



PROGRESSIVE REDUCTION OF CAPTIONS IN LANGUAGE LEARNING

Mónica S. Cárdenas-Claros*	Pontificia Universidad Católica de Valparaíso. Viña del Mar, Chile	monica.cardenas@pucv.cl
Daniela Ramírez-Orellana	Colegio Padre Hurtado y Juanita de los Andes. Santiago, Chile	ramirezdaniela.ro@gmail.com

* Corresponding author

ABSTRACT

Aim/Purpose	This exploratory qualitative case study examines the perceptions of high-school learners of English regarding a pedagogical intervention involving progressive reduction of captions (full, sentence-level, keyword captions, and no-captions) in enhancing language learning.
Background	Recognizing the limitations of caption usage in fostering independent listening comprehension in non-captioned environments, this research builds upon and extends the foundational work of Vanderplank (2016), who highlighted the necessity of a comprehensive blend of tasks, strategies, focused viewing, and the need to actively engage language learners in watching captioned materials.
Methodology	Using a qualitative research design, the participants were exposed to authentic video texts in a five-week listening course. Participants completed an entry survey, and upon interaction with each captioning type, they wrote individual reflections and participated in focus group sessions. This methodological approach allowed for an in-depth exploration of learners' experiences across different captioning scenarios, providing a nuanced understanding of the pedagogical intervention's impact on their perceived language development process.
Contribution	By bridging the research-practice gap, our study offers valuable insights into designing pedagogical interventions that reduce caption dependence, thereby preparing language learners for success in real-world, caption-free listening scenarios.

Accepting Editor Peter Blakey | Received: December 19, 2023 | Revised: February 5, February 29, 2024 | Accepted: March 1, 2024.

Cite as: Cárdenas-Claros, M. S., & Ramírez-Orellana, D. (2024). Progressive reduction of captions in language learning. *Journal of Information Technology Education: Innovations in Practice*, 23, Article 2.

<https://doi.org/10.28945/5263>

(CC BY-NC 4.0) This article is licensed to you under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). When you copy and redistribute this paper in full or in part, you need to provide proper attribution to it to ensure that others can later locate this work (and to ensure that others do not accuse you of plagiarism). You may (and we encourage you to) adapt, remix, transform, and build upon the material for any non-commercial purposes. This license does not permit you to use this material for commercial purposes.

Findings	Our findings show that learners not only appreciate the varied captioning approaches for their role in supporting text comprehension, vocabulary acquisition, pronunciation, and on-task focus but also for facilitating the integration of new linguistic knowledge with existing background knowledge. Crucially, our study uncovers a positive reception towards the gradual shift from fully captioned to uncaptioned materials, highlighting a stepwise reduction of caption dependence as instrumental in boosting learners' confidence and sense of achievement in mastering L2 listening skills.
Recommendations for Practitioners	The implications of our findings are threefold: addressing input selection, task design orchestration, and reflective practices. We advocate for a deliberate selection of input that resonates with learners' interests and contextual realities alongside task designs that progressively reduce caption reliance and encourage active learner engagement and collaborative learning opportunities. Furthermore, our study underscores the importance of reflective practices in enabling learners to articulate their learning preferences and strategies, thereby fostering a more personalized and effective language learning experience.
Recommendations for Researchers	Listening comprehension is a complex process that can be clearly influenced by the input, the task, and/or the learner characteristics. Comparative studies may struggle to control and account for all these variables, making it challenging to attribute observed differences solely to caption reduction.
Impact on Society	This research responds to the call for innovative teaching practices in language education. It sets the stage for future inquiries into the nuanced dynamics of caption usage in language learning, advocating for a more learner-centered and adaptive approach.
Future Research	Longitudinal quantitative studies that measure comprehension as captions support is gradually reduced (full, partial, and keyword) are strongly needed. Other studies could examine a range of individual differences (working memory capacity, age, levels of engagement, and language background) when reducing caption support. Future research could also examine captions with students with learning difficulties and/or disabilities.
Keywords	caption reliance, pedagogical innovation, listening comprehension, language learning, Chile

INTRODUCTION

Integrating multimodal input in language education has created new opportunities for enhancing language learning. Among these resources, video captioned materials have garnered significant attention for their potential to provide visual text in sync with spoken words. While research broadly supports the positive impact of captions on comprehension, indicating that their use is generally more beneficial than not using them, there remains a concern that an over-reliance on captions may limit learners' ability to adapt to environments without them.

In our work, we extrapolate the concept of scaffolding, where support is gradually withdrawn as learners become more proficient, promoting independence and confidence in the target language to make a case for progressive caption reduction. Caption reduction is understood as the systematic decrease in the amount or complexity of text displayed as captions on video content. We contend that this approach helps transition learners from relying heavily on written text to enhance comprehension towards greater reliance on listening skills and directly understanding spoken language.

This study examines the perceptions of a group of high-school EFL learners regarding a pedagogical intervention involving progressive reduction of captions (full, partial, or sentence-level, keyword captions, and no-captions) in listening comprehension tasks. It is motivated by both theoretical and contextual reasons. Theoretically, firstly, caption research has grown exponentially as Günter Burger's frequently updated database (<https://www.fremdsprache-und-spielfilm.de/Captions.htm>) of research on captions shows, as well as recent reviews (Montero Perez, 2022; Rodgers & Webb, 2017; Vanderplank, 2019a). Recent research has examined caption benefits from the perspective of comprehension (Avello & Muñoz, 2023; Pujadas & Muñoz, 2020; Rodgers & Webb, 2017), vocabulary learning (Montero Perez, 2022), grammar (M. Lee & Révész, 2020; Pattemore & Muñoz, 2022), and pronunciation (Mohsen & Mahdi, 2021; Wisniewska & Mora, 2020). The focus has been primarily on quantitative and comparative designs and in studies conducted in laboratory-like conditions, contributing to the research-practice gap in language education (Cárdenas-Claros et al., 2023; Sato & Loewen, 2019). Moreover, explanations of why and how students use different types of captions remain top-down and do not often represent the L2 learners' views.

Secondly, while existing studies generally support the benefits of captions and agree that the use of captions, irrespective of caption type, is better for comprehension than the non-use of captions (Guillory, 1998; P.-J. Lee et al., 2021; Montero Perez, 2022; Montero Perez, Peters, Clarebout, & Desmet, 2014; Rodgers & Webb, 2017), concerns persist about potential over-reliance hindering students' adaptability to caption-free environments (Leveridge & Yang, 2013, 2014; Rooney, 2014; Vanderplank, 2016; Yeldham, 2018). Yeldham (2018) summarizes this view:

[T]he danger that [captions] could lead viewers to overly rely on [them]. They could develop entrenched viewing strategies dependent on including the use of captions to the extent that they are largely unable to transfer these abilities to settings where captions are unavailable (p.15).

Thirdly, research on students' perceptions regarding captions is scarce (Montero Perez, 2022; Stewart & Pertusa, 2004; Vanderplank, 2019b; Wang, 2012). So, we know that captions do aid comprehension problems, but we do not know much about students' acceptance of captions for language learning purposes. Recent work has investigated students' perception of caption reliance (P.-J. Lee et al., 2021; Leveridge & Yang, 2013, 2014) and found that students rely on captions. Notably, no studies have specifically examined informed pedagogical interventions aimed at reducing students' reliance on captions – a gap this study addresses.

Contextually, the study is situated in the challenging landscape of EFL listening comprehension in Chile (Cárdenas-Claros & Dassonville, 2022). The pressing need to address the low proficiency levels of high school students, as evidenced by national standardized examinations (TOEIC, KET, and APTIS for Teens), provides a meaningful backdrop for investigating the perceptions of gradual reduction in caption reliance as an alternative approach to helping L2 listeners.

We contend that a way forward to enable L2 listeners to successfully operate in caption-free environments is gradually reducing caption reliance to the point of absence. In our pedagogical innovation, we followed closely Vanderplank's (2016) suggestion that for language learners to benefit from interaction with captions, at least twofold measures should be set in place: (1) teachers should propose pedagogical ways to engage students through the full orchestration of tasks, strategies, and focused viewing, and (2) learners should be actively involved and not simply "sitting back" and watching captioned materials.

LITERATURE REVIEW

CAPTION TYPES

Caption use assumes that language learners may read and understand aural input at a reasonable speed, so various types of captions may influence listeners differently. For learners with strong reading skills, full captions or verbatim written transcriptions in video synchronized with the audio may

ease language learning; indeed, simultaneous verbal and visual representations of the same input can compensate for learners' lack of ability to map written forms into spoken ones (Montero Perez, Peters, Clarebout, & Desmet, 2014). Partial or sentence-level captions require learners to actively focus their attention to connect the text they see with the spoken words they hear shortly after, a process that demands cognitive engagement to align visual and auditory information. However, because partial captions are limited to key sentences rather than the full text, they may help learners concentrate more on the listening aspect of the material as they reduce the overall amount of reading required. Keyword captions focus on content words that convey the most meaning from the text and have been shown to likely aid learner concentration on aural messages over visual text (Abobaker, 2017; Vanderplank, 2016), although some learners have also found them distracting (Montero Perez, Peters, Clarebout, & Desmet, 2014).

RESEARCH ON CAPTIONS

Research on the use of captions for language learning has predominantly been comparative in nature, with few studies exploring learners' perceptions of caption use. We constructed this state-of-the-art review following the work of Montero Perez (2022). We identified two areas in caption research: (1) the effects of different caption types on language learning and (2) studies investigating students' perceptions of caption use.

Effects of caption types on language learning

Research efforts have focused on identifying the types of captions that promote listening comprehension. Guillory (1998) studied 202 adult French learners under three caption conditions: full-caption, keyword, and none. Results indicated both full-caption and keyword captions improved comprehension over no captions, with negligible differences between the full and keyword types. Guillory suggested keyword captions might similarly aid learning by focusing attention on auditory comprehension. In a related study, Abobaker (2017) employed a quasi-experimental design to assess the impact of keyword captions, full captions, and transcripts on 80 English learners at beginner and advanced levels. Findings showed beginners benefited more from full captions, whereas advanced learners did better with keyword captions. Teng (2019) examined 182 students watching English short story videos under three conditions: fully-captioned, keyword-captioned, and non-captioned. The study revealed that full captions led to superior global comprehension across all learners, and higher English proficiency among learners improved more significantly when viewing the video twice.

Other work focused on caption types and vocabulary learning gains. Montero Perez, Peters, Clarebout, and Desmet (2014) compared the impact of full captions, keyword captions, full captions with highlighted keywords, and non-captioned materials on vocabulary learning with 133 high-intermediate learners of French. Results showed that the keyword and full captions with highlighted keywords groups performed as well as the full captions group on vocabulary recognition tests and outperformed the control group. The keyword groups also outperformed the control group on the multiple-choice meaning recognition test (i.e., choosing correct translations of target words). According to Montero Perez, Peters, Clarebout, and Desmet (2014), the findings suggest that keyword captions and/or full captions with highlighted keywords might not only encourage learner attention but also beneficially affect the initial acquisition of word meanings.

The effects of different caption types on grammar and pronunciation have also been examined. Yang and Chang (2014) examined three types of captions and their contribution to learning reduced forms. Forty-four EFL learners from Taiwan worked in three groups: full caption, keyword-only, and annotated keyword. The keyword caption adds symbols to the captions to make reduced forms more salient. Each group was exposed to 51 short video clips, which only differed in the provision of caption types, and used a pre-test and post-test design. Through the comparison of results, Yang and Chang (2014) reported that the participants improved their performance from the pre-test to the post-test

and that the annotated keyword caption group showed the best ability to recognize reduced forms across groups.

When investigating the effects of caption use on pronunciation, Mahdi (2017) analyzed the effects of full captions and keyword captions on pronunciation with 34 Arabic speakers learning English and found that both full and keyword captions improved learners' pronunciation, with no significant differences between the two types of captions. Although one of the only studies reporting pronunciation gains, Mahdi's report lacks information on the proficiency of the learners, the description of pre- and post-tests, and the criteria to determine whether a word was correctly pronounced or not. Mohsen and Mahdi's (2021) study with 55 Arab learners of English viewed videos with full, partial, or no captions. In addition to watching the videos, their opinions on learning with captioning were collected through a survey. The outcomes of the pre-tests and post-tests revealed that the groups receiving captioning (either full or partial) outperformed the group without captioning in pronunciation tests.

The above studies seem to suggest that any amount of captions is better than non-caption materials for comprehension (Abobaker, 2017; Guillory, 1998; Montero Perez, Peters, Clarebout, & Desmet, 2014) and vocabulary learning (Montero Perez, Peters, Clarebout, & Desmet, 2014), that the type of captions preferred and how these are used depend on listener proficiency (Abobaker, 2017; Montero Perez et al., 2013; Sydorenko, 2010; Teng, 2019; Vanderplank, 2016), and that pronunciation improves, irrespective of the type of captions used (Mahdi, 2017; Mohsen & Mahdi, 2021).

