Research

On the effectiveness and limitations of captioning in L2 listening

Saadet Korucu-Kış [1]

Abstract:

Listening is often perceived to be the most challenging skill by second/foreign language (L2) learners. Due to its real-time nature, L2 listeners experience several comprehension problems related to the processing of aural input. To scaffold L2 listening, captioning is commonly used since the dual coding of aural and written stimuli is expected to make L2 input more comprehensible leading to more in-depth processing. However, a survey of the extant literature precludes us from drawing firm conclusions about the effectiveness of captioning since in some circumstances captions were found to have no significant effect on listening comprehension. So, the question of whether captions function as a comprehension aid in L2 listening remains inconclusive. Hence, adopting a narrative literature review methodology, the present study aims to contribute to this inconsistent research area by clarifying some of these issues answering the following questions: (1) Is captioning really effective in L2 listening?, (2) Does captioning always work for L2 listening?, and (3) Why is research on captioning in L2 listening still inconclusive? Based on the insights gained, it is concluded that the mere presence of captions does not necessarily lead to improved comprehension. Captioning effectiveness is influenced by learner, material, measurement, task, and L1/L2 characteristics. Implications arising are discussed.

Keywords: Captioning, L2 listening, comprehension, audiovisual materials

---

[1] Assistant Professor, Necmettin Erbakan University, Faculty of Education, Konya, Turkey. skorucukis@gmail.com,
Orcid ID: https://orcid.org/0000-0002-1128-1747

INTRODUCTION

Listening has been a neglected skill within applied linguistics for a long time (Vandergrift, 2007). Traditionally, it was regarded as a natural skill that developed over time and did not need to be taught. It was reckoned to be an effortless process through which a listener passively received the incoming input (LeLoup et al., 2007). However, with advances in the fields of recording technology, telecommunications and information processing, computational science and cognitive psychology, transpersonal psychology, globalism and anthropology, organizational behavior, computer technology, and digital networking, the characterization of listening has changed markedly. Moving away from being perceived as a mere act of passive reception, it has come to be viewed as an active process through which listeners construct, represent, negotiate, and create meaning individually or in collaboration with others through imagination, participation or empathy (Rost, 2011). In other words, it has now been well recognized that listening is an interpretive process through which L2 learners actively engage in "constructing meaning based on expectations, inferences, intentions, prior knowledge, and selective processing of the input" (Richards, 2005, p. ix).

Listening is critical to the process of language acquisition as the amount and quality of aural exposure to the target language influences language learners' overall L2 proficiency. Vandergrift (2003) points out that prioritizing listening comprehension especially in the early stages of language instruction provides a more naturalistic approach to acquire the target language. According to Brett (1997), this will involve learners in an implicit process which enables them to internalize the new language and automatically reproduce it. Similarly, Gary and Gary (1981) state that engaging language learners in substantial listening comprehension practices offers several benefits for language development such as cognitive, affective, communicative, and utility benefits. The authors further maintain that developing substantial listening competence in learners before requiring production (a) prevents cognitive load and prepare learners to produce the material without language interference, (b) reduces stress, (c) increases the potential for communicative acts, and (d) enhances self-confidence, which may, in turn, results in autonomous study practices. In similar veins, Vandergrift (2007) states that language learners' competency in listening plays a central part in the development of other skills. To illustrate, Richards (2008) indicates that listening can be viewed as the initial step of communication and much of the input needed for oral production is provided by it. In his review study, Jones (1997) contends that listening plays a critical part in pronunciation development and will continue to be one of the effective means of pronunciation training through different tasks with a diversity of accents. Highlighting the reciprocal relationship between listening and vocabulary knowledge, Stæhr (2009) concludes that there is a strong association between listening and vocabulary knowledge. He furthers that while vocabulary size may significantly contribute to learners' listening comprehension, use of different listening strategies may increase their lexical level. As to the relationship between listening and grammatical knowledge, De Jong (2005)

concludes that substantial listening can support students' implicit learning of linguistic structures.

In view of these observations, listening can be said to hold a central position in the language acquisition process. Yet, due to its multidimensional and complex nature, it poses serious challenges to language learners while processing auditory input (Yang & Chang, 2014). Wipf (1984) states that listening requires discriminating between sounds, recognizing words and linguistic forms, identifying intonation and stress, recalling, and interpreting meaning within the immediate and broader context of the utterance. According to Goh (2000), this real-time processing of data causes several comprehension problems related to the cognitive processes of perception, parsing and utilization. L2 listeners may not understand grammatical structures or lexical items, or they may forget what is heard or they may understand the individual words but not the intended message. In a similar sense, Yang and Chang (2014) point out that the real-time nature of listening does not allow learners to control the speed of ongoing streams of speech at their own pace. Feeling precipitous to understand the aural input, learners get stressed. Their working memory is overloaded and disables them to process the continuing acoustic messages. As a result, L2 listening is often perceived to be one of the most difficult language skills in practice. Particularly, in foreign language environments, it may become a source of frustration for learners given the lack of a real-world English practicing context. After all, since L2 learners' listening skills can be improved through practice, one way to help L2 learners overcome the listening difficulties is to scaffold their listening skills development by means of richer and authentic learning sources.

Audiovisual media such as feature films, drama series, sitcoms, documentaries, news, game shows and so forth can help create rich and authentic learning environments for learners (Danan, 2004). Moreover, the input provided by this media meets the five criteria outlined by Nation and Newton (2007) for suitable input. In that, audiovisuals can offer (a) familiar content, (b) interesting and engaging topics, (c) comprehensible materials, (d) contextual hints to construct meaning, and (e) large quantities of input for processing. Providing learners with multisensory experiences, they can have an attentional, emotional and motivational influence on them (Baltova as cited in Danan, 2004). The contextual, visual and non-verbal cues available in audiovisual materials can assist L2 listeners to compensate for any comprehension deficiency resulting from the provision of audio-only input (Brett, 1995).

