

Research paper

How Do Different Keyword Captioning Strategies Impact Students' Performance in Oral and Written Production Tasks? A Pilot Study

Long He
Iowa State University, USA

lhe@iastate.edu

Abstract

As an increasingly popular format of input, the affordances of audio-visual materials have been widely studied. Past research has provided evidence that audio-visual input combined with different captioning strategies could benefit learners in terms of vocabulary learning, listening comprehension, and the development of grammatical knowledge. However, there is a lack of research on how manipulating captioning conditions could help learners use their own linguistic resources to produce L2. Comparing the effects of three captioning techniques, L1 glossed keyword captioning, keyword captioning, and no captioning on English learners' oral and written recall of a short video, this pilot study aims to test the instruments and the data collection methods. The tentative results suggest that L1 glossed keyword captioning might have worked better in facilitating students' oral and written production of the keywords than keyword captioning and no captioning. The study also shows that L1 glossed keyword captioning might be more useful than keyword captioning and no captioning in helping students comprehend and reproduce the content of the video. Suggestions for further research on this topic are presented in the final part of this paper.

Keywords: Audio-visual input, keyword captioning, gloss, recall, oral and written production tasks.

1. Introduction

Though Krashen's (1985) argument that second language (L2) learners just need comprehensible input to activate their built-in syllabus and that L2 acquisition relies entirely on input proved to be controversial, researchers have widely accepted the essential role of exposure to L2 input in second language acquisition (SLA). L2 input is especially crucial for implicit learning. As Ellis (2015) puts it, "Implicit learning is a slow process that requires massive exposure to the second language" (p. 36). Previous studies have investigated the effect of different types of input (e.g., audio, written, and visual) on learners' L2 acquisition. One type of input, audio-visual input, has attracted sustainable interest from researchers in SLA.

A main strand of research on audio-visual input centers on the effect of using native language (L1) or L2 subtitles or captions to enhance language learning. Markham (1999) defines subtitles as "on-screen text in the native language combined with the second language soundtrack" and captions as "on-screen text in the second language combined with the second language soundtrack." In this study, L1 caption refers to native language captioning, and L2 caption refers to second language captioning.

Multiple studies have examined the effectiveness of L1 and L2 captions in facilitating learners' vocabulary acquisition and listening comprehension. Koolstra and Beentjes (1999) compared the effects of watching L1 captioned television programs and watching

English television programs without captions on Dutch children's incidental acquisition of L2 English vocabulary. The results showed that the children scored higher in the vocabulary test and word recognition test in the L1 caption condition. The fact that the audio-visual input combined with L1 captions enables learners to hear the English words, read the Dutch translation, and infer meanings from the visual images facilitates Dutch children's vocabulary acquisition. Rodgers and Webb (2017) conducted a similar study but used 10 episodes of a TV series. Their results revealed that L2 captions were especially useful in aiding comprehension when the content was difficult. Focusing on L2 Spanish, Markham, Peter, and McCarthy (2001) investigated how three different caption conditions, i.e. L1 English captions, L2 Spanish captions, and no captions, influenced learners' performance on a written summary task and a multiple-choice task. The results again showed that learners under the L1 captions condition performed considerably better than the other two pairs. Winke, Gass, and Sydorenko (2010) explored the use of captioned videos in listening activities and concluded that L2 captioned videos were more effective in facilitating novel vocabulary recognition and overall comprehension.

The previously mentioned studies have documented positive effects of captioning on L2 learning, but they did not include procedures to draw learners' attention to target words or phrases. In his Noticing Hypothesis, Schmidt (2001) claims that "people learn about the things they attend to and do not learn much from the things they do not attend to" (p. 30). In other words, it might be easier for learners to acquire more salient language features in the input. Hypothesizing that keyword captioning presents less information and thus could keep students' attention on the linguistic message, Guillory (1998) investigated how different types of captions, L2 full captioning, L2 keyword captioning, and no captioning, impact learners' comprehension. The results showed that both full captioning and keyword captioning had a positive effect on comprehension. Montero Pérez et al. (2014) studied the effects of two types of captioning, namely L2 full captioning and L2 keyword captioning. They found that the full captioning pair scored higher on the global comprehension questions than the no captioning and the keyword captioning pair. They also reported the participants' preference for full captioning. In a later study, Montero Pérez et al. (2018) compared the effects of three captioning techniques, full captioning, keyword captioning, and L1 glossed keyword captioning, on vocabulary learning. They found that the students in the L1 glossed keyword captioning pair performed the best in both the form recognition test and the meaning recall test. The findings suggested that the access to meaning through L1 glossed keyword captioning could help students to make form-meaning connections. In their eye-tracking study, Lee & Révész (2018) enhanced the captions by boldfacing the target grammatical structure and observed the advantage of textual enhancement in directing learners' attention to the grammatical feature. The results suggested that the enhanced captions could facilitate learners' development of grammatical knowledge.