Research on students' perceptions of captions

Research on perceptions about captions has been predominantly done in two areas: students' attitudes toward the use of captions and students' perceptions of caption reliance. Research on students' perceptions about captions is generally positive for various reasons. Stewart and Pertusa (2004) reported that 75% of their participants agreed or strongly agreed that they learned new vocabulary thanks to captions and noted a preference for L2 captions over L1 subtitles. Winke et al. (2010) found that L2 learners use captions like a 'crutch,' "just to double-check to make sure that what you are hearing is correct" (p. 79). Vanderplank (2019b) reported participants' positive attitudes as a result of working with captions. The participants in his study noted that captions were "a reassuring and often helpful presence" and "useful for the parts in which there were new words" that students did not know (p. 112). More recently, Mirzaei et al. (2017) examined keyword captions with learners of English from Japan. They found that partial and synchronized captions were preferred, theoretically and pedagogically, because they encouraged students to focus more on the aural input than on the written one. Montero Perez, Peters, and Desmet (2014) compared the effects and perceptions of two caption types: full captions and keyword captions. Montero Perez, Peters, and Desmet (2014, p. 35) observed that full captions aided in vocabulary recognition, speech segmentation, and comprehension, especially with fast speech, accents, and unclear pronunciation, while keyword captions were seen as potentially distracting and confusing. Mohsen and Mahdi (2021) reported that approximately 80% of their participants noted a perceived improvement in their pronunciation after watching the captioned videos. Video captioning was reported to facilitate independent learning of word pronunciation, assist in understanding correct pronunciation, and aid in correcting pronunciation errors,

Students' reliance on captions

Researchers have also investigated students' perception of caption reliance. Leveridge and Yang (2014) examined 139 Taiwanese learners of English as they worked with full captions, completed the Caption Reliance Test (CRT), and were interviewed. They reported that the overall perceived reliance on captions "was significantly related to actual reliance as assessed by the CRT" (p. 545), which suggests that, even if listeners may face some challenges in reflection tasks, the qualitative reports from learners themselves can still be reliable. In a more recent study, P.-J. Lee et al. (2021) explored learners' self-reported caption reliance when comparing the effects of full, partial, and no captions on L2 learners' comprehension. They reported that learners in a high captioning reliance group performed

best on a comprehension test after watching an 11-minute TED talk with full captions (as compared with partial and no captions). However, learners in the low captioning reliance group scored best under a partial captioning condition.

As noted previously, the above studies have primarily followed comparative designs to favor one caption type over the other. Also, while the studies mention learners' positive attitudes towards captions and their reliance on them for vocabulary acquisition and comprehension, there is room for more in-depth qualitative research to understand learners' experiences, challenges, and preferences in using captions beyond quantitative measures of reliance and comprehension. Moreover, no studies to our knowledge have examined informed pedagogical interventions aimed at progressively reducing students' reliance on captions. Our research questions are:

1. What are the perceptions of high-school learners of English regarding a pedagogical innovation that includes three types of captions (full, partial, and keyword) in video-based materials?
2. What are the perceptions of high-school EFL learners regarding the absence of captions after progressive caption reduction?

METHOD

We used an exploratory qualitative case study design to investigate perceptions of gradually reducing captions to absence with 11 high-school learners of English. Following research guidelines from Miles et al. (2014), a series of measures were taken to enhance study dependability, trustworthiness, and quality. These measures also informed the selection of the participants, materials design, and the data collection and analysis procedures. This study adheres to Creswell's (2013) definition of 'case' within case study research as a 'contemporary bounded system' (p. 97). In this work, such a system is understood to be a group of Chilean learners of English from the same school.

PARTICIPANTS

The participants were recruited using criterion-based sampling. In our study, we looked for participants who were able to commit to a five-week intervention, as the study took place after regular school hours. Twelve participants were accepted for the study under the premise that a qualitative study's depth and richness of data are not necessarily determined by the number of participants but by the quality of information they provide (Creswell & Poth, 2018). As one of the participants dropped out, there were 11 language learners (10 female, 1 male) (Table 1).

Table 1. Participants

Participants	Age
Pedro, Noelia, Sofía*	14
Carol, Nelly, Kath, María, Ambar, Belén, and Camila	16
Julia	17

*Pseudonyms

The observed gender imbalance reflects the volunteer pool rather than an intentional selection, highlighting the need for future studies to strive for more gender-balanced sampling. Learners ranged in age from 14 to 17 years old and belonged to the A2 level according to the CEFR. Proficiency was determined with the Key English Test (Cambridge English, 2016) and an in-house test given to students at the start of the year. Participants with pre-intermediate proficiency were selected because research has shown that while beginners tend to read off captions, advanced proficiency listeners use them as mere text comprehension backups (Leveridge & Yang, 2014; Yeldham, 2018). Other than

watching captioned movies and singing along with lyrics, participants had no formal instruction on how to use captioned videos for language learning.

PEDAGOGICAL INTERVENTION

The pedagogical intervention was built to orchestrate tasks, focus viewing, and active student participation. First, to engage students through tasks, the pedagogical intervention was built upon previous work by Ramírez-Orellana (2014). A needs analysis was used to identify students' interests and preferred media formats. Accordingly, the four-unit syllabus and associated sample lessons revolved around the topic of 'teenage issues' as portrayed in different media: music, animated cartoons, TV series, and movies (Table 2). Also, each lesson was structured into pre-, while-, and post-listening formats. Second, to engage students through focused viewing, the goal for the design of captioned materials was that these did not become a crutch on which listeners relied too heavily but instead enabled listeners to perform activities as they advanced in listening competence. Third, to promote active student participation, each lesson was designed to encourage students to listen to the texts and construct meaning collaboratively through games and role-playing. In each lesson, students were assigned roles and each role conveyed different purposes for listening.

Table 2. Lesson sequencing, objectives, and input text characteristics

Units	Objectives	Types of texts
Unit 1: Music in my teens!	Recognize the speaker's use of language (formal, informal) with full captions . Predict the speaker's message and meaning with full captions.	Song: Fireworks Length: 03:50 mins WPM: 89.48
Unit 2: Teenagers and family life in cartoons	Recognize context and main ideas with partial captions . Rank order the main events with partial captions.	TV series snippet: Daria Length: 02:10 mins WPM: 133.38
Unit 3: Teenager stereotypes in movies	Recognize specific information with keyword captions . Compare the literal and implied meaning of the message with keyword captions.	TV series snippet: Nerds and Geeks in film and television Length: 02:00 mins WPM: 143
		Movie snippet: Mean Girls Length: 01:58 mins. WPM: 167.
Unit 4: Teenage issues in TV series	Summarize the main ideas and supporting details. No caption support.	TV series snippet: Sabrina the Geek Length: 02:08 mins WPM: 202

MATERIALS

For input text selection, we followed the principle of "grade the task, not the text," as suggested in Vandergrift and Goh's (2012) work to favor the use of authentic input. This decision also aligns with language teaching practices in Chile, where teachers resort to authentic input that addresses the content with no further text manipulation (Cárdenas-Claros, 2020). Moreover, the selected texts addressed students' needs and interests, as found in the needs analysis by Ramírez-Orellana (2014). Also, as this was not a comparative study, we selected texts with different text lengths and speed of

reproduction. Input texts were deemed appropriate for the targeted proficiency level, given their length, speed of reproduction, and vocabulary complexity. The selected videos were captioned using the DivXL and media subtitler freeware version 2.1.2, which allows users to create and edit external subtitle or caption files for AVI, MPG, WMV, and all types of videos (Divxland, 2017). Full captions in our study were a verbatim written transcription of the content of the video synchronized with the audio. The captions were displayed at the bottom of the video, as seen in Figure 1.

