

ARTICLE



Promoting grammatical development through multimodal digital recasts in videoconferencing tasks

Yeonwoo Jung, University College London

Andrea Révész, University College London

Abstract

This study investigated the effectiveness of multimodal recasts (the simultaneous presentation of oral and written recasts) relative to oral recasts and written recasts for L2 grammatical development in the context of videoconferencing. We employed a pretest–posttest–delayed posttest design, with 60 Korean learners of English randomly assigned to three experimental groups and a control group: an oral recast group, a written recast group, a multimodal recast group, and a no recast group. Each group completed a series of communicative videoconferencing treatment activities, during which they received recasts according to their respective feedback conditions. The target L2 construction was English wh-question formation. Participants' gains were assessed using oral, written, and spontaneous oral production tests. We found that the presence of recasts, regardless of modality, facilitated participants' development in wh-question formation, as reflected in their performance across the three outcome measures. Multimodal recasts, however, proved more effective in promoting gains than their oral and written counterparts.

Keywords: Corrective Feedback Modality, Recasts, Multimodal Feedback, Synchronous Computer-Mediated Communication (SCMC)

Language(s) Learned in This Study: English

APA Citation: Jung, Y., & Révész, A. (2024). Promoting grammatical development through multimodal digital recasts in videoconferencing tasks. *Language Learning & Technology*, 28(1), 1–29.

<https://hdl.handle.net/10125/73601>

Introduction

In instructed second language acquisition (SLA) research, considerable effort has been devoted to investigating the effects of corrective feedback (CF) on second language (L2) development. Among various feedback techniques, recasts have received the most attention (Li & Vuono, 2019). Recasts constitute a partial or full reformulation of an erroneous learner output into a target-like form while retaining its central meaning (Long, 2007), as in the example below from the current study:

Learner: When the incident occur?

Researcher: When did the incident occur?

The beneficial role of recasts in L2 acquisition is well-established in both face-to-face (e.g., Li, 2010) and synchronous computer-mediated communication (SCMC) contexts (Ziegler & Mackey, 2017). Most SCMC studies, however, have examined the effects of recasts in text-based SCMC (e.g., Sauro, 2009; Yilmaz, 2012; Yilmaz & Yuksel, 2011; Kourtali, 2022), with limited exploration of video-based recasts (e.g., Monteiro, 2014; Rassaei, 2017, 2019). Moreover, research on video-based recasts has primarily explored the effects of oral recasts. Consequently, the feedback affordances of videoconferencing, which allow for multimodal recasts (combining oral and written feedback), remain unexplored. To address these

gaps, our study aimed to investigate the potential benefits of multimodal recasts in fostering L2 grammatical development in a videoconferencing setting. We operationalized multimodal recasts as oral recasts accompanied by written recasts, presented as captions on the screen (i.e., captioned oral recasts).

Multimodal recasts offer several advantages over their unimodal counterparts. According to the Cognitive Theory of Multimedia Learning (Mayer, 2014), learning is enhanced when information is constructed from both aural and visual channels rather than from a single channel as each channel has capacity limitations. Engaging both systems facilitates information processing and helps establish meaningful mental connections, potentially leading to better recall. Furthermore, multimodal recasts combine the benefits of oral and written feedback. Compared to oral feedback, written CF offers visual salience, re-readability, and longer processing time (e.g., Smith, 2004; Smith & Sauro, 2009). Meanwhile, unlike written feedback, oral CF provides verbal and paralinguistic cues (e.g., intonation, gestures), along with the contingency of feedback immediately following a learner's error.

Taken together, these factors are assumed to increase learners' attention to multimodal recasts, which is a prerequisite for L2 development (Schmidt, 2001). Given these potential benefits and the proliferation of language learning and teaching via videoconferencing, exploring the usefulness of multimodal recasts appears to be a timely research endeavor.

Background

Recasts in SLA

The interactionist perspective on SLA (e.g., Long, 1996) holds that drawing learners' attention to linguistic properties during meaningful interaction facilitates L2 development. One way to achieve this is through CF. CF is considered beneficial as it highlights learners' erroneous production and aids in noticing the gap between their interlanguage form and the target-like model (Schmidt, 2001), thereby promoting acquisition (Gass et al., 1998). This view is supported by meta-analyses, affirming that CF promotes L2 development (Brown, 2016; Li, 2010; Lyster & Saito, 2010). Among CF techniques, recasts have been identified as the most frequent feedback strategy in various instructional contexts (e.g., Li & Tan, 2022; Lyster et al., 2013), prompting extensive exploration of their potential benefits.

Recasts offer various advantages, including the ability to provide learners with both positive evidence (target form) and negative evidence (signaling an error), which helps them notice and correct their errors by juxtaposing incorrect forms and target-like ones (Leeman, 2003; Long, 2007). This process can lead learners to reject non-target-like forms. The contingency of recasts also creates favorable conditions for L2 acquisition, as the meaning of the recast is already familiar to learners (at least partly), being based on their own utterance. This allows them to allocate more attentional resources to the linguistic information in the feedback (Long, 2007). Furthermore, recasts facilitate attention to L2 features, providing optimal conditions for learners to establish and develop form-meaning connections without disrupting the focus on meaning (e.g., Long, 2007).

Despite the view that recasts may not necessarily lead to acquisition (e.g., Lyster, 2004; Lyster & Ranta, 1997), their effectiveness in promoting L2 development is well-documented in several meta-analyses (e.g., Li, 2010; Rassaei, 2024). Relevant to the present study is research demonstrating a positive relationship between recasts and the development of English question formation (e.g., Li & Iwashita, 2021; McDonough, 2005; McDonough & Mackey, 2006), the target construction in this study.

Recasts in Previous SCMC Research

A growing number of scholars have explored the potential of recasts in facilitating L2 grammatical development in various SCMC contexts. However, most of this research has focused on text-chat interaction (see Ziegler, 2016; Ziegler & Mackey, 2017 for reviews), involving comparisons between text-based recasts and more explicit CF (e.g., Sauro, 2009; Yilmaz, 2012) or face-to-face and text-based recasts (e.g., Kourtali, 2022; Yilmaz, 2012; Yilmaz & Yuksel, 2011).

For instance, Yilmaz and Yuksel (2011) observed greater noticing of Turkish locative and plural morphemes when participants received recasts during text-chat than face-to-face interaction, while in a study comparing text-based recasts with metalinguistic feedback targeting zero article use, Sauro (2009) found that only the metalinguistic group showed significantly greater immediate gains than the control group. Sauro attributed these findings to two main factors. First, lengthy, full recasts were used, which might have been perceived as confirmation checks or repetitions rather than corrections. Second, the frequent turns between recasts and the erroneