hide the interface once again by tapping or clicking on the video.

4.1.1 The 9-Button Menu

The 9-button menu contains the buttons that allow the user to change settings and control video playback. It is a rearrangement of the control bar that can be found at the bottom of traditional

Figure 4.16. The 9-button menu.

video players. We designed the menu to based on a touch-friendly design menu suggested by Wang and Ren [2009]. Since the menu resembles a pie menu (a menu whose options are located radially along a circle [Callahan et al. 1988]), it is also mouse-friendly.

Behaviours

When a learner triggers the menu, either by clicking or tapping on the video, the menu appears at the tapped/clicked location, with the play button (▷) centered at the point of the initial press. This behavior allows the user to pause or resume the player by double tapping or double clicking. The other menu buttons have the following behaviors:

Turn off the subtitles. The button is pink if the subtitles are currently off. In Figure 4.16, it is black, indicating that subtitles are on.

fr **Change subtitles to the language the user is learning.** In Figure 4.16, the button has "fr" on it, because the user has indicated that French is their learning language. The button is black, because French is currently not selected.

en **Change subtitles to the language in which the user is proficient.** In Figure 4.16, the button has "en" on it, because the user has indicated that English is their proficient language. The button is pink, indicating that English subtitles are currently on.

Subtitle-aware rewind. If subtitles are on, the button will rewind the video to the point where the previous subtitle begins. This functionality is inspired by the video players presented by Schoeffman et al. [2010]. If subtitles are off, the button ( ) instead

rewinds the video for 4 seconds — slightly longer than the typical subtitle-reading time (2-3 seconds) identified by Bisson et al. [2014].

Play button. Resume the video. The button assumes this appearance if the video is playing. Clicking on pauses the video and changes the button to .

Subtitle-aware fast-forward. The button will fast-forward the video to the beginning of the next subtitle. If the subtitles are off, the button ( )fast-forwards for 4 seconds.

Turn on previous subtitles. When this option is on, the video player will also display the previous subtitles as a smaller gray line (Figure 4.15-D) just above the current subtitle (Figure 4.15-C). The button is pink, because the option is currently on. Otherwise, it will be black. If the user turns off the main subtitles, the previous subtitles also disappear.

1x **Playing the video at full speed.** When this option is on, the video player will play at full speed. Since this option is currently on, the button is pink.

½x **Playing the video at half speed.** When this option is on, the video playback rate is reduced to 0.5x. Since this option is currently off, the button is black.

4.1.2 Interactive Subtitles

The learner can translate a subtitle word by tapping or clicking on it. When they do so, the word's definition appears in a yellow box. The learner can then click or tap the box to hide it. Figure 4.17 and 4.15-C contain screenshots of the interactive subtitles.

4.1.3 Previous Subtitles

Often, subtitles move too quickly for new learners, leaving them with too little time to read or interact with the text and forcing them to frequently pause and rewind the video. In response, Kalgan includes a feature that displays the previous round of subtitles as small gray text above the current one. As with the regular subtitles, the learner can translate words by tapping or clicking on them. Figure 4.17 and 4.15-D contain the screenshots of previous subtitles.

4.1.4 Word History Box

This panel sits at the edge of the screen and records a list of all the words a learner has translated while watching the video. Whenever the learner taps or clicks on a word in the subtitles, both the

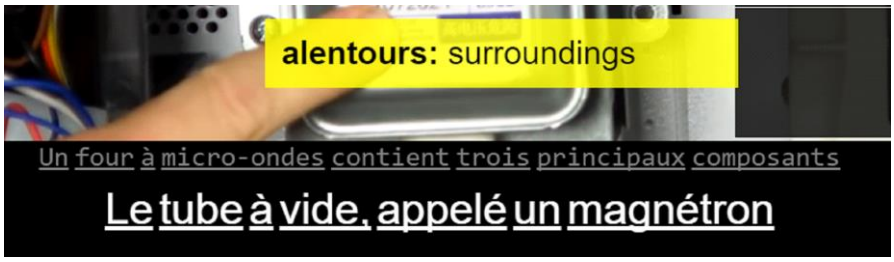


Figure 4.17. The interactive subtitles (Close-up of Figure 4.15).

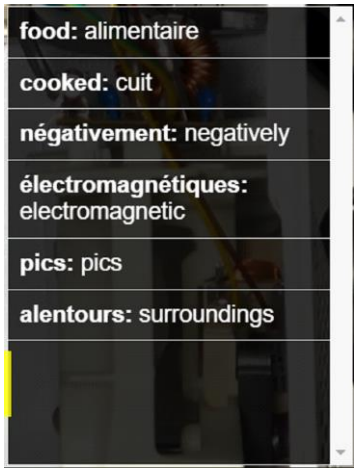


Figure 4.18. Word history box (Close-up of Figure 4.15).

word and its translation are added to the box. In Figure 4.17, the French word “alentours” and its English meaning “surroundings” were added after the learner tapped or clicked on “alentours”. This feature allows learners to track and revisit new words during and after the video. Figure 4.15-E and 4.18 contain screenshots of this feature.

4.1.5 Annotated Timeline

In contrast to the horizontal timelines used by most video players, Kalgan uses a vertical timeline positioned at the left edge of the screen. When the learner taps or clicks to translate a word in the subtitles, that word is added to the timeline at the current position of the playhead. The word appears in pink if comes from the language the learner already knows, and blue if it comes from the language they are trying to master. The added words serve as a record of what the learner has tried to learn so far. This feature is similar to the visualization feature found in a video player created by Kim et al. [2014]. Figure 4.15-A contains a screenshot of this feature.

4.2 Changes from Kalgan I

4.2.1 Features

We created Kalgan II after the evaluation of Kalgan I. Although Kalgan II appears to be a completely different video player, it contains many of the features found in Kalgan. Below is the list of features that are common to both Kalgan I and Kalgan II and how we upgraded them for Kalgan II.

- **Interactive Subtitles.** We replaced the Yandex Dictionary API with the more reliable Microsoft Translator API for Kalgan II.
- **Word history box.** While this feature has a different appearance in Kalgan II, it is functionally the same.

- **Quick rewind/fast-forward.** Instead of rewinding or fast-forwarding the video for five seconds, it now rewinds/fast-forwards for four seconds. We discovered that a person probably needs around two to three seconds to parse a subtitle [Bisson et al. 2014]. However, since Bisson et al. only conducted an experiment using English and Dutch subtitles, we did not know if the results would generalize to other languages. Therefore, we only reduced the time by one second for our rewind/fast-forward system, erring on the side of caution. Also, this feature is only accessible when the subtitles are off. When the subtitles are on, the learner must instead use subtitle-aware rewind/fast-forward.
- **Easy subtitle change.** The learner can still make a quick subtitle change. However, since Wang and Ren [2009] suggest that swiping is not efficient, the learner now changes the subtitles with the 9-button menu.

We also removed some features from Kalgan I, because research [Wang and Ren 2009] demonstrates that they were inefficient or the participants from the evaluation of Kalgan I did not like them.

- **Flagging.** We removed the feature, because the participants did not like the feature.
- **Swiping and On-screen Graphics.** Wang and Ren [2009] argue that swiping is not efficient and it can introduce fatigue. We created on-screen graphics to complement swiping gestures. Since the learner now always have a visual clue when executing a task, the player no longer needs the graphics.

4.2.2 Implementation

Because we wanted Kalgan II to support remote online studies, we made extensive changes to how we implement Kalgan II. Instead of using HTML5 to implement the video playback functionalities, we used the YouTube API. The API allows Kalgan II to access YouTube videos

which makes Kalgan II much more suitable for a remote study. We also replaced the Yandex Translator API with the more reliable Microsoft Translator API.

4.3 Comparing Kalgan II with YouTube Player Using Formal Methods

Dix [2003] discusses using formal methods as a tool to simplify usability research in HCI. He argues that by using logic and formalism, we can anticipate how someone uses an interface with less evaluation. He advocates for the use of flowcharts to represent how users interact with an interface. Additionally, we can quickly communicate the complexity of the interface as well as find ways to simplify it. Here, we use flowcharts to quickly compare how Kalgan II is better than the YouTube video at allowing its users to (1) quickly adjust the level of difficulty, (2) easily recover missed information, and (3) conveniently track their learning progress.

4.3.1 Adjusting the Difficulty Level

First, Kalgan II provides several different ways for learners to adjust the level of difficulty while a video is playing. By putting most of the features in a single menu (the 9-button menu), Kalgan II allows learners to turn off the subtitles or change the subtitle language with less tapping or clicking. Although traditional video players like the YouTube player also allow the learner to adjust the same settings, the menus are not as accessible and require more steps to reach. Figure 4.19 demonstrates the numbers of steps required to adjust the subtitles and the playback rate, both in Kalgan II and in the YouTube player. We can see that for Kalgan II, each action takes only one or two presses (depending on whether the menu is already visible). For the YouTube player, changes like turning on the subtitles and reducing the playback rate, can instead take as many as four steps to complete. If the YouTube player's menu is already visible, the actions will require fewer steps. However, the YouTube player will often automatically hide its menu, so the

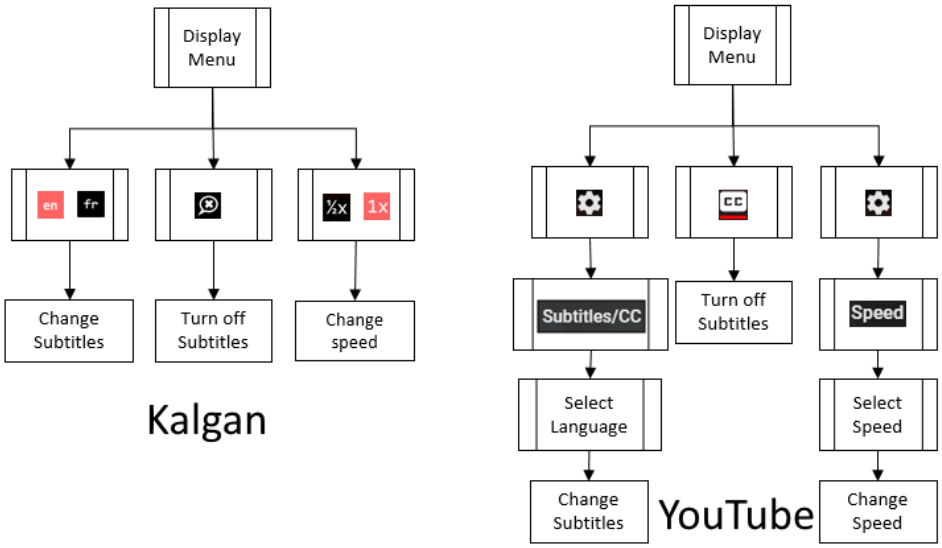


Figure 4.19. The flowcharts comparing adjusting the difficulty level in both players.

learner will frequently need to display it. While it is possible to turn on the subtitles with only two steps on the YouTube player, the user cannot quickly change the subtitle language.

