Language Learning & Technology
ISSN 1094-3501 CC BY-NC-ND

2024, Volume 28, Issue 2
pp. 1–25

ARTICLE

Audiovisual input in language learning: Teachers' perspectives

*Tetyana Sydorenko, Portland State University

Mónica S. Cárdenas-Claros, Pontificia Universidad Católica de Valparaíso, Chile

Elizabeth Huntley, University of Colorado, Denver

Maribel Montero Perez, Ghent University, Belgium

Abstract

A substantial body of research shows that various types of audiovisual (AV) input such as videos and videos with second language (L2) subtitles can facilitate language learning. However, language teachers' day-to-day practices with regard to multimodal input is less understood. To bridge the gap in language education, this study investigates teachers' perceived use of four types of AV input (video only, video with subtitles, video with captions and video with enhanced captions) and factors influencing teachers' perspectives on these types of input for in-class and out-of-class learning. Questionnaire data were collected from 193 L2 teachers across the globe about their perceived use of AV input. Teachers reported that they use video and captioned video most frequently in both classroom and out-of-class contexts. Logistic regression analyses revealed that teachers' perceived importance and comfort using specific AV input types were the two most important factors explaining teachers' reported use. Complementarily, open-ended responses were analyzed qualitatively to identify teachers' additional reasons for non(use) of such input.

Keywords: *Captions, Subtitles, Audiovisual Input, Teacher Perceptions*

Language(s) Learned in This Study: *English, Japanese, Chinese, French, German, Russian, Arabic, Italian, Spanish, Portuguese, Dutch, isiZulu*

APA Citation: Sydorenko, T., Cárdenas-Claros, M. S., Huntley, E., & Montero Perez, M. (2024). Audiovisual input in language learning: Teachers' perspectives. *Language Learning & Technology*, *28*(1), 1–25. https://hdl.handle.net/10125/73587

Introduction

Multimodal input, i.e., input that provides information through different modes (written, auditory and visual), is pervasive in everyday life. In particular, audiovisual (AV) input—dynamic videos with or without on-screen text of various forms—is increasingly present across streaming platforms, websites, and social media. As such, the field of Second Language Acquisition (SLA) has become increasingly interested in learning from AV input over the last decade. Research has consistently shown that AV input provides opportunities for vocabulary learning (e.g., Peters & Webb, 2018; Rodgers & Webb, 2020), but there is also beginning evidence for the learning of grammar (e.g., Cintrón-Valentín et al., 2019; Muñoz et al., 2021), pronunciation (Wisniewska & Mora, 2020), and listening/speech decoding (Mitterer & McQueen, 2009). Several language learning tools have been designed on the basis of such empirical research (e.g. the Language Reactor extension (https://www.languagereactor.com/), (n.d.), and IdiomsTube (Lin, 2022)). Additionally, many resources are available to teachers and students for creating and editing videos (e.g., Godwin-Jones, 2015).

Despite research findings on the benefits of AV input for learning and the corresponding increase in

*** Corresponding Author:** Tetyana Sydorenko, tsydorenko@pdx.edu

learning applications and teaching resources, AV input use by language teachers appears to be under investigated (cf. Vanderplank, 2016). For example, do language teachers know about the specific learning benefits of various types of AV input? If so, do they have time to create, adapt, or find materials corresponding to these AV input types?

To narrow the gap between research and practice, it is important to know not only whether teachers use the types of AV input that are commonly investigated by researchers, but also the reasons behind (non)use. In language teaching in general, it has been shown that teachers' beliefs can directly influence classroom practices (Borg, 2018). However, teacher beliefs are only one among many contextual factors (e.g., institutional support, curriculum, etc.) further shaping their decisions (Wesely et al., 2024). These findings likewise apply to technology implementation: as "the lynchpin around which successful online learning events revolve" (Guichon & Hauck, 2011, p. 188), instructors decide which technologies are implemented and how (Arnold & Ducate, 2015). We know of only two consequential studies that have begun to examine teachers' reported use of AV input and possible factors influencing their decisions: Kaderoğlu and Esquerré (2021) and Mariotti (2015). To contribute to the research-practice dialogue (Sato & Loewen, 2022) and address Vanderplank's (2016) call for more research on AV input use by language teachers, we examined second language (L2) teachers' reported use of four types of AV input (video, subtitles, captions, and enhanced captions) and their reasons for using them. Although reported practices do not always directly correlate with actual practices (cf. Al-Hoorie et al., 2022), in this exploratory study we take one more step to shedding more light on AV input from the perspective of practitioners from multiple teaching contexts globally.

Literature Review

Affordances of AV Input for L2 Learning

Although the present study focuses on language teachers' perceived use of AV input, it is important to contextualize these practices within their corresponding empirical findings in SLA (see Montero Perez, 2022, for a full review). The commonly-investigated AV input types are video only (video without on-screen text), subtitles (video with subtitles in L1), captions (video with subtitles in L2), and enhanced captions (with specific words or structures in captions typographically enhanced).

Given the widespread availability and use of L1 subtitles, several studies have examined their affordances for L2 learners. Findings show that L1 subtitles typically lead to higher comprehension scores than captions (e.g., Birulés-Muntané & Soto-Faraco, 2016; Pujadas & Muñoz, 2020) because learners can rely on their L1 knowledge to understand the input. However, they were found to be less effective than captions for vocabulary learning in both one-off and extended viewing conditions (e.g., Frumuselu et al., 2015; Peters et al., 2016). Additionally, L1 subtitles may negatively affect L2 speech parsing (Mitterer & McQueen, 2009).

Captioned video typically leads to higher vocabulary gains than uncaptioned video (e.g., Peters, 2019). It is hypothesized that captions promote learners' noticing of unknown words (Winke, et al., 2010) because they provide access to the written account of the aural input. Captioned video has also been found to be more effective than uncaptioned video for grammar learning (e.g., Muñoz et al., 2021) and speech decoding (Mitterer & McQueen, 2009).

Research on AV input has recently extended into the effects of captions with textual enhancement. Target words and constructions can be enhanced through underlining (Puimège et al., 2023) or color (Lee & Révész, 2020). Results from several studies suggest that enhanced captions are successful in directing L2 learners' attention to the specific target items in the captions (e.g., LaBrozzi, 2016; Puimège et al., 2023). However, research comparing enhanced captions to regular captions has produced mixed results. While there are indications that textually enhanced captions may facilitate learners' vocabulary learning gains (e.g., Cintrón-Valentín et al., 2019; Puimège et al., 2023), their effects on grammar learning are less clear-cut and might also be influenced by the specific grammatical structure under investigation (e.g., Cintrón-

Valentín et al., 2019; Lee & Révész, 2020). Recent research also suggests that textual enhancement might be beneficial only under certain conditions, such as when learners can fluently read captions and are able to effectively distribute their attention over different features of multimodal input (e.g., Puimège et al., 2023).

The majority of the studies on AV input focus on one-shot viewing designs (Montero Perez, 2022), which may be more ecologically valid for in-class learning. However, the type of viewing which typically occurs outside of class—extensive viewing—may offer specific benefits. Extensive viewing of AV input in the L1 or the L2 is a very common activity in everyday life, with 81% of survey respondents reporting watching TV (almost) every day (European Commission, 2020). With regard to AV input, out-of-class exposure to this type of input has been identified as an important predictor of vocabulary learning in particular (Peters et al., 2019). Out-of-school use of AV input has also been linked with better receptive L2 skills (Lindgren & Muñoz, 2013). In this respect, Webb (2015) argued that extensive viewing should initially take place in the classroom to prepare learners for regular out-of-class viewing. Language teachers thus play a crucial role not only by introducing viewing activities and preparing learners for them, but also by encouraging learners to watch videos (with or without on-screen text) in out-of-class contexts (Webb, 2015).

