Optimization Strategy of Interaction Design for Foreign Language Asynchronous Online Learning

Mengni Yang and Mengfei Liu

School of Design, Hunan University, Changsha, HN 430000, China

ABSTRACT

With the development of mobile communication technology and the outbreak of COVID-19, the much-anticipated online learning has ushered in a new growth peak. Among them, online language learning, a segmented field of online learning with a relatively long development, has a vast market scale and is expanding at a high growth rate. Nowadays, numerous participatory video websites offer high-quality foreign language online courses. Learners favor these courses because they can use the fragmented time for conducting asynchronous online learning regardless of space and time. However, foreign language learning has strong interactive and practical characteristics, while the current websites have insufficient functional support, resulting in low learner engagement, less sense of gain and unsustainable learning. Therefore, it is necessary to explore the needs of foreign language learners to design more targeted website features and optimize the asynchronous online learning experience. We select a typical website (Bilibili) in China and summarize four types of foreign language courses on the website according to the teaching method, teaching duration, teaching content, and interface: lecture course, circumstance course, demonstration course and resource course. Through behavior analysis, sentiment analysis, and topic detection, we analyze the danmaku data of each type of course to obtain the learners' behavior patterns, text sentiments, and hidden topics effectively. Combining the course type and data analysis, three learners' needs are summarized: the learning need to imitate and follow the teacher, the interactive need for frequent communication with teachers or study partners, the emotional need for a sense of ritual in class and the desire for participation and presence. With this as the foundation support, we propose three directions of targeted interaction design optimization strategies for participatory websites. Increase the multimodal interaction function, such as writing and voice following, and dual channels to help learners imitate language learning. Increase teacher-student and student-student interaction functions such as voting selection and optimizing the form of information feedback from teachers and learning partners. Enhance the ritual of online learning and improve the learning function system of the website such as designing interactive buttons at the beginning and end of the course, and visualizing course learning records. It is expected to provide suggestions for developing and reforming participatory websites to optimize the learning experience of foreign language asynchronous learning and to provide a scalable perspective for the optimization of other learning subjects.

Keywords: Asynchronous online learning, Foreign language learning, Participatory website, Danmaku, Learning analytics, Interaction design

INTRODUCTION

With the development of mobile communication technology and the outbreak of COVID-19, the much-anticipated online learning has ushered in a new growth peak. Among them, online language learning, a long-developed segment of online learning, has a considerable market size and is expanding at a high growth rate. Nowadays, the better development of language learning is inseparable from the support of network resources and information technology (Chen, 2019). Asynchronous online learning is one of the main modes of online learning, which means that the teachers and students are separated in space and time (Liu, 2020). Students can achieve 4A (Anyone, Anytime, Anywhere, Anything). They could choose learning content and self-regulate their learning time and pace according to their life and learning situation (Liu, 2021).

Today, participatory video websites such as YouTube and Bilibili offer many online language courses for the public to learn online asynchronously. Participatory websites (Walther and Jang, 2012) have been an emerging class of computer-mediated communication (CMC) systems in recent years, which always present a mix of messages from different sources of authors on the same page. The ultimate message mixing is online video sites offering a danmaku system, where the site publishes videos and users contribute danmaku to the videos (Qian et al., 2022). However, language learning differs from other subjects and is more practical, communicative, contextual, and consolidative, requiring learners to use language for practical use and communication (Nguyen, 2008). Current websites have insufficient functional support, resulting in low learner engagement, less sense of gain and unsustainable learning. Therefore, it is necessary to explore the needs of foreign language learners to design more targeted website features and optimize the asynchronous online learning experience.

The video's interaction behavior data is an essential basis for learner analysis and demand mining, and danmaku is learners' primary data in participatory websites. Danmaku originated in Japan as a series of time-synchronized annotations (TSCs) scrolling across the top of a video, an interaction between users across time and space, and a channel for them to express their opinions and attitudes (Bai et al., 2019). Some researchers have seen the potential of danmaku in online learning. Yao et al. (2017) used information based on danmaku to improve user engagement and interaction in online video learning. Leng et al. (2016) studied the effect of danmaku on learning outcomes based on eye gaze data. However, there is less literature on TSC analysis of foreign language online courses. This paper aims to explore TSCs from language courses to analyze the needs of foreign language asynchronous online learning and propose interaction design strategies that are essential for optimizing participatory websites.