Kalgan II's interactive subtitles also provide an easy mechanism for adjusting the level of difficulty, allowing learners to selectively translate difficult words or phrases between languages. This avoids the dichotomous choice offered by traditional players like YouTube, where learners must choose between seeing subtitles in the language they know or the one they are learning. Instead, beginners can choose to watch using subtitles in a familiar language and translate individual words to gradually increase their fluency in the new one. Similarly, more advanced learners can focus on subtitles in the new language, translating difficult words as they appear. This scaffolds the learning process and helps learners avoid over-relying on subtitles in the language they already know.

4.3.2 Recovering Missed Information

Kalgan's subtitle-aware rewind, previous subtitles, and playback rate adjustment all make it easier for learners to recover information that they might have otherwise missed in a traditional player. Subtitle-aware rewind allows learners to jump backward one subtitle at a time so they can precisely catch words or phrases they may have missed. In contrast, learners using the YouTube player must guess when the previous piece of dialogue began and try to rewind to that point. This process (Figure 4.20) can be cumbersome and disruptive, especially with long videos, where it can be difficult to make precise adjustments using the timeline.

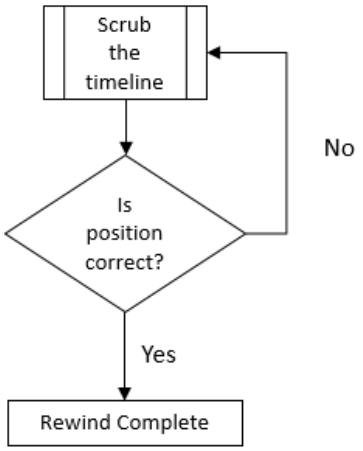


Figure 4.20. Emulating subtitle-aware rewind in the YouTube player.

Several of Kalgan II's other features, including previous subtitles and playback rate adjustment, do not directly help the learner to recover information, but instead, prevent them from missing it in the first place. Both give viewers more time to read and react to subtitles without pausing or rewinding the video. Interactive subtitles also help prevent disruptions by simplifying auxiliary tasks like looking up a word. In a traditional player like YouTube, these kinds of lookups require the viewer to navigate outside the player to consult a separate

translation service or dictionary (Figure 4.21). These transitions may require learners to either pause the video or cause them to miss subsequent content.

4.3.3 Tracking Learning Progress

Kalgan also contains several tools including the word history box and timeline that provide learners with additional information they can use to assess and review their progress. These have no analogue in traditional players like YouTube. Figure 4.21 contains a flowchart that emulates using the YouTube player to track learning progress.

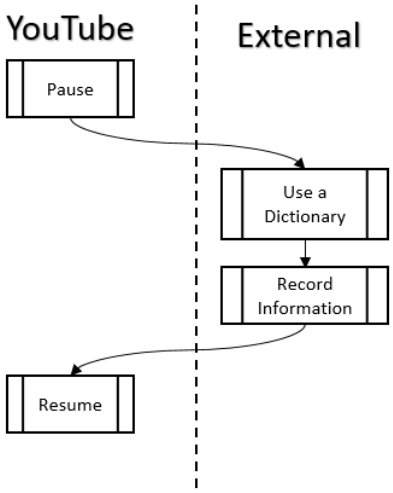


Figure 4.21. Emulating word look up in the YouTube player.

Chapter Five: **Evaluation of Kalgan II**

We evaluated Kalgan II using multiple methods including a remote online study, interviews, and a think-aloud protocol. We chose this approach, because we wanted to be able to track participants over a long period of time, and to collect some "in-the-wild" data. We also wanted to have multiple types of data so we could better triangulate our results. By having only one group of participants, we also did not need to deal with the kinds of ordering effects that we encountered in our evaluation of Kalgan I. The two research questions for the study were: (1) how do the features of Kalgan II assist in language learning, and (2) how do people use Kalgan II. The results of the evaluation demonstrate that the subtitle-based features are well-liked and are quite important for language learning. Furthermore, the results show while there are many factors that influence how people use Kalgan II, there are also some predictable patterns as well.

5.1 Differences from the Evaluation of Kalgan I

The evaluation of Kalgan II was very different from the one for Kalgan I. First, we included the use of an online study and a think-aloud protocol which allowed us to triangulate and obtain more accurate results. We also aimed to reduce the effect that we observed in the evaluation of Kalgan I. To do so, instead of asking our participants to use Kalgan II for less than an hour, each participant would use Kalgan II for a week or longer. Since Kalgan II used the YouTube API, we no longer needed to pre-select videos for the evaluation. We also did not need to target specific languages. Hence, we could recruit any language learner. We also eliminated language proficiency quizzes, because we found them to be unreliable and we wanted to focus more on understanding participants' behaviours.

Table 5.2. Participant information.

	Total Time Watching Videos hh:mm:ss	# of Videos Watched	Gender	First Languages	Languages Learned
P1	00:47:48	4	F	English	French, Japanese
P2	01:15:02	10	M	English, Polish	German
P3	02:17:07	17	M	Spanish	French
P4	01:17:35	12	F	Farsi	French
P5	00:50:38	3	F	Korean	English
P6	03:26:28	22	F	Spanish	English
P7	01:06:08	5	F	Farsi	English
P8	N/A	9	F	English	Japanese
P9	00:37:33	2	M	Mandarin	English
P10	02:07:04	12	M	English	French

5.2 Evaluation with Human Participants

5.2.1 Participants

We recruited 10 participants from University of Calgary using posters and online advertisements on social media. To be eligible for the study, the participants had to indicate that they were interested in language learning. The participants could be from any language background and many of them did not have English as their first language. Each participant who completed the study received $20. Additional participant details including gender, and number of watched videos, are available in Table 5.1.

5.2.2 Procedure

The evaluation had two phases: a remote online study phase and an interview phase. In the online study phase, the participants used Kalgan II remotely for at least a week. They could use the video player at any time and at any location. During this period, the server tracked the actions that they performed. After this phase, we invited each participant into the lab for an exit

interview. At the end of the interview, each participant performed a short (approximately 5-minute) think-aloud protocol where they used Kalgan II in our presence.

5.2.3 Apparatus

We conducted the online study using a server hosted at University of Calgary. We used Sinatra, a Ruby web server library, to create the server backend. The server provided many facilities for the evaluation. First, it provided a system for a potential participant to book an interview and to give their consent. Second, it provided a training material to the participants. Third, it provided a list of recommended videos to the participants and an option for the participants to launch custom videos. We used a convertible tablet to record the interviews and to conduct the think-aloud protocol study.

5.2.4 Results

We collected a variety of data. During the interviews, we asked our participants to state their language learning strategies, the devices that they used for the study, and their feature preferences. We also derived some of our results using the server log and the think-aloud-protocol. Since our sample size is small and not normally distributed, we used bootstrap with 10,000 to generate the 95% confidence intervals.

General Statistics.

Learning Strategies. Most of the participants (6 out of 10 participants) indicated that they already spent time casually watching videos for language learning. This high number indicated that most of the participants were already comfortable using a traditional player for this kind of task. A smaller number of participants (4 out of 10) reported taking language classes, reading foreign texts (4 out of 10) or using language learning apps (4 out of 10). A few also

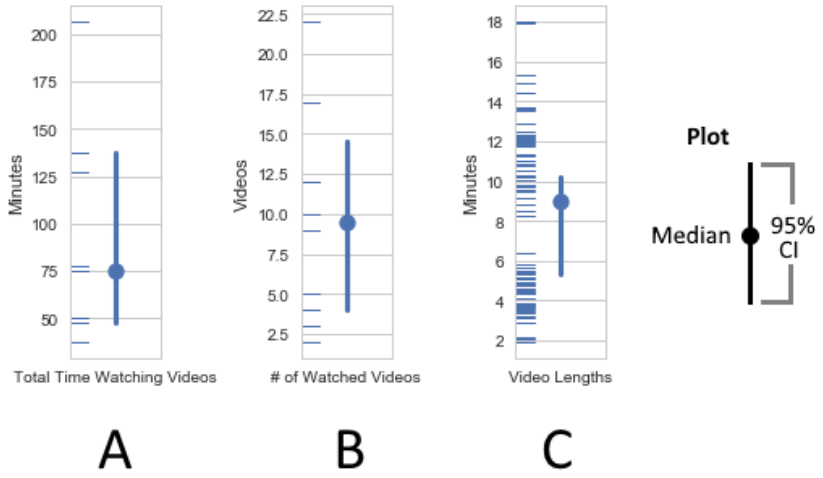


Figure 5.22. A: The distribution of the total time each participant spent watching a video. Each tick represents a participant. B: The number of unique videos watched by each participant. Each tick represents a participant. C: The distribution of the video lengths. Each tick represents a video.

reported listening to foreign music (2 out of 10), conversing with foreign language speakers (2 out of 10), and using flashcards (2 out of 10).

Devices and Inputs. Most of the participants (P2-P10) used laptop or desktop computers to complete the online component of the study. P1 used a convertible tablet and used a combination of mouse and touch inputs. P2, P3, P8, and P10 interacted with the video player using laptop touchpads. All other participants used mice.

Video Consumption. We computed total time spent by a participant by summing the lengths of the videos that the participant had watched, excluding re-watches (Figure 5.22-A). The median total time was 75 minutes (95% CI: [47.8, 137.1]). We also computed the number of unique videos that a participant watched (Figure 5.22-B). We excluded P8 from the calculation, because some of the videos that he watched were no longer available. The median of the number of videos watched was 9.5 videos (95% CI: [4, 14.5]). The median length of each video

consumed by the participants (Figure 5.22-C) was 9 minutes (95% CI: [5.3, 10.2]). Interestingly, the distribution of the video lengths seems to be bimodal.