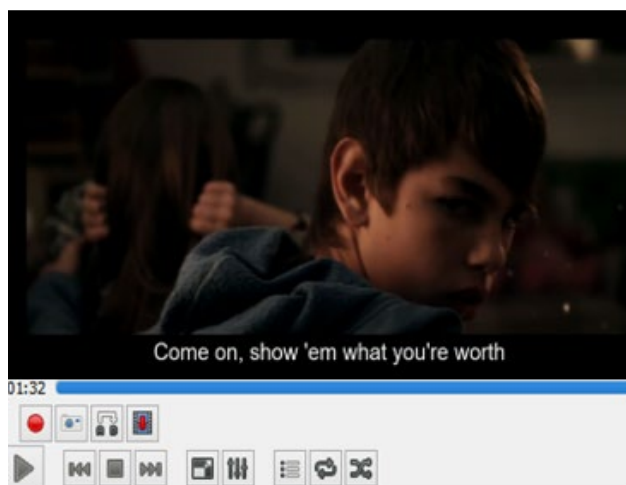


Figure 1. Full captions

Aware that in most research designs, keywords are one type of partial captions (P.-J. Lee et al., 2021; Montero Perez, Peters, Clarebout, & Desmet, 2014), in this study, we distinguish from partial captions or sentence-level captions and keyword captions. Partial captions entail the verbatim transcription of full sentences that convey idea units. We opted for partial or sentence-level captioned materials for learners to make the transition to keyword caption smooth and staged. Moving from full captions to keyword captions seemed an abrupt measure, as in full captions, reliance can be entirely from text to audio and in keyword captions from audio to text. The sentence-level captions appeared in synchrony with the spoken text and were also displayed at the bottom of the video. To select sentences for partial captions, we employed expert judgment from two of the students' regular language teachers. These experts, with more than ten years of teaching experience, assessed together with the authors which video segments contained essential information and key idea units for inclusion in the captions (see Figure 2). For that, they were given the transcripts and were asked to (1) underline sentences that convey the main points or central ideas of the video segment, (2) highlight information that provided essential details or context necessary for understanding the main idea, and (3) prioritize sentences that introduced target vocabulary words. Their suggestions were then discussed with the researchers, and decisions were made accordingly.

As for keyword captions, similar to the work of P.-J. Lee et al. (2021), content words that conveyed the most meaning appeared once they were heard in the text and vanished as new content was heard. Analyses were performed using the HKU CAES vocabulary profiler to determine keywords. Keywords were part of either the 1k or 2k most common words list.



Figure 2. Partial or sentence-level caption

DATA COLLECTION MATERIALS AND PROCEDURES

The pedagogical intervention was implemented in a five-week period after consent and assent forms were received and checked to comply with ethical procedures for the protection of minors. Classes met twice a week in 90-minute sessions. In the first session, students shared their expectations with the pedagogical intervention and completed an entry survey. This was a paper-based instrument that gathered demographic information and assessed participant experience and familiarity with captioned materials.

The sessions in which students were exposed to captioned materials (sessions 2 to 9) were classes lasting about 60 minutes each. As our study focused on perceptions, we did not control any variables but modeled our activities based on students' regular class activities. Therefore, our participants developed pre-listening exercises and watched the captioned materials (sessions 2 and 3 – full-captions; sessions 4 and 5 – partial caption; sessions 6 and 7 – keyword captions; and sessions 8 and 9 – no-caption at all), completed comprehension questions, and did vocabulary exercises as part of the post-listening exercises. At the beginning of each new caption session, students were explained about its corresponding characteristics and then were given time to write a 3-minute reflection addressing the prompt: “My expectations for this session are.”

Captioned videos were displayed via a data projector on a communal screen. Students worked collaboratively with assigned roles, discussed the topics of the listening texts, and completed associated pre-listening, while-listening, and post-listening comprehension exercises. In the last 15 minutes of each second session, participants were asked individually to complete paper-based written reflections around three main prompts: (1) perceptions of the caption type used, (2) advantages/disadvantages of the caption type used, and (3) whether their initial expectations have been met.

In the last session, we conducted a focus group in which participants reflected on their overall experience of the course. To avoid reflections being influenced by others' opinions, participants were asked to write their answers individually on Post-it notes. Then, Post-it notes were collected and displayed on a board, and participants were invited to elaborate on their answers.

DATA ANALYSIS

The data set consists of 11 entry-surveys, 44 written reflections, and a 65-minute focus group session. To comply with procedures for qualitative data, reflections and focus group transcripts were read several times to identify functions, themes, and explaining factors (Miles et al., 2014). Data analysis was conducted in three rounds. In an initial round, the first author followed open coding techniques

to identify students' perceptions regarding benefits and drawbacks across caption types. Relying primarily on data from the entry-surveys and the written reflections, the initial round of coding helped the researcher name and classify perceptions that served for the coding protocol construction. Then, the same author grouped students' perceptions and established linkages from the data. The coding protocol was then extended to account for the information derived from the focus group session. In the third round, the second author coded 20% of the data to calculate inter-coder reliability, which was 0.97. Triangulation of data sources (one entry-survey per participant, four reflections per participant, and a focus group session) and coders (two coders) was done to enhance dependability.

RESULTS

Analyses of listener perceptions about full, partial, and keyword captions show that the three types of captions served two functions: (1) to assist – text comprehension, on-task focus, vocabulary learning, and pronunciation, and (2) to integrate – written-word-to-sound mapping and background knowledge into new knowledge. Findings also show that listeners positively valued transitioning from full-captioned materials to non-captioned materials.

In this section, we elaborate on each of the functions and associated clusters that were identified across data sets. Tables 3 and 4 summarize the function, provide definitions and sample data, report the total number of times a cluster was mentioned, and list the number of participants reporting it. Then, we discuss how progressive caption reduction, to the complete absence thereof, was perceived by the participants. Also, to enhance transparency in the analysis, we provide full information about the origin of the data; that is, a quote associated with “Kath, WR1-Q1, L2” should be understood as a comment from Kath, stemming from the written reflection 1, question 1, line 2.