Past research has provided evidence that different captioning strategies could benefit learners differently in terms of vocabulary learning, listening comprehension, and the development of grammatical knowledge. However, there is a lack of research on how manipulating captioning conditions could help learners use their own linguistic resources to produce L2. Swain's (1985) Output Hypothesis proposes that language acquisition requires not only comprehensible input but also output production. When discussing the importance of output, Nava & Pedrazzini (2018) provided further explanation:

While exposure to input that is made comprehensible for a learner is a needed starting point for SLA, it is in itself insufficient to satisfy all the demands of acquisition. Engaging in second language production, through both speaking and writing, is thus held to be crucial for acquisition, particularly if a learner wishes to increase their proficiency towards more native-like accuracy (p. 156).

Considering the benefits of audio-visual input combined with captions in assisting language learning and the importance of output in improving language proficiency, it is worthwhile to explore how different caption strategies, combined with audio-visual input, could aid students' oral and written production. The current study aims to investigate how utilizing different captioning options could influence learners' output task performance. In other words, this study compares the effects of L1 glossed keyword captioning, keyword captioning, and no captioning on English learners' oral and written recall of a short video.

The pilot study also aims to test the instruments and data collection methods. The purpose of using keyword captioning was to draw learners' attention to those words that would pose a challenge to learners' comprehension. The purpose of using L1 glossed keyword captioning was to help learners make form-meaning connections (Lee & Révész, 2018). The following two research questions guided this study.

Research Question 1: How do the three captioning conditions influence the students' use of the keywords in their oral and written production?

Research Question 2: How do the three captioning conditions impact the overall quality (based on correctly produced idea units) of the students' oral and written production?

2. Method

2.1. Participants

The participants included a female high school English teacher and six 11th grade high school students with an average age of 15.5. The teacher, a native Chinese, had been teaching English at the same school for about 12 years. The students were native speakers of Mandarin Chinese and were enrolled in the same English class. All the participants had received four years of classroom English instruction. The English teacher selected the six participants because they had similar scores from the achievement test taken at the beginning of the semester. Prior to the study, the students took the bilingual mandarin version Vocabulary Size Test developed by Nation and Beglar (2007). The results suggested that the students' vocabulary size was comparable, averaging 1,500 word families. Based on the students' performance on the achievement test and the vocabulary size test, the English proficiency level of the students was close to B2 level in the Common European Framework for Reference (CEFR). The teacher randomly assigned the students into pairs to complete the production task under three captioning conditions, L1 glossed keyword captioning, keyword captioning, and no captioning.

2.2. Video selection

The audio-visual input used in this study was a two-minute video on the cultural differences between China and the UK. To select a video that could spark the students' interest, the researcher provided the students four topics to choose from. The four topics included how to improve memory, global warming, the best way to practice English, and cultural differences between China and the UK. The students needed to select two topics of their interest. The last topic was selected for this study because all the students chose that one.

The video was recorded by a native speaker of British English, and it contained 394 word tokens. The Vocabulary Profiler, which was developed by the University of Hong Kong and based on Paul Nation's Word Frequency Lists, was used to determine the difficulty level of the vocabulary. After running the video transcription in the web-based software, it was found that about 88 percent of the words were from the first 2,000 word families. Therefore, it was anticipated that the video should be mostly comprehensible to the students. However, given the speech rate, 197 words per minute, the video should still be challenging to the participants.

2.3. The two types of captions

Two types of captioning strategies were used in this study; keyword captioning and L1 glossed keyword captioning. Figure 1 and 2 are screenshots of the two types of captions. Montero Pérez et al. (2018) defined *keyword* as one word or a string of no more than four words that are essential for the meaning making of a sentence. In this study, the researcher worked with the teacher to select 31 keyword types. iMovie was used to combine the audio-visual input and the captioning. In the keyword-captioned video, the keyword appeared at the lower right corner of the video. In the L1 glossed keyword captioning condition, the keyword and its L1 translation appeared at the lower right corner of the video. In both conditions, the keyword was synchronized with the speech, meaning each keyword appeared when spoken. The presentation duration of the keyword ranged from one to two seconds depending on its length.