Languages in Evaluation. Four of the participants attempted to learn French. Five participants who were likely international students attempted to learn English. Only one participant tried to learn German. Two of the participants who learned French told us that they originally tried to learn other languages (P3 planned to learn Danish and P10 planned to learn Korean.). However, since it was difficult for them to find videos in those languages, they decided to learn French instead. This outcome is not surprising because we conducted the study in Canada, a country whose official languages are English and French. P1 was an interesting case. She reported to us that as a linguistics student, she had to acquire rudimentary knowledge in many languages. As a result, while she stated that she knew nine languages, she was only fluent in a few of the languages.

Feature Preferences

To collect data from our participants regarding their preferences and dislikes toward Kalgan II, we asked our participants to perform two tasks during the interview. In the first task, we asked our participants to rate ten features of the player. The participants assigned a number between one to five to each feature – with one being the least useful and five being the most useful. In the second task, we asked our participants to rank-order the ten features in terms of their usefulness. A feature that receives a higher rank means that the participants deemed them to be more useful. The results of the two tasks are available in Figure 5.23. We also used the results from the open-ended questions (The questions are available in Appendix B.) and the think-aloud protocol to explain why the participants liked or disliked the features.

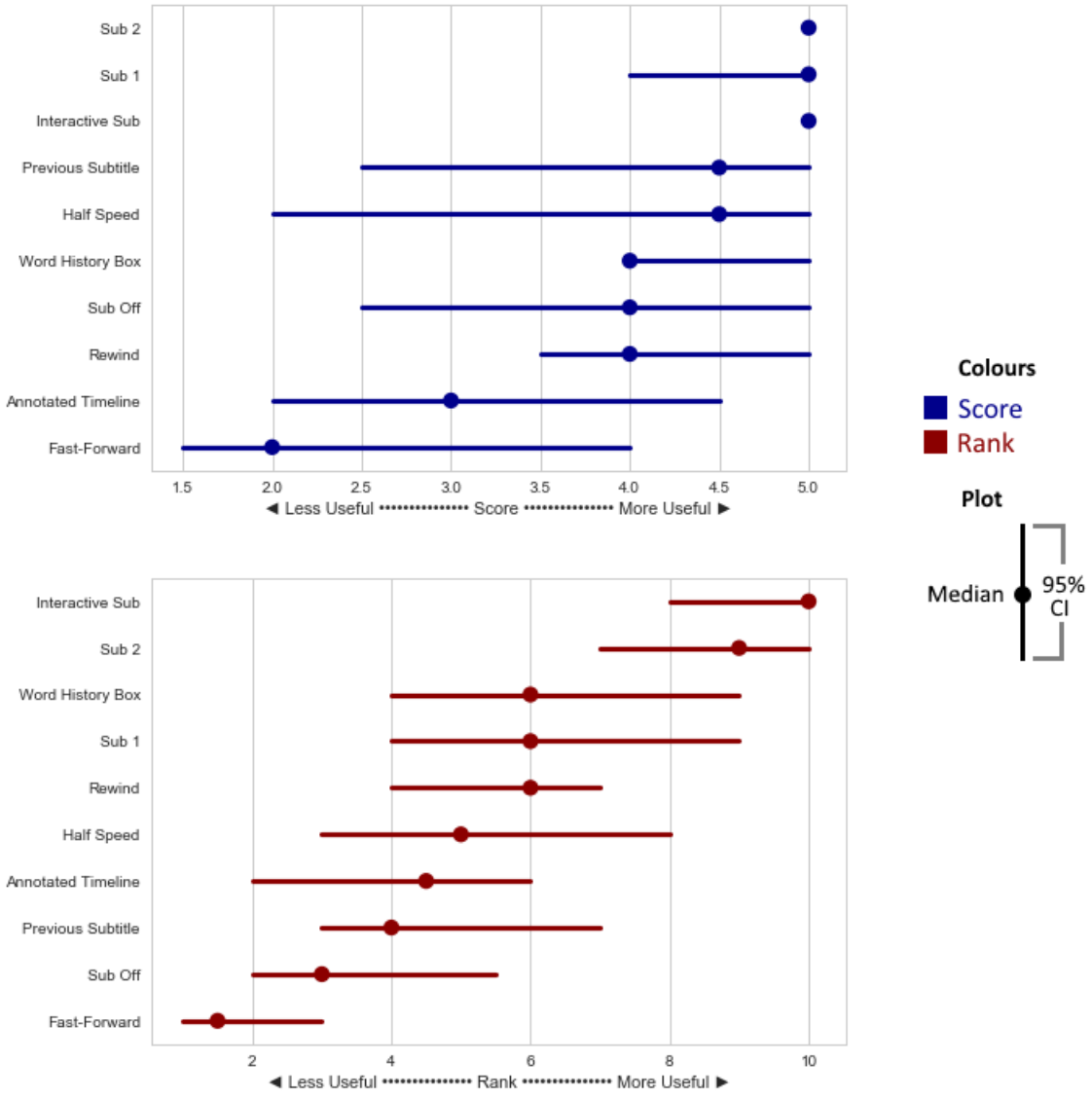


Figure 5.23. TOP: The distributions of scores for features in Kalgan II. BOTTOM: The distribution of ranks for features in Kalgan II.

- **Interactive subtitles.** The median score for this feature was 5 (95% CI: [5, 5]) and the median rank was 10 (95% CI: [8, 10]). Most of the participants identified the interactive subtitles as the most useful feature in the player. They found it convenient to have a language learning tool built right into the video player itself.
- **Previous subtitles.** This feature had the median score of 4.5 (95% CI: [2.5, 5]) and the median rank of 4 (95% CI: [3, 7]). We suspect this feature might not have been sufficiently discoverable since at least two participants (P3, P4) remarked that they were unaware of it.
- **Subtitle-aware rewind/fast-forward.** The median score and rank for subtitle-aware rewind were 4 (95% CI: [3.5, 5]) and 6 (95% CI: [4, 7]) respectively. The median score and rank for subtitle-aware fast-forward were 2 (95% CI: [1.5, 4]) and 1.5 (95% CI: [1, 3]) respectively. Interestingly, while we observed many participants (P1, P3, P5, P7, and P8) using subtitle-aware rewind during the think-aloud protocol, they had not mentioned it during the interviews. P4 commented that the rewind system could have been more precise. Subtitle-aware fast-forward, on the other hand, received very low scores and ranks. None of the participants mentioned it in the interviews nor used it during the think-aloud protocol.
- **Playing the video at half-speed.** The median score for this feature was 4.5 (95% CI: [2, 5]) and the median rank was 5 (95% CI: [2.5, 8]). All participants except P6 used this feature during the online study. P2 did not find this option useful because the video became slurred due to a flaw in the YouTube API. P5, on the other hand, found it to be helpful. One participant (P3) asked about the possibility of reducing the rate to 1/3 of the original speed.
- **Word history box and annotated timeline.** Word history box received a relatively high median score (M = 4, 95% CI: [4, 5]) and moderate median ranking (M = 6, 95% CI: [4, 9]). However, some participants (P2, P4, P7, and P9) told us that they wished word history box

could have been more powerful; for example, word history box could have a button that allowed learners to easily export its content. Meanwhile, the annotated timeline had a lower score median (M = 3, 95% CI: [2, 4.5]) and a lower rank median (M = 4.5, 95% CI: [2, 6]).

- **Subtitle settings.** The median score and rank for turning on the subtitles in the learner's proficient language (Sub 1 in Figure 5.23) were 5 (95% CI: [3.5, 5]) and 6 (95% CI: [4, 9]) respectively. The median score and rank for turning on the subtitles in the learning language (Sub 2 in Figure 5.23) were 5 (95% CI: [5, 5]) and 9 (95% CI: [7, 10]) respectively. We find that these scores and ranks are higher than the ones for turning off the subtitles (Sub Off in Figure 5.23). The median score and rank for turning off the subtitles were 4 (95% CI: [2.5, 5]) and 3 (95% CI: [2, 5.5]) respectively.

Intrusiveness of the Player

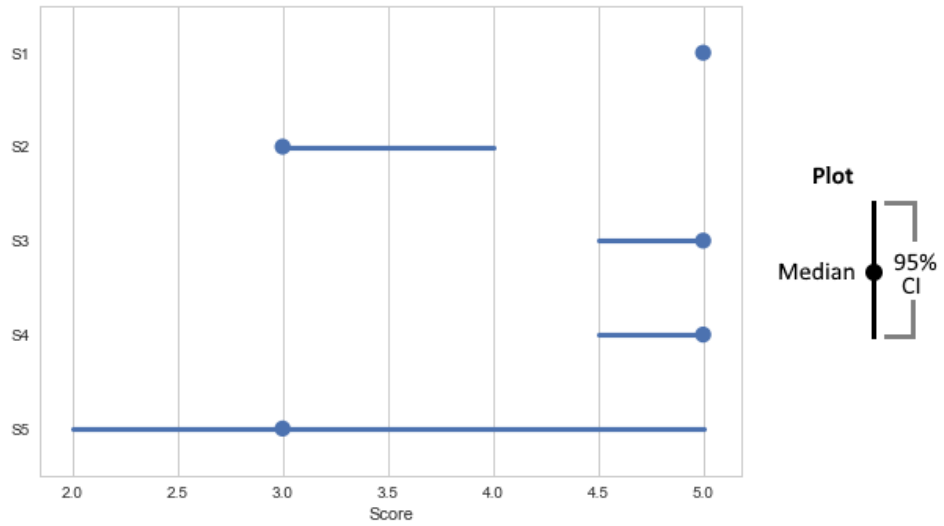
A traditional video player usually aims not to be intrusive to its users. Therefore, it has a minimalistic design that sometimes can hide itself very quickly. Kalgan II, on the other hand, is not as minimalistic. Even though learners can hide the interface when they do not need it, we still believed that it could be distracting. As part of the study, we asked the participants to indicate whether Kalgan II's interface was intrusive to them or not. All participants, except P5, P6, and P10, indicated that the intrusiveness of the player was not relevant since they could simply hide the menu. Furthermore, P6 remarked specifically that she could tolerate the intrusiveness of the player because the features the player provided were useful. P10 indicated that he wished he could move the word history box, unaware that he could already reposition it. Only P5 strongly indicated that she found Kalgan to be intrusive.

Surprisingly, despite the 9-button menu's unconventional appearance, most of the participants (6 out of 10) did not have any strong feeling toward it. Only one participant (P9)

expressed an explicitly negative opinion toward it. He stated that the menu felt unfamiliar to him and he wished he could use a more traditional menu layout instead.

