

# MangoSub: Learnersourced Video Subtitling

Sungmin Choi\*

KAIST

Daejeon, Korea

smc9601@kaist.ac.kr

Wonkeun Choi\*

KAIST

Daejeon, Korea

anders@kaist.ac.kr

Muso Khon\*

KAIST

Daejeon, Korea

muso@kaist.ac.kr

Zhanto Orynbassar\*

KAIST

Daejeon, Korea

zhanto97@kaist.ac.kr

## ABSTRACT

The lack of availability of subtitles for little-known TV shows and movies is a frustrating problem encountered by many non-native English speakers. Also, while watching TV shows and movies is a common method for non-native speakers to expose themselves to English dialogues, this can be quite ineffective as they tend to become increasingly dependent on subtitles. In this paper, we introduce MangoSub, a platform that generates subtitles for user-requested videos by crowdsourcing transcription from the same crowd of English-learning users who request them. MangoSub is unique in that it provides non-native English speakers an interactive way to familiarize with English by actively interpreting English dialogues, while also taking part in generating subtitles. We present the Generate-Fix-Verify workflow which allows an accurate generation of subtitles from a relatively unskilled crowd. While the results of our deployment indicate our prototype was working as intended to generate subtitles for the most part, we explain the problems encountered including some confusions with the user interface and inefficient quality control.

## Author Keywords

Crowdsourcing; Learnersourcing; Workflow; Subtitle Generation

## ACM Classification Keywords

H.5.2. Information interfaces and presentation: User interfaces

## INTRODUCTION

Listening comprehension is one of the most difficult tasks for non-native English speakers. Thus, being regularly exposed to English dialogues is the most crucial aspect of mastering English listening, which is often a luxury unavailable for those who live in non-English speaking countries. For these non-native speakers, a common replacement for having real conversations is watching TV shows and movies in English. While subtitles for popular shows and movies can be easily obtained online, there are frustrating instances in which subtitles are needed for less well-known shows and movies, documentaries, lectures, speeches, or even self-recorded videos. This is only the surface of the problem. While watching videos with subtitles does help non-native speakers to gradually become more exposed to English, they are not utilizing the full

potential of their time when they are mindlessly reading subtitles instead of actively trying to interpret the dialogues.

In an attempt to address the above-mentioned problems, we motivate and introduce MangoSub, a platform in which users can 1) request videos to be subtitled by other users and 2) contribute to the generation of subtitles for videos requested by other users. While the primary motivation for the users to work is to improve listening comprehension by being exposed to English dialogues, we included features to further incentivize the workers: the credit system and the leaderboard. The users can only request videos with credits earned by completing tasks that contribute to the generation of subtitles. After repetitions of working on generating subtitles for other users and watching subtitled videos, the users will eventually improve their listening comprehension to the point that they no longer need subtitles.

Later, we extensively describe the Generate-Fix-Verify workflow model which allows us to create accurate subtitles with an unskilled crowd. Thus, we will not go into the details of the model in this section.

In the rest of the paper, we explore the existing tools for obtaining subtitles and discuss our unique contribution to this field. Then, we explain the technical details of the features in MangoSub. Finally, we discuss our deployment with its results, as well as possible improvements to our platform.

## BACKGROUND AND RELATED WORK

There are many online resources which non-native speakers can use to obtain subtitles. One of the most prevalent platforms is OpenSubtitles [1], where users can upload a complete subtitles or freely download subtitles. One problem that arises from requiring users to only upload completed subtitles is that this tends to create a shortage of subtitles for less well-known shows and movies. Our platform has a different, narrower user-base; it is more oriented towards non-native English speakers who wish to actively improve their English, while OpenSubtitles tends to be used by those who simply want to enjoy popular TV shows and movies.