Figure 1. Keyword captioning.



Figure 2. L1 glossed keyword captioning.

2.4. The task

The task in this study required the learners to watch a 2-minute video clip twice and then discuss with a partner to produce a written recall of the content of the video in English. Ellis (2018) reemphasized that "...learners must notice new features in the input and also notice the gap between what they attend to in the input and their current interlanguage systems in order to learn" (p. 202). This provided the rationale for watching the video twice. For the first watching, the students were expected to focus on the general meaning of the video clip and notice what might be new (the keywords) to them. During the second watching, the students had the opportunity to pay more attention to the gap between the new information and their own interlanguage systems so that they could deepen their understanding of the video.

In the discussion phase, the students needed to mobilize their own linguistic resources to communicate with each other regarding what information they each had gleaned from the video. During this phase, the student could interact with the partner to negotiate meaning. In his Interaction Hypothesis, Long (1983) claims that meaning negotiation facilitates L2 acquisition because learners obtain comprehensible input when they negotiate meaning. Meaning negotiation also allows learners more time to process the input (Ellis, 2018). While producing the written recall, the students needed to co-construct meaning and achieve a communicative outcome. Since the teacher needed to record the discussion, the three pairs of students completed the task separately in the teacher's office. The total time for the task was 25 minutes.

2.5. Procedure

Two days before the teacher invited the students to her office to do the task, the researcher sent the teacher the following table and discussed the questions she had about the steps via a Zoom meeting. After the Zoom meeting, she completely understood how to direct the students to complete the task.

Table 1. Step by step instructions for the teacher.

Step 1	Step 2	Step 3	Step 4	Step 5
Tell the students that they need to watch a video, discuss in pairs, and reconstruct the content of the video on paper as a pair (1 min.)	Play the video for the students for the first time (no notes; 2 mins.)	Play the video for the students for the second time; ask students to take notes. (2 mins.)	Ask students to work together to reproduce the content on paper. Encourage them to use their own linguistic resources and provide as much detail as possible. (record the discussion; 15 mins.)	Collect the notes and written work from the students

In the first step, the teacher briefly introduced the task and informed the students that they would need to discuss the content and produce a written recall. The rationale behind

informing the students about the oral and written production task beforehand was that they could be more focused on the audio-visual input. The teacher invited the first pair of students to the office where they watched the video under the L1 glossed keyword captioning (L1GKC) condition. The teacher asked the students to pay attention to the global meaning of the video during the first watching, and instructed them to take notes during the second watching. After spending five minutes watching the video, the teacher asked the students to spend another five minutes to discuss what had been going on in the video. At the same time, the teacher encouraged the students to use their own linguistic resources and started to record the discussion. Lastly, the students spent ten minutes to complete a written recall together. After the first pair of students completed the task, the teacher invited the keyword captioning (KC) pair to her office to do the task and then the no captioning (NC) pair. The teacher followed the same steps for all three pairs of students.

2.6. Data collection and data analysis

There were three sets of data in this study, namely the notes after the second watching, the recording of the discussion, and the written recall. The teacher recorded the discussion using her phone and collected the notes and written recall after the students completed the task. Then she put the data from each pair into a separate zip file and sent me the data. After receiving the data, the researcher transcribed the recordings.

To answer the first research question, the researcher read through the transcription and the written recall and counted the places where the students correctly used a keyword or paraphrased a keyword. The notes were to check the students' uptake of keywords they noticed and help interpret the data. To evaluate the overall quality of the students' oral and written production, Riley and Lee's (1988) idea unit analysis method was adopted. According to Riley and Lee, an idea unit refers either to a simple sentence, a basic semantic proposition, or a phrase. Based on Riley and Lee's criteria, the researcher divided the transcription into 35 idea units. Then the same criteria were used to count the correct idea units in the students' oral and written production. If the students paraphrased the idea units, those idea units were also counted as correct. If the idea units produced were correct but not mentioned in the video, those idea units were not counted.

3. Results and discussions

This section presents the results of this pilot study. After analyzing the notes, transcription of the students' discussion, and the written recall, it was found that the L1GKC pair was able to produce and paraphrase more keywords than the other two pairs in both oral and written production. Though the KC pair noticed more keywords than the NC pair, the two pairs' keyword use in the discussion and written recall was similar. The overall quality of the oral and written production follows the same trend with the L1GKC pair producing more correct and accurate idea units than the other two pairs.