Teachers' Use of AV Input

Despite the ubiquity of AV input in our lives and its clear benefits for language learning, few studies have investigated teachers' reported use of AV input in L2 teaching. Mariotti's (2015) study, which included 45 language teachers in Europe, provides the first comprehensive look at this issue. In this study, the teachers completed a survey consisting of Likert-scale and open-ended questions that focused on their use of video with L1 subtitles, captions, and reversed subtitles (L1 audio + captions). Teachers in the study reported using AV materials to supplement regular class activities for motivational reasons, and they perceived media with on-screen text to be beneficial for oral comprehension. More specifically, L1 subtitles were perceived to be beneficial for beginners, while captions were believed to be more suitable for advanced learners. Teachers varied in their reported use of different types of on-screen text, depending on students' proficiency levels, teachers' teaching philosophy, accessibility of materials with on-screen text in specific L2s, time needed for creating materials, and access to teacher training on how to use subtitled media for language learning.

Kaderoğlu and Esquerré (2021) investigated the use of captioned videos in class by 63 EFL teachers in Turkey. The results of the survey indicated that teachers held overwhelmingly positive attitudes towards captions, with 98% of teachers reporting using captions "always," "frequently," or "sometimes" when playing videos in class. Teachers found captions beneficial for oral comprehension and vocabulary development, which directly aligns with research findings. Teachers' use was predominantly influenced by their own language learning experiences with captions. Regarding specific ways of using captions, most teachers reported turning captions on in class. They suggested their students do the same outside of class, and provided them with viewing strategies for captioned media. These studies suggest that learners' L2 and their proficiency level (Mariotti, 2015) as well as teachers' own learning experiences (Kaderoglu & Esquerré, 2021) influence teachers' use of AV input.

While few studies have explored factors influencing teachers' use of AV input, research in Computer-Assisted Language Learning (CALL) has investigated teachers' use of technology more broadly (see Table 1). Previous research reported that perceived comfort in using different CALL materials can influence teachers' use of them (e.g., Hegelheimer, 2006; Wang et al., 2004). Additionally, prior studies indicate that teachers' beliefs considerably influence their behaviors and decisions around implementing technology (Arnold & Ducate, 2015; Hubbard, 2008; Cárdenas-Claros & Oyanedel, 2016). Teaching experience can likewise be a contributing factor when teachers consider the implementation of technologies (e.g., Kartchava & Chung, 2015). Finally, the general availability of technology in and outside of class (i.e., technological resources) has been found to be an important barrier to CALL use (e.g., Sadaf et al., 2016).

Table 1

Factors That May Influence AV Input or Technology Use

Factor	Relevant Study(s)	Context of Study(s)
L2	Mariotti (2015)	AV input use by teachers
Learner proficiency level	Mariotti (2015)	AV input use by teachers
Comfort of use/perceived self-efficacy	Hegelheimer (2006) Wang et al. (2004)	CALL use by teachers
Beliefs (e.g., on importance for learning)	Arnold & Ducate (2015) Cárdenas-Claros & Oyanedel, (2016) Hubbard (2008)	CALL use by teachers
Teaching experience	Kartchava & Chung (2015)	CALL use by teachers
Availability of technology	Sadaf et al. (2016)	CALL use by teachers

Motivation for the Study and Research Questions

Extensive research has consistently identified important benefits of various AV input types for language learning. Few studies, however, have been conducted to understand whether teachers use AV input in their practices both in and out of the classroom, and what factors might affect their usage. This imbalance in research is potentially problematic; as Sato and Loewen (2022) put it, "without a bidirectional, collaborative, constructive dialogue, L2 researchers may conduct studies that are irrelevant to and out of touch with real-world teaching issues" (p. 509).

To gain a better understanding on whether there is a research-practice gap within the area of AV input, we set up an exploratory study in which the following research questions were addressed:

RQ 1: How often do teachers report utilizing AV input in the form of video only, subtitles, captions, and enhanced captions inside and outside the classroom?

RQ 2: Which factors influence how L2 teachers utilize different types of AV input in L2 teaching inside and outside the classroom?

Methodology

Participants

The participants were 193 teachers (135 female, 58 male) of second or foreign languages in 42 countries from all continents except Antarctica (see Supplemental Material). They taught a variety of proficiency levels (beginning through advanced) and learner age groups (children through adults). 70% of teachers taught English, and 30% taught other languages (i.e., Japanese, Chinese, French, German, Russian, Arabic, Italian, Spanish, Portuguese, Dutch, and isiZulu). Respondents had a wide range of teaching experiences (M = 14.32; SD = 9.75; range: 1-45 years). Most teachers reported easy access to technology in class (87%) and outside of class (85%), which suggests that our sample represents contexts where technologies are relatively easily accessible.

Data Collection Instrument

We used a two-part online survey that sought to (a) identify teachers' frequency of use of various multimodal input types and reasons behind such use, and (b) assess how well teachers' beliefs about the

perceived benefits for specific language aspects, such as vocabulary development, grammar, pronunciation, listening, speaking, and reading, correlate with research findings. The research questions for this paper are answered with data from part 1 of the survey, which included demographic questions and questions about teaching context (e.g., learner populations taught, availability of technology). This part of the survey also evaluated (a) teachers' frequency of use of various AV input types and (b) possible factors that may influence frequency of use (see Appendix A). The survey was made up of a combination of binary (e.g., yes/no), Likert-scale, and open-ended items. Survey design was informed by guidelines offered by Dörnyei and Csizér (2012), Dörnyei and Dewaele (2022), and Iwaniec (2019) and by prior research on teachers' use of AV input and technology (Table 1). There were two rounds of pilot testing with five experienced teachers (Dörnyei & Dewaele, 2022). To help narrow the research-practice divide (Sato & Loewen, 2022), we asked the teachers to comment on the relevance of our questions to their classroom teaching experiences. Teachers' feedback allowed us to improve the clarity of items and confirmed that open-ended questions allowed teachers to more fully express the reasons for using various types of AV input across a wide range of contexts.

At the end of the survey, we asked participants to provide their name and email address if they were interested in participating in a structured interview. The interview questions were on the broader topic of multimodal input and the factors that influence teachers' views on it. The analysis of the interview data is outside the scope of the current study but is reported on in Cárdenas-Claros et al. (2023).

Procedures

We administered the survey via Qualtrics. We utilized convenience sampling and recruited L2 teachers via mailing lists from professional teaching organizations, personal contacts, social media, and snowball sampling. To comply with ethical considerations, the call for participation prompted teachers to read the consent form and to ask any remaining questions before making the decision to participate. Participants took approximately 30 minutes to complete the survey.

Analysis

Quantitative Analysis

When answering research question 1 on the frequency of use of various AV input types, we provided descriptive statistics on the full data set ($N = 193$). When answering research question 2 on factors influencing the frequency of use of various AV input types inside and outside of class, we conducted logistic regressions ($n = 163$ due to missing data).

Frequency of use of the four types of AV input was originally measured on a 4-point Likert scale. However, because in several cases there were too few responses to certain Likert-scale options, we collapsed the 4-point Likert scale into the 2-point scale (see Table B_1 in Appendix B for scales of measurement of all variables). Such transformations were done to minimize the number of cells with expected count of less than 5, which is one of the assumptions of a logistic regression (Field et al., 2012). Although prior studies have similarly collapsed categories within variables when assumptions of statistical tests were not met (e.g., Diaubalick & Guijarro-Fuentes, 2019; Loewen, 2004), we recognize that doing so renders some imbalances within sub-groups and thus distorts the models. As such, we emphasize that the study is exploratory in nature; future confirmatory analyses would be desirable.

Given that the outcome variables are categorical, the data were modeled using logistic regressions. Binary logistic regression models illustrate not only the statistical significance of each predictor, but also how likely each predictor leads to one outcome versus another of the dependent variable (the odds ratio) (Field et al., 2012).

We conducted four logistic regressions, one for each AV input type as the dependent variable, with regard to frequency of use *in class*, and four logistic regressions with same variables for AV input use *outside of class*. The AV input types we examined were frequency of use of video only, subtitles, captions, and enhanced captions.