Google provides the YouTube automatic captioning [2] feature, whose performance has dramatically increased lately. Its machine-generated subtitles are extraordinarily accurate in videos with simple, one-to-one conversations. However, it is still far from adequate, as dialogues are not

the only things happening in most scenes; there are almost always sounds in the background, which degrades the accuracy of the subtitles, or even worse, prevents the machine from detecting the dialogue at all. However, because of this algorithm's usefulness in certain cases, we are considering including it in our workflow to reduce the amount of raw generation required from the crowd.

Finally, there is a crowdsourced subtitling platform called Amara [3], which generates subtitles with professional linguists, online educators, and other experts. Since our platform uses unskilled, non-native English speakers to generate subtitles, it requires a more rigorous crowdsourcing workflow to account for the innate inaccuracy of the human computation. In addition, we believe that our user base and their incentives to continue using our platform is more robust; our workers strive to improve themselves, and they must continue to work if they wish to request. On the other hand, Amara's expert user-base is incentivized solely from humanitarianism.

**SYSTEM**

In aiming to create a platform that can provide both subtitle-creation service and educational environment, we came up with three different tasks that we have to tackle:

1. How can users ask for subtitles?
2. How can we create an engaging workflow?
3. How can we create the profile of an user for better immersion in the platform?

**Request**

Requested videos are a crucial part of the platform because they are what keeps users busy with work and helps to improve English listening comprehension. For that purpose, the platform provides subtitling services to users. However, to request subtitles, user has to have enough credits, i.e. platform charges 10 credits per minute of video requested. The justification behind charging users with credit is that people should not use the platform for simply getting subtitles. We encourage people to work and learn, which is the primary goal of this platform.

**Workflow**

Inspired by workflow used in Soylent [4], we divided the process of subtitling a video into three separate tasks: Generate, Fix, and Verify. Our team believes that this systematic approach most effectively harnesses crowd intelligence since separate tasks require little work and makes the learning process more engaging. It also introduces quality control that is necessary for the unskilled crowd.



Figure 1. Workflow model

**Generate**

In the first stage of subtitle-making, users collectively create first draft of the subtitles. To make the experience more engaging, platform provides users an opportunity to freely choose parts of the video they want to create subtitles for. Additional advantage of this design choice is that less experienced users do not struggle with difficult tasks, which in turn results in better quality of the subtitles. Platform notifies users about portions of the video for which subtitles are already created via a bar above the video progress bar as can be seen on Figure 2. The intervals for which subtitles are already created are colored red. Platform does not allow users to create subtitles for a portion that overlaps with another portion for which subtitles are already created, which means that each portion of the video is subtitled by only one user.



Figure 2. Interface of the Generate page

The user has to move Start and End flags to tell the platform which portion of the video he or she is creating the subtitles for. These flags are located under the video progress bar as one can see on Figure 2. After typing the subtitles in the text bar, user is asked to press the Show button and watch the selected portion so that he or she can check whether Start and End flags are in correct positions for text and speech to be perfectly aligned.

**Fix**

Following the Generate stage, platform continues subtitle-creation process by giving users Fixing work. In this particular type of task, users are asked to listen to the portion of the video starting at Start flag and ending at End flag and to provide better version of subtitles for that portion if they think subtitles are incorrect. If users deem the subtitles to be correct, they can press the button with Tick image (Figure 3) and submit.

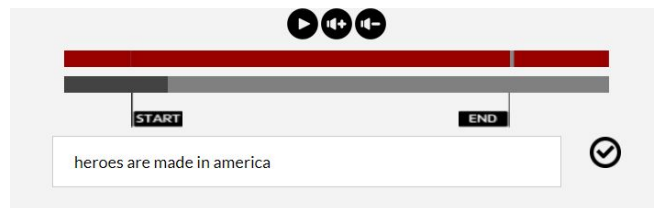


Figure 3. Interface of the Fix Page

For each portion of the video, we accept only two fixes from users so that in the following stage of the workflow, the platform has enough number of choices to provide to users.