Comparing the Players

We also asked our participants to compare Kalgan with a traditional video player. We provided them with five statements. For each statement, they could choose how much they agree with the statements with one being “mostly disagree” to five being “mostly agree.” The statements are available on Figure 5.24. On average, the participants agreed that Kalgan II’s interface made language learning easier (M = 5, 95% CI: [5, 5]) and more enjoyable (M = 5, 95% CI: [4.5, 5]) than a traditional video player. They also agreed that it was more likely to elicit interaction (M = 5, 95% CI: [4.5, 5]). However, the participants were divided as to whether Kalgan made language learning less stressful (M = 3, 95% CI: [3, 4]) or whether its interface was generally



S1: Kalgan II makes language learning easier than a traditional video player.
S2: Kalgan II makes video watching less stressful than a traditional video player.
S3: Kalgan II makes language learning more enjoyable than a traditional video player.
S4: Kalgan II makes me want to interact more with the video than a traditional player.
S5: Kalgan II has a better interface than a traditional video player.

Figure 5.24. The distribution of scores for the statements.

better than a conventional player (M = 3, 95% CI: [2, 5]).

Common Patterns from the Server Log

During the study, we logged all participant interactions with the player. However, due to a technical glitch, we could not collect the data for rewind, fast-forward, and turning on subtitles. We found that overall the participants tended to use the video player quite differently from each other. These differences often manifest based on the individual participant, the language that they were learning, and the videos themselves. Despite of this, we still observed several some emergent patterns:

- **Turning on subtitles.** During the think-aloud protocol, we found that 9 out of 10 of the participants (all but P7) turned on the subtitles immediately for each new video. This suggests that most participants preferred to have the subtitles turned on by default.
- **Menu repositioning.** The log data indicated that all participants repositioned the menu at some point during their viewing. More specifically, the logs and our observations from the think-aloud protocol suggest that most participants tended to reposition the menu in the upper right-hand corner of the screen. This position may be a good default since it avoids major occlusions with the video, yet is easily accessible with the learner's right hand.
- **Re-watching Video.** We found that half of the participants (P1 – P4, P9) actually watched the same videos more than once. One participant (P4) initially watched videos passively but became more interactive in later viewings. On the other hand, P9 did not become more interactive during the re-watches. The remaining participants only watched each video once.
- **Repeated word lookup.** Many participants (P1 – P4, P6 – P8) made frequent word lookups during certain sections of the videos that they were watching. As prior research suggests that

frequent interactions indicate that the video is more difficult [Li et al. 2015], it is likely that the participants found these sections of the videos to challenging.

5.2.5 Discussion

Interactions with Subtitles

Prior research demonstrates the power and the potential of subtitles [Kovacs and Miller 2014; Grgurovic and Hegelheimer 2007; Cubbison 2005]. The results of our study reaffirm this as our participants gave subtitle-based features higher scores and ranks. However, we see additional opportunities for extending the kinds of subtitle interactions introduced in Kalgan II:

- **More Information for Word Lookup.** Yang and Wu [2015] argue that people often learn better if they receive more than just the translation of a word. In their study, they found that people remembered words better if they had an option to receive words' antonyms, synonyms, example images, and more. Future video players can do well to provide additional information on word translations. These features could also potentially be useful in more general eLearning tools, allowing learners can obtain additional information about technical terms and jargon.
- **Highlighting.** Parts of the subtitles can be highlighted to provide additional information to the learner. A video player can highlight words that the learner has translated in the past. By doing so, it can encourage repetition, which may help with acquiring new vocabulary [Webb 2007; Zhu et al. 2017]. The video player can also use highlighting to draw learners' attention to new and difficult words, encouraging additional exploration [Zhu et al. 2017].
- **Optimized Positioning.** Prior research demonstrates that a person typically needs between two to three seconds to process a subtitle [Jensema et al. 2000; Bisson et al. 2014]. Hu et al. [2015] argue that it may be possible to reduce the time by automatically positioning

subtitles close to the speaking character. Reducing subtitle reading times is useful because it can give the learner additional time to interact and learn. However, there is a caveat: The subtitling system assumes that the learner will not interact with the subtitles. If someone interacts with the subtitles, they will need to move either their finger or the mouse cursor around the video player. This may influence how they parse each subtitle.

- **Language-Specific Features.** Some languages use writing systems that are far too complex for the learner to learn from a simple subtitling system alone. Even if their characters are rendered correctly on the screen, the learner may still struggle to make sense of them. One example of this is Mandarin. Although deciphering the pronunciation of a written English word is difficult (as attested in a poem called "The Chaos" [Trenité 2014]), it is even more challenging in Mandarin. For example, even though the English words "break" and "bleak" do not rhyme, we still have enough clues to pronounce them without having to learn an alternative representation. In Mandarin, this is not the case. For example, the word pairs "胡" and "壶" do not provide any hint that they have the same pronunciation. This means that unlike in English, if a Mandarin learner acquires a word's pronunciation, it does not mean that they will be able to identify the corresponding character, and vice-versa. To facilitate learning, Mandarin has a Latin-based pronunciation guide called Pinyin (For example, the Pinyins for the Chinese characters are: 胡= hú, 壶= hú). To efficiently learn Mandarin, the learner will need a subtitling system that allows them to learn both words and their Pinyins. While some initial work has examined specialized subtitling systems to help people learn Mandarin [Kovacs and Miller 2014; Zhu et al. 2017], there are many more languages and writing systems that might benefit from having specialized subtitling systems.

Learning and Enjoyment

Casual language learning tools like Kalgan must strike a precarious balance between active and passive viewing, allowing viewers to treat video content both as an opportunity for learning and as entertainment. Experiences that require too much interaction may compromise the experience of watching, especially if they disrupt the flow of the video. However, experiences that fail to challenge viewers may result in fewer opportunities for learning.

Research on Cognitive Load Theory (CLT) [Sweller 1994] highlights the relationship between mental load and learning and may have major implications for the design of these systems. Work on CLT has identified three different types of cognitive loads: germane loads (which can be beneficial to learners), extraneous loads (which hinder learning and should be eliminated), and intrinsic loads (which hinder learning but are also unavoidable) [Schnotz and Kürschner 2007].

Given this framing, the core challenge of creating tools for casual learning involves designing interactions which allow viewers to challenge themselves by introducing germane loads into the experience while mitigating extraneous ones. In our observations, we saw a strong preference for subtitle-based features that allow viewers to reduce the difficulty of comprehending a foreign language video.

At the same time, we also saw enthusiasm for tools like interactive subtitles which made it possible for viewers to adapt to the difficulty of the video and challenge themselves without disrupting the flow of the video. Tools like these provide lightweight opportunities for viewers to introduce germane loads as they inquire about new vocabulary and grammar. Furthermore, these kinds of tools can eliminate repetitive, mundane, and disruptive side tasks like dictionary lookups that introduce considerable extraneous load.

Our experiences with Kalgan II suggest that lightweight interactions like these, which dovetail nicely with existing casual watching practices, represent a clear opportunity for future language learning tools. Moreover, they point the way towards a rich space of simple, game-like interactions that allow viewers to more explicitly test and expand their knowledge in the context of their everyday viewing.

Chapter Six: **Discussion and Future Work**

In this section, we discuss the main themes that we observed during the development and evaluation of Kalgan I and Kalgan II. Although we created Kalgan I and II to be video players for language learning, we find that some of these themes are also applicable to video players designed for learning in general. Afterward, we provide a low-fidelity prototype of Kalgan III, which illustrates how we might apply the results from the evaluations to future designs.

6.1 Discussion

6.1.1 The Power of Subtitles

The results from the evaluations of Kalgan I and Kalgan II demonstrate that subtitles can be powerful. Many features in Kalgan I and II also heavily rely on the subtitles. We observed in the evaluations of both Kalgan I and Kalgan II that the participants tended to give high scores to subtitle-based features. Some of our participants in the evaluations also wished for more subtitle-based features such as the ability to display two subtitles at once. However, one should note that subtitles are not always perfect. Sometimes, people who supply the subtitles may not be professional translators [Cubbison 2005] and may not accurately translate the video. Additionally, there are also cultural elements that may become altered or lost during translation. For example, Thai language, unlike English, always uses “it” to represents an animal regardless of their gender although Thai language has third-person gendered pronouns.

6.1.2 Types of Learning Intentions

Learning intention is an element that influence informal language learning. During the evaluations of Kalgan I and II, we collected some personal information from the participants

including the reasons of learning a foreign language. Although we found that individuals had their own personal reasons or intention of learning a language, there are some emergent common patterns that many participants share. This means that in the future, we might develop learning tools that target specific learning intentions. Being aware of the common patterns might also allow us to understand how someone may approach a language learning tool.

- **Low-Stakes versus High-Stakes Learning.** For some people, such as those who learn Japanese due to their interest in Japanese animation, learning a foreign language is simply a hobby. They do not need to know the language to function in the society. Therefore, for these people, knowing the language is a low-stakes matter. However, others find language learning to be high-stakes and vital to their function in the society. For example, international students in Canada need to have a very good mastery of English to complete their programs.
- **Mandatory versus Voluntary Learning.** Learners sometimes must learn a language, simply because the government forces them to do so. For example, in many countries, students must learn English as a second language. Voluntary learners, on the other hand, learn a language simply because they enjoy doing so.
- **Learning for Specific Purposes versus Learning for Overall Proficiency.** We found that some people in the studies learned certain languages for specific purposes. For example, one participant from the evaluation of Kalgan II learned many languages, because she was a linguistics student. However, she was not proficient in many of the languages that she learned. She simply needed to know enough to be able to analyze the languages' features. Meanwhile, other people may want to have an overall proficiency in a foreign language. For example, people who settle in a foreign country should try to become proficient in all aspects of its official language so that they can function well in its society.

6.1.3 Interactivity, Immersion, and Cognitive Load Theory

Prior research demonstrates that there is a relationship between learning and interactivity. Guo et al. [2014] argue that higher numbers of interactions in a video means that the video may be more difficult to consume. Interactivity can also increase the mental workload of the learners which can reduce the entertainment values of videos themselves [Homer et al. 2008]. On the other hand, Schwan and Riempp [2004] argue that if the interaction design of a video player is well thought out, interactions can also enhance the learning experience. Most video players are designed to discourage interactions so viewers can have the most immersive experience possible.