The potential benefits of audiovisual materials for the facilitation of input processing can also be increased by means of captioning technology. Due to the visual support captions provide, L2 listeners can link what they hear to the on-screen text. They can easily break down continuous speech streams, recognize words, understand form-meaning connections, and identify meaningful speech units. The dual coding of aural and written stimuli makes L2 input more comprehensible leading to more in-depth processing. This, in turn, enhances confidence, allays anxiety and provides motivation, which is in the long run expected to

develop L2 learners' listening comprehension (Hsieh, 2020, Leveridge & Yang, 2013; Winke et al., 2013). In this respect, there is a commonsense assumption that captions improve performance and have a positive impact on listening skills (Markham & Peter, 2003) However, a comprehensive review of the extant literature (e.g. Caimi, 2006; Robin, 2007; Taylor, 2005) precludes us from drawing firm conclusions about the effectiveness of video captioning since in some circumstances captions were stated to have no significant effect on listening comprehension due to such reasons as concentration on reading the text rather than listening to the audio, heavy reliance on the text, overloaded working memory, and learner perception of captions as a source of distraction (e.g. Diao et al., 2013; Hui, 2007; Montero-Perez et al., 2013; Zanon, 2006). So the question of whether captions function as an aid or hindrance to the development of listening comprehension remains controversial. Based on a survey of primary research in the area, this brief article aims to contribute to this inconsistent research field by clarifying some of these issues answering the following questions: (1) Is captioning really effective in L2 listening?, (2) Does captioning always work for L2 listening?, and (3) Why is research on captioning in L2 listening still inconclusive?

METHOD

A narrative literature review was conducted to summarize primary research on the effectiveness of captioning in L2 listening. Baumeister and Leary (1997) state that one of the goals of narrative literature reviews is "to reveal problems, weaknesses, contradictions, or controversies in a particular area of investigation" (p. 312). According to the authors, literature reviews that focus on problem identification do not require full-size review of literature. Based on a representative cover of past work, the researcher "may venture some tentative solutions to the problems he or she identifies but is more concerned with simply informing the field that some difficulty exists", brings up more questions than s/he can respond to and leaves "it to future researchers to straighten out the mess" (p.312). Accordingly, drawing on a number of key studies—mostly referring to the works of leading researchers and most frequently cited articles in the field—, the current study aimed to identify the issues underlying the inconclusiveness in the area of captioning effects on L2 listening.

RESULTS AND DISCUSSION

Is Captioning Really Effective in L2 Listening?

Captions, also known as uni-lingual (Vanderplank, 1990), intralingual (Williams & Thorne, 2000), or same-language subtitles (Bird & Williams, 2002) refer to "on-screen text in a given language combined with a soundtrack in the same language" (Markham et al., 2001, p. 440). Primarily designed for individuals with hearing disabilities, captions have readily come to be used for language learners who are "hard of listening" since the early 1980s (Vanderplank, 1988, p. 272). One of the reasons that captioned viewing was heralded in L2 listening contexts emanates from information processing theories and SLA hypotheses.

Regarding the former, Paivio's (1986) dual coding theory and Mayer's (2005) cognitive theory of multimedia learning suggest that the use of two different modes of stimuli activate two separate systems (verbal and visual) that interact. The activation of both systems is expected to enhance information processing which subsequently results in better comprehension. From this viewpoint, onscreen text and soundtrack in captioned videos help learners code the information dually leading to more in-depth processing and higher level of recall. As to the latter, Schmidt's (1990) noticing hypothesis, and Krashen's input and affective filter hypotheses can particularly be referred to. The noticing hypothesis proposes that attending to certain features in input is a prerequisite for converting input into intake. In this sense, captions can help L2 listeners notice some rules, word forms or meaning units, lead them from unknown to known, and subsequently contribute to their listening comprehension fluency. According to Krashen's (1985) aforementioned hypotheses, a language can be acquired through understanding messages in low-anxiety environments. In light of this, Yang and Chang (2014) state that captions make incomprehensible continuous streams of speech comprehensible by the scaffolding effect they bring into auditory processing and alleviate some of the anxiety experienced by learners when listening to "native speakers each with their own slang, reduced speech, stress, accents and dialects" (Seferoğlu, 2008, p. 1).

In addition to theoretical support, empirical research also highlights the effectiveness of captions in L2 listening comprehension. To start with, Garza (1991) carried out an experiment with 70 English as a second language and 40 Russian as a foreign language university students to compare the comprehension test results of caption and no-caption groups. Having viewed a number of video segments with differing lengths, participants were tested through a multiple choice comprehension quiz. Results revealed that captions facilitated the comprehension of video segments for caption group and they significantly outperformed the no-captions group. In his study, Markham (1999) inquired into the effects of captioned viewing on 118 advanced ESL learners' aural word recognition. The materials included excerpts from two different TV programs lasting between 12 to 13 minutes. Having watched the videos once, the participants took a 50-item multiple choice test administered orally. The findings showed that captions enabled learners to identify more words and improved their listening skills to recognize words. Similarly, Huang and Eskey (1999) investigated the effects of captioned TV watching on intermediate level ESL learners' listening comprehension. The study included 30 participants and they were equally and randomly distributed among closed-captioned TV condition and no-captioned TV condition. Each group watched an episode from a TV series designed for ESL classrooms and took a listening test consisting of sixteen multiple-choice items. The researchers found that students who watched the program with full captions outscored their counterparts in the listening comprehension test.

Unlike previous studies, Guillory (1998) set out to explore how different modes of captioning (i.e. full captions, keyword captions and no captions) would impact learners'

listening comprehension. The participants consisted of 202 beginning second language learners. They were randomly assigned to three groups: full text captions treatment, keyword captions treatment and no captions group. The results of the experiment revealed the positive effects of both full captions and keyword captions treatment on listening comprehension compared to no caption condition. However, the subsequent post-hoc analysis demonstrated no significant difference between full caption and key-word caption groups. The researcher concluded that since keyword captioning presents words essential to the meaning, it provides beginning learners with less to read and does not overload their listening processing capacities.

