A Review of On-Screen Texts in Multimedia and Audiovisual Inputs for Second Language Learning

Fang Liu

Department of Literature, Capital Normal University, Beijing, China
Corresponding author: 1220106135@cnu.edu.cn

Abstract:

As an essential component of multimedia input, on-screen text effectively enhances second language vocabulary, listening, and grammar learning. The on-screen text also benefits second language learning in on-screen texts+audio and on-screen texts+audio+video combined modes as an important part of audiovisual input. The principle that on-screen texts function in the two modes is supported by different theories, especially the Cognitive Load Theory and the redundancy principle within the cognitive theory of multimedia learning. After distinguishing the differences between animation in multimedia input and video in audiovisual input, this paper summarizes the characteristics of on-screen texts in promoting comprehension, avoiding redundant input, and reducing cognitive stress. This paper introduces three types of subtitles commonly used in second-language teaching, emphasizes the importance of keyword-highlighted subtitles, and reviews the development status of several other types of on-screen texts in the current research field.

Keywords: On-screen texts; multimedia input; audiovisual input; second language learning.

1. Introduction

On-screen text is widely used in multimedia teaching with the advancement of educational technology. It plays a crucial role in multimedia and audiovisual input modes and is generally considered beneficial to second language learning [1]. There is a subtle difference in the role of on-screen text between the two modes, which mainly derives from the difference in the features of animation in the multimedia input and video in the audiovisual input [2]. Given this, it is necessary to review how on-screen text works in different modes and, on that basis, to distinguish its impact on second language acquisition (SLA).

1.1 Multimedia and Audiovisual Input Modes

There are several elements of multimedia input: text, image, audio, animation, and subtitles/captions [1]. In SLA, multimedia input is a tool used to present instructional content. Considering that the core of this paper is a review of the function of screen texts for second language learning, the instructional content is mainly shown in a model combining three elements: audio + animation + subtitles/captions [1].

Audiovisual input mode can be seen as a part of multimedia input mode, which mainly focuses on three main elements: audio, video, and on-screen text. Compared to traditional multimedia input modes, on-screen text uses video, expands the range, and refines the variety of subtitles and captions. More definitions are given to dual subtitles, reversed subtitles, keyword captions, and so forth.

1.2 On-screen Text

On-screen text means text relating to the content of a video that is displayed on the screen either synchronously or asynchronously [2]. Three types of on-screen text are used frequently in SLA: subtitles, captions, and dual subtitles. During instructional processes in first language (L1) teaching, subtitles are considered more effective than one singular video [3, 4]. In this regard, captions are more influential than subtitles in promoting second language learning, according to previous studies [2]. Dual subtitles are also considered beneficial for second language acquisition, but related research is limited.

1.3 Cognitive Theory of Multimedia Learning

The cognitive theory of multimedia learning is associated with two theories. The dual coding theory suggests that language learners cope with diverse multimedia input forms through auditory (audio) and visual (video) channels. The integration of auditory and visual channels is more effective in enhancing human information processing than a single channel. [1, 2]. The Cognitive load theory asserts that the cognitive load is related to redundant inputs and the limited capacity of human memory [5].

Based on these two theories, three assumptions were hypothesized. First, the modality principle states that learners process instructional content within auditory and visual channels [6]. Second, the redundancy principle contends that learners have limited cognitive resources, and redundant input within one channel can distract attention [1, 7, 8]. Third, the coherence principle argues that learners process information coherently, and irrelevant information can disrupt the processing of instructional content [1].

1.4 Research Gaps

Although many investigations of on-screen texts have been conducted in multimedia and audiovisual input modes, shortcomings still exist. First, traditional multimedia input has not considered the role of video as one of the multimedia elements. In second language (L2) education, videos often convey less semantic content than animations, resulting in lower cognitive load [3]. Second, the ideas of dual subtitles, reversed subtitles, and other subtitle forms have not been thoroughly researched in audiovisual input. Third, the positioning, size, color, and duration of different types of subtitles and captions should also be considered within the scope of future research [3, 7].

2. The Effectiveness of On-screen Texts

2.1 On-screen Texts within Multimedia Input

Subtitles and captions are considered as a unit in the concept of multimedia input. L1 subtitle (L1 text, L2 video) is an interlingual subtitle that translates the foreign language into the learner's native language [2, 6]. L2 subtitle/caption, on the other hand, is an intralingual subtitle that provides learners with the original foreign language subtitles [2, 6].