ANALYSIS PROCESS

Sample Screening of Foreign Language Videos

This paper selects a typical participatory website in China (Bilibili), which attracts young people to use the platform for entertainment and learning with its rich video resources and TSC system (Bo, 2020). After looking

through the videos on the site, we find that English and Japanese courses make up most of the language courses on the site. Language courses can be divided into four types: lecture course, circumstances course, demonstration course and resource course, based on the dimensions of the teaching method, teaching duration, teaching content, and interface. Table 1 shows the details.

The lecture course is usually in the form of live replays or screencasts, with the longest average teaching time of any type. Its most significant feature is that the teacher often interacts with the learners during the courses. They ask questions to learners, write with an electronic stylus or type with a keyboard to guide their interaction and synchronized thinking. The circumstances course is beautifully produced and rich in knowledge presentation, with text, pictures, videos, and real-life conversations to give learners a sense of presence. Usually, the teacher will be integrated into the courseware through green screen keying to show the learner the complete knowledge content without simultaneous modifications. The demonstration course is mainly in speaking, pronunciation, and writing, with the teacher or a foreign teacher demonstrating the application. Most of them are live demonstrations with subtitles and annotations. The resource course is mostly about listening, text reading aloud and animations collected by the uploader and does not contain any additional explanations or commentaries. Such resources allow learners to use them freely for purposes such as shadow-following, listening tests, and developing a sense of language.

Table 1. Type foreign language courses in Bilibili.

Type	Average duration	Teaching content	Teaching method	Interface
Lecture course	More than 40 minutes	vocabulary, grammar, pronunciation	Teach, display courseware, typewrite, electronic handwrite	Courseware, teacher images, TSCs during live streaming
Circumstances course	10–20 minutes	Vocabulary, grammar, text, pronunciation, culture	Teach, display courseware	Courseware, teacher images, pictures, videos, animation
Demonstration course	Less than 10 minutes	Vocabulary, pronunciation, writing	Demonstrate pronunciation or writing	Teacher images, subtitles, pictures
Resource course	Less than 15 minutes	Pronunciation, listening	Play resources	Various materials, subtitles

TSC Mining Analysis Model

A behavior-sentiment-topic TSC mining analysis model is developed, as shown in Figure 1. Based on user data, behavior analysis is a statistical method for analyzing how people behave or act. For danmaku video viewers, sending TSC is essential in watching videos. Their behavior can be inferred from the relationship between the number of TSCs and time. We define the timeline as the running time of a video, so the relationship is the change in the number of comments during the video playback (Bai et al., 2019). Sentiment analysis is a textual computational process that classifies the perceived author's attitude as positive, negative, or neutral (Medhat et al., 2014). According to official information, the text uses the sentiment tendency analysis service provided by the Baidu AI Open Platform, which can achieve an accuracy rate of 95%. Topic detection is an automatic technique for discovering topics in extensive text data (Mottaghinia et al., 2021). LDA is one of the most popular models (Jelodar et al., 2019), which is a generative probabilistic model for discrete data collections (e.g., text corpora (Blei et al., 2003)) that can be used to discover a user-specified number of topics shared by documents in a text corpus.