With their study in 2001, Markham et al. attempted to broaden the scope of research on captioning by comparing the effects of intralingual subtitles (i.e. captions), interlingual subtitles (i.e. L1 on-screen text and L2 soundtrack) and no subtitles on listening comprehension. The participants of the study were 169 intermediate-level university students learning Spanish as a second language. The material comprised a 7-minute listening passage. The participants were randomly divided among interlingual subtitles condition, intralingual subtitles condition and no subtitles condition. The subsequent listening-based comprehension tests aimed to measure both global understanding through a written summary of video content and detailed understanding via a 10-item multiple choice test. The students in treatment groups outscored significantly higher than their counterparts in non-treatment group. However, the results also showed that the interlingual subtitles group outperformed the intralingual subtitles group. In view of this, the authors advocate for the graded use of multilingual captions to enhance L2 learners' listening comprehension. In a recent study, Hayati and Mohmedi (2011) also conducted a study to analyze the effects of intralingual subtitles, interlingual subtitles and no subtitles on students' listening comprehension. The participants were 90 intermediate level learners of English. The material consisted of six segments (each approximately five minutes in length) from two episodes of a documentary film. The study lasted for six weeks and students took a ten-item multiple choice test after each session. The researchers found that both treatment groups had significantly higher scores than the no caption group; yet, the intralingual group performed at a considerably higher level than the interlingual group. The researchers conclude that the additional process of translation while viewing captioned videos may cause listeners to lose track of second language audio track and therefore lowers the effectiveness of interlingual subtitles.

Winke et al. (2010) investigated the impact of captioning on the comprehension of other languages than English. The authors also examined the effects of captioning order and proficiency differences on the effectiveness of captioning. The participants were a total of second or fourth year foreign language learners of Russian and Spanish and second year learners of Chinese and Arabic. The materials included three documentaries which were 3-5 minutes long. After watching the videos twice in a randomized order, the participants sat for a multiple choice comprehension test. As a further note, the learners of Spanish were

divided into two additional groups in the study. While one group watched the videos twice with captions, the other watched them without captions. In line with their aims, the authors came up with four main findings. First, Spanish learners who watched the videos twice with captions performed significantly higher than the no caption group in the listening test. Second, the effect of the order of captions was not found to be influential in improving learners' listening comprehension. Yet, when the orthographic differences were considered, the ordering of captions was found significant. While Russian and Spanish learners did better when the captions were presented first, Arabic and Chinese learners seemed to process aural input better during their second exposure to the captions despite a lack of consistency in test scores. Finally, the study revealed that the impact of captioning order on learners' performance was not influenced by proficiency differences.

In their experimental study, Yang and Chang (2014) focused on exploring whether different modes of captioning might influence L2 learners' listening ability of reduced forms frequently appearing in colloquial English. Assigning 44 intermediate-level learners of English to full, keyword-only, and annotated keyword captions "(a format similar to pictorial presentation)" (Yang & Chang, 2014, p. 51) groups randomly, the researchers also examined pre- and post-test scores of the participants to measure their overall comprehension improvement. The materials included 51 video clips whose duration ranged from 30 to 120 seconds. Analysis results indicated that all three groups improved on their score from pre- to post-test, with the annotated key word caption group outscoring the full caption and the keyword-only caption groups, particularly in the recognition of reduced forms. The findings of the study suggested that the use of keyword captioning modes can work better than full captioning in the instruction of reduced forms which is closely related to the improvement of learners' overall comprehension competence.

In a 2017 study, Rodgers and Webb drew attention to the fact that existing research on captioning often leveraged short videos and the use of full-length TV programs that students most often watch in their out-of-class time remained largely untapped. To address this gap, the researchers conducted an experiment with 372 pre-intermediate to intermediate level university students. The L2 learners assigned to full-caption and no-caption conditions watched ten 42-minute episodes of a TV program and took a total of ten comprehension tests consisting of true/false, multiple-choice and sequencing items. The results indicated that the full-caption group had slightly better test scores compared to no-caption group across all episodes; yet, a significant difference was observed only for the episodes of one, four and seven between the test scores of two groups. These findings imply that captions can be especially useful when viewing the material for the first time and video content is comparatively difficult.

More recently, Teng (2019) examined the effects of different types of captioning (i.e. full captions, keyword captions and no captions) on 182 primary school students' comprehension of video content. The researcher also investigated the effects of proficiency level (high vs. low), and viewing frequency (once or twice) on video content comprehension.

The students were randomly distributed among three groups. The materials were two short videos from two different stories, each approximately 10 minutes in length. The measures included a written recall protocol instrument for testing global comprehension and a multiple-choice test for testing detailed comprehension. Overall, the results revealed that full captioning helped primary school students comprehend better than keyword or no captioning since full captions allowed children to construct better connections between events. Second, higher proficiency learners outperformed low proficiency learners in each condition. Finally, viewing frequency significantly influenced the comprehension of the video content.

Taken as a whole, these studies reveal that captions positively and significantly influence listening comprehension. Although one mode (i.e. full captioning, keyword captioning and annotated keyword captioning) or type (i.e. interlingual and intralingual) of captions may show superiority over others under different conditions, L2 learners in captioned viewing groups generally have substantial gains in listening comprehension compared to learners who are not exposed to any captioning condition.

Does Captioning Always Work for L2 Listening?