FUNCTION ONE: CAPTIONS ARE PERCEIVED TO ASSIST COMPREHENSION

In our analysis, assistance is understood as a perceived value of captions as explained by four clusters: text comprehension, on-task focus, vocabulary learning, and pronunciation improvement (Table 3).

The first cluster, *text comprehension*, is understood as how students perceive that the use of different types of captions helps them understand text meanings. This cluster appeared across data sets 29 times and was mentioned at some point by all the participants.

Six participants stated that comprehension was eased when exposed to full captions. They were able to confirm what they were listening to without feeling lost trying to understand the audio. Kath signaled that she was able to “absorb” the whole content of what she was listening to and felt satisfied with her performance. Ambar added that using full captions “helped her to understand perfectly what she was listening to.”

Six participants also stressed that partial captions served as ancillary elements in assisting comprehension. Pedro described that partially captioned material helped him understand the text better. Noelia and María used the term ‘guide’ to refer to the type of support provided by partial captions. Noelia went on to describe that, although she was unfamiliar with this type of video-based material, she became comfortable with it as she progressed in the course. Other participants found partial captions more challenging than full captions. Ambar, for instance, noted that with partial captions, she had to attentively listen to aural input and synchronize ‘her ears and sight.’ Inadvertently, she was constantly ‘waiting for the textual support to appear on screen,’ and this affected comprehension. Ten participants also acknowledged partial captions as facilitators of listening comprehension.

Table 3. Function one: Assistance

Cluster	Definition	Sample data	# of occurrences	# of participants
Text comprehension	The use of different types of captions eases the understanding of aural input.	“When I use captions, I understand better English, but it should be a gradual process ...” (Kath, WR1-Q1, L2*).	29	11
On-task focus	The notion that using captions requires listeners to be focused on the listening exercise.	“... although captions start being removed, you are already familiarized with them and have acquired new vocabulary, so you also start paying more attention to the audio ...” (Pedro, FG-Q2, L7).	17	9
Vocabulary learning	The perception that the use of captions aids vocabulary recognition and learning.	“... You can learn the meaning of the words ...” (Nelly, FG-Q1, L41-42).	19	11
Pronunciation improvement	The notion that pronunciation is learned	“... When you use captions, you learn how to pronounce the words ...” (Belen, FG-Q1, L15-16).	8	6

* WR = Reflection; FG = Focus group; Q# = Question number; L# = Line in the transcription

The first cluster, *text comprehension*, is understood as how students perceive that the use of different types of captions helps them understand text meanings. This cluster appeared across data sets 29 times and was mentioned at some point by all the participants.

Six participants stated that comprehension was eased when exposed to full captions. They were able to confirm what they were listening to without feeling lost trying to understand the audio. Kath signaled that she was able to “absorb” the whole content of what she was listening to and felt satisfied with her performance. Ambar added that using full captions “helped her to understand perfectly what she was listening to.”

Six participants also stressed that partial captions served as ancillary elements in assisting comprehension. Pedro described that partially captioned material helped him understand the text better. Noelia and María used the term ‘guide’ to refer to the type of support provided by partial captions. Noelia went on to describe that, although she was unfamiliar with this type of video-based material, she became comfortable with it as she progressed in the course. Other participants found partial captions more challenging than full captions. Ambar, for instance, noted that with partial captions, she had to attentively listen to aural input and synchronize ‘her ears and sight.’ Inadvertently, she was constantly ‘waiting for the textual support to appear on screen,’ and this affected comprehension. Ten participants also acknowledged partial captions as facilitators of listening comprehension.

When working with keyword captions, eight participants perceived them as tools for scaffolding comprehension. Ambar explained that keyword captions helped reinforce her listening comprehension skills since she was unknowingly “forced” to relate the keywords to the text to which she was listening. Kath emphasized how listener attention is drawn to the most relevant words when concentrating on the input text. Camila and Sofia also claimed that keyword captions helped them “better

develop” listening comprehension over full captions. Sofia acknowledged that using keyword captions did not initially contribute much to her text comprehension but gradually became easier for her to understand. Four other participants also felt comfortable with keyword captions, which were seen as “scaffolds” to listening comprehension.

For the second cluster, *on-task focus*, nine participants agreed that captions helped them focus more on the oral message. For instance, Carol and Kath stated that the dual input of full captions forced them to focus on the text since they were “kind of expected” to read and listen to the text simultaneously. María described full captions as elements that helped them focus better on the task. She reflected on how using full captions mimicked real-life tasks like watching movies. Carol and María reported feeling at ease when using full captions and described overcoming the frustration experienced in regular classes when exposed to audio-only listening exercises.

Belén and Carol highlighted the value of partial captions in aiding selective task attention. They posited that partial captions helped them to concentrate more on the task since the lack of other words kept them alert to what was coming next.

For Julia, Camila, Noelia, Kath, and Nelly, keyword captions demand that listeners pay closer attention to aural input since only the most important words are provided. They felt that paying attention to the on-screen keywords was crucial for keeping them on-task and not having their minds wander, as when working with audio-only materials.

A third cluster, *vocabulary learning*, refers to the perception that the use of captions aids vocabulary recognition and learning. This cluster appeared across data sets on 19 different occasions and was mentioned by the 11 participants. Carol, Nelly, and Julia considered full captions a useful tool for acquiring new words and learning how to write them. All participants pointed to learning new vocabulary as a benefit of interacting with partial captions. Camila summarized this view: “Using partial captions was difficult during the first session, as I had to familiarize myself with this new type of captions; as lessons advanced; however, it was easier to remember and learn the words displayed” (Camila WR2, L2-3).

Being exposed to keywords was also indicated as a boon to vocabulary learning. Camila and six other participants saw keyword captions as a potential vocabulary-building technique. Nelly found that keyword captions allowed her to learn more vocabulary since the most important words were displayed on the screen. Belén remarked that keyword captions allowed her to not only recognize but also learn more vocabulary. Kath also supported this idea and acknowledged that her L2 vocabulary was greatly enriched after the course.

The fourth and last theme, *pronunciation*, was mentioned in eight instances and by six participants. *Pronunciation* in this work refers to listeners’ perceptions of how interactions with captions foster their awareness of pronunciation cues and how this contributes to potentially improving their own pronunciation. Six participants described full captions as useful in learning how target language words are pronounced. They highlighted that, with full captions, listeners remember the pronunciation of previous words, remarking that the pronunciation of difficult words was aided through simultaneous seeing and hearing of words. In the reflection data, the participants did not specifically state the value of partial or keyword captions for pronunciation purposes.