Although breaking a viewer's immersion can be bothersome for people who are viewing videos for entertainment, we found that the participants of the evaluation of Kalgan II did not mind the disruptions. This led us to believe that people who try to learn while watching video care less about immersion. They know that since they may not be able to completely understand everything, they might need to perform activities that will break the immersion such as looking up a meaning of a foreign word.

We believe that future work of creating a video player for language learning should apply Cognitive Load Theory (CLT) in some way. We find that CLT can explain some of the results of the evaluations of Kalgan II. DeLeeuw and Mayer [2008] provide some suggestions on how to design an experiment that measures cognitive load. There is also some existing work that applies CLT to design; for example, Mayer and Moreno [2003] provide suggestions on how to design a video content that incurs less cognitive load. Additionally, Feinberg and Murphy [2000] discuss how to apply CLT to web design. Finally, Hollender et al. go further by suggesting that CLT should be integrated into HCI itself [2010].

6.1.4 Evaluation

Evaluation is one of the challenges that loomed large during the research. While evaluation is integral to the research, it is also difficult to perform it correctly. During the evaluation of Kalgan I, we accidentally introduced an effect that confounded the results. Fortunately, with the use of formal methods introduced by Dix [2003], we were able to simplify some of the processes in evaluating Kalgan II. For example, by formalizing the steps required to perform some actions on Kalgan II and the YouTube player, we demonstrated that Kalgan II is simpler to use in some situations such as changing the subtitle language.

We found that the use of mixed methods to be immensely beneficial. This was very clear when the server failed to log the actions of turning on the subtitles until near the end of the experiment. Through the observations from the think-aloud protocol, we could show most participants were turning on the subtitles as soon as the videos started, even though the server did not register most of their actions. The qualitative feedback also gave us additional insight into why the participants gave low scores and ranks to some of the features.

Meaningfully evaluating a participant's language performance is also a significant challenge. During the evaluation of Kalgan I, we believed that despite the potential pitfall, we could evaluate how the participants improved their proficiency. However, evaluation of language proficiency can be extremely difficult. There are many factors, such as the difficulty level and the language of the videos, that are not easy to isolate and are difficult to identify. Since they are hard to identify, they may also manifest as hidden variables during an experiment.

6.2 Kalgan III

Kalgan III is a hypothetical and unimplemented version of the Kalgan video players. The design is based on the feedback that we received from the evaluation of Kalgan I and Kalgan II. It illustrates how a future design might incorporate some of the findings.

6.2.1 Motivation

Kalgan III further augments the design of Kalgan II. As such, while it does not look much different from Kalgan II, we improved some of its features. Kalgan III also emphasizes better touch and more consistent interactions.

6.2.2 Design

There are some small incremental changes. We increased the size of the playhead so they will be 1.1 cm wide and 1.1 cm high. This eliminates the “fat-finger” problem [Wang and Ren 2009].

Figure 6.25. A low-fidelity prototype of Kalgan III.

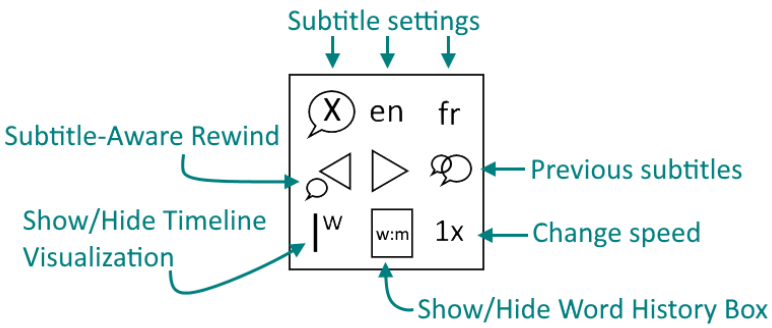


Figure 6.26. The redesigned 9-button menu with descriptions.

Upgraded 9-Button Menu

We removed two buttons, the fast-forward button and the half-speed button in order to streamline the interface. Since the evaluation of Kalgan II demonstrates that subtitle-aware fast-forward is not useful, we eliminated it as well for Kalgan III. We also realized that since there are only two playback speeds, we could reduce the number of speed buttons from two to one. This frees up spaces for more buttons. We decided to add buttons to show or hide the word history box and the annotation on the timeline, giving the learner more control over the interface. Figure 6.26 summarizes the new button layout.

Additionally, if the learner is using a mouse, the menu will initially appear at the bottom-right corner instead of where the learner clicks on the screen. This behaviour comes from the evaluation of Kalgan II where we observed the participants moved the menu to the corner of the screen to prevent it from blocking the content. If the learner is using a touchscreen device, they could also perform an action by swiping from the centre of the menu to the button that represents their intended action. We reintroduced the swipe gesture, because the results of the evaluation of Kalgan I suggest that people liked swiping. Figure 6.27 shows that the learner is trying to rewind the video by swiping into the rewind button. Although Wang and Ren [2009] argue that swiping

can be exhausting, we believe that swipe gesture in Kalgan III is small and fast enough that fatigue is not an issue.

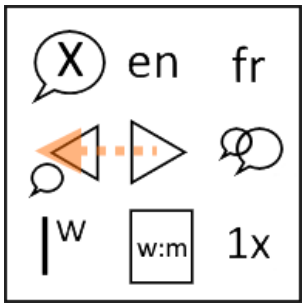


Figure 6.27. An example of swiping to rewind.

Upgraded Word History Box

The participants of the evaluation for Kalgan II indicated that they wished they could export the words and their meanings from the word history box for later uses. Therefore, we decided to upgrade the word history box so that the learner can easily perform this task. The word history box would have a button that allows the user to export the content. Kalgan III might package the content into a simple text file which the learner can then download.

Ruby for Subtitles

We found that there were some participants who were learning Asian language such as Japanese. In Japanese, there are three writing systems that learners have to master. Hiragana and Katakana, the first and the second systems are phonetic and represent certain sounds. However, Kanji, the

Figure 6.28. Japanese and Chinese words accompanied by rubies.

third system, is not phonetic. In order to comprehend a Japanese text, a learner must be able to work with all systems. To help with learning, there is another system called Romanji. Romanji describes sounds using Latin characters which makes learning much simpler to foreigners. There is a text annotation system called the ruby annotation [Sawicki et al. 2008] which appends pronunciation aids either on the top or at the bottom of a text. In Kalgan III, the learner may summon the ruby annotation by holding on the subtitle option button. Such system is very helpful for those who are learning Japanese and Chinese languages.

Chapter Seven: **Conclusion**

Although videos themselves are a convenient, popular, and affordable way to learn a language, they can also be negatively affected by the interfaces of video players. We created Kalgan I and Kalgan II to explore the how to improve the casual language learning experience. Since there is no clear consensus on what constitutes casual language learning, we choose to define it as a type of language learning where the learner is engaging in an activity that can be both leisure and educational at the same time. However, the learner prioritizes leisure, and acquiring new linguistic knowledge is merely a secondary goal.

While we designed both Kalgan I and Kalgan II to be video players for casual language learning, they both have different design philosophies. We designed Kalgan I based on our own experience as language learners. Since we wished to make it simple and approachable, Kalgan I has an interface similar to that of a traditional video player. We based the design of Kalgan II on the observations we made during the evaluation of Kalgan I and on prior research. It has a somewhat unconventional interface, because we decided to prioritize creating an effective interface over creating an interface that everyone would be more familiar with. The traditional video player is inefficient in some respects. For example, the horizontal menu located at the bottom resembles a linear menu which is less efficient than the numpad grid layout and the pie menu. Since we prioritized efficiency, we decided to adopt the numpad grid, thus, giving Kalgan II a less conventional interface.

We also evaluated Kalgan I and Kalgan II using different methodologies. We evaluated Kalgan I in a much more quantitative manner. On the other hand, we evaluated Kalgan II in a more qualitative manner; we used a combination of an online study, interviews, and a think-

aloud protocol. By deploying a diverse set of tools, we obtained richer results. We found that a traditional video player makes it difficult to: (1) adjust the difficulty level, (2) to recover missed information is difficult, and (3) to keep track of learning progress. We also found that subtitles are a powerful language learning tool. Despite their simplicity, we used them to create some of the most well-liked features in Kalgan I and Kalgan II.

During the evaluations of Kalgan I and Kalgan II, we also learned about the strengths and weaknesses of various evaluation methods – from controlled studies to think-aloud protocols. We also learned about formal methods, a way to simplify the evaluation processes. We designed Kalgan III based on what we learned from Kalgan I and Kalgan II. We removed some features that the participants did not like. At the same, we also improved features that were popular among our participants. The interface of Kalgan III represents an incremental improvement based on the feedback that we received. Instead of introducing a dramatic redesign, Kalgan III's goal is to provide better consistency. For example, we planned to enlarge the size of the playhead so it would be at least 1.1 cm x 1.1 cm large to eliminate the "fat-finger" problem and to be consistent with the size of an individual button in the 9-button menu. We also planned to upgrade some features that the participants from the evaluations of Kalgan I and II liked.

While the Kalgan video players share many similarities with existing video players designed for language learning, the Kalgan players have a different design goal. Instead of adding specific learning tasks into video watching experience, we designed the Kalgan players to support casual learners and whatever they want to do while they are watching videos. We want the Kalgan players to be good players even if the learner is simply watching videos without trying to learn from them.

During the evaluations, we also came to appreciate the diversity of our participants. While it is obvious that there are many reasons why a person learns a language, we still felt somewhat surprised by the reasons given by the participants. We found that for some people, the reasons can be very high-stakes. For example, some of our participants are international students who must learn English so that they can have a better quality of life in Canada. Although our research is rather small in the large and diverse field of CALL, it may still have a potential to play an important role in many people's lives in the future.