Although there is ample evidence that captioning facilitates aural comprehension process, it should be evident that captioning is not a panacea. Some research on captioned viewing has revealed that captions do not consistently profit L2 listeners in all cases. One of the theories often referred to in order to explain why captions can be a hindrance to aural input processing is Mayer's (2005) cognitive theory of multimedia learning. The redundancy principle of this theory suggests that the concurrent presentation of audio, imagery, and captions causes learners to divide their attention across three information channels and increases the decoding load (Montero-Perez et al., 2013). In a similar vein, Diao et al. (2007) state that captions function as a source of distraction and retard the development of L2 listening since the simultaneous presentation of multiple stimuli imposes a heavy demand on short-term memory. From a different perspective, Hui (2007) contends that captions causes text reliance in L2 listeners. Likewise, Zanon (2006) notes that students are largely inclined to focus on reading captions rather than listening to the soundtrack during captioned viewing. Furthermore, in a review study, Danan (2004) indicates that heavy dependence on on-screen texts creates a form of laziness and may not contribute to improved listening abilities.

To illustrate these issues with reference to more empirical studies, Hsieh (2020), for instance, investigated how different types of video captioning affected English as a foreign language (EFL) students' listening comprehension. The participants were 105 low-intermediate university students. The learners were randomly assigned to the following captioning conditions: no captions, full captions with no audio, full captions, full captions with highlighted target words and full captions with highlighted target words and L1 gloss. The materials were two 4-5 minute long English animations. For each video, they completed

a comprehension test consisting of ten multiple-choice items. The findings of the study showed that there was no significant difference between treatment groups in terms of listening comprehension. The author explains this result by video difficulty. Even though captions provide the textual support, low-intermediate learners may get difficulty in processing the aural input. The results also demonstrated that despite a lack of statistical difference between groups, no captions group was slightly better at listening comprehension than their counterparts. The author accounts for this finding referring to modality effect. More specifically, the absence of on-screen text had reduced the extra load in the visual system and the aural input was solely processed by the auditory channel.

Adding annotations to captioned animated videos, Aldera and Mohsen (2013) examined whether learners' listening comprehension differed under the following three conditions: animation-only, animation and captions, and animations, captions and keyword annotations. A total of 50 high-beginner EFL university students were recruited for the study. The material included an animated story. The participants took a five-item multiple-choice test and listening recall test just after watching the animation. After four weeks, the instructors administered delayed tests identical to the previous tests taken by the participants. The results revealed that the participants in the animation-only condition had outperformed the other treatment groups in both of the listening comprehension tests and recall tests. The authors attribute the negative effects of captions and annotated keywords to the attention split between three different types of stimuli and imposing cognitive burden, which result in comprehension decrements.

In their study, Bairstow and Lavaur (2012) explored whether proficiency and different types of captions (interlingual, intralingual and no subtitles) would make any difference to participants' comprehension. The participants were a sample of 90 secondary school students and they were divided into three groups. The material was an approximately 9-minute excerpt from a full-length film. Immediately after watching the video, the participants took a comprehension test consisting of multiple-choice questions. The authors stated that while interlingual subtitles facilitated video comprehension for low proficiency learners, the advanced-level learners found both intralingual and interlingual subtitles distracting. These findings suggest that captions may be found distracting if they are not needed; yet, they are perceived beneficial if needed.

Montero-Perez et al. (2014a) examined how different captioning types would affect L2 learners' comprehension of video content. The participants were 133 (higher-) intermediate undergraduates and divided among four conditions: full captions, partial captions, no captions, and full captions with highlighted keywords. The materials included three short clips. Three comprehension tests were developed based on the content of the video clips. The tests aimed to measure global understanding, detailed understanding and inference ability through open-ended, true/false and combination items. The results suggested no significant difference between the test scores of captioning and no captioning groups. These findings were accounted for by the content and difficulty level of the tests. The authors

further explained that unchallenging test items inquiring into factual information reduced the comprehension difference between the groups.

In a 2010 study on captioned viewing, Sydorenko examined the effects of captions, visuals and audio on learners' attention to input. Twenty-six beginning level university students were recruited for the study. They were divided into three condition groups: video with captions, video with audio, and video with audio and captions. The materials included three video clips from a comedy series, each 2-3 minute in length. After watching the video clips, the students were asked to rate the amount of attention they directed to captions, audio and video. They were also requested to rate the usefulness of these different types of stimuli. The authors found that the video with captions and the video with audio and captions groups focused more on captions than on video. The video with audio group, on the other hand, paid equal amount of attention to audio and video. As to the utility of the captions, visuals and audio for video comprehension, the video with audio and captions group found the audio to be the least useful. These findings indicate that learners largely relied on reading captions to comprehend the video, closely followed by their use of visual images, with listening being the least-utilized channel. Winke et al. also obtained similar results in their 2013 study conducted with 33 (low-) intermediate English learners of Arabic, Chinese, Spanish and Russian. The study investigated caption reading behaviors of L2 learners while also analyzing the relationship between the target and native language that influences this behavior. The materials consisted of two 3-5 minute long video clips from two documentaries. While students were watching the videos, their eye movements were tracked. The authors analyzed the duration of fixation on on-screen texts. The findings revealed that these low proficiency L2 learners paid attention to captions for 68% of the time, which indicates that they often refer to captions when viewing videos. The researchers also found out that L2 learners of Arabic spend more time on reading captions then L2 learners of Spanish and Russian. L2 learners of Chinese, on the other hand, showed the greatest heterogenity in their use of captions. In addition to proficiency differences, this result is also explained by the fact that "distance between the L1 and L2 affects the way in which learners are able to use captions at a given time in the trajectory of learning" (Winke et al., 2013, p. 268).