Research indicates that there are two modes beneficial for second language learning within multimedia elements incorporating subtitles/captions. The first is a dual-element combination model: audio+subtitles/captions [1]. The second is a triple-element combination model: audio+animation+subtitles/captions [1]. Since animation/video is an important carrier of on-screen text, the dual-element model is not considered in the scope of this review.

The triple-element combination model audio+animation+subtitles/captions are beneficial for vocabulary, listening, and grammar learning in terms of second language learning [2]. For vocabulary learning, subtitles facilitate the activation of vocabulary representations and enable learners to focus on capturing key information from the subtitles [5]. For listening learning, subtitles provide textual support, reduce the pressure on learners to process the audio information, lower the cognitive load through learning,

and enable learners to better process the information in depth [9]. For grammar learning, subtitles also provide more grammatical information, making learners more inclined to focus on the learning activity [1, 2].

The major difference between animation in multimedia input and video in audiovisual input is the amount of information carried [3]. Teaching videos usually focus on filming the main teacher, which limits the amount of teaching information the videos carry. However, animation is often a visual presentation of the teaching content. Without considering the existence of on-screen texts, the animation itself carries more information than the video and, at the same time, may generate more redundant information [7]. Therefore, compared with multimedia input, placing on-screen texts in audiovisual input is more conducive to reducing the cognitive pressure on learners [1, 3]. Consequently, animations with more information are easier for learners to understand without on-screen texts [3, 7].

2.2 On-screen Texts within Audiovisual Input

2.2.1 Combination of different elements containing on-screen texts within audiovisual input

Within the combination of three elements in audiovisual input, the on-screen texts and video modes cannot allocate the learner's attention from the auditory channel to supplement additional information. Therefore, the combination of the three elements requires integration with audio. Research indicates that the mode of on-screen texts+audio is more advantageous for learners to comprehend audio content than the video+audio mode [7]. The dual-element mode of video+audio, compared to a single mode of audio or on-screen texts+audio, relies more on learners' cognitive abilities and affects the allocation of learners' limited cognitive resources [5]. Subtitles guide attention, reducing learners' reliance on cognitive abilities and alleviating cognitive pressure. In summary, adding subtitles is more beneficial for readers to extract key information from audio content that provides important information [8].

In the triple-element mode of audio + video + on-screen texts, video complements audio information with scenes and textual content. The material combines audio and visual elements, which makes it less challenging to produce the same content as pure audio material. Its vivid, realistic, and contextualized content aids learners in understanding the material [9]. When supplemented with explanatory subtitles, unnecessary, redundant information is significantly reduced. However, if there is a significant gap between the difficulty of the audiovisual material and the learners' proficiency level, such as providing high-difficulty material to low-achieving learners or low-difficulty material to high-achieving learners, reading subtitles may

interfere with audiovisual comprehension [1]. Therefore, audiovisual materials should be tailored to the learners' proficiency levels, opting for moderately challenging materials to prevent learners from becoming distracted or encountering redundant information.

2.2.2 Different types of on-screen texts

The on-screen text includes different types of subtitles and captions. Current research demonstrates that subtitle (L1) is beneficial for novice learners of a second language, while caption (L2) is more effective for learners with some proficiency [6]. In addition to these two types of subtitles, dual subtitles are also widely used in educational videos. The dual subtitle refers to two lines of subtitles appearing simultaneously on the screen, with the top subtitle being in L1 and the bottom one in L2. Compared to L1 subtitles, dual subtitles will establish a bond between the literal and oral modes of words incidentally, aiding lower-level learners in better recognizing the meanings of words and building relationships between L1 meanings and L2 forms [10].

In education, three common types of subtitles are used: full subtitles, keyword subtitles, and keyword-highlighted subtitles [11]. Full subtitles appear synchronously with the teacher's speech, providing consistent subtitles. Keyword subtitle extracts important information from the teacher's speech content. Keyword-highlighted subtitle combines the two types mentioned above. Keyword-highlighted subtitles are more advantageous for second-language learning within these three subtitles. Learners noticing inconsistencies between visual and auditory content stimulates their learning motivation, prompting them to match the two types of information conceptually. Keyword-highlighted subtitles provide complete teacher speech content, reducing the cognitive load on the auditory channel during learners' process of concept matching using both channels, deepening learning and memory [5, 7, 9].

Among other subtitle types, glossed captions have a significant impact. Glossed captions (L1 gloss, L2 text, L2 video) are believed to enhance learners' understanding of vocabulary and supplement their lack of knowledge of cultural background [2, 12]. However, excessive annotations can impose cognitive load, hindering learners from building connections and transferring knowledge.