3.1. Use of keywords in oral and written production

The first research question concerns how different captioning strategies impact students' use and paraphrasing of keywords in the discussion and written recall. Table 2 presents students' notes after the second watching. The researcher transferred the notes directly to the table without correcting the misspelling or translating the words written in Chinese. Table 3 is a summary of keywords used or paraphrased in oral and written production by the three pairs, and keywords in the video. The notes were used to help interpret the data in Table 3.

Table 2 shows that the L1GKC pair wrote down 19 of the 31 keywords appeared in the video. The KC pair registered 15 keywords, while the NC had only 6 keywords. This indicates that keyword captioning, with or without L1 gloss, might have facilitated students' noticing of the keywords. It is also worth mentioning that both the L1GKC and KC pair noted down only 5 words that are not keywords in the video, but the NC pair wrote down 7. To put it into perspective, non-keywords account for 20 percent and 25 percent of the notes by the L1GKC and KC pair respectively, while they constitute 54 percent of the notes by the NC pair. This suggests that keyword captioning could effectively draw students' attention to the target feature. Another interesting finding is that one of the students in the L1GKC pair wrote down some of the keywords in L1 instead of L2. This signals that the student was paying attention to the meaning of the keywords.

Table 2. Summary of students' notes.

Pair 1 (L1GKC)	SA: China, build dense, food, massive, quaint, bowls, manners, chopstick, complete, food waste, ju..., spit, queuing, finish, host, adible, instinct	n=19
	SB: Billion, China, massive 巨大, food, different, doesn't sit will with, inrect with, 小册子 (pamphlet), queuing, 懊恼 (frustrate), 发脾气 (lash out)	
Pair 2 (KC)	SA: Check out, quaint, complete, lead to, host, edible, sit well with me, improve, manner, government, spit, don't mind	n=15
	SB: Billion, check out, UK, China, way to eating, chopstick, food waste, host, manners, government, spitting, queuing	
Pair 3 (NC)	SA: 80,000, food, chopstick, finish, hostess, manners, don't mind, queuing	n=6
	SB: People, village, 80,000, food, chopstick, round, table, manners, queuing, skeap	

Table 3. Keywords used or paraphrased in oral and written production by three pairs, and keywords in the video.

Pair	Keywords used/paraphrased in oral production	Keywords used/paraphrased in written production	Keywords in the video
L1GKC	SA: manners, many people, big bowls, wasted food, communicate with others, government, spread the thin book	Crowded, big bowl, host, don't know how much food the people will have, wasted, manners, government, spread the thin book, spitting, queuing, angry	dense, flats, billion, check out, quaint, massive, communal bowl, interact with, complete, lead to, food waste, judge, finish, host, edible, doesn't sit well with, improve, manners, common, government, release a pamphlet, inform, spit, throw litter, don't mind, frustrate, queuing, skip to the front, control my british instinct, lash out, queue jumper
	SB: queuing(wrongly pronounced), can't stand		
KC	SA: check out, don't mind, improve	Manners, food waste, spitting, government	
	SB: manners		
NC	SA: don't mind, manners, finish, host(er)	host(er), manners, skeap the queuing, throw rubbish	
	SB: manners, queuing, skip the queuing		

According to Table 3, the L1GKC pair used or paraphrased 9 keywords in their oral production and 11 keywords in their written production. In contrast, the KC pair used only 4 keywords in both the oral and written production. For the NC pair, 6 keywords were used in oral production and 4 in written production. Even though the KC pair noticed more keywords based on their notes, the students under that condition either were not able to or at least did not use or paraphrase most of the keywords in their production. The tentative results of this pilot study show that L1 glossed keyword captioning might have worked better in facilitating students' oral and written production of the keywords than keyword captioning and no captioning. A more detailed analysis of the transcription and

written recall revealed that the access to meaning provided by L1 gloss enabled the students to paraphrase some of the keywords. For example, the L1GKC pair paraphrased “dense” as “crowded” in their written production and used “spread the thin book” in the place of “release a pamphlet”, for which one student used Chinese in the notes, in both oral and written production. The pair also used “angry” for “lash out”. In comparison, in the KC and NC pair, no students paraphrased any of the keywords.

3.2. Overall quality of oral and written production

The second research question examines whether the overall quality of the oral and written production by the three pairs differs. The overall quality of the discussion and written recall was assessed based on how many correct idea units (35 in total) the students produced. Figure 3 shows that the L1GKC pair produced about twice as many idea units as the other two pairs. The KC pair and the NC pair, however, did not differ in terms of idea units in both oral and written production.