To sum up, research achieving positive results with the use of captions demonstrates that because one of the preconditions of acquiring a language is being exposed to large amounts of authentic and comprehensible input (Krashen, 1985), captioned viewing aids learners to understand native speech. Captions allow learners to visualize the aural input and process natural English filled with colloquial expressions, natural pauses and reduced forms (Danan, 2004). Captioned viewing also increases attention and subsequently helps learners develop proficiency in segmenting acoustic messages, recognizing aural forms, and identifying structure-meaning connections (Markham, 1999; Yang & Chang, 2014). In a way, captions serve as an invaluable aid providing learners with the immediate confirmation of ongoing aural streams (Winke et al., 2010). As a result of comprehending what is heard, L2

listeners feel relieved and confident, which might otherwise cause demotivation for L2 listening. On the other hand, studies obtaining no significant difference by means of captions suggest that captioning cannot contribute to the development of listening comprehension due to such reasons as concentration on reading the text rather than listening to the audio, heavy reliance on the text, overloaded working memory, and learner perception of captions as a source of distraction (e.g. Diao et al., 2013; Hui, 2007; Montero-Perez et al., 2013; Zanon, 2006)

In view of all these observations, it can be concluded that although captions are often promoted as a means to aid L2 listeners, all learners do not make equal uses of captions. In other words, captioning benefits language learners with varying degrees (Leveridge & Yang, 2013). As a result, these findings suggest that the focus in this research area should go beyond a naïve search of *whether captioning is effective or not* to a systematic and detailed examination of what factors underlie these contradicting findings in order to move from inconsistency to consistency in this research field. To this end, the following section discusses variables that potentially affect the utility of captions for L2 listening.

Why is Research on Captioning in L2 Listening still Inconclusive?

Existing research on captioning in L2 listening is replete with inconsistent findings. Although most researchers promote captions as an aid to augment L2 listener comprehension (Danan, 2004; Teng, 2017; Yang & Chang, 2014) some have argued that captions do not provide appropriate support to develop L2 listening (e.g. Diao et. al., 2007; Taylor, 2005; Zanon, 2006). There are also others contending that the potency of captioning for improved listening can be compromised, if learners are taught how to take advantage of on-screen texts (e.g. Danan, 2004). One of the ways to increase the effectiveness of captioning as an instructional aid is contingent upon identifying the variables affecting captioning effects. Although, some individual factors (e.g. test difficulty, script differences) are discussed in separate studies, the available literature lacks a comprehensive list of factors influencing the utility of captions for L2 listening. In an effort to fill this gap, we surveyed literature (e.g. Bairstow & Lavaur, 2012; Behroozizad et al., 2015; Bianchi & Ciabattoni, 2008; Chai & Erla, 2008; Hayati & Mohmedi, 2011; Hwang et al., 2019; Latifi et al., 2011; Lee, 2021; Liversidge, 2000; Leveridge & Yang, 2014; Markham, 2001; Markham, 2003; Mayer, Lee & Peebles, 2014; Montero-Perez et al., 2013; Montero-Perez et al., 2014a; Montero-Perez et al. 2014b; Pujadas & Munoz, 2020; Pujola, 2002; Rodgers & Webb, 2011; Stewart & Pertusa, 2004; Taylor, 2005; Teng, 2019; Vanderplank, 2016; Winke et al. 2010; Winke et al. 2013) and distilled research results into a list of characteristics that influence captioning effects. We have generated five main categories with 13 sub-categories: listener-related factors (e.g. proficiency level, background knowledge, cognitive variables and affective variables), material-related factors (e.g. types of captioning, modes of captioning, genre and video difficulty), test-related factors (e.g. test types and test items), task-related factors (e.g. frequency of viewing and task purpose) and language-related factors (e.g. L1-L2 script

differences). We now turn to a discussion of these variables to explain why research on captioned viewing is still inconclusive.

Listener-related factors: Research shows that listener characteristics including proficiency level, cognitive differences, affective variables, and background knowledge have a decisive influence on the effectiveness of captioning in L2 comprehension.

Regarding proficiency level, Vanderplank (2016) states that "the relative effectiveness of captioned viewing varies according to language level" (p.4). In line with this, several researchers (e.g. Montero-Perez et al. 2013; Leveridge & Yang, 2014) suggest that less proficient listeners generally rely on on-screen texts more than higher level learners. While more proficient learners use captions as "a backup to their listening activity" (Pujola, 2002, p. 254), low proficiency levels view it essential for better comprehension. Nevertheless, others indicate that advanced-level learners perceive captions distracting (see Bairstow & Lavaur, 2012) and lower-proficiency learners find it difficult to process three types of stimuli simultaneously (Taylor, 2005).

Cognitive differences such as learning styles and caption reliance are also highlighted as significant variables affecting captioning effects. Hwang et al. (2019) note that active-style learners prefer learning through interaction and experiencing. Reflective-style learners, on the other hand, prefer learning through thinking and observation. Accordingly, while active learning style tends to benefit more from keyword captioning, reflective learning style prefers full-captioned viewing. In a similar line of research, Lee et al. (2021) state that learners' cognitive profiles affect how they utilize from captioned videos. Less-caption-reliant learners have best comprehension outcomes via partial-captioned videos. Whereas, more-caption-reliant learners need more textual support to decode the incoming aural input. These studies suggest that the simultaneous presentation of visual, auditory and textual stimuli increases active-style and less-caption-reliant learners' cognitive load and debilitates their decoding process.

Affective variables are among other factors impacting captioning effects on listening comprehension. Drawing on extant research, Behroozizad et al. (2015) suggest that if learners can confirm what is heard by means of captions, they can feel confident and efficient which may, in turn, lead them to make more strategic uses of captions to improve their comprehension. From a different standpoint, Vanderplank (2016) states that learners respond to different video genres in different ways and furthers that captioning may offer "cognitive counterweight to the affective pull of well-constructed program designed for entertainment and easy viewing" (p.239). Additionally, if students lack motivation or interest in the topic and do not attempt to interpret the listening text, then it is unlikely that captioning can make a difference.

Background knowledge can also influence L2 listeners' caption viewing. Previous knowledge of L2 learners affects their ability to understand the new information. According to Winke et al. (2013) learners with less content familiarity spend more time to read captions.