Possible predictors (see Table 1) were factors related to teachers' experience (years of teaching and comfort of use of specific AV input types), how important they considered each AV input type for learning, and contextual factors (technology availability and L2 being taught - English or other). Although student proficiency level can also be an important variable in teachers' AV input use (Mariotti, 2015), we did not include student proficiency levels as a predictor in the present study because it was a non-exclusive categorical factor (i.e., some teachers taught at all levels, others at only one or two of them).

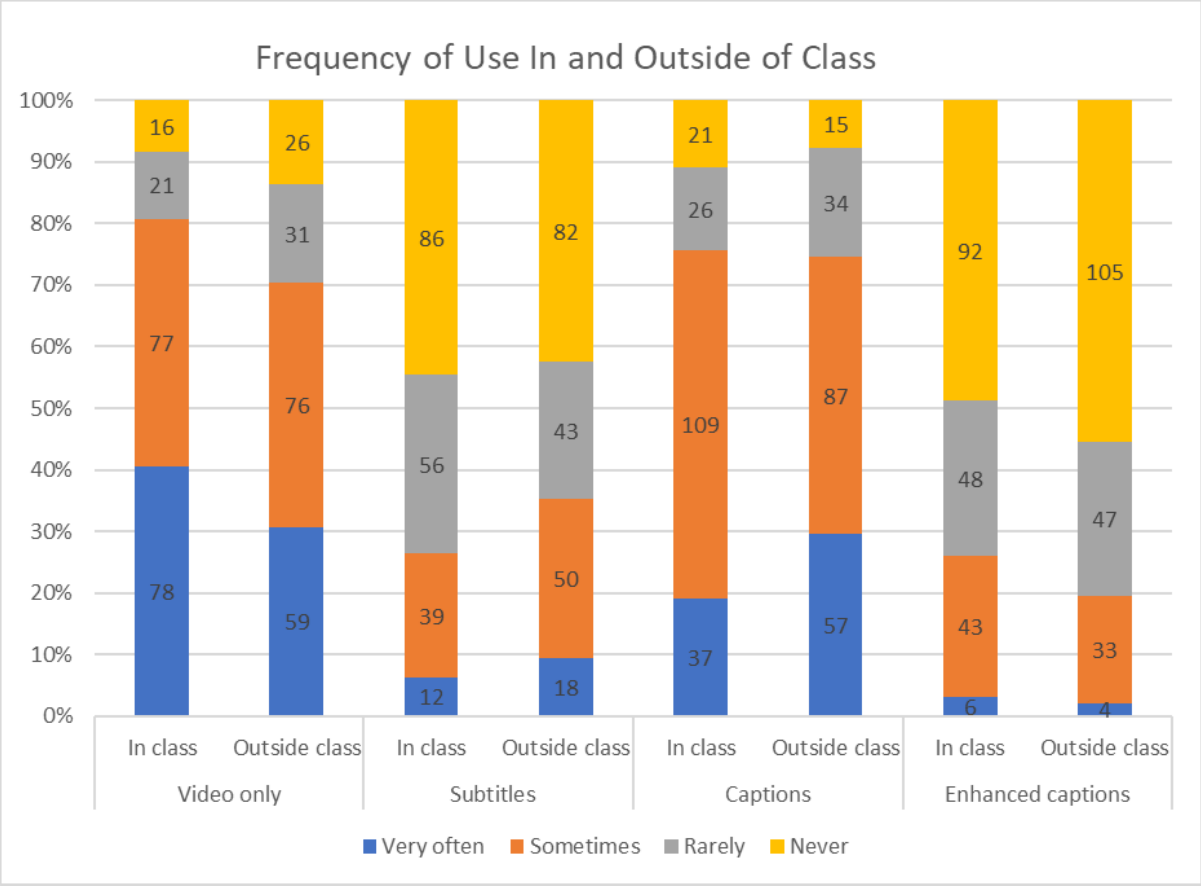
All assumptions for logistic regressions were met, and a forced entry method was used (see Appendix B).

Analysis of Open-Ended Questions

Using a bottom-up approach for qualitative content-analysis (Miles et al., 2014), answers to open-ended questions were analyzed to complement the quantitative analysis. To do so, one of the researchers read the open-ended answers in the survey and put together a coding protocol that identified reasons for the use/non-use of four types of AV input. Next, this coding scheme was tested through a trial coding conducted by two other authors of the study, incorporating any reasons that were initially missed into the revised coding scheme. Then, a second coder coded 10% of the data. Discrepancies were resolved via discussion; inter-coder reliability used the Miles et al. (2014) formula and reached .90. The final codes, together with the number of mentions, are provided in the Supplemental Material.

Figure 1

Frequency of Using Four AV Input Types Inside and Outside Class



Note. Numbers on bars indicate raw frequency counts.

Results

RQ1: How often do teachers use/assign different types of AV input inside and outside of class?

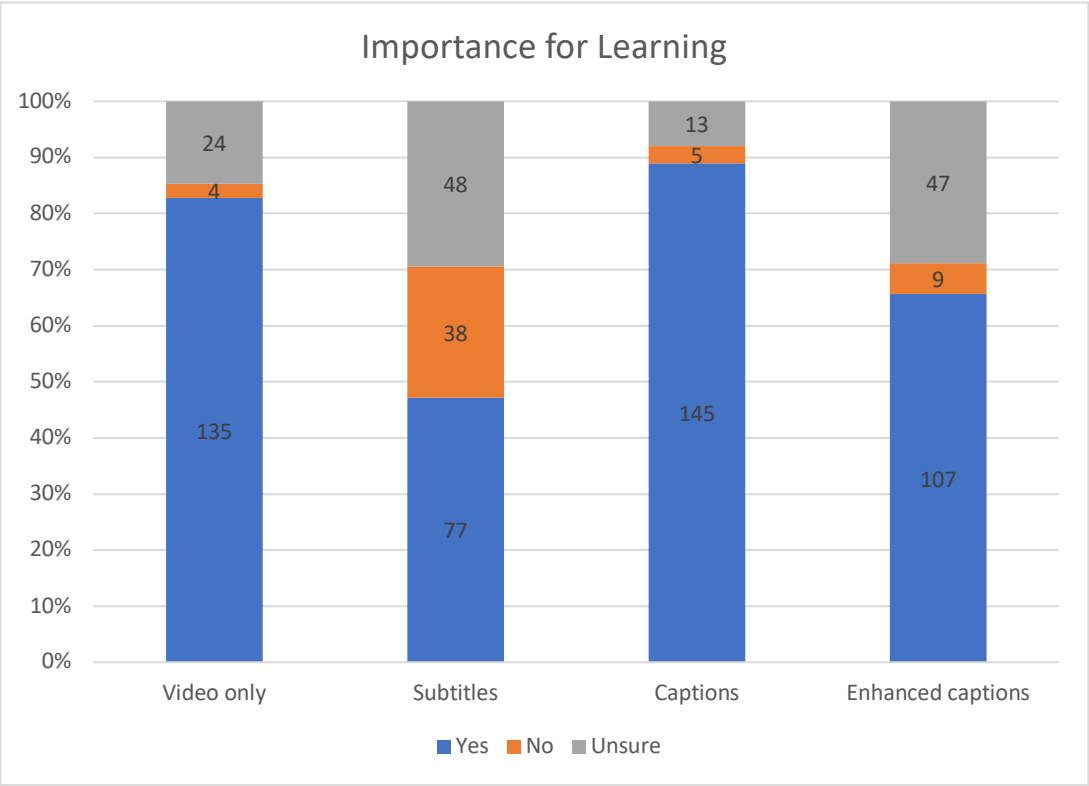
The results presented similar patterns for teachers' use or assignment of specific AV input types inside and outside the class, and for this reason both contexts for this research question are presented together. Figure 1 shows that the frequency of use of video only and captions followed one pattern, that is, rather frequent use both in and outside of class, while subtitles and enhanced options were, in general, rarely used both inside and outside of class.

RQ2: Which factors influence teachers' frequency of use of different types of AV input inside and outside of the classroom?