We select two Japanese and English videos for each category, 16 samples in total. Numbers 1 to 4 are lecture courses, 5 to 8 are circumstances courses, 9 to 12 are demonstration courses, and 13 to 16 are resource courses. Due to the linguistic specificity of foreign language TSCs, we separate Chinese text from other text, and comments containing Chinese and foreign language are classified as other text. Finally, we obtained 41,742 valid TSCs (as of January 06, 2023).
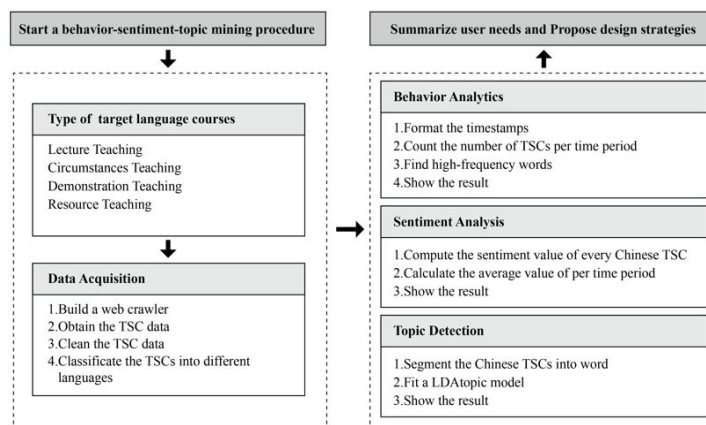

Figure 1: Behavior-sentiment-topic TSC mining analysis model.

DATA ANALYSIS

Behavior Analytics

Because the video length differs, we unify 60 time points to calculate the number of Chinese TSCs and foreign language TSCs in each period. The trend change in Chinese TSCs is shown in Figure 2, and the change in foreign
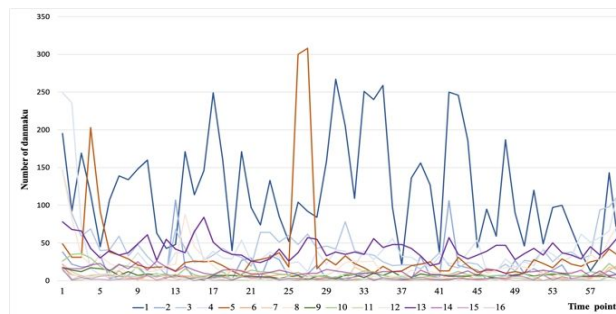
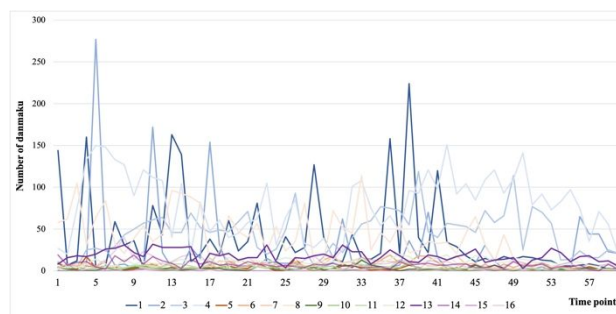**Figure 2:** Trend change in Chinese TSCs.

Figure 3: Trend change in foreign language TSCs.

language TSCs is shown in Figure 3. The TSCs word frequency analysis is performed in the period with noticeable change, and the corresponding course content is recorded. The partial result is shown in Table 2.

From Figure 2, Figure 3 and Table 2, we can see that the Chinese TSCs of all types of courses significantly increase at the beginning and end of the course. This is because learners like to leave records during these two time periods. The high-frequency words " daily attendance" and "second time" indicate the marks of their learning. The words "thank you" and "88 (bye-bye in Chinese)" express their gratitude to the teacher and farewell to the teacher and other learners. This also maps the rituals at the beginning and end of traditional offline courses to online.

The number and magnitude of fluctuations of Chinese and foreign language TSCs for lecture and circumstances courses are higher than others because these two types cover more content and are longer. There are two main reasons for the high amplitude fluctuation of TSCs, and one is that learners like to use TSCs to respond to the teacher's questions and requests. For example, at 1014 seconds of sample 1, the teacher asked, "send a yes if you're right or a no if you're wrong." At 694.2 seconds of sample 5, the teacher said, "If you understand, please type it on the public screen." Such leading questions and requests can trigger many learners to send TSCs for interaction. Another reason is that learners imitate the teacher's demonstration or repeat the material's content. For example, when the teacher demonstrates vowel

Table 2. Partial result of TSCs high-frequency words and course content.