Similarly, Rodgers and Webb (2011) indicate that learners' background knowledge can compensate for their limited linguistic knowledge and decreases heavy reliance on captions.

Material-related factors: Types of captions, modes of captions, video genre and video difficulty can be included in the material-related factors.

Type of captions refers to the language of captions —L1 or L2— and can have differential effects on listening comprehension. While some studies (e.g. Latifi et al., 2011; Markham, 2001; Pujadas & Munoz, 2020) reveal that L1 captioning results in better comprehension scores, others (e.g. Hayati & Mohmedi, 2011; Stewart & Pertusa, 2004) suggest that L2 captions are more useful than native language captions since they encourage learners to practice L2 listening skills.

Modes of captions primarily point to full captioning and partial captioning. As full captioning allows learners to decode the ongoing streams of speech better, many researchers (Markham, 2013; Montero-Perez et al., 2014b; Teng, 2019) promote the use of full captions for improved comprehension. Several other researchers, on the other hand, support the use of partial captioning in L2 listening since it provides learners with key information and decreases decoding load.

Video genres indicate news, documentaries, series, sitcoms, movies, and so on. Numerous studies (e.g. Mayer, Lee & Peebles, 2014; Vanderplank, 2016) suggest that listeners find captions intrusive for programs which provide easy viewing or place a heavy focus on visuals.

As to video difficulty, it indicates novel content, unfamiliar lexis, speaker speed, reduced forms, and so on. According to Bianchi and Ciabattoni (2008), captions "cannot compensate for an excessively wide gap" (p. 70). Similarly, many researchers (e.g. Hsieh, 2020) contend that video difficulty can significantly predict the utility of captions.

Test-related factors: One of the variables concerned with the influence of testing on the effectiveness of captions is test type. Montero-Perez et al. (2013) state that captioning effect is found largest when comprehension is measured through receptive tests. Contending for an opposite finding, Liversidge (2000) posits that captioning effect on comprehension is found highest if learners are assessed by means of productive tests.

Non-significant or significant results as to captioning effectiveness can also be attributed to test difficulty (Montero Perez et al., 2014b; Pujadas & Munoz, 2020; Teng, 2019). The level of complexity or easiness of the test may lead to inaccurate evaluations with regard to the utility of captioned viewing.

Task-related factors: Frequency of viewing and task purpose have also been highlighted to influence the role of captions in learners' comprehension performance. Several researchers (e.g. Teng, 2019; Winke et al., 2010) suggest that tasks involving repeated captioned viewing lead to better scores in comprehension tests. As to the task purpose, listening activities requiring learners to divide their attention between comprehending the video content and

focusing on language areas (pronunciation, grammar and vocabulary) cause learners to rely on reading captions and neglect the comprehension task, which negatively influences the effectiveness of captions in L2 listening (Chai & Erlam, 2008; Pujadas & Munoz, 2020).

Language-related factors: Ortographic differences between one's native language and the target language impact how learners make use of captions. Research (e.g. Winke et al., 2010; Winke et al., 2013) has shown that when there is a great distance between L1 and L2 scripts, learners tend to process auditory input as initial intake rather than utilizing the on-screen-text.

LIMITATIONS AND RECOMMENDATIONS

These findings reveal that listening with captioned viewing is a rather complicated task influenced by a myriad of factors. The different variables at play demonstrate why research on captioning is still inconclusive although there is a commonsense conviction that captioned viewing leads to better performance in L2 listening. In view of all these inconsistent findings, we suggest that the mere presence of captions does not necessarily lead to improved comprehension. In order to fulfill the potency of captions as a promising tool in L2 listening, practitioners need to design their courses considering these factors and addressing related issues. Moreover, students should be taught how to use captions strategically to improve their L2 listening skills. They should be made aware of cognitive and metacognitive strategies for more active captioned viewing. Practitioners should support this process by incorporating some activities during and after viewing, and making equal uses of both captioned and uncaptioned videos.

Although this study is significant in that it attempts to provide insights into why research on captioning in L2 listening is inconclusive by distilling results of primary research into a list of characteristics that influence captioning effects, some limitations should be noted. First, this narrative literature review drew on a limited number of available studies mostly referring to frequently cited articles and works of leading researchers in the field. While surveying the literature, we might have missed potential studies. Second, despite pooling a variety of factors affecting the utility of captions for L2 listening, the list proposed in the study is inevitably constrained by the findings of studies discussed in this paper. Finally, adopting a meta analysis method or extensive systematic review could have enabled us to draw more rigorous conclusions about the effectiveness of captions in L2 listening and might have yielded a more exhaustive list of variables at work.

CONCLUSION

Based on a survey of primary research on the effectiveness of captioning in L2 listening, this paper aimed to find responses to the following questions: (1) Is captioning really effective in L2 listening?, (2) Does captioning always work for L2 listening?, and (3) Why is research on captioning in L2 listening still inconclusive? While some of the included studies evidenced that captioning can be an invaluable aid for improving L2 listening

comprehension, others have revealed that the sole availability of captions does not warrant success nor does it impede the comprehension process since there are other variables at play. Pooling individual variables discussed in different studies, this paper suggests that the efficiency of captioning in L2 listening can be increased by taking account of learner, material, measurement, task, and L1/L2 characteristics in unique L2 contexts. Of course, the list proposed in the study is not exhaustive and more research is needed to identify other potential variables influencing captioning effectiveness. Further research is also needed to see to what extent addressing these variables can promote the utility of captioning effects for L2 listening comprehension.

REFERENCES

Aldera, A. S., & Mohsen, M. A. (2013). Annotations in captioned animation: Effects on vocabulary learning and listening skills. *Computers & Education*, *68*, 60–75. https://doi:10.1016/j.compedu.2013.04.018

Bairstow, D., & Lavaur, J. (2012). Audiovisual information processing by monolinguals and bilinguals: Effects of intralingual and interlingual subtitles. *Approaches to Translation Studies*, *36*, 273–293. https://doi.org/10.1163/9789401207812_016

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology, 1*(3), 311–320. https://doi.org/10.1037/1089-2680.1.3.311.