Before presenting the logistic regression results, we first provide descriptive statistics to contextualize these results (n = 163). Figure 2 shows how teachers perceived the importance of the four AV input types for learning. Like in the findings for RQ1, video only and captions followed a similar pattern in that most teachers considered them to be useful for learning, and, as shown in Figure 1 above, these were also the most frequently used AV input types. On the other hand, only about 50% of the teachers thought that subtitles are useful for learning, and subtitles were less frequently used than video only or captions. Thus, an overall trend that we observed is that teachers' beliefs about utility of each AV type are reflected in their corresponding frequency of use. Enhanced captions were the only input type that did not follow such a pattern: while about 65% of the teachers thought that enhanced captions are useful, 75% of the teachers never or rarely used them in class.

Figure 2

Teachers' Beliefs on Whether the Four AV Input Types Are Important for Learning

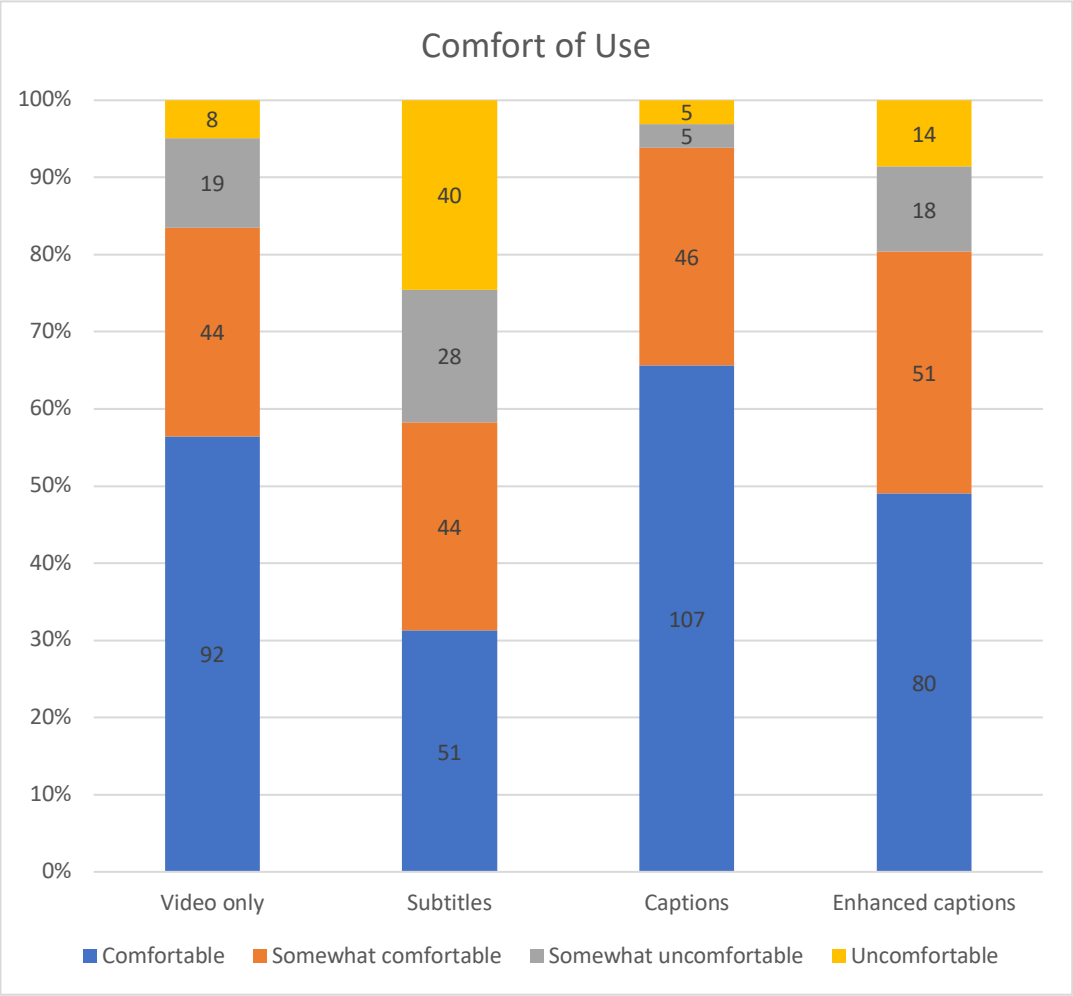


Note. Numbers on bars indicate raw frequency counts.

With regard to comfort of use of various AV input types, 80% (or more) of the teachers were comfortable or somewhat comfortable using video only, captions, and enhanced captions (see Figure 3). On the other hand, slightly less than 60% of the teachers were (somewhat) comfortable using subtitles in class.

Figure 3

Teachers' Comfort of Using the Four AV Input Types



Note. Numbers on bars indicate raw frequency counts.

Next, we present the results of logistic regressions and complement them with additional data from open-ended responses. In the interest of space, we discuss the most frequently mentioned topics from these responses and how they relate to quantitative findings; the complete results of the open-ended responses (i.e, all codes and number of mentions) are presented in the Supplemental Material.

Table 2 presents the results of the logistic regressions modeling which factors influenced frequency of use *in class*. Model fit was adequate (see Appendix B).

Table 2

Logistic Regressions: Frequency of Use in Class by AV Type

Dependent Variable & Predictors	**Predictor Statistics**					**Model Statistics**			
	B	*SE*	*p*	Odds Ratio[a]	95% CI for Odds Ratio	Overall % correctly predicted	x^2(df)	Cox & Snell R^2	Nagel kerke R^2
Video only						89%	59.89 (5)	.31	.50
Importance	2.56	.58	<.001*	12.98	4.19, 40.25				
Comfort using	2.44	.61	<.001*	11.44	3.49, 37.48				
L2	-1.22	.68	.074	.30	.08, 1.13				
Tech availability	-.85	.66	.197	.43	.12, 1.55				
Years of teaching	-.03	.03	.303	.97	.92, 1.03				
Intercept	-.68	.79	.393	.51					
Subtitles						79.8%	59.14 (5)	.30	.45
Importance	2.54	.61	<.001*	12.73	3.86, 42.05				
Comfort using	1.83	.70	.009*	6.22	1.58, 24.49				
L2	-.71	.48	.140	.49	.19, 1.26				
Tech availability	1.02	.65	.115	2.78	.78, 9.86				
Years of teaching	.03	.03	.220	1.03	.98, 1.08				
Intercept	-4.35	1.02	<.001*	.01					
Captions						83.4%	33.65 (5)	.19	.28
Importance	2.09	.57	<.001*	8.08	2.65, 24.65				
Comfort using	2.57	.86	.003*	13.12	2.44, 70.60				
L2	.69	.46	.132	1.98	.81, 4.83				
Tech availability	-.67	.60	.268	.51	.16, 1.67				
Years of teaching	-.03	.02	.221	.97	.93, 1.02				
Intercept	-2.88	1.15	.012*	.06					
Enhanced captions						74.8%	44.95 (5)	.24	.36
Importance	3.56	1.09	.001*	35.14	4.16, 296.87				
Comfort using	.98	.87	.258	2.66	.49, 14.54				

L2	1.17	.53	.028*	3.23	1.14,	9.16
Tech availability	1.52	.61	.013*	4.55	1.36,	14.98
Years of teaching	.02	.03	.345	1.02	.98,	1.08
Intercept	-6.40	1.54	<.001*	.01		

Note. For each model, $p \leq .001$. Video only = frequency of reported use of video only, etc. Importance = important for learning.

[a] Probability of outcome relative to the reference category 0. Reference categories: frequency of use of each type of AV input (lower use [i.e., rarely or never] = 0, higher use [i.e., very often or sometimes] = 1), class L2 (not English = 0, English = 1), Tech availability in class (yes = 0, no = 1), comfort of use of specific AV input type ([somewhat] uncomfortable = 0, [somewhat] comfortable = 1), beliefs about importance of specific AV input type (not important /unsure = 0, important = 1).

* $p < .05$.