### Verify

The last stage in the process of subtitle-making is Verify. Users are provided with three different versions of subtitles for particular portion of video and are asked to vote for the best one, as shown in Figure 4 (1 choice from the Generate stage, 2 from Fix). For each portion of the video, only one user votes for the best choice and this choice is considered to be the final subtitle for that portion.

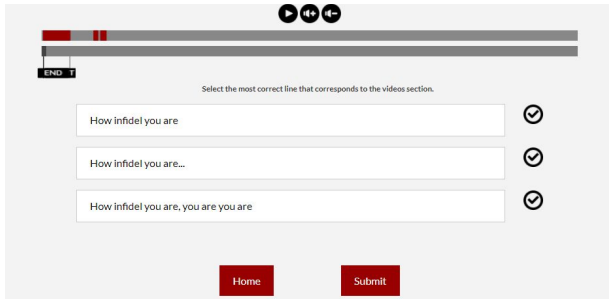


Figure 4. Interface of the Verify Page

### Finalizing subtitling

When all of three stages have completed, subtitles are written to an .srt file and sent to the database. After that, the user who requested subtitles is notified and provided a link to download subtitles in their Profile page.

### Profile

A credit system is used for rankings and video requests. The platform gives 5, 3, and 2 credits for Generate, Fix and Verify tasks, respectively.

Users can track the history of the credits received for working on each task, and credits used for requesting videos in the Profile page. They can also download completed subtitles in the Profile page. In addition, users can see the leaderboard, which shows their global ranking.

### EVALUATION

We advertised our platform among university students and by the time of writing this paper 49 users registered, 18 users requested videos to get subtitled and 16 users worked on making subtitles. From our Generate, Fix and Verify stages 74 lines of subtitles were generated, 93 lines of subtitles were fixed and 46 lines of subtitles were verified. Out of 18 videos requests 1 of them had corrupted file format and 3 videos were vandalized. Vandalized videos are those which had different language, but English. You can see example of vandalized videos in Figure 5.

From remaining 14 videos, 8 subtitles were returned for users who requested videos. Current Leaderboard standings can be seen in Figure 6.

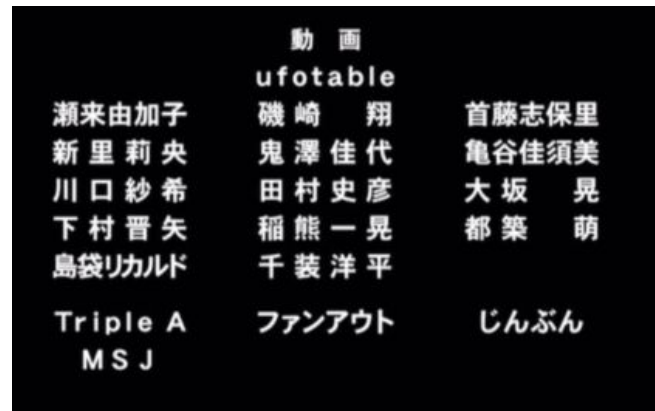


Figure 5. Example of a vandalized video

LEADERBOARD		
RANK	USERNAME	POINTS
1	noidwasavailable	112
2	anamel	98
3	username	91
4	ayylmao	87
5	asdfq	71
6	Anders	59
7	1231	54
8	aq	48
9	LoveDonaldTrump	35

Figure 6. Leaderboard standings of users

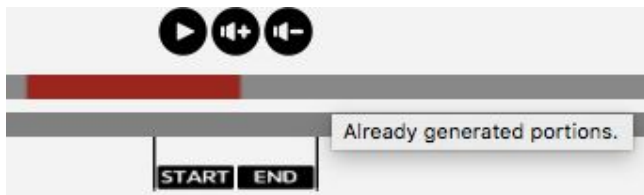
### DISCUSSION

As our platform gives opportunity for users to improve their English skills, a lot of users were interested in our platform and were intrinsically motivated to work on our platform. Not only improving their English skills, users can also gain more credits to be able to request subtitles and this also motivates our users to work more. After using our platform for some time, both for requesting and working on subtitles, users were generally satisfied with our platform. While most of them were happy to use our platform without any struggle, some faced with difficulties and limitations while working on making subtitles. We obtained feedbacks to further improve our system. We will now walk through some of the difficulties that users faced and possible improvements.