Behroozizad, S., & Majidi, S. (2015). The effect of different modes of English captioning on EFL learners' general listening comprehension: Full text vs. keyword captions. *Advances in Language and Literary Studies*, *6*(4), 115–121. https://www.journals.aiac.org.au/index.php/alls/article/view/1626

Bianchi, F. & T. Ciabattoni (2008). Captions and subtitles in EFL learning: An investigative study in a comprehensive computer environment. In A. Baldry, M. Pavesi & C. Taylor Torsello (eds.), *From didactas to ecolingua* (pp. 69-80). Edizioni Universit`a di Trieste.

Bird, S. A. & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics 23*(4), 509–533. https://doi.org/10.1017/S0142716402004022

Brett, P. (1995). Multimedia for listening comprehension: The design of a multimedia-based resource for developing listening skills. *System, 23*(1), 77-85. https://doi.org/10.1016/0346-251X(94)00054-A

Caimi, A. (2006). Audiovisual translation and language learning: The promotion of intralingual subtitles. *The Journal of Specialised Translation 6*, 85–98. https://www.jostrans.org/issue06/art_caimi.php

Chai, J., Erlam, R., 2008. The effect and the influence of the use of video and captions on second language learning. *New Zealand Studies in Applied Linguistics, 14*(2), 25-44. https://search.informit.org/doi/10.3316/INFORMIT.997746379301329

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta 49*(1), 67–77. http://www.erudit.org/revue/meta/2004/v49/n1/009021ar.html.

De Jong, N. (2005) Can second language grammar be learned through listening? An experimental study. *Studies of Second Language Acquisition, 27*(2), 205–34 https://doi.org/10.1017/S0272263105050114

Diao, Y., Chandler, P., & Sweller, J. (2007). The effect of written text on comprehension of spoken English as a foreign language. *American Journal of Psychology, 120*(2), 237–261. https://www.jstor.org/stable/20445397

Gary, J., & Gary, N. (1981). Caution: Talking may be dangerous to your linguistic health. *International Review of Applied Linguistics in Language Teaching, 19*(1), 1–14. https://doi.org/10.1515/iral.1981.19.1-4.1

Garza, T. J. (1991) Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Language Annals, 24*(3), 239–258. https://doi.org/10.1111/j.1944-9720.1991.tb00469.x

Goh, C.C., 2000. A cognitive perspective on language learners' listening comprehension problems. *System 28*(1), 55-75. https://doi.org/10.1016/S0346-251X(99)00060-3

Guillory, H. G. (1998). The effects of key word captions to authentic French video in foreign language instruction. *CALICO Journal, 15*(1-3), 89-108. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.494.3324&rep=rep1&type=pdf

Hwang, G. J., Hsu, T. C., & Hsieh, Y. H. (2018). Impacts of different smartphone caption/subtitle mechanisms on English listening performance and perceptions of students with different learning styles. *International Journal of Human–Computer Interaction, 35*(4–5), 1–12. https://doi.org/10.1080/10447318.2018.1543091

Hayati, A. and Mohmedi, F. (2011), The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology, 42*(1), 181–192. https://doi.org/10.1111/j.1467-8535.2009.01004.x

Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning, 33*(5-6), 567–589. https://doi.org/10.1080/09588221.2019.1577898

Huang, H. C. & Eskey, D. E. (1999) The effects of closed-captioned television on the listening comprehension of intermediate English as a second language (ESL) students. *Journal of Educational Technology Systems, 28*(1), 75–96. https://doi.org/10. 2190/RG06-LYWB-216Y-R27G

Hui, W. (2007). The effects of captions on Chinese EFL students' incidental vocabulary acquisition. *Chinese English Language Education Association Journal, 30*(4), 9–16. http://en.cnki.com.cn/Article_en/CJFDTotal-WYDH200502011.htm

Jones, R. H. 1997. Beyond 'listen and repeat': pronunciation teaching materials and theories of second language acquisition. *System, 25*(1), 103–12. https://doi.org/10.1016/S0346-251X(96)00064-4

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* Laredo Publishing Company.

Latifi, M., Mobalegh, A., & Mohammadi, E. (2011). Movies subtitles and the improvement of listening comprehension: Does it help? *The Journal of Language Teaching and Learning, 1*(2), 18-29. https://dergipark.org.tr/en/download/article-file/209072

Lee, P.-J., Liu, Y.-T. & Tseng, W.-T. (2019) One size fits all? In search of the desirable caption display for second language learners with different caption reliance in listening comprehension. *Language Teaching Research*, *25*(3), 400-430. https://doi.org/10.1177/1362168819856451

LeLoup, J. W. & Ponterio, R. (2006). Listening: You've got to be carefully taught. *Language Learning & Technology, 1*(11), 4–15. http://dx.doi.org/10125/44082

Leveridge, A. N. & Yang, J. C. (2013) Testing learner reliance on caption supports in second language listening comprehension multimedia environments. *ReCALL, 25*(2), 199–214. https://doi.org/10.1017/S0958344013000074

Leveridge, A. N. & Yang, J. C. (2014) Learner perceptions of reliance on captions in EFL multimedia listening comprehension. *Computer Assisted Language Learning, 27*(6), 545–559. https://doi.org/10.1080/09588221.2013.776968

Liversidge, G. B. (1999). *The Role of closed captioning in second language acquisition*. Bell & Howell Company.