| Number	Language	Time(s)	High Frequency Words	Video content
1	English	1014	Yes, I, kiss	Judgment questions, the teacher said, "send a yes if you are right or a no if you are wrong"
	Chinese	3354	88, goodbye, fighting	Teacher concludes and says goodbye to students
2	Japanese	4095	Nihongo, no, hon	The teacher explained Japanese grammar, using the example of "Nihongo no Hon".
	Chinese	105	For, coming, Daily attendance	The teacher says the opening speech and introduces the course content
5	English	694.2	Ok, simple, get it	After the teacher finished the summary, he said, "If you understand please type it on the public screen"
10	English	99.45	Ei, ok, gay	Teacher demonstrates reading aloud the vowel "ei"
	Chinese	345.15	Thank you, teacher, excellent	Teacher summary of the lesson
13	English	111.3	Allergy, similar, it	Listening Play: "it was similar to my childhood allergy."
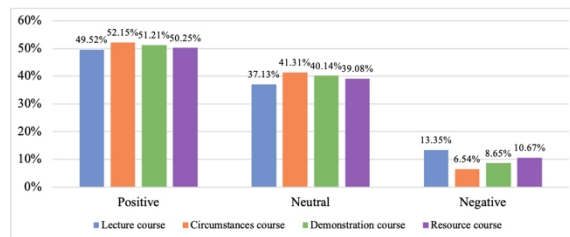
Figure 4: Average percentage of sentiment value.

pronunciation in sample 10 and the listening answers are shown in sample 13, learners send the same TSCs as the lesson content. Learners would translate their phonological knowledge into textual TSCs sent in the pragmatic course to practical knowledge. From this, we can find that guided interaction and imitative application of knowledge are critical behaviors of learners.

Sentiment Analytics

sentiment values were calculated for each Chinese TSC text on a scale of positive (2 points), neutral (1 point), and negative (0 points). Figure 4 shows the average percentage of sentiment types. Figure 5 shows the trend of the sentiment values.
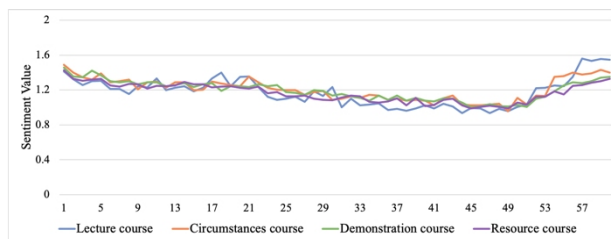
Figure 5: Trend of the sentiment values.

As shown in Figure 4, the sentiments are primarily positive or at least neutral, and the negative sentiment is below 20%, which indicates that most learners feel good. However, some students still have difficulties with the course content. The trend of sentiment change is roughly the same for all four types of courses. The sentiment is initially relatively high due to learners' excitement about the new study. Over time, they quickly became frustrated by the difficulty of the content and the reduction in their attention span. In the middle of the course, their mood fluctuated, most often due to the teacher asking questions for interaction or interesting topics in the class. Towards the end of the period, the value rose significantly because the learners felt satisfied that they were about to complete the course. The decreasing trend of lecture courses is most significant because of the longer duration and more content, which burdens students more. Demonstration and resource courses have fewer types of knowledge and are shorter in time, so the magnitude of affective change is smoother than other types.

Topic Detection

We use the LDA implementation in the Gensim Topic Model tool library and sets the number of topics K = 2, K = 3, K = 4,..., K = 8. The experimental results show that the clustering results are good when the number of topics is set to 4 (hyperparameters $\alpha = 0.25$, $\beta = 0.25$). As shown in Table 3, there is enough differentiation among the topics, and the topic words can fully reflect the semantics of the topics.

The topic of sign-in shows that learners are not only leaving their learning marks but also greeting other learners watching with words such as "10 people" and "hello." The topic of expressing learning feelings shows that learners use TSCs to keep track of their feelings while learning, whether positive or negative. In expressing emotion, the TSCs of learner-teacher interaction are mostly positive, and the whole online learning atmosphere is good. The highest percentage of topics is the discussion of course knowledge. For language learning, learners focus on "pronunciation", "intonation" and "grammar." They are highly motivated to discuss language knowledge with teachers and other learners, answer questions initiated by teachers in the videos, and take the initiative to correct mistakes for other viewers.