### Difficulties and possible solutions

Some users didn't know the purpose of the red/gray bar located above the progress bar. (Figure 7). We showed this

bar to indicate which parts of the video is already subtitled (the red segments). Users were only allowed to make subtitles for the parts that does not have subtitle (the gray segments). Although this information was provided when the mouse was hovered over the bar, many users were not able to fully understand this functionality and later were not able to submit subtitles. A possible solution for such complication would be an on-screen tutorial of how to use our system. Even though we provide written instructions of each step before the users start working, users find it quite boring and will be reluctant to read it. Having an on-screen tutorial would solve this issues and users will not face difficulties on how to use our system in the future.



**Figure 7. Red/gray bar and progress bar for any video**

One other problem we witnessed was users in Generate section had too much freedom while generating subtitles. One user could simple make one subtitle line and making timings of it as much as he wanted. For example, for a video of 1 hour long, user could just set make one subtitle and set duration for the subtitle same as whole video of 1 hour long. This resulted with poor quality of subtitles. One possible solution is to make upper time threshold for duration of a single line of subtitle. Apart from this, in our future improvements we are also introducing “Reporting User” so that if such subtitles were generated users of the later stages can report it and we could punish such mishandled subtitles.

The last problem that we faced was with vandalized videos. As currently our platform is a place where users can improve their English Skills, we are restricting the requested videos to be only in English. We think that with the introduction of a report system, these vandalization can be solved.

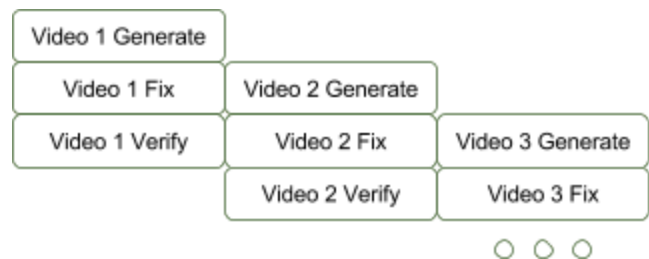
**Other improvements**

We are planning to upgrade our workflow in the future. Currently, the ordering of subtitle generation is done by FIFO. When a second video gets requested by some user, its generation will not begin until the generation for the first video gets completed. In addition, within the process of each video, a stage begins only after its previous stage is complete. For example, when Video 1 is requested, every worker will work only on the Generate part. Fix will only begin after Generate is complete, and so on. Full ordering of the workflow can be seen on Figure 8.



**Figure 8. Current workflow**

We believe that although a video requested earlier should have higher priority than the later ones, users will lose interest if the platform forces them to work only on one type of task repetitively until the entire stage is complete. So by creating a more randomized workflow, we can expect higher motivation and participation from users. We are planning to introduce higher levels of workflow in which at any given time users can be given any type of task. For example, when Video 2 is requested and some workers are working on Fixing part of Video 1, others will be allowed to work on generation part of Video 2. This improved workflow can be seen on Figure 9.



**Figure 9. Future workflow**

**REFERENCES**

1. OpenSubtitles. Retrieved December 17, 2017 from <http://www.opensubtitles.org/>
2. YouTube. Use automatic captioning. Retrieved December 17, 2017 from <http://support.google.com/youtube/answer/6373554/>
3. Amara. Retrieved December 17, 2017 from <http://amara.org/>
4. Bernstein, M.S., Little, G., Miller, R.C., et al. Soylent: a word processor with a crowd inside. UIST’10, ACM Press (2010).