As seen in Table 2, teachers' perceived importance of a particular AV input type as well as their comfort in using it were generally the only statistically significant variables predicting how frequently that AV input type would be used in class. In logistic regression, the odds ratios can be interpreted as the probability of an outcome occurring relative to the baseline probability of the reference category (see caption of Table 2 for list of reference categories). For example, if teachers considered video without on-screen text to be important for learning, they were 12.98 times more likely to report higher ("very often / sometimes") rather than lower (reference category = "rarely/never") use of it in their classrooms. Likewise, teachers were 11.44 times more likely to report higher use of videos without on-screen text if they indicated that they were (somewhat) comfortable (rather than the reference category "(somewhat) uncomfortable") with using them in class. Similar trends are observed for the use of subtitles and captions.

The results were somewhat different for enhanced captions. Here, the three significant predictors were perception of importance for learning, L2 being taught (English or other), and the availability of technology in class. The odds ratios indicate that teachers were about 35.14 times more likely to use enhanced captions if they considered this type of input to be important for learning. Teachers were 3.23 times more likely to use enhanced captions if they taught English rather than a different L2, and they were 4.55 times more likely to use enhanced captions if technology was not easily accessible. Unlike the other three input types, comfort of use was not a significant predictor for enhanced captions.

As for factors influencing frequency of use outside of class (RQ2), Table 3 shows that the overall percentage accuracy of each model was between 73% and 80.4% (slightly lower than the accuracy of the models for in-class use). The significant predictors for out-of-class use were similar to those for in-class use, with the perceptions of importance for learning and comfort of use being the most common predictors. However, the odds ratios were somewhat lower than for the previous set of models (e.g., 6.68 as compared to 12.98 for the importance of learning with video only). The L2 being taught was a significant predictor only for captions; the odds ratio indicates that teachers are 2.36 times more likely to report using captions more frequently if they teach English as opposed to another language. Technology availability outside of class was not a significant predictor for using any of the AV input types outside of class. Similarly to reported in-class use, comfort with using was a significant predictor for all input types except enhanced captions.

Table 3

Logistic Regressions: Frequency of Use Outside of Class by AV Type

Dependent Variable & Predictors	**Predictor Statistics**					**Model Statistics**			
	B	*SE*	*p*	Odds Ratio[a]	95% CI for Odds Ratio	Overall % correctly predicted	x^2 (df)	Cox & Snell R^2	Nagel kerke R^2
Video only						80.4%	47.93 (5)	.26	.37
Importance	1.90	.52	<.001*	6.68	2.40, 18.59				
Comfort using	1.90	.54	<.001*	6.69	2.33, 19.25				
L2	-.02	.48	.975	.99	.38, 2.53				
Tech availability	-.94	.54	.081	.39	.14, 1.13				
Years of teaching	.02	.02	.433	1.02	.97, 1.07				
Intercept	-2.16	.75	.004*	.12					
Subtitles						73%	45.86 (5)	.25	.34
Importance	1.98	.44	<.001*	7.22	3.05, 17.09				
Comfort using	1.05	.47	.026*	2.85	1.13, 7.15				
L2	.06	.43	.899	1.06	.45, 2.46				
Tech availability	-.17	.53	.746	.84	.30, 2.39				
Years of teaching	.02	.02	.391	1.02	.98, 1.06				
Intercept	-2.81	.70	<.001*	.06					
Captions						77.9%	20.20 (5)	.12	.17
Importance	1.26	.55	.021*	3.53	1.21, 10.28				
Comfort using	1.93	.75	.010*	6.85	1.57, 29.84				
L2	.86	.42	.039*	2.36	1.05, 5.34				
Tech availability	-.90	.50	.070	.41	.15, 1.08				
Years of teaching	.01	.02	.56	1.01	.97, 1.06				
Intercept	-2.41	1.04	.02*	.09					
Enhanced captions						80.4%	25.19 (5)	.14	.23
Importance	3.00	1.07	.005*	20.15	2.45, 165.55				
Comfort using	.37	.85	.661	1.45	.28, 7.62				

L2	.91	.55	.098	2.48	.85,	7.24
Tech availability	-.18	.58	.753	.83	.27,	2.62
Years of teaching	.03	.03	.320	1.03	.98,	1.08
Intercept	-5.27	1.37	<.001*	.01		

Note. For each model, $p \leq .001$. Video only = frequency of reported use of video only, etc. Importance = important for learning.

[a] Probability of outcome relative to the reference category 0. Reference categories: frequency of use of each type of AV input (lower use [i.e., rarely or never] = 0, higher use [i.e., very often or sometimes] = 1), class L2 (not English = 0, English = 1), tech availability in class (yes = 0, no = 1), comfort of use of specific AV input type ([somewhat] uncomfortable = 0, [somewhat] comfortable = 1), beliefs about importance of specific AV input type (not important /unsure = 0, important = 1).

* $p < .05$.

Although the models had a good fit, as indicated by the overall percentage correctly predicted in Tables 2 and 3, some models contained more influential cases than others[1], which suggests that the results need to be interpreted with caution. Additionally, confidence intervals (CI) for each odds ratio varied in width (and thus precision of the associated estimate). We observed particularly wide CIs for the variables predicting use of enhanced captions both in and outside of class, which suggests there might be other relevant variables affecting teachers' reported use of enhanced captions.

To shed more light on other potential predictors not captured by the quantitative analyses, we also presented teachers with open-ended questions. As shown by the results of the statistical analyses, factors influencing teachers' use of different AV types did not seem to be context-dependent (in class vs. out-of-class). This was also confirmed by teachers' open-ended comments: teachers do not use/assign these AV types in drastically different ways in versus outside of class. Teachers did, however, report that they might select longer videos when used in out-of-class contexts whereas shorter videos are typically preferred in the classroom (n = 18, "Learners are suggested to work with L1 subtitles at home, when watching longer videos, documentaries or movies," see Supplemental Material).

Teachers' comments also highlight specific uses of different types of AV input. For instance, teachers report using captions to engage and motivate students (n = 37). Captions were also found to be helpful to activate background knowledge or new content (e.g., trigger conversation/writing topics) (n = 10), as well as establish sound-text connections and focus learners' attention on form (vocabulary, spelling, grammar) (n = 19). However, the drawbacks are that specific resources with captions may not be available for given proficiency levels (n = 3) and some teachers fear that students may rely on captions (as opposed to audio) too much when processing input (n = 4).

Regarding video only, teachers reported finding it useful to help students assess how ready they are for real-life situations (n = 8). They find it helpful to challenge students' comprehension of the input without any support (n = 13). Some also reported using videos because they are readily available and easy to use (n = 4). However, teachers also reported that students can get frustrated if they do not understand the input (n = 5).

Although less commonly used, L1 subtitles were nevertheless found helpful for beginning learners to navigate authentic materials or aid their confidence. They are, however, not practical when multiple L1s are represented in the classroom (n = 11), they can be distracting or learners might use them as a crutch when processing L2 input (n = 12). Finally, some teachers used enhanced captions occasionally to draw learners' attention to pronunciation, vocabulary, and grammar (n = 9) but most teachers reported that they rarely or never used them because they did not know about their existence or did not have time to find or create them (n = 59).

As such, the open-ended responses suggest that there may be additional factors influencing teachers' use of various AV input types. Teachers stated that they use subtitles with lower-level learners (n = 19), a factor that we suspected but could not include in regression models because teachers reported teaching learners at a variety of proficiency levels. Overwhelmingly, teachers reported that enhanced captions are not easy to find; this further clarifies the quantitative analyses which revealed a mismatch between teachers' beliefs on enhanced captions' importance versus their reported frequency of use. Additionally, teachers' comments suggest that there are situational variables (e.g. many L1s in class) and pedagogical variables (e.g., varied phases of instruction, such as brainstorming a topic vs. focusing on form) that may further complicate the identification of factors affecting frequency of use of AV input if measured quantitatively. As one teacher summarized: "It all depends on the activity I'm doing because I believe that we can use the same apps, for example, for different classes, different levels, it all depends on your aim" (ID 118).