Table 3. TSC content features classification.

Topic	Percentage	Example of topic words	Example of TSCs
Sign	25.96%	Daily attendance, time, hello, 10 people, 2021	First day of daily attendance 2021/3/29
Express the feeling of learning	19.12%	Can't understand, difficult, 1.5x speed, simple, got it	It's all simple sentences Don't understand the meaning at all
Express the emotion	25.08%	Ha-ha, smile, awesome, thank you, like	Ha-ha, the teacher is so affable Thank you, teacher
Discuss course knowledge	29.83%	Pronunciation, Intonation, grammar, words, listen	The tenth vowel is pronounced without the force of the lips Not sore, the word has no verb

DESIGN STRATEGY

Combining the course type and data analysis, three learners' needs are summarized: the learning need to imitate and follow the teacher, the interactive need for frequent communication with teachers or study partners, the emotional need for a sense of ritual in class and the desire for participation and presence. With this as the foundation support, we propose three directions of optimization strategy.

Increase Multimodal Interaction Function

According to Harmer (2008), human memory is divided into short-term, working, and long-term memory. The ultimate goal of the learner's imitation of language is to transform knowledge from short-term to long-term memory. Multimodal learning that uses different senses such as vision, auditory and haptic to integrate information and absorb knowledge is increasingly emphasized (Gilakjani et al., 2011). Nowadays, multi-touch and other haptic technologies such as keyboard input, gesture operation and electronic handwriting have been widely used (Lin and Song, 2010). Notably, numerous cognitive psychological and neurological studies (Grabowski, 2007) have shown that handwriting can help encode information better and help learners with content recollection and recall (Glenberg and Robertson, 1999). When a teacher explains words or phrases, the platform can provide haptic interaction features to enhance memory. In addition to the existing keyboard typing, a touch screen handwriting area can be added to convert written text into TSCs or notes.

In a conventional classroom, the teacher would ask students to follow along as they demonstrate reading aloud, which combines "listening" and "reading." The information is input by the auditory organ and output by the oral articulatory organ, and this process fully engages multiple senses and objectively improves learners' memory of foreign speech (Tergujeff, 2012).

Experts have discovered that after following and imitating ten standard recordings, students' pronunciation and phonetic intonation can vastly improve, as can their ability to recognize sounds (Ji, 2008). In demonstration and resource courses, the platform can configure shortcut buttons to allow students to independently select which segments to repeat and how many times to repeat them to follow along. The voice can also be converted to TSCs.

Increase the Interaction Function of Teacher-Student and Student-Student

In the asynchronous online learning, the interaction between learners and teachers is inter-temporal but this behavior can increase learners' sense of presence and participation. For long courses like lecture courses, exercises or votes could be inserted in the critical content to guide learners to post their understanding actively. When many TSCs with the same perspective as theirs fly by, students experience a strong sense of satisfaction and even elation (Guo, 2021). This can also alleviate the emotional decline and inattentiveness of learners due to long hours of study. At the same time, these interactive data can also help producers understand the effectiveness of learners' learning. The error rate of the exercises and the ranking of the votes can help teachers determine the acceptance and difficulty of the teaching content. Learners can communicate with each other simultaneously, and the current site only shows the number of people watching simultaneously. It is possible to further differentiate the style of TSCs sent by learners watching at the same time so that learners are more aware of their peers' presence. Alternatively, temporary anonymous voice chat rooms can be set up to give learners a space to use the language in a practical way and communicate with each other.

Enhance the Ritual of Online Learning

The interactive features of the ritual increase the motivation and focus of the learner. The platform can design quick TSC buttons, such as "daily attendance" at the beginning and "thank you, teacher" at the end of the course. The website could formally build a learning management system, such as adding record classification to visualize learning progress. Especially for demonstration and resource courses that need to be watched repeatedly, it displays the number of times the learner has protected them, aiding learning planning while satisfying the learner's sense of accomplishment.