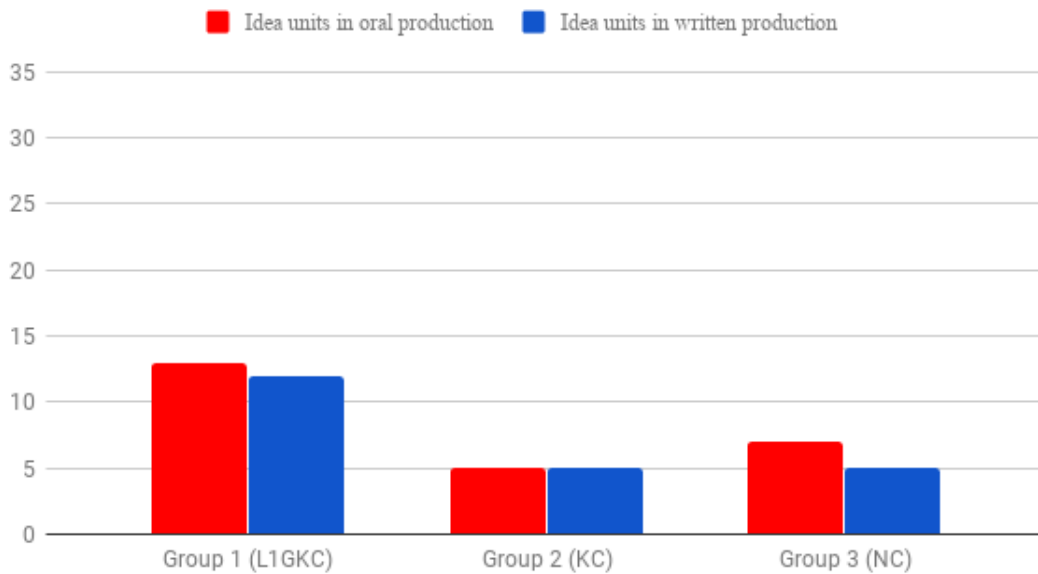


Figure 3. Idea units in oral and written production.

A closer look at the oral and written production data revealed that the L1GKC pair had a better comprehension of the video compared with the other two pairs. In the oral discussion, the students in the L1GKC pair had the following conversation:

B: *He says it can be more...talk with...*

A: *During the eating, they will communicate with others, right?*

Even though student B misspelt “interact with” as “incret with” under L1GKC, the student understood the meaning and used “talk with” in the discussion. That seemed to help student A to produce “During the eating, they will communicate with others...” which corresponded to “Everyone has to interact with each other in order to complete the meal...” in the video. However, the students in the KC pair did not mention this at all in their discussion. The NC pair produced the following utterance “In the UK, the people more outgoing than Chinese. When they meeting, they could say hello each other” which might have resulted from either pure guessing or misunderstanding of the content.

The L1GKC pair not only had more idea units but also had more accurate production in the written recall. In correspondence to “I really like this way of eating...but on the flip side it does lead to more food waste because it's much harder to judge just how much food you should actually cook,” the L1GKC pair wrote “I like the way to eat, but it will waste food. Because the host don't know how much food the people will have.” Though the students did not use the word “judge” and only wrote down “ju” in their notes, they were still able to reproduce the meaning. In comparison, the KC pair put down “And

Chinese can't allow food waste," and the NC pair wrote "when you go to others' house, the hoster would make you eat the food."

The results suggest that L1 glossed keyword captioning might be more useful than keyword captioning and no captioning in helping students comprehend and reproduce the content of the video. Though having successfully drawn student's attention to the keywords, keyword captioning did not increase students' understanding of the video. The only difference between the L1GKC pair and the KC pair was that students in the first pair had access to the meaning of the keywords through L1 gloss. This might have provided the much-needed information for the learners in the L1GKC pair to decode the speech and construct meaning, leading to a better grasp of the global meaning of the content.

4. Limitations and future research

Considering the purpose of the study was to test the instruments and data collection methods and there were only one pair of students in each captioning condition, the power of any statistical test will be very limited, so no statistical analyses were conducted in this pilot study. As a result, the findings of this pilot study should be interpreted with caution. The future study (In progress) will involve more participants and add the statistical tests to compare the data. Another limitation of the pilot study is that some students might have prior knowledge about the topic chosen, making it possible that these students might have performed better because of their familiarity with the topic rather than the different viewing condition. In the future study, a survey on students' prior knowledge of the video topic will be carried out to eliminate this effect. Another factor to consider is the difficulty level of the input itself. Even though the L1GKC pair did the best among the three pairs, the learners in that pair only produced a little over one third of the total idea units in the input. The L1GKC pair did capture the main ideas of the video, but their oral and written production lacked details. The fast speech rate (around 198 words per minute) of the video might have caused some trouble for the students. When selecting the video for the future study, both vocabulary and speech rate will be considered.