FUNCTION TWO: CAPTIONS ARE PERCEIVED TO AID INTEGRATION OF CONTENTS AND KNOWLEDGE

The second function, *integration*, refers to how captions allow listeners to combine one set of information with another to form a whole. This integration was perceived in two ways: written-word-to-sound mapping and previous content knowledge to new content knowledge (Table 4).

Written-word-to-sound mapping is the representation of sound and written text in the listeners’ minds. Nine participants referred to this theme with different amounts of caption support, claiming they

were able to relate words read with the aural stream of what they listened to. Noelia and Julia posited that full captions allowed them to relate the sound heard in the video with the written form of the text. Carol and four other participants agreed that partial captions enabled them to immediately corroborate what they were listening to with the written texts. Camila remarked that if she was unable to understand the audio, she was able to recognize the word sounds thanks to simultaneous exposure to written texts.

Table 4. Function two: Integration of content and knowledge

Theme	Definition	Sample data	# of occurrences	# of participants
Written word-to-sound mapping	Mental representation of written text-and-sound	"... When you use captions, you can relate the audio with the sound, so you can understand the content of the videos better ..." (Julia, FG-Q1, L2).	9	7
Previous content knowledge integration	Integration of previous content to develop the listening comprehension tasks.	"... when using captions, if you don't understand something, you can relate it with previous words from previous exercises ..." (Julia, FG-Q3, L3).	7	7

Keyword captions were also identified as a means to help listeners map sound-to-written forms of the language. Sofia and Kath remarked that as listeners, they were able to listen to the text and relate the words that they had seen written.

The second theme, *previous content knowledge integration*, or how content knowledge the listener brings to the listening episode is integrated with incoming information from the input text. Seven participants referred to this theme. Julia mentioned that when using captions, though some ideas may not be understood, such ideas could still be related to content developed from previous videos and exercises. At the end of the course, Kath remarked that she could better understand videos without captions because previous classes had shown similar content aided by captions. Noelia emphasized that, due to working with captions from previous classes, content materials seemed more familiar to her, which helped her integrate what she had learned with the new information. This gave her a sense of achievement and proficiency attainment. María emphasized that she was able to relate what she had learned using full captions with new content provided in partially captioned materials. Ambar praised the transition from full captions to partial captions and noted that it was easier for her to understand content with partial captions since full captions had been used previously with similar topics. Camila also commented that partial captions allowed her to build up content knowledge, given that the selected videos all addressed the same teenage issues and thus were on similar topics.

TRANSITIONING FROM DIFFERENT CAPTION TYPES TO NO-CAPTION

The second research question inquired about participants' perceptions of their experience transitioning from working with captioned materials to non-captioned ones. Seven students signaled that they were better able to understand input that had been seen in preceding sessions because they were trained to work with full, partial, and keyword captions. The lack thereof was no longer an issue for them. In their view, this gradual reduction of captions made the process of understanding "less challenging." Ambar explained that using full captions was almost a "necessary step" to aid comprehension, given her previous experience with written texts. This, in her view, scaffolded subsequent encounters with progressive caption reduction.

Carol described that working with videos without captions was only appropriate and useful for listening comprehension because the work had been “stepwise and done with the other three types of captions” (FG-Q2, L25). She also stressed how, in the last session, although the videos had no captions, she did not experience the regular frustration she experienced when working with listening materials in the classroom. In her view, the gradual reduction of captions made the materials more engaging, and, most importantly, she was able to overcome comprehension problems. Kath remarked that captions “are the best way to understand spoken English,” and she suggested that it should be a gradually reduced process so as not to create reliance in the long run. Kath recommended using videos without captions, but to ensure that students really understood, she noted that “it should be a gradual process” and that “teachers should not expect it to happen overnight.”

Nelly claimed that it is more difficult to understand aural input when textual support is not part of a video. However, she admitted that work done in previous lessons helped her gradually recognize and learn vocabulary, allowing her to understand the text messages. Belén and Noelia also remarked that they were able to understand the video with no captions more easily due to the recycling of contents and the exposure to the three caption types. María also added that, after having worked with the final non-captioned videos, she gained confidence and experienced first-hand that listening comprehension development was possible. Pedro, in comparing his first and last experiences between the full caption and no caption sessions, stated that he was able to better understand the videos played in the last class due to the training to let go of captions. Pedro noted that working with full, partial, and keyword captions should be essential prior steps to exposing students to uncaptioned videos. Exceptionally, Julia did not find much value in transitioning from full captions, partial, and keyword captions to no captions. In the non-caption sessions, she still felt lost and struggled to understand the text. To aid text comprehension, she had to “rely on L1 mental translation” (WR4, L3).

In an overall evaluation of the pedagogical intervention, students valued the input text selected and the collaborative listening activities. María noted:

In our class, we mostly do grammar and complete worksheets ... when we do the listening [activities], we always work alone, and we only listen from the book, we never watch TV programs or songs ... Here, we watched videos ... and videos that you watch in Spanish like Sabrina ... and yes, it is not that you can tell your ears not to listen to other parts, but we all had to watch the whole thing, so we talked and compared what we understood, and one remembers words from captions ... And that was also very helpful (FG-Q3 41).

Similarly, participants valued the systematically structured lessons where they knew in advance that they had to reflect on the types of captions. Belén summarized this view better: “I found myself constantly thinking if captions were useful or not ... I knew that after each lesson I had to write my evaluation of the class” (FG-Q3, 47).

DISCUSSION

This study explored listeners’ perceptions of a pedagogical intervention that examined three types of captions as they were gradually reduced to the point of absence. As noted in the results section, participants reported relying on captions to assist comprehension, on-task focus, vocabulary learning, and pronunciation awareness. The choice of input that was meaningful, engaging, and familiar to participants possibly aided the perception of improvement of text comprehension and aligned with the work of Abobaker (2017), in which listeners relied on different types of captions to ease comprehension. Six participants also perceived that working with partial captions was more challenging than working with full captions. Possibly, when working with full captions, the auditory signal seems to have been disregarded altogether, as the participants perceived that attention resources were not necessarily employed to understand the aural input but in waiting for the textual support offered by partial captions (Taylor, 2005).

In this study, keyword captions were perceived to afford learners valuable opportunities to focus more on listening rather than on reading, a benefit also reported in Guillory's (1998) work. The results show that captions were perceived to assist *on-task* focus and highlighted the role of partial and keyword captions in 'forcing' listeners to stay on task. This perception may be influenced by selective attention processes afforded through partial and keyword captions. Similarly, Montero Perez, Peters, Clarebout, and Desmet (2014) and Yang and Chang (2014) reported that keyword captions help listeners increase their attention on the task.