Markham, P. (1999). Captioned videotapes and second language listening word recognition. *Foreign Language Annals, 32*(3), 321-328. https://doi.org/10.1111/j.1944-9720.1999.tb01344.x

Markham, P. L., Peter, L. A. & McCarthy, T. J. (2001) The effects of native language vs. target language captions on foreign language students' DVD video comprehension. *Foreign Language Annals, 34*(5), 439–445. https://doi.org/10.1111/j.1944- 9720.2001.tb02083.x

Markham, P. & Peter, L. (2003) The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems, 31*(3), 331–341. https://doi.org/10.2190/ BHUH-420B-FE23-ALA0

Mayer, R. E. (2005). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In Mayer, R. E. (ed.), *Cambridge handbook of multimedia learning* (pp. 169-182). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.012

Mayer, R. E., Lee, H., & Peebles, A. (2014). Multimedia learning in a second language: A cognitive load perspective. *Applied Cognitive Psychology, 28*(5), 653–660. http://dx.doi.org/10.1002/acp.3050

Montero Perez, M., Noortgate, W. V. D. & Desmet, P. (2013) Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*(3), 720–739. https://doi.org/10.1016/j.system.2013.07.013

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014a). Effects of captioning on video comprehension and incidental vocabulary. *Language, Learning & Technology, 18*(1), 118–141. https://doi.org/10.1080/09588221.2017.1375960

Montero Perez, M., Peters, E., & Desmet, P. (2014b). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL, 26*(1), 21–43. https://doi.org/10.1017/S0958344013000256

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.

Pujola, J.-T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL, 14*(2), 235–262. https://doi.org/10.1017/S0958344002000423

Pujadas, G. & Munoz, C. (2019) Extensive viewing of captioned and subtitled TV series: A study of L2 vocabulary learning by ˜ adolescents. *The Language Learning Journal, 47*(4), 479–496. https://doi.org/10.1080/09571736.2019.1616806

Robin, R. (2007). Commentary: Learner-based listening and technological authenticity. *Language Learning & Technology,1*(11), 109–115. http://dx.doi.org/10125/44092

Paivio, A. (1986). *Mental representation: A dual-coding approach*. Oxford University Press.

Richards, J. (2005) Second thoughts on teaching listening. *RELC Journal, 36*(1), 85–92. https://doi.org/10.1177/0033688205053484

Richards, J. (2008). *Teaching speaking and listening*. Cambridge University Press.

Rodgers, M. R. H., & Webb, S. (2017). The effects of captions on EFL learners' comprehension of English-language television programs. *CALICO Journal, 34*(1), 20–38. https://doi.org/10.1558/cj.29522

Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Pearson Education Ltd.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158. https://doi.org/10.1093/applin/11.2.129

Seferoğlu, G. (2008). Using feature films in language classes. *Educational Studies, 34*(1), 1–9. https://doi.org/10.1080/03055690701785202

Staehr, L.S., (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition 31*(4), 577-607. https://doi.org/10.1017/S0272263109990039

Stewart, M. A., & Pertusa, I. (2004). Gains to language learners from viewing target language closed captioned films. *Foreign Language Annals, 37*(1), 438–447. https://doi.org/10.1111/j.1944-9720.2004.tb02701.x

Syodorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology, 2*(14), 50–73. http://dx.doi.org/10125/44214

Taylor, G. (2005) Perceived processing strategies of students watching captioned video. *Foreign Language Annals, 38*(3), 422–427. https://doi.org/10.1111/j.1944-9720.2005.tb02228.x

Teng, F. (2017). Flipping the classroom and tertiary level EFL students' academic performance and satisfaction. *The Journal of Asia TEFL, 14*(4), 605–620. http://dx.doi.org/10.18823/asiatefl.2017.14.4.2.605

Teng, F. (2019). Maximizing the potential of captions for primary school ESL students' comprehension of English-language videos. *Computer Assisted Language Learning, 32*(7), 1–27. https://doi.org/10.1080/09588221.2018.1532912

Vandergrift, L. (2003). Orchestrating strategy use: Toward a model of the skilled second language listener. *Language Learning, 53*(3), 463–496. https://doi.org/10.1111/1467-9922.00232

Vandergrift, L. (2007) Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(3), 191–210. https://doi.org/10.1017/S0261444807004338

Vanderplank, R. (1988). The value of teletext subtitles in language learning, *ELT Journal 42*(4), 272–281. https://doi.org/10.1093/elt/42.4.272

Vanderplank, R. (1990). Paying attention to the words: Practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) subtitles. *System 18*(2), 221–234. https://doi.org/10.1016/0346-251X(90)90056-B

Vanderplank, R. (2016). "Effects of" and "effects with" captions: How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching, 49*(2), 235–250. https://doi.org/10.1017/S0261444813000207

Williams, H. & D. Thorne (2000). The value of teletext subtitling as a medium for language learning. *System 28*(2), 217–228. https://doi.org/10.1016/S0346-251X(00)00008-7

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 65–86. http://dx.doi.org/10125/44203

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors influencing the use of captions by foreign language learners: An eye-tracking study. *The Modern Language Journal, 97*(1), 254–275. https://doi:10.1111/j.1540-4781.2013.01432.x

Wipf, J. (1984) Strategies for Teaching School Language Listening Comprehension. *Foreign Language Annals, 17*(4), 345–348. https://www.proquest.com/openview/a8137d6608a0f452daae72fd1a73afe2/1?cbl=1821600&pq-origsite=gscholar

Yang, J. C., & Chang, P. (2014). Captions and reduced forms instruction: The impact on EFL students' listening comprehension. *ReCALL, 26*(1), 44–61. https://doi.org/10.1017/S0958344013000219

Zanon, N. T. (2006). Using subtitles to enhance foreign language learning. *Porta Linguarum, 6*, 41-52.
https://www.researchgate.net/publication/28178514_Using_Subtitles_To_Enhance_Foreign_Language_Learning

Biographical notes:

Name and Surname of the Author 1: Saadet Korucu Kış is an Assistant Professor in the English Language Teaching Department of Necmettin Erbakan University. Her research interests relate to computer-assisted language learning, teacher education, instructional design, and reflective practice.