Discussion

While numerous studies have investigated the linguistic benefits of different types of AV input, little is known about L2 teachers' use of AV input (RQ1) as well as about the factors that influence teachers' use of them (RQ2). To address these gaps, we set up an exploratory study in which survey data were collected from L2 teachers across the globe about their use of four types of AV input. Our first research question asked how frequently teachers use different types of AV input inside the classroom or recommend its use in out-of-class contexts. Findings show that teachers report using video only and captioned video most often, whereas most teachers indicate they rarely or never use subtitled video or video with enhanced captions. These patterns are found for both in-class and out-of-class use.

Regarding the use of captioned video, approximately 75% of the teachers in our study reported using this type of AV input very often or sometimes (see Figure 1). This rate is comparable to findings in previous studies. For example, approximately 90% teachers in Turkey reported using captioned video at least sometimes in the classroom or encouraged students to use it outside of class (Kaderoglu & Esquerré, 2021). Both results appear to be higher than in Mariotti's study (2015) on the use of AV input with on-screen text where 64% of the teachers used captions or subtitles a lot or sometimes, though such comparisons are difficult to make because in Mariotti's study teachers' frequency of use of video with on-screen text was not separated by captions vs. subtitles. Our study expands on previous research by showing that findings on captioned videos and video only show similar patterns when it comes to the use of these types of input outside of class. More particularly, results indicate that teachers tend to recommend the use of these types of AV input outside of class or use in the context of assignments. These findings are encouraging for at least two reasons. First, video and captioned video are types of input that are easily accessible thanks to the Internet and streaming services. There are also numerous freely available tools which provide the possibility to activate captions or to manually (or automatically) add captions. Second, these findings indicate that the basic principles of extensive viewing programs (see Webb, 2015), which develop in-class viewing strategies to scaffold independent out-of-class use, seem to be operationalized by the L2 teachers who participated in this study. This may subsequently aid L2 learners to engage in more extensive AV input exposure which may in turn positively influence L2 learning gains (e.g., Lindgren & Muñoz, 2013).

With regard to subtitled video or video with enhanced captions, it was found that these are less frequently used than the previously discussed types of AV input. However, just like captioned video and video only, their use does not seem to differ inside and outside of class. In addition, the widespread availability of subtitled video through, for instance, streaming services, does not seem to influence teachers' use. This finding was surprising given that out-of-class subtitled AV input exposure is associated with better comprehension (e.g. Pujadas & Muñoz, 2020) even though, on the other hand, L1 subtitles may be detrimental for speech perception (cf. Mitterer & McQueen, 2009). Even though studies have addressed the role of enhanced captions (see Literature Review), teachers do not tend to use enhanced captions. This

is not surprising as these types of materials are typically less easily accessible or involve timely manipulation of captions.

To understand the factors that influence L2 teachers' use of AV input (RQ2) in both contexts (in-class use and out-of-class assignments), a regression analysis per type of AV input was conducted. Below, we will discuss the role of each predictor that was included in the model: teachers' perceived importance, comfort in using, technology availability in the context of L2 teaching, the L2, and the length of teachers' experience teaching the L2.

With regard to the role of *teachers' perceived importance*, it can be seen that this predictor positively and significantly influenced L2 teachers' use both inside and outside of class for all four types of AV input. This indicates that when people find the type of input important for L2 learning, they are more likely to use it in their teaching practice. Teacher beliefs thus seem to play a consistent and important role in their use of AV input. This is in line with previous research on teachers' beliefs which were found to considerably influence teachers' behaviors and decisions in class in general (Borg, 2018) and for technology implementation in particular (Arnold & Ducate, 2015; Hubbard, 2008).

The predictor *comfort using the technology* is the second main factor that plays an important role in explaining teachers' use, even though its effect tends to be smaller as compared to the role of teachers' perceived importance. This finding aligns with prior research indicating that perceived self-efficacy influences teachers' willingness to adopt various technologies (e.g., Hegelheimer, 2006; Wang et al., 2004). Notably, this factor was only found to influence L2 teachers' use of video, subtitled video and captioned video, which are the types of AV input that are most easily accessible. This finding therefore seems to indicate that teachers are willing to use AV input if they do not face specific technological challenges related to the retrieval or creation of these materials. The predictor did not, however, significantly influence teachers' use of enhanced captions. This may indicate that teachers are not familiar with this type of input and therefore have no specific idea about how comfortable they are using it. This is in line with teachers' comments on the survey which indicate that while teachers think enhanced captions could be useful, many are not sure how to access them or do not have time to modify captions themselves. The accessibility of this particular input type, while not measured quantitatively in our survey, might thus be a more important predictor of teachers' use of it.

As for the remaining predictors (L2 and technology availability), findings are less clear-cut. With regard to the role of the L2 (English as L2 vs. Other language), it was found that this factor influenced L2 teachers' use of enhanced captions but only for in-class use. More particularly, they were 3.23 times more likely to use enhanced captions if they taught English as opposed to a different L2. This might be because materials for English as L2 with this particular type of on-screen text are easier to find than for other L2s (cf. Mariotti, 2015). For use in out-of-class contexts, findings indicate that teachers are 2.36 times more likely to report higher use of captions if teaching English as compared to other languages. Finally, the availability of technology did not surface as a significant predictor in our analyses (except for enhanced captions in the classroom). This might indicate that teachers who completed the survey have access to the basic technology needed for AV input (e.g., computer in combination with a beamer and a screen), as the descriptive statistics show. In addition, this might suggest that AV input is relatively easy to implement without too many technological challenges which might further promote its use in the classroom but also in out-of-class contexts. We were surprised by the statistically significant result that teachers were 4.55 times more likely to use enhanced captions if technology was not easily accessible; however, this result might be influenced by the findings from open-ended responses in which it was reported that enhanced captions are not easily available ($n = 59$) – a factor we did not specifically measure quantitatively. The results for enhanced captions need to be interpreted with caution because we observed large CIs for odds ratios both in and outside of class. However, taken together, results from this study do seem to point to a potential difficulty to implement this type of captioning. Even though it presents an interesting case to study attention allocation and learning, our findings suggest that expectations with regard to pedagogical implications might be less straightforward.

Finally, the fifth factor that we included in all models, but which never surfaced as a predictor was teachers' years of experience teaching the L2. Similar to technology availability, this might suggest that the AV input types we examined are so widespread (cf. Kaderoğlu & Esquerré, 2021) that the length of teaching experience might not significantly impact frequency of their use.

In sum, our findings indicate that teacher-related factors including teachers' perceived importance and their comfort using AV materials played an important role in how frequently they reported using them, as was also found in previous CALL research (see Table 1). As such, our study further illustrates the crucial role that teachers play in material selection and learners' input exposure to AV input, both inside and outside of the classroom.

Concluding Remarks

A number of limitations need to be acknowledged. First, frequency of use was not measured in quantitative terms, e.g. "rare" for one teacher might mean 1 hour per week, while for another it might mean once per semester. Also, in instructional contexts teachers may use specific (repeated) viewing sequences (e.g., watch without captions, then captions, then again no captions), as mentioned by some teachers, which makes it difficult to objectively quantify frequency of use. Second, there may be other factors, not included in the regression analyses, that might influence the frequency of use of various AV input types. Data from our open-ended questions suggest, for instance, that students' proficiency levels, availability of relevant high-quality media with specific AV input types, especially enhanced captions, and teaching contexts (e.g., variety of L1s in class) may be additional contributing factors. These findings are also in line with the results of Cárdenas-Claros et al. (2023) on factors influencing teachers' views of multimodal input in that a variety of meso, macro, and micro factors (e.g., reliability of technology, Covid-19 related contexts of teaching) can affect teachers' perceptions and use of multimodal input. Finally, although participants spoke a variety of L1s, both our call for participation as well as our instrument was in English. This may have limited the scope of respondents (and their responses) to those who feel comfortable expressing themselves in English. In addition, while our data represents a varied group of participants, most of them taught English as L2. As prior research indicates (Mariotti, 2015), there may be more relevant AV materials in more commonly studied languages.