CONCLUSION

In this paper, we summarized four types of foreign language courses in Bilibili based on the dimensions of teaching method, teaching duration, teaching content and interface. Then a behavior-sentiment-topic analysis model was established and applied to analyze the TSC data of the language learning courses. We drew three needs of learners in asynchronous online learning of foreign languages, thus further proposing three directions of interaction design optimization strategy. This paper enriches the gap between TSC

analysis and online language courses and advances the development of asynchronous online learning of languages. In the future, driven by technological developments, the online platform will have more special interactive features for languages and other subjects, giving learners more autonomy and flexibility for asynchronous online learning.

ACKNOWLEDGMENT

REFERENCES

Bai, Q., Hu, Q. V., Ge, L.and He, L. (2019). Stories that big danmaku data can tell as a new media. IEEE Access, 7, 53509–53519.

Blei, D. M., Ng, A. Y.and Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3, 993–1022.

Bo, Y. Learning Alone yet Together: An Exploratory Study of the Interaction through Online Danmaku Videos. 2020 IEEE Learning With MOOCS (LWMOOCS), 2020. IEEE, 100–104.

Bullet Screen. Journal of Guizhou Radio & TV University, 7, 417–430.

Chen, J. (2019). Advantages of online courses in foreign language teaching. Youth& Society, 105.

Gilakjani, A. P., Ismail, H. N. and Ahmadi, S. M. (2011). The effect of multimodal learning models on language teaching and learning. Theory & Practice in Language Studies, 1.

Glenberg, A. M.and Robertson, D. A. (1999). Indexical understanding of instructions. Discourse processes, 28, 1–26.

Grabowski, J. (2007). The writing superiority effect in the verbal recall of knowledge: Sources and determinants. STUDIES IN WRITING, 20, 165.

Guo, X. (2021). Research on the Identity Phenomenon of Bilibili Users under the Barrage Culture. Journal of YUlin College, 31, 101–104.

Harmer, J. (2008). How to Teach English (Second Edition). ELT Journal, 62, 313–316.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y.and Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools and Applications, 78, 15169–15211.

Ji, Q. (2008). Research on the Application of Follow-up Reading Training in College English Listening Teaching. Journal of Beijing Vocational College of Labour and Social Security, 56–58.

Leng, J., Zhu, J., Wang, X.and Gu, X. Identifying the potential of Danmaku video from Eye Gaze Data. 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT), 2016. IEEE, 288–292.

Lin, M.and Song, R. (2010). A stroke-segment-mesh (SSM) glyph description method of Chinese characters. Journal of Computer Research and Development, 47, 318–327.

Liu, X. (2021). "Synchronous" and "Asynchronous": Thoughts on Two Online Teaching Modes of Distance Education. Journal of Guizhou Radio & Television University, 29, 11–15.

Liu, Y. (2020). Historical retrospect and status analysis of synchronous/asynchronous online teaching mode. Modern English, 10–12.

Medhat, W., Hassan, A.and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams engineering journal, 5, 1093–1113.

Mottaghinia, Z., Feizi-Derakhshi, M.-R., Farzinvash, L.and Salehpour, P. (2021). A review of approaches for topic detection in Twitter. Journal of Experimental & Theoretical Artificial Intelligence, 33, 747–773.

Nguyen, L. V. (2008). Computer mediated communication and foreign language education: Pedagogical features. International journal of instructional technology and distance learning, 5, 23–44.

Qian, Y., Yajing, D., Jiaming, L.and Shilun, G. (2022). Research on User Interaction Experience in Participatory Websites Based on the.

Tergujeff, E. (2012). English pronunciation teaching: Four case studies from Finland. Journal of language teaching and research, 3, 599–607.

Walther, J. B.and Jang, J.-w. (2012). Communication processes in participatory websites. Journal of Computer-Mediated Communication, 18, 2–15.

Yao, Y., Bort, J.and Huang, Y. Understanding Danmaku's potential in online video learning. Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems, 2017. 3034–3040.