Exposure to multimodal input may have also influenced listeners' perception of vocabulary learning (Hayati & Mohmedi, 2011; Montero Perez, Peters, Clarebout, & Desmet, 2014; Sydorenko, 2010; Vanderplank, 2019a; Yang & Chang, 2014). Text salience afforded through partial and keyword captions, in particular, may have directed students' attention to keywords, an essential condition for vocabulary learning. Additionally, participants perceived that interaction with full captions assisted pronunciation awareness. Similar findings were reported by Mohsen and Mahdi (2021) and Mahdi (2017), who noted that the use of full captions "bridges the gap between reading and listening skills, helping learners with pronunciation" (p.834). Interestingly, in this study, participants did not perceive many benefits of using partial or keyword captions for improving pronunciation. It is possible that when presented with keyword captions, listeners were unable to distinguish word boundaries as they did not see the written words that came before and after. Moreover, a combination of a fast speech rate and a rapid showing of keywords may not have given them enough time for listeners to stop and think about pronunciation cues. Perhaps if they worked in a self-regulated environment, having to decide when to use playback controls and for what purposes could have encouraged further reflection. Also, the participants were at a pre-intermediate level of proficiency in which automaticity is developing and attentional resources are distributed for other processes, thus affecting reflection.

The data also shows that listeners reported captions to be integrative tools that encouraged written-word-to-sound mapping and integration of new content knowledge with previous knowledge. The participants also reported the perceived benefits of full, partial, and keyword captions to assist written word-to-sound mapping. These results need to be interpreted with consideration of the strong literacy skills of the participants, given their previous instruction experience based on reading development.

The provision of different caption types was seen as a content schemata activation device that listeners relied on to integrate previous content knowledge with new knowledge. This finding aligns with that reported by Winke et al. (2010), who found that caption use facilitates content knowledge integration. Also, participants reflected on how even if some ideas were not understood, captions helped them to relate previous contents to missing information and thus construct new knowledge.

Additionally, the study results show that students valued transitioning from full caption to non-caption materials, given the progressive support provided through captions. Two key findings have to do with students' perception that progressive reduction of captions aids text comprehension and vocabulary learning. Similar findings are described by Vanderplank (2019b), who noted that participant "HL used the captions more in the first half of the Spanish film, *El Niño*, since the number of unfamiliar words was much greater during the first half" (p. 413). These perceptions were likely influenced by the carefully designed syllabus and the informed materials selection. Videos were not only engaging for the learners but also revolved around a single topic, "teenage issue," which yielded vocabulary and content recycling from one lesson to the next, and this may have influenced the perception of comprehension improvement. Also, the multimodal input afforded through captions may have supported recognition and implicit learning of words that re-occur later without captions (Sydorenko, 2010; Vanderplank, 2019b). Despite the stepwise intervention, the participants in this study did not perceive partial and keyword captions to aid pronunciation awareness. They did not have the option of regressing as the videos were projected on a communal screen. If they got distracted, the video simply kept advancing to new frames, making the identification of pronunciation cues challenging. All the participants, except for one, praised the gradual and stepwise

reduction of captions. This reduction helped listeners overcome comprehension breakdowns and gave listeners a sense of accomplishment.

Most participants also valued the opportunities provided for reflection both through the collaborative listening activities and the reflection logs they completed after each session. While the collaborative listening activities ‘forced’ listeners to self-monitor and peer-monitor text comprehension, the written reflections prompted them to reflect, monitor, and keep track of whether and how each type of caption contributed to listening competence development.

IMPLICATIONS, CAVEATS, AND FUTURE RESEARCH

The findings stem from the reflections of 11 high-school learners of English. However, pedagogical implications can be extended to broader language-learning communities. We extend Vanderplank’s (2016) suggestion to note that to reap the benefits of gradual reduction of captions, input selection, orchestration of task design, strategies, focused viewing, and opportunities for reflection are crucial. With regard to input selection, teachers and curriculum designers should consider contextual factors and address students’ needs and interests. Input texts need to be meaningful, cater to listener interests, and be presented through a variety of media formats and genres, as noted in the needs analysis that served as input for the design of the pedagogical intervention for this study.

Task design orchestration is also imperative to promote the active participation of the students as well as the effective use of captioned video materials. This study noted three key aspects: a progressive reduction of captions, sequencing of lessons that followed a pre-, while-, and post-listening format, and provision of opportunities for collaborative listening. Designing a pedagogical intervention around full lessons, not data-collection-only sessions, gave participants the opportunity to interact with different caption types and to perceive the benefits of caption reduction for language learning. Structuring lessons using a pre-, while, and post-listening format helped learners prepare for the task, activate background knowledge, and establish a benchmark of what they were able to do with the support of each caption type. Including activities that promoted collaborative listening, where listeners not only activated previous knowledge or confirmed text comprehension but were given different roles and thus different listening purposes, helped listeners develop a sense of community and helped perceive caption reduction as beneficial for comprehension. Complementarily, providing opportunities for reflection was also crucial. Giving students time to stop, think, and reflect on if and how they benefited from each type of caption helped them construct informed opinions that were later shared in the focus group discussion.

The findings presented here should be interpreted with some caveats in mind. Success in L2 listening with captioned materials relies on student literacy (Avello & Muñoz, 2023). Therefore, interventions like the one presented here are most appropriate for contexts where students have a foundation in reading and sound and written systems are similar to English. With L2 learners whose L1s writing scripts are very different (as with Arabic, Hebrew, Japanese, Chinese, and Korean), the perceptions might be different, and a more active intervention would be needed to reap the benefits of progressive caption reduction for L2 listening comprehension. Also, results should be interpreted in the context from which they originated. As we favored the use of authentic input, the captioned materials used for this study have inevitably differing lengths and WPMs. We acknowledge that text characteristics may pose different challenges to the learners, and this may have an impact on their perception.

Avenues for further research are multiple. As this is a qualitative study, longitudinal quantitative studies that measure comprehension and/or vocabulary learning as captions support is gradually reduced (full, partial, and keyword) are strongly needed. This would entail using research designs that include larger sample sizes or control for different variables, using texts with similar characteristics, and these decisions might be at the expense of authentic text use. These studies could examine a range of individual differences (working memory capacity, age, levels of engagement, and language background).

Future research could also examine captions with students with learning difficulties and/or disabilities. We contend that progressively reducing captions not only aids language learning but may offer a means of making language classrooms more inclusive, a goal of many teachers and schools these days. Future research initiatives could also look at the effects of caption reduction when displayed on different media. With the increasing use of handheld computing devices and the extent of commuting times for language learners, extensive listening projects in self-regulated environments that combine different types of captions could be examined.