The current study did not solicit the students' and the teacher's opinions about the task. The learners' and teacher's feedback could provide insights into how they interact with the task and how the task should be modified to suit their needs. For example, after analyzing the survey questions, Montero Pérez et al. (2014) found that learners perceived the keyword as too distracting because they focused too much on the keywords and missed what was being said. Given the scope of the study, the researcher only investigated three captioning conditions. It will be beneficial to explore how other types of captioning, e.g., full captioning and L1 glossed full captioning, influence students' understanding of the content and their performance in the oral and written recall task.

Another research direction could be to rearrange the timing for the second watching of the video. This study adopted an input-input-output sequence, meaning the students watched the video the second time immediately after the first watching and then completed the production task. However, Nguyen and Boers (2018) argue that using an input-output-input sequence, where the learners work on the production task immediately after watching the video and then watch the video the second time, could help students notice the gaps between their production and the input content. As a result, they could focus on the information they need during the second watching. Thus, it is worthwhile to test whether using the input-output-input sequence could generate results that are different from using the input-input-output sequence.

5. Conclusion

In this study, learners under the L1 glossed keyword captioning condition better used and paraphrased the keywords in their discussion and written recall than learners under the other two captioning conditions. Learners under L1 glossed keyword captioning condition also produced more correct and accurate idea units than learners under the other two conditions. The results of this study indicate that L1 glossed keyword captioning has the potential to better promote learners' performance in the oral and written production task after watching a video clip. One implication of the study is that by integrating L1 glossed keyword captioning into the audio-visual input, the teacher might be able to facilitate students' understanding of the keywords and comprehension of the video content and promote learners' oral and written production. Considering the growing popularity of

audio-visual materials in L2 teaching and learning, further research concerning how to effectively integrate audio-visual input into L2 classrooms is needed. To achieve more accurate and generalizable results, the future study will recruit more pairs of participants, select a video whose topic is not familiar to the participants, and consider the vocabulary level and speech rate of the video.

References

- Anthony, L. (2014). AntWordProfiler (Version 1.4.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/software>.
- Ellis, R. (2015). *Understanding Second Language Acquisition* (2nd Ed.). Oxford: Oxford University Press.
- Ellis, R. (2018). *Reflections on task-based language teaching*. Bristol: Multilingual Matters.
- Guillory, H. G. (1998). The Effects of Keyword Captions to Authentic French Video on Learner Comprehension. *CALICO Journal*, 15(1-3), 89-108.
- Koolstra, C. M. & Beentjes, J. W. J. (1999) Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology, Research and Development*, 47(1): 51-60.
- Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implications*. New York: Longman.
- Lee, M. & Révész, A. J. (2018). Promoting Grammatical Development through Textually Enhanced Captions: An Eye-Tracking Study. *Modern Language Journal*. <https://doi.org/10.1111/modl.12503>.
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126-141.
- Markham, P. L. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, 32(3), 321-328.
- 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.
- Montero Pérez, M., Peters, E. & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, 31(1-2), 1-26.
- Montero Pérez, M., Peters, E. & Desmet, P. (2014). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL*, 26(1): 21-43.
- Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nava, A., & Pedrazzini, L. (2018). *Second language acquisition in action: Principles from practice*. London; New York, NY: Bloomsbury Academic.
- Nguyen, C. D. & Boers, F. (2018). The Effect of Content Retelling on Vocabulary Uptake from a TED Talk. *TESOL Quarterly*, 52 (1), 1-25. doi: 10.1002/tesq.441.
- Riley, G. L., & Lee, J. E. (1996). A comparison of recall and summary protocols as measures of second language comprehension. *Language Testing*, 13(2), 173-98.
- Rodgers, M.P.H. & Webb, S. (2017). The Effects of Captions on EFL Learners' Comprehension of English-Language Television Programs. *CALICO Journal*, 34(1), 20-38.
- Schmidt, R. (2001). Attention. In Robinson, P. (ed.): *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, pp. 3-32.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Massachusetts: Newbury House.

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.