This study has a number of implications. First, it shows alignment between the types of input that had been studied in previous research and reported frequency of use by L2 teachers. However, it is important to continue to invest in transfer of research insights to teaching practice since teachers' perceived importance of the AV input has been identified as a major predictor of their use in class and for out-of-class assignments. Second, if we want to make sure that other types of AV input, such as videos with enhanced captions, find their way to the classroom, material designers or video-based L2 learning platforms that provide L2 teachers with these materials will be needed. Despite their potential for L2 learning, these materials might not be implementable because teachers feel less comfortable using them or simply do not have access to them.

Research into the role of AV input has grown exponentially over the past decade. However, to the best of our knowledge, this study is the first to investigate L2 teachers' perceptions of the use of the four frequently researched types of AV input both inside and outside the classroom. Overall, findings suggest that captioned video and video only are the two types of AV input that are reported to be most frequently used by L2 teachers in the classroom or assigned for out-of-class use. Teachers' perceived importance and comfort using specific AV input types were identified as the two most critical factors explaining teachers' reported use. These are important aspects to consider both in future research and in materials development if we want to further bridge the practice-research gap in AV use. Avenues for future research could result in confirmatory analyses building upon the factors identified in the present study. As research on specific benefits of different types of AV input types continues to evolve (e.g., glossed captions, cf. Teng, 2022), the research-practice dialogue between researchers and teachers may provide additional nuanced insights.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Notes

1. In the inside-class analyses, residuals >3 were found for video (7), captions (5) and enhanced captions (1). In the outside-class analyses, residuals >3 were found in the video (2), subtitles (2) and enhanced captions (1) analyses.

References

Al-Hoorie, A. H., Oga-Baldwin, W. L. Q., Hiver, P., & Vitta, J. P. (2022). Self-determination mini-theories in second language learning: A systematic review of three decades of research. *Language Teaching Research*. https://doi.org/10.1177/13621688221102686

Arnold, N., & Ducate, L. (2015). Guest editor commentary. *Language Learning & Technology, 19*(1), 1–9. http://dx.doi.org/10125/44394

Birulés-Muntané, J., & Soto-Faraco, S. (2016). Watching subtitled films can help learning foreign languages. *PLoS ONE, 11*(6), 1–10. https://doi.org/10.1371/journal.pone.0158409

Borg, S. (2018). Teachers' beliefs and classroom practices. In P. Garrett & J. Cots (Eds.), *The Routledge handbook of language awareness* (pp. 75–91). Routledge.

Cárdenas-Claros, M., & Oyanedel, M. (2016). Teachers' implicit theories and use of ICTs in the language classroom. *Technology, Pedagogy and Education*, *25*(2), 207–225. https://doi.org/10.1080/1475939X.2014.988745

Cárdenas-Claros, M. S., Sydorenko, T., Huntley, E., & Montero Perez, M. (2023). Teachers' voices on multimodal input for second or foreign language learning. *Language Teaching Research*. https://doi.org/10.1177/13621688231216044

Chatterjee, S., & Hadi, A. S. (2012). *Regression analysis by example* (5th ed.). Wiley.

Cintrón-Valentín, M., García-Amaya, L., & Ellis, N. C. (2019). Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal*, *47*(4), 439–459. https://doi.org/10.1080/09571736.2019.1615978

Diaubalick, T., & Guijarro-Fuentes, P. (2019). The strength of L1 effects on tense and aspect: How German learners of L2 Spanish deal with acquisitional problems. *Language Acquisition*, *26*(3), 282–301. https://doi.org/10.1080/10489223.2018.1554663

Dörnyei, Z., & Csizér, K. (2011). How to design and analyze surveys in second language acquisition research. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 74–94). Blackwell. https://doi.org/10.1002/9781444347340

Dörnyei, Z., & Dewaele, J. M. (2022). *Questionnaires in second language research: Construction, administration, and processing* (3rd ed.). Routledge. https://doi.org/10.4324/9781003331926

European Commission. (2020). Media use in the European Union: Report. https://data.europa.eu/doi/10.2775/80086

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.

Frumuselu, A. D., De Maeyer, S., Donche, V., & Colon Plana, M. D. M. G. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education*, *32*, 107–117. https://doi.org/10.1016/j.linged.2015.10.001

Godwin-Jones, R. (2015). Contributing, creating, curating: Digital literacies for language learners. *Language Learning & Technology 19*(3), 8–20. Retrieved from http://llt.msu.edu/issues/october2015/emerging.pdf

Guichon, N., & Hauck, M. (2011). Editorial: Teacher education research in CALL and CMC: More in demand than ever. *ReCALL, 23*, 187–199. http://dx.doi.org/10.1017/S0958344011000139

Hegelheimer, V. (2006). When the technology course is required. In P. Hubbard & M. Levy (Eds.), *Teacher education in CALL* (pp. 117–133). John Benjamins. https://doi.org/10.1075/lllt.14

Hubbard, P. (2008). CALL and the future of language teacher education. *CALICO Journal, 25*(2), 175–188. https://www.jstor.org/stable/10.2307/calicojournal.25.2.175

Iwaniec, J. (2019). Questionnaires: Implications for effective implementation. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 324–335). Routledge.

Kaderoğlu, K., & Esquerré, F. R. (2021). Captions in L2 learning from language teachers' perspective: What do teachers believe and do? *Computer Assisted Language Learning Electronic Journal (CALL-EJ), 22*(3), 86–106. https://callej.org/index.php/journal/article/view/354

Kartchava, E., & Chung, S. (2015). Pre-service and in-service English as a second language teachers' beliefs about the use of digital technology in the classroom. *Studies in English Language Teaching, 3*(4), 355–383. https://doi.org/10.22158/selt.v3n4p355

LaBrozzi, R. M. (2016). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research*, *20*(1), 75–91. https://doi.org/10.1177/1362168814561903

Language Reactor. (n.d.). Language Reactor extension. https://www.languagereactor.com

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

Lin, P. (2022). Developing an intelligent tool for computer-assisted formulaic language learning from YouTube videos. *ReCALL*, *34*(2), 185–200. https://doi.org/10.1017/S0958344021000252

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1), 105–129. https://doi.org/10.1080/14790718.2012.679275

Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, *54*(1), 153–188. https://doi.org/10.1111/j.1467-9922.2004.00251.x

Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *The Modern Language Journal*, *90*(4), 536–556. https://doi.org/10.1111/j.1540-4781.2006.00465.x

Mariotti, C. (2015). A Survey on stakeholders' perceptions of subtitles as a means to promote foreign language learning. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 83–103). Peter Lang. https://doi.org/10.3726/978-3-0351-0719-7

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE*, *4*(11), 1–5. https://doi.org/10.1371/journal.pone.0007785

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

Muñoz, C., Pujadas, G., & Pattemore, A. (2021). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Language Research, 39*(1), 13–37. https://doi.org/10.1177/02676583211015797

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audio-visual input. *TESOL Quarterly, 53*(4), 1008–1032. https://doi.org/10.1002/tesq.531

Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System, 63,* 134–148. https://doi.org/10.1016/j.system.2016.10.002

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/S0272263117000407

Peters, E., Noreillie, A-S., Heylen, K., Bulté, B., & Desmet, P. (2019). The impact of instruction and out-of-school exposure to foreign language input on learners' vocabulary knowledge in two languages. *Language Learning, 69*(3), 747–782. https://doi.org/10.1111/lang.12351

Puimège, E., Montero Perez, M., & Peters, E. (2023). Promoting L2 acquisition of multiword units through textually enhanced audiovisual input: An eye-tracking study. *Second Language Research*, *39*(2), 471–492. https://doi.org/10.1177/02676583211049741

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition, 42*(3), 525–549. https://doi.org/10.1017/s027226311900055x

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

Rodgers, M. P. H., & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics, 171*(2), 191–220. https://doi.org/10.1075/itl.18034.rod

Sadaf, A., Newby, T. J., & Ertmer, P. A. (2016). An investigation of the factors that influence preservice teachers' intentions and integration of Web 2.0 tools. *Educational Technology Research and Development*, *64*, 37–64. https://doi.org/10.1007/s11423-015-9410-9