References

Richard W. Backs, Larry C. Walrath, and Glenn A. Hancock. 1987. Comparison of Horizontal and Vertical Menu Formats. *Proc. Hum. Factors Soc. Annu. Meet.* 31, 7 (September 1987), 715–717. DOI:https://doi.org/10.1177/154193128703100705

Hülya Felek Başaran and Gül Durmuşoğlu Köse. 2013. The Effects of Captioning on EFL Learners' Listening Comprehension. *Procedia - Soc. Behav. Sci.* 70 (January 2013), 702–708. DOI:https://doi.org/10.1016/j.sbspro.2013.01.112

Marie-José Bisson, Walter J.B. van Heuven, Kathy Conklin, and Richard J. Tunney. 2014. Processing of native and foreign language subtitles in films: An eye tracking study. *Appl. Psycholinguist.* 35, 2 (March 2014), 399–418. DOI:https://doi.org/10.1017/S0142716412000434

Shelley Buchinger, Ewald Hotop, Helmut Hlavacs, Francesca De Simone, and Touradj Ebrahimi. 2010. Gesture and touch controlled video player interface for mobile devices. In *Proceedings of the international conference on Multimedia - MM '10*. New York, New York, USA: ACM Press, 699. DOI:https://doi.org/10.1145/1873951.1874055

Adolfo Alfredo Carrillo Cabello. 2013. Website Review: LomásTV. *CALICO J.* 30, 1 (January 2013), 142–153. DOI:https://doi.org/10.11139/cj.30.1.142-153

Jack Callahan, Don Hopkins, Mark Weiser, and Ben Shneiderman. 1988. An Empirical Comparison of Pie vs. Linear Menus. *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.* (1988), 95–100. DOI:https://doi.org/10.1145/57167.57182

Christine Canning-Wilson. 2000. Practical Aspects of Using Video in the Foreign Language Classroom. *Internet TESL J.* 6, 11 (2000).

Laurie Cubbison. 2005. Anime Fans, DVDs, and the Authentic Text. *Velv. Light Trap* 56, 1 (2005), 45–57. DOI:https://doi.org/10.1353/vlt.2006.0004

Yann Dall'Aglio. 2012. *Comment sauver l'amour?*, France: TEDx Talks.

Krista E. DeLeeuw and Richard E. Mayer. 2008. A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *J. Educ. Psychol.* 100, 1 (2008), 223–234. DOI:https://doi.org/10.1037/0022-0663.100.1.223

Alan Dix. 2003. Upside-Down ∀ s and Algorithms—Computational Formalisms and Theory. In John M. Carroll, ed. *HCI Models, Theories, and Frameworks : Toward a Multidisciplinary Science*. San Francisco: Elsevier Science, 381–429.

Susan Feinberg and Margaret Murphy. 2000. Applying cognitive load theory to the design of Web-based instruction. *18th Annu. Conf. Comput. Doc. ipcc sigdoc 2000. Technol. Teamwork. Proceedings. IEEE Prof. Commun. Soc. Int. Prof. Commun. Conf. an* (2000), 353–360. DOI:https://doi.org/10.1109/IPCC.2000.887293

Ignacio Garcia. 2013. Learning a Language for Free While Translating the Web. Does Duolingo Work? *Int. J. English Linguist.* 3, 1 (January 2013), 19–25. DOI:https://doi.org/10.5539/ijel.v3n1p19

Goscinny, René, Morris, Henri Gruel, and Pierre Watrin. 1987. *La Ballade de Dalton*, France: Productions René Goscinny.

Maja Grgurovic and Volker Hegelheimer. 2007. Help options and multimedia listening: Students' use of subtitles and the transcript. *Lang. Learn. Technol.* 11, 1 (2007), 45–66.

Lourdes Guàrdia, Marcelo Maina, and Albert Sangrà. 2013. MOOC Design Principles. A Pedagogical Approach from the Learner's Perspective. *eLearning Pap.* 33, May (2013), 1–6.

Nazli Gündüz. 2005. Computer Assisted Language Learning. *J. Lang. Linguist. Stud.* 1, 2 (2005), 193–214. DOI:https://doi.org/10.17263/jlls.24391

Philip J. Guo, Juho Kim, and Rob Rubin. 2014. How video production affects student engagement. *Proc. first ACM Conf. Learn. @ scale Conf. - L@S '14* , March (2014), 41–50. DOI:https://doi.org/10.1145/2556325.2566239

Ralph Haefner. 1932. Casual Learning of Word Meanings. *J. Educ. Res.* 25, 4–5 (April 1932), 267–277. DOI:https://doi.org/10.1080/00220671.1932.10880263

Bill Hammack. 2012a. *Anodizing (Or the beauty of corrosion)*, United States.

Bill Hammack. 2017. *DLP Projector Stereolithography 3D Printer*, USA, USA: YouTube.

Bill Hammack. 2012b. *How a Microwave Oven Works*, United States: YouTube.

Michel Hazanavicius and Dominique Mézerette. 1993. *La Classe américaine*, France, France.

Nina Hollender, Cristian Hofmann, Michael Deneke, and Bernhard Schmitz. 2010. Integrating cognitive load theory and concepts of human-computer interaction. *Comput. Human Behav.* 26, 6 (2010), 1278–1288. DOI:https://doi.org/10.1016/j.chb.2010.05.031

Bruce D. Homer, Jan L. Plass, and Linda Blake. 2008. The effects of video on cognitive load and social presence in multimedia-learning. *Comput. Human Behav.* 24, 3 (May 2008), 786–797. DOI:https://doi.org/10.1016/j.chb.2007.02.009

Yongtao Hu, Jan Kautz, Yizhou Yu, and Wenping Wang. 2015. Speaker-Following Video Subtitles. *ACM Trans. Multimed. Comput. Commun. Appl.* 11, 2 (January 2015), 1–17. DOI:https://doi.org/10.1145/2632111

Wolfgang Hürst and Konrad Meier. 2008. Interfaces for timeline-based mobile video browsing. In *Proceeding of the 16th ACM international conference on Multimedia - MM '08*. New York, New York, USA: ACM Press, 469. DOI:https://doi.org/10.1145/1459359.1459422

Carl J. Jensema, Ramalinga Sarma Danturthi, and Robert Burch. 2000. Time Spent Viewing Captions on Television Programs. *Am. Ann. Deaf* 145, 5 (2000), 464–468. DOI:https://doi.org/https://doi.org/10.1353/aad.2012.0144

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. *Proc. 27th Annu. ACM Symp. User interface Softw. Technol. - UIST '14* (2014), 563–572. DOI:https://doi.org/10.1145/2642918.2647389

Geza Kovacs and Robert C. Miller. 2014. Smart subtitles for vocabulary learning. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. New York, New York, USA: ACM Press, 853–862. DOI:https://doi.org/10.1145/2556288.2557256

Z. Krzysztof et al. 2015. Leveraging Video Interaction Data and Content Analysis to Improve Video Learning. (2015).

Mike Levy. 2015. The role of qualitative approaches to research in CALL contexts: Closing in on the learner's experience. *CALICO J.* 32, 3 (May 2015), 554–568. DOI:https://doi.org/10.1558/cj.v32i3.26620

Nan Li et al. 2015. How Do In-video Interactions Reflect Perceived Video Difficulty ? In-Video Interaction Analysis. *Proc. Third Eur. MOOCs Stakehold. Summit* 1, 17 (2015), 112–121.

Anthony J. Liddicoat and Angela Scarino. 2013. Technologies in Intercultural Language Teaching and Learning. In *Intercultural Language Teaching and Learning*. West Sussex, UK, UK: John Wiley & Sons, 107–122.

Aleksandra Łuczak. 2017. Using Memrise in Legal English Teaching. *Stud. Logic, Gramm. Rhetor.* 49, 1 (January 2017), 141–152. DOI:https://doi.org/10.1515/slgr-2017-0009

Hideki Maki, Jessica Duton, and Carlyn Obringer. 2003. What Grade Would I Be In If I Were Japanese? 岐阜大学地域科学部研究報告 12 (2003), 109–118.

Richard E. Mayer and Roxana Moreno. 2003. Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educ. Psychol.* 38, 1 (2003), 43–52. DOI:https://doi.org/10.1207/S15326985EP3801_6

a McAuley, B. Stewart, G. Siemens, D. Cormier, and Creative Commons. 2010. The MOOC model for digital practice. *Massive Open Online Courses Digit. ways knowing Learn.* (2010), 1–64. DOI:https://doi.org/10.1016/j.im.2011.09.007

Sandra Meunier. 2014. *happiness is a matter of choice*, France: TEDx Talks.

Howard Nicholas, Patsy M. Lightbown, and Nina Spada. 2001. Recasts as feedback to language learners. *Lang. Learn.* 51, 4 (2001), 719–758. DOI:https://doi.org/Doi 10.1111/0023-8333.00172

Moonyoung Park. 2014. A task-based needs analysis for mobile-assisted language learning in college ESL contexts. In Jeong-Bae Son, ed. *Computer-Assisted Language Learning*. Cambridge, UK: Cambridge Scholar Publishing, 47–68.

M. Peterson. 2010. Computerized Games and Simulations in Computer-Assisted Language Learning: A Meta-Analysis of Research. *Simul. Gaming* 41, 1 (February 2010), 72–93. DOI:https://doi.org/10.1177/1046878109355684

Maja Pivec. 2007. Editorial: Play and learn: potentials of game-based learning. *Br. J. Educ. Technol.* 38, 3 (May 2007), 387–393. DOI:https://doi.org/10.1111/j.1467-8535.2007.00722.x

Abolfazl Qassemzadeh and Hassan Soleimani. 2016. The Impact of Feedback Provision by Grammarly Software and Teachers on Learning Passive Structures by Iranian EFL

Learners. *Theory Pract. Lang. Stud.* 6, 9 (September 2016), 1884. DOI:https://doi.org/10.17507/tpls.0609.23

Gonzalo Ramos and Ravin Balakrishnan. 2003. Fluid interaction techniques for the control and annotation of digital video. In *Proceedings of the 16th annual ACM symposium on User interface software and technology - UIST '03*. New York, New York, USA: ACM Press, 105–114. DOI:https://doi.org/10.1145/964696.964708

Peter J. Rich and Michael Hannafin. 2008. Video Annotation Tools: Technologies to Scaffold, Structure, and Transform Teacher Reflection. *J. Teach. Educ.* 60, 1 (November 2008), 52–67. DOI:https://doi.org/10.1177/0022487108328486

N. Sakunkoo and P. Sakunkoo. 2009. GliFlix: Using Movie Subtitles for Language Learning. (2009), 7–8.

Marcin Sawicki, Michel Suignard, Masayasu Ishikawa, Martin Dürst, and Tex Texin. 2008. Ruby Annotation. (2008). Retrieved September 28, 2017 from https://www.w3.org/TR/ruby/

Wolfgang Schnotz and Christian Kürschner. 2007. A reconsideration of cognitive load theory. *Educ. Psychol. Rev.* 19, 4 (2007), 469–508. DOI:https://doi.org/10.1007/s10648-007-9053-4

Klaus Schoeffmann. 2010. Video browsing interfaces and applications: a review. *J. Photonics Energy* 1, 1 (February 2010), 18004. DOI:https://doi.org/10.1117/6.0000005

Stephan Schwan and Roland Riempp. 2004. The cognitive benefits of interactive videos: learning to tie nautical knots. *Learn. Instr.* 14, 3 (June 2004), 293–305. DOI:https://doi.org/10.1016/j.learninstruc.2004.06.005

G. Sockett. 2014. Informal Learning and Languages. In *the Online Information Learning of*

English. London, UK: Palgrave Macmillan UK, 7–30.