On a final note, the findings reported here invite teachers and researchers to continue to examine how to best reduce reliance on captions and thus enable L2 listeners to cope with L2 listening in caption-free environments.

REFERENCES

- Abobaker, R. (2017). Improving ELLs' listening competence through written scaffolds. *TESOL Journal*, 8(4), 831-849. <https://doi.org/10.1002/tesj.339>
- Avello, D., & Muñoz, C. (2023). The development of receptive language skills from captioned video viewing in primary school EFL learners. *Education Sciences*, 13(5), 479. <https://doi.org/10.3390/educsci13050479>
- Cambridge English. (2016). *A2 Key for Schools paper-based sample listening test* [Sample examination paper]. Cambridge University Press. <https://www.cambridgeenglish.org/exams-and-tests/key-for-schools/preparation/>
- Cárdenas-Claros, M. (2020). Spontaneous links between help option use and input features that hinder second language listening comprehension. *System*, 93, 102308. <https://doi.org/10.1016/j.system.2020.102308>
- Cárdenas-Claros, M., & Dassonville, K. (2022). A first step to task design in computer-based L2 listening: Task characteristics elicitation. In J. Colpaert, & G. Stockwell (Eds), *Smart CALL*. (pp. 266-292). Castletown Publications. <https://doi.org/10.29140/9781914291012-12>
- Cárdenas-Claros, M., Sydorenko, T., Huntley, L., & Montero Perez, M. (2023). Teachers' voices on multimodal input for language learning. *Language Teaching Research*. <https://doi.org/10.1177/13621688231216044>
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage.
- Divxland. (2017). *Divxland media subtitler (Version 2.2.0)*. <https://divxland-media-subtitler.en.softonic.com/?ex=RAMP-1768.1>
- Guillory, H. (1998). The effects of keyword captions to authentic French video on learner comprehension. *CALICO Journal*, 15(1/3), 89-108. <http://www.jstor.org/stable/24147690>
- Hayati, A., & 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>
- Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition*, 42(3), 625-651. <https://doi.org/10.1017/S0272263120000108>
- Lee, P.-J., Liu, Y.-T., & Tseng, W.-T. (2021). 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>
- Leveridge, A. N., & Yang, J. C. (2013). Testing learner reliance on caption supports in second language listening comprehension multimedia environments. *ReCALL*, 25(2), 199. <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>
- Mahdi, H. (2017). The effect of keywords video captions on vocabulary learning through mobile-assisted language learning. *Australian Educational Computing*, 32(1), 1-12. <https://doi.org/10.5539/ijel.v7n4p1>
- Miles, M., Huberman, M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. Sage.
- Mirzaei, M. S., Meshgi, K., Akita, Y., & Kawahara, T. (2017). Partial and synchronized captioning: A new tool to assist learners in developing second language listening skills. *ReCALL*, 29(2), 178-199. <https://doi.org/10.1017/S0958344017000039>
- Mohsen, M., & Mahdi, H. (2021). Partial versus full captioning mode to improve L2 vocabulary acquisition in a mobile-assisted language learning setting: Words pronunciation domain. *Journal of Computing in Higher Education*, 33(2), 524-543. <https://doi.org/10.1007/s12528-021-09276-0>
- Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, 55(2), 163-192. <https://doi.org/10.1017/S0261444821000501>
- Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, 18(1), 118-141.
- Montero Perez, 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. <https://doi.org/10.1017/S0958344013000256>
- Montero Perez, M., Van Den Noortgate, W., & 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>
- Pattemore, A., & Muñoz, C. (2022). Captions and learnability factors in learning grammar from audio-visual input. *The JALT CALL Journal*, 18(1), 83-109. <https://doi.org/10.29140/jaltcall.v18n1.564>
- Pujadas, G., & Muñoz, C. (2020). Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles. *Studies in Second Language Acquisition*, 42(3) 551-575. <https://doi.org/10.1017/S0272263120000042>
- Ramírez-Orellana, D. (2014). *Enhancing listening comprehension through video captions* [Unpublished undergraduate thesis, Pontificia Universidad Católica de Valparaíso, Chile].
- Rodgers, M., & 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>
- Rooney, K. (2014). The impact of keyword caption ratio on foreign language listening comprehension. *International Journal of Computer-Assisted Language Learning and Teaching*, 4(2), 11-28. <https://doi.org/10.4018/ijcall.2014040102>
- Sato, M., & Loewen, S. (2019). Do teachers care about research? The research-pedagogy dialogue. *ELT Journal*, 73(1), 1-10. <https://doi.org/10.1093/elt/ccy048>
- Stewart, M., & Pertusa, I. (2004). Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals*, 37(3), 438-442. <https://doi.org/10.1111/j.1944-9720.2004.tb02701.x>
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, 14(2), 50-73.
- 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. (2019). Maximizing the potential of captions for primary school ESL students' comprehension of English-language videos. *Computer Assisted Language Learning*, 32(7), 665-691. <https://doi.org/10.1080/09588221.2018.1532912>
- Vandergrift, L., & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

- 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>
- Vanderplank, R. (2019a). Video and informal language learning. In M. Dressman, & R. W. Sadler (Eds.), *The handbook of informal language learning* (pp. 181-201). John Wiley & Sons. <https://doi.org/10.1002/9781119472384.ch12>
- Vanderplank, R. (2019b). 'Gist watching can only take you so far': Attitudes, strategies and changes in behaviour in watching films with captions. *The Language Learning Journal*, 47(4), 407-423. <https://doi.org/10.1080/09571736.2019.1610033>
- Wang, Y.-C. (2012). Learning L2 vocabulary with American TV drama: From the Learner's perspective. *English Language Teaching*, 5(8), 217-225. <https://doi.org/10.5539/elt.v5n8p217>
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning and Technology*, 14(1), 65-86.
- Wisniewska, N., & Mora, J. C. (2020). Can captioned video benefit second language pronunciation? *Studies in Second Language Acquisition*, 42(3) 549-624. <https://doi.org/10.1017/S0272263120000029>
- Yang, J., & 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>
- Yeldham, M. (2018). Viewing L2 captioned videos: What's in it for the listener? *Computer Assisted Language Learning*, 31(4), 367-389. <https://doi.org/10.1080/09588221.2017.1406956>

AUTHORS



Mónica S. Cárdenas-Claros is a full professor at the Instituto de Literatura y Ciencias del Lenguaje at Pontificia Universidad Católica de Valparaíso in Chile. Mónica is the author of a number of articles on computer-based L2 listening, technology integration in language classrooms, and the development of research competencies of pre-service teachers.



Daniela Ramírez-Orellana is an English teacher at a private school in Chile. She has worked across different areas of education and has also worked as a research assistant. Her areas of interest are L2 listening instruction and the use of captions to improve comprehension.