Sato, M., & Loewen, S. (2022). The research-practice dialogue in second language learning and teaching: Past, present, and future. *The Modern Language Journal*, *106*(3), 509–527. https://doi-org.proxy.lib.pdx.edu/10.1111/modl.12791

Teng, F. (2022). Vocabulary learning through videos: Captions, advance-organizer strategy, and their combination. *Computer Assisted Language Learning, 35*(3), 518–550. https://doi.org/10.1080/09588221.2020.1720253

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

Wang, L., Ertmer, P. A., & Newby, T. J. (2004). Increasing preservice teachers' self-efficacy beliefs for technology integration. *Journal of Research on Technology in Education, 36*(3), 231–250. http://dx.doi.org/10.1080/15391523.2004.10782414

Webb, S. (2014). Extensive viewing: Language learning through watching television. In D. Nunan & J. C. Richards (Eds.), *Language learning beyond the classroom* (pp. 159–168). Routledge. https://doi.org/10.4324/9781315883472

Wesely, P. M., Vyn, R., & Neubauer, D. (2024). Teacher beliefs about instructional approaches: Interrogating the notion of teaching methods. *Language Teaching Research*, *28*(1), 29–51. https://doi.org/10.1177/1362168821992180

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 1*4(1), 65–86. http://dx.doi.org/10125/44203

Wisniewska, N., & Mora, J. C. (2020). Can captioned video benefit second language pronunciation? *Studies in Second Language Acquisition, 42*(3), 599–624. https://doi.org/10.1017/S0272263120000029

Appendix A: Survey Questions That Were Analyzed in the Present Paper

Note. While in the paper we use the terms *subtitles* and *captions* because they are more widely used in research, in the survey we used the terms *L1 subtitles* and *L2 subtitles* (respectively) because they might be more familiar to teachers.

Background Section

Gender

- Male
- Female
- Other

Age _____

Language Teaching Section: Experience, Context, and Technology

Which second/foreign language are you teaching? (If you taught more than one, which language have you taught the longest?) _____

Years of teaching this language? _____

Learners' proficiency level(s): select all that apply.

- Beginner
- Intermediate
- Advanced

Learner age group(s): select all that apply.

- Children
- Teenagers
- Adults

Name of the country(s) where you taught this language. _____

Technology

Several questions below ask about technology in general.

By technology, we mean anything that you yourself would consider technology (it could include email, document camera, the Internet, etc.).

Is technology easily available or accessible in the classrooms where you normally teach?

__ Yes __ No

Is technology easily available or accessible to your learners outside of class time (i.e., for autonomous learning)? __ Yes __ No

The next set of questions will ask about videos without subtitles.

(Note: Analogous sets of questions were designed for other AV input types).

(An image to illustrate each AV type was provided for context and clarity).

I believe videos without L2 subtitles are important in language learning.

- Yes
- No
- Unsure

Please explain why. *(The answers to this open-ended question were not included in the results because the answers were similar to those on the open-ended question below)*.

_____

	Never	Rarely	Sometimes	Very often
I use videos (without subtitles) during class				
I assign this technology as homework or suggest students use it outside of class				

Please explain how or why you use this technology in and/or out of class. We appreciate as many details as you are able to provide. _____

Please rate how comfortable you are using the following multimedia materials in the classroom.

	Uncomfortable	Somewhat uncomfortable	Somewhat comfortable	Comfortable
Videos with L2 subtitles (i.e., text in the same language as the video)				
Videos with L1 subtitles (i.e., text in L1, video in L2)				
Videos with modified/enhanced L2 subtitles (e.g., certain words highlighted)				
Videos without subtitles				

Appendix B: Logistic Regression: Assumptions and Methods

Table B_1

Categorical Variables (DVs and IVs) for Regression Analyses

Dependent (DV) and independent (IV) variables	Exact wording of the survey question	Transformations for non-dichotomous variables
DV: Frequency of use of each AV input type in class	I use videos (without subtitles) during class Never Rarely Sometimes Very often	Never & rarely → non-use Sometimes & very often → use
DV: Frequency of assignment of each AV input type outside of class	I assign this technology as homework or suggest students use it outside of class. Never Rarely Sometimes Very often	Never & rarely → non-use Sometimes & very often → use
IV: Technology availability in class	Is technology easily available or accessible in the classrooms where you normally teach? Yes No	
IV: Technology availability outside of class	Is technology easily available or accessible to your learners outside of class time (i.e., for autonomous learning)? Yes No	
IV: L2	Which second/foreign language are you teaching? (If you have taught more than one, which language have you taught the longest?) _____	English → English All other → Not English
IV: Years of teaching L2	Years of teaching this language _____	
IV: Comfort of using each AV input type	Please rate how comfortable you are using the following multimedia materials in the	Comfortable & somewhat comfortable →

	classroom. Videos without subtitles Uncomfortable Somewhat uncomfortable Somewhat comfortable Comfortable	comfortable Somewhat uncomfortable & uncomfortable → uncomfortable
IV: Belief about whether each AV input type is important for learning	I believe videos without subtitles are important in language learning. Yes No Unsure	Yes → Yes No & Unsure → No/unsure

Note. This table includes questions about video only. The questions for other input types (videos with subtitles, captions, and enhanced captions) were worded in an analogous manner.

We assumed a binomial distribution because a) the number of observations is fixed; b) each observation is independent; c) each observation represents one of two possible outcomes of "success" or "failure" (in our case, the collapsed categories of use or non-use of each AV input type; d) the probability of "success" is the same for each outcome.

Other assumptions of binary logistic regressions were also checked. With regard to multicollinearity, we looked at correlation matrices – there were no correlations above .41, thus indicating no multicollinearity (Loewen & Philp, 2006). Teachers' age, however, was removed as a predictor variable because it correlated highly with years of teaching. Additionally, average VIF was no greater than 1.28 (i.e, close to 1, as suggested by Field et al., 2012) and the tolerance statistic (1/VIF) was always above .71 (i.e., more than 0.2, as suggested by Field et al., 2012).

A forced entry method was used where all predictor variables were entered into the model at the same time. This was done for two reasons: 1) the predictor (or independent) variables were substantiated by prior (albeit limited) research and 2) stepwise methods have been found to be highly problematic (e.g., Field et al., 2012).

Because variance does not occur with binary variables, logistic regression model fit is reported using the overall percentage accuracy rather than the amount of variance accounted for (as is done with linear regression). A higher proportion of correct classifications indicates a better performing model (Chatterjee & Hadi, 2012). For our models, the overall percentage accuracy was between 73% and 89%. Additionally, the model fit was assessed via Hosmer-Lemeshow test, which was adequate ($p > .05$) for all models.

About the Authors

Tetyana Sydorenko is an associate professor in the Department of Applied Linguistics at Portland State University, Portland, Oregon, USA. Her research interests include computer-assisted language learning (multimodal input, spoken dialogue systems, chatbots) and the learning of L2 vocabulary and pragmatics. Tetyana Sydorenko is the corresponding author.

E-mail: tsydorenko@pdx.edu

ORCiD: https://orcid.org/0000-0002-1173-657X

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.

E-mail: monica.cardenas@pucv.cl

ORCiD: https://orcid.org/0000-0003-4031-5392

Elizabeth (Lizz) Huntley is an instructor of Arabic and Linguistics at the University of Colorado Denver. Her research uses psycholinguistic approaches to explore the acquisition of language variation. She is the managing editor for the journal Studies in Second Language Acquisition.

E-mail: elizabeth.huntley@ucdenver.edu

ORCiD: https://orcid.org/0000-0002-3861-656X

Maribel Montero Perez is an assistant professor in the Department of Linguistics at Ghent University (Belgium). She teaches courses on language teaching methodology and research methods for second language acquisition. Her research primarily focuses on second and foreign language learning and teaching, with a particular emphasis on the role of various input types in formal and informal contexts.

E-mail: maribel.monteroperez@ugent.be

ORCiD: https://orcid.org/0000-0002-0868-588X