John Sweller. 1994. Cognitive load theory, learning difficulty, and instructional design. *Learn. Instr.* 4, 4 (January 1994), 295–312. DOI:https://doi.org/10.1016/0959-4752(94)90003-5

Kuo-Hao E. Tang. 2001. Menu Design with Visual Momentum for Compact Smart Products. *Hum. Factors J. Hum. Factors Ergon. Soc.* 43, 2 (2001), 267–277. DOI:https://doi.org/10.1518/001872001775900913

Alan M. Taylor. 2013. CALL versus Paper: In Which Context Are L1 Glosses More Effective? *CALICO J.* 30, 1 (January 2013), 63–81. DOI:https://doi.org/10.11139/cj.30.1.63-81

The Editors of Encyclopædia Brittanica. 1998. Kalgan. *Encycl. Brittanica* (1998).

Gerard Nolst Trenité. 2014. The Chaos. *Word Ways* 47, 3 (2014), 219–228.

Carsten Ullrich, Ruimin Shen, Ren Tong, and Xiaohong Tan. 2010. A Mobile Live Video Learning System for Large-Scale Learning - System Design and Evaluation. *IEEE Trans. Learn. Technol.* 3, 1 (January 2010), 6–17. DOI:https://doi.org/10.1109/TLT.2009.54

Robert Vanderplank. 2010. Déjà vu? A decade of research on language laboratories, television and video in language learning. *Lang. Teach.* 43, 1 (January 2010), 1. DOI:https://doi.org/10.1017/S0261444809990267

Feng Wang and Xiangshi Ren. 2009. Empirical evaluation for finger input properties in multi-touch interaction. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, 1063. DOI:https://doi.org/10.1145/1518701.1518864

Stuart Webb. 2007. The effects of repetition on vocabulary knowledge. *Appl. Linguist.* 28, 1 (2007), 46–65. DOI:https://doi.org/10.1093/applin/aml048

Paula Winke, Susan Gass, and Tetyana Sydorenko. 2010. The effects of captioning videos used

for foreign language listening activities. *Lang. Learn. Technol.* 14, 1 (2010), 65–86.

Fang-chuan Ou Yang and We Vivian Wu. 2015. Using mixed-modality learning strategies via e-learning for second language vocabulary acquisition research background. *Educ. Technol. Soc.* 18 (2015), 309–322.

Dongsong Zhang, Lina Zhou, Robert O. Briggs, and Jay F. Nunamaker. 2006. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Inf. Manag.* 43, 1 (2006), 15–27. DOI:https://doi.org/10.1016/j.im.2005.01.004

Yeshuang Zhu et al. 2017. ViVo: Video-Augmented Dictionary for Vocabulary Learning. *Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI '17* (2017), 5568–5579. DOI:https://doi.org/10.1145/3025453.3025779

Appendix A: **Kalgan I Evaluation Forms**

A.1 Consent Form

UNIVERSITY OF CALGARY

CONSENT FORM

Name of Researcher, Faculty, Department, Telephone & Email:

Sathaporn Hu, Faculty of Graduate Studies, Department of Computer Science
Telephone: [redacted]
Email: [redacted]

Supervisor:

Professor Wesley Willett, Department of Computer Science

Title of Project:

Learning A Language with Enhanced Video Player

This consent form, a copy of which has been given to you, is only part of the process of informed consent. If you want more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.

Purpose of the Study

The purpose of the study is to learn how people learn a foreign language through watching a video with subtitle.

What Will I Be Asked to Do?

You will be asked to watch a video in order to prepare for a quiz which will test you about the vocabulary in the video and the plot. You will be doing this for 6 times. During the experiment, you will also be asked to complete some other questions as well.

What Type of Personal Information Will Be Collected?

Should you agree to participate, you will be asked to provide some information about your academic background and linguistic background.

Are there Risks or Benefits if I Participate?

There are no immediate or significant risks for you to participate in the study. The recruitment material should have already informed you that the video clip contains a small amount of humorous profanity.

However, the clip should not be anymore offensive than any other video in the mainstream media. You will receive $20 for your participation even if you decide to leave early. Your participation will help to enhance language learning experience in the future.

What Happens to the Information I Provide?

The study is completely voluntary and your data will be made confidential. The paper data may be either stored in a cabinet or digitized and stored in a computer. The digital data will be stored on the computer that is only accessible to the investigators.

The video tape record will be stored on a machine that only the investigators can access. If the investigators decide to use a specific video clip in public, any feature that can be used to identify you (such as face) will be blurred out or covered up. The investigators will also remove the audio in the video so you do not need to worry about your voice being heard in public.

You will be assigned a number or a pseudonym for the purpose of the study and your real name will not appear in the data collected. Therefore, no identifying data will ever be made public.

Since the data does not contain any information that can be used to identify you, it is impossible to withdraw your data once you have completed the experiment. Therefore, if you want your data to be removed, please leave the experiment early. If you decide to leave the experiment early, your data will not be used for the study and you do not need to make any explicit request. It should be noted that you are still eligible for the reward even if you decide to leave early.

Signatures

Your signature on this form indicates that 1) you understand to your satisfaction the information provided to you about your participation in this research project, and 2) you agree to participate in the research project.

In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from this research project at any time. You should feel free to ask for clarification or new information throughout your participation.

Participant's Name: (please print) _____

Participant's Signature: _____ Date: _____

Researcher's Name: (please print) _____

Researcher's Signature: _____ Date: _____

Questions/Concerns

If you have any further questions or want clarification regarding this research and/or your participation, please contact:

Mr. Sathaporn Hu,
Department of Computer Science
[redacted]
and Professor Wesley Willett

If you have any concerns about the way you've been treated as a participant, please contact the Research Ethics Analyst, Research Services Office, University of Calgary at [redacted]; email [redacted].

A copy of this consent form has been given to you to keep for your records and reference. The investigator has kept a copy of the consent form.

A.2 Group I Study Package

COVER PAGE

Participation Number: _____

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

BACKGROUND QUESTIONS

1. Please indicate your level of French by ticking the words below:
   □ Beginner □ Intermediate □ Advanced □ First Language
2. Are you currently taking a French class? Tick an appropriate box and write additional information.

   □ Yes – please enter your current class's name (Please use generic name instead of code. For instance, "2nd year French" is preferable to "FREN369):

   □ No – if you have taken French class(es) in the past, please write the highest level of French class that you have taken (Please use generic name instead of code. For instance, "2nd year French" is preferable to "FREN369):

3. Please describe your general French learning strategy:

4. Please describe how use videos to study French.

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

QUIZ #1

Vocabulary: Translate these words to English. If you already know the word before the study, please tick the box right of the word.

1. thérapeute: ☐ I know this word before study.
2. interviens: ☐ I know this word before study.
3. soufflé: ☐ I know this word before study.
4. cancérologie: ☐ I know this word before study.

Story: Briefly answer the question using English.

1. What was the occupation of the speaker?

2. What kind of disease did the patients have?

3. What did the speaker ask French people to stop doing?

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

QUIZ #2

Vocabulary: Translate these words to English. If you already know the word before the study, please tick the box right of the word.

1. héritage: ☐ I know this word before study.
2. tonton: ☐ I know this word before study.
3. marchand: ☐ I know this word before study.
4. file: ☐ I know this word before study.

Story: Briefly answer the question using English.

1. How did the Daltons escape the prison?

2. What was the name of the shortest Dalton?

3. What did the merchant sell?

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

QUIZ #3

Vocabulary: Translate these words to English. If you already know the word before the study, please tick the box right of the word.

1. digérer: ☐ I know this word before study.
2. patron: ☐ I know this word before study.
3. bâfrer: ☐ I know this word before study.
4. ballonné: ☐ I know this word before study.

Story: Briefly answer the question using English.

1. What made the journalist feel sick?

2. Why did the journalist have to disguise himself?

3. Apart from l'Ouische Lorraine, what kind of food could Jose cook very well?

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

CONTROL PLAYER QUESTIONS

Please indicate how much you agree with the statements below. The statements pertain to the player used for the first three videos.

		Most Disagree						Most Agree
1.	I can understand the videos' content well.	1	2	3	4	5	6	7
2.	I feel I have learned new French words.	1	2	3	4	5	6	7
3.	I feel I have learned something new about French grammar.	1	2	3	4	5	6	7
4.	I feel I have learned something new about French pronunciation and/or accents.	1	2	3	4	5	6	7
5.	I feel that this video player is helping me to learn French.	1	2	3	4	5	6	7
6.	I feel that this video player has enough functionalities to help me learn French.	1	2	3	4	5	6	7
7.	I want to learn French with this video player.	1	2	3	4	5	6	7

If you have any other comment or anything else in mind, please write it here:

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

QUIZ #4

Vocabulary: Translate these words to English. If you already know the word before the study, please tick the box right of the word.

1. astre: ☐ I know this word before study.
2. entremêlés: ☐ I know this word before study.
3. dévaloriser: ☐ I know this word before study.
4. angoisse: ☐ I know this word before study.

Story: Briefly answer the question using English.

1. What was the eternal problem of love?

2. What was one of the 3 processes that form the modern concept of "individual"?

3. How do we deal with the "Anxiety of the Modern Man"?

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

QUIZ #5

Vocabulary: Translate these words to English. If you already know the word before the study, please tick the box right of the word.

1. hâte: ☐ I know this word before study.
2. déchéances: ☐ I know this word before study.
3. bienfaiteur: ☐ I know this word before study.
4. ingrat: ☐ I know this word before study.

Story: Briefly answer the question using English.

1. What was the cost of the elixir?

2. What did Lucky Luke suggested to the Dalton to do in order to deal with the doctor?

3. What did the doctor drink in the end?

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

QUIZ #6

Vocabulary: Translate these words to English. If you already know the word before the study, please tick the box right of the word.

1. sans-gêne: ☐ I know this word before study.
2. connerie: ☐ I know this word before study.
3. de droit: ☐ I know this word before study.
4. grabuge: ☐ I know this word before study.