Reversed subtitles (L2 text, L1 video) are also an emerging type of educational subtitles [2, 6]. Studies have found that the reading process with subtitles and captions is smoother and more systematic than with reversed subtitles [6]. The extent to which the irregularity of reversed subtitles affects second-language learning requires further research.

3. Discussion

Few other subtitle types have been introduced in multimedia input modes. Since animation communicates more information, it might be necessary to consider the implementation of keyword subtitles and keyword annotated subtitles [3]. While experiments on facilitating second-language learning tend to address objective aspects such as the effect of subtitle adjustment techniques on second-language learning, eye-capture techniques for directing learners' attention, and so on, the learner's idiosyncrasies are rarely considered [13]. Raising language learners' awareness of their learning achievements might better guide learners' use of subtitles and captions [14].

4. Conclusion

Generally, on-screen texts play a beneficial role in both multimedia input and audiovisual input in second language learning. Combining different elements, such as audio and video, on-screen texts invoke both auditory and visual channels to help learners process learning content and reduce redundant information input. Nonetheless, its effectiveness is constrained by learners' cognitive capacities and proficiency levels. The limited cognitive resources should be reasonably allocated in the learning content carried on the dual channels to avoid distracting learners' attention and interrupting consistency in information processing. The complexity of the learning material should likewise be matched to the learner's level of proficiency to ensure that the learner absorbs the instructional information effectively. Hence, different types of subtitles should be selected under different conditions and learning targets. Subtitles should be provided for beginners, captions for basic learners, dual subtitles to construct the relationship between the learner's mother tongue and the second language, and keyword-highlighted subtitles to deepen the learner's memory. In summary, integrating on-screen texts in multimedia and audiovisual inputs represents a potent strategy for enhancing second language learning. By leveraging the synergies between auditory, visual, and textual elements, educators could create a conducive learning environment that maximizes cognitive engagement, ultimately empowering learners in their language learning.

References

[1]Zhang, R., & Zou, D. (2022). A state-of-the-art review of the modes and effectiveness of multimedia input for second and foreign language learning. Computer Assisted Language Learning, 35(9), 2790-2816.

[2]Wei, R., & Fan, L. (2022). On-screen texts in audiovisual input for L2 vocabulary learning: A review. Frontiers in Psychology, 13, 904523.

[3]Wang, C. (2016). Are subtitles necessary in online learning video? The modification of redundant effect. e-Education Research, 3,59-65+85.

[4]Diao, Y., Liu, M., & Yang, H. (2014). Study on online video open courses learning behavior in the eye tracker test. Modern Educational Technology, 11,81-87.

[5]Zheng, Y., Ye, X., & Hsiao, J. H. (2022). Does adding video and subtitles to an audio lesson facilitate its comprehension? Learning and Instruction, 77, 101542.

[6]Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. System, 63, 134-148.

[7]Meng, X., & Xu E. (2020). Do animated online teaching videos require presenting subtitles? - experimental analysis based on the comparison of college students' academic performances. Modern Educational Technology, 30(9), 42-48.

[8]Chan, W. S., Kruger, J. L., & Doherty, S. (2019). Comparing the impact of automatically generated and corrected subtitles on cognitive load and learning in a first-and second-language educational context. Linguistica Antverpiensia, New Series - Themes in Translation Studies, 18.

[9]Shaojie, T., & Samad, A. A. (2022). Systematic literature review on audio-visual multimodal input in listening comprehension. Frontiers in Psychology, 13, 980133.

[10]Wang, A., & Pellicer-Sánchez, A. (2022). Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study. Language Learning, 72(3), 765-805.

[11]Yang, J., Chen, H., Yang, W., Li, Y., Yu, Q., & Pi, Z. (2022). How to design subtitles in instructional videos? - A meta-analysis based on 44 experimental and quasi-experimental research literatures from 2000 to 2021. e-Education Research, 8,59-67.

[12]Teng, F. (2022). Vocabulary learning through videos: captions, advance-organizer strategy, and their combination. Computer Assisted Language Learning, 35(3), 518-550.

[13]Hu, K., O'Brien, S., & Kenny, D. (2020). A reception study of machine translated subtitles for MOOCs. Perspectives, 28(4), 521-538.

[14]Pattemore, A., & Muñoz, C. (2024). Perceptions of learning from audiovisual input and changes in L2 viewing preferences: The roles of on-screen text and proficiency. ReCALL, 1-17.