Story: Briefly answer the question using English.

1. Why was Prof. Hammond upset at the journalist?

2. Why did the dinosaur attack the cowboys?

3. What did Joel Hammond do while everyone was fighting the prehistoric animal?

DO NOT TURN THE PAGE UNLESS INSTRUCTED TO DO SO!

ENHANCED PLAYER QUESTIONS

Please indicate how much you agree with the statements below. The statements pertain to the player used for the last three videos.

	Most Disagree						Most Agree
1. I can understand the videos' content well.	1	2	3	4	5	6	7
2. I feel I have learned new French words.	1	2	3	4	5	6	7
3. I feel I have learned something new about French grammar.	1	2	3	4	5	6	7
4. I feel I have learned something new about French pronunciation and/or accents.	1	2	3	4	5	6	7
5. I feel that this video player is helping me to learn French.	1	2	3	4	5	6	7
6. I feel that this video player has enough functionalities to help me learn French.	1	2	3	4	5	6	7
7. I want to learn French with this video player.	1	2	3	4	5	6	7
8. The ability to flag the video is useful.	1	2	3	4	5	6	7
9. The ability to rewind/forward 5 seconds is useful.	1	2	3	4	5	6	7
10. The interactive subtitle system is useful.	1	2	3	4	5	6	7
11. The word history is useful.	1	2	3	4	5	6	7
12. I like swiping to control the video more than tapping on the video controls.	1	2	3	4	5	6	7
13. I like that the video temporarily pause when I tap on the screen.	1	2	3	4	5	6	7

PLEASE TURN THE PAGE TO COMPLETE ALL QUESTIONS

The features below have swipe gestures associated to them.

- **Tick "appropriate" if you think the gesture is appropriate.**
- **Tick "modify" if you think the gesture should be modified.**
- **Tick "remove" if you think the feature should not have any touch-gesture associated with it.**

1. Change subtitle [swipe up]: □ appropriate □ modify □ remove
2. Flag [swipe down]: □ appropriate □ modify □ remove
3. Rewind 5 seconds [swipe left]: □ appropriate □ modify □ remove
4. Forward 5 seconds [swipe right]: □ appropriate □ modify □ remove

Please answer the questions

1. What do you like most about the player?

2. What do you dislike most about the player?

3. What features do you want to see in the player?

4. Did you struggle with the player? If so, please describe your struggle:

5. Please write down any comment or anything that you have in mind:

A.2 Group II Study Package

The second group used almost the same package with the first one. The only difference was the "Control Player Questions" page and the "Enhanced Player Questions" page were swapped.

Appendix B: **Kalgan II Evaluation Forms**

B.1 Consent Form

UNIVERSITY OF CALGARY

Name of Researcher, Faculty, Department, Telephone & Email:

Hubert (Sathaporn) Hu, Department of Computer Science, [redacted]

Supervisor:

Professor Wesley Willett, Department of Computer Science, [redacted]

Title of Project:

Enhanced Video Player for Language Learning

This consent form, a copy of which has been given to you, is only part of the process of informed consent. If you want more details about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.

Participation is completely voluntary and you are free to discontinue participation at any time during the study. Since the most of the experiment is done primarily online, we cannot completely ensure your confidentiality. However, we work to make sure that your data is secure. We will ensure complete your confidentiality and anonymity in the non-online portion of the study.

Purpose of the Study

The study is to evaluate the performance of a video player which has been designed for language learning over a period of one week. The investigators would like to see how the video player is being used in home setting, and how to model the usage pattern.

What Will I Be Asked To Do?

You will be asked to use a special video player to watch YouTube videos for at least one week. Afterwards, you will participate in an in-person interview. The in-person interview will last at most 60 minutes. Our pilot study indicates that a normal interview length is around 30-45 minutes. During the interview, we will ask you about your language background and your experience using the video player. We will ask you to rate the video player. We will also ask you to use the video player during the interview while we record your on-screen activities. At the beginning of the interview, we will seek your permission to audio-tape you. If you allow us to audio-tape you, your voice will be recorded. Otherwise, no audio will be recorded.

We ask that you try to watch three hours or more of YouTube videos in total before the interview. While we will provide you with a list of suggested videos, they are not mandatory. You may watch any video that you want.

What Type of Personal Information and Data Will Be Collected?

1

We will require your email address during the experiment to provide you with the support on using the video player. It is important to note that the video player will log your activities and transmit data over the Internet. The data that will be transmitted are:

- The YouTube ID for the video that you are watching. Because the ID is transmitted, if you do not want the investigators know that you are watching certain videos, please watch those videos using the default player (such as the website itself, the mobile player, etc.) instead.
- Your browser information.
- Your email address.
- The type and details of actions that you have performed while watching the video.
- The time when an action is performed.

During the in-person interview, we will ask about your education background and your gender.

Are there Risks or Benefits if I Participate?

Since most of the experiment will be done remotely, your log data will be transmitted over the Internet. This means that there will be some risks that your data may be intercepted by a malicious party. However, the data will be sent through an encrypted protocol (HTTPS). Therefore, it should be difficult for a malicious party to intercept and decrypt the data. Participating in the study should be no riskier than performing an online activity such as using a social media website or doing online banking.

The in-person portion should not present you with any risk. You will receive $20 if you are physically present at the interview. You will receive this amount even if you decide to end the interview early. You will not be reimbursed if you are not present at the interview.

What Happens to the Information I Provide?

Only we will have the information that you have provided. The digital data will be stored in a computer that only the investigators can access. The data on paper will be stored in a place that only we can access. None of the audio recordings will ever be made public. All data will be destroyed within five years after your participation. If you decide to leave your study early or if you ask us to not use your data, we will attempt to destroy data associated to you. However, this is not guaranteed due to various circumstances. We suggest that if you want your data to not be used in the study or if you want your data to be destroyed, you should contact us as soon as possible.

Questions/Concerns

If you have any further questions or want clarification regarding this research and/or your participation, please contact:

Mr. Hubert (Sathaporn) Hu,
Department of Computer Science
[redacted]

Prof. Wesley Willett,
Department of Computer Science,
[redacted]

If you have any concerns about the way you've been treated as a participant, please contact the Research Ethics Analyst, Research Services Office, University of Calgary at [redacted]; email [redacted]. Once you sign up, you will be given a link to a website where you can download and print a copy of this consent form.

In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from this research project at any time. You should feel free to ask for clarification or new information throughout your participation.

Signatures

By completing the sign-up form below and clicking the submit at the bottom the screen, you indicate that 1) you understand to your satisfaction the information provided to you about your participation in this research project, and 2) you agree to participate in the research project.

You are encouraged to save this consent form for your reference. If you want a copy of the consent form with actual signatures, please either send us an email or ask for one during the interview.

B.2 Study Package

Interview Booklet

Participant Number: ___

Ask the participant if he or she wants to review the consent form.

Please also re-confirm whether the participant wants to be audio-taped or not during the interview. Circling "YES" means he/she consent to being audio-taped. Circling "NO" means he/she do not consent to being audio-taped:

YES NO

Background

1. What is the gender that you identify as: Male Female Other

2. Please tell us your languages that you know and you are trying to learn. Please also report your level of fluency as well with 1=novice, 2=intermediate, 3=advanced, 4=fluent. Also, please give us details on each language.

[The interviewer fills out the form below]

1.	_____	Level:	1	2	3	4
2.	_____	Level:	1	2	3	4
3.	_____	Level:	1	2	3	4
4.	_____	Level:	1	2	3	4
5.	_____	Level:	1	2	3	4

Comments:

3. Please tell us how difficult it is for you to watch a video for language learning in general:

Easy Somewhat easy Neither easy nor difficult Somewhat difficult Difficult

4. Please tell us about your language learning strategy.

Bugs Technical Report

5. Please report any bug and technical issues that you have encounter. We will allow you to report any non-technical issues afterward.

Like and Dislikes

6. What do you like about the player?

7. What do you dislike about the player?

8. Please tell us if the video player is helpful to language learning.

9. Please tell us if the video player's interface is intrusive.

10. Please compare and contrast this video player with other conventional video players.

Mode of Use

11. Please tell us the languages that you are trying to learn with the video player.

12. Please describe the device that you use for the video player. Please also tell us whether you normally use touch or mouse to interaction with the video player.

Rating and Ranking

13. Please rank the usefulness of the following the functionalities. (The interviewer should produce the image of the interface and point out each component to the participant) [The interviewer helps the participant with filling the questionnaire]:

	Useless	Somewhat Useless	Neither Useful nor Useless	Somewhat Useful	Useful
Turn off subtitle					
Turn on 2nd language subtitle					
Turn on 1st language subtitle					
Rewind button					
Fastforward button					
Past subtitle					
Playing at half speed					
Interactive subtitle					
Word history box					
Words on the timeline					

14. Please rank the functions in the order of their importance. The table below provides an example of how to rank this. The filling of the table below means that the person feels that A is the most important, B is the second most important, C is the third most important and D is the least important.

Functionality	1	2	3	4
A	X			
B		X		
C			X	
D				X

You cannot assign the same rank to different functionality. The table below provides an example of a wrongly filled table:

Functionality	1	2	3	4
A	X			
B		X Forbidden!		
C		X		
D			X	

Now, please rank these functionalities [The interviewer helps the participant with filling the questionnaire]:

	1	2	3	4	5	6	7	8	9	10
Turn off subtitle										
Turn on 2nd language subtitle										
Turn on 1st language subtitle										
Rewind button										
Fastforward button										
Past subtitle										
Playing at half speed										
Interactive subtitle										
Word history box										
Words on the timeline										

15. Please indicate how much you agree with each statement:

	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat Agree	Agree
The experimental video player makes language learning easier than a conventional video player.					
The experimental video player makes video watching less stressful than a conventional video player.					
The experimental video player makes language learning more enjoyable than a conventional video player.					
The experimental video player makes me want to interact more with the video than a conventional video player.					
The experimental video player has a better interface than a conventional video player.					

16. If you have any additional comments on the order of function use, please tell us:

Think Aloud Protocol

[This part will not be performed if the participant does not consent to be audio-taped.]

17. Please attempt to watch a YouTube video using the experimental player. When you are encountering the part that you don't understand, please attempt to use the functions in the player to aid you. While attempting to use the functions, please announce steps aloud. For example, if you are looking up a word in the subtitle, you should declare "I am tapping on a word in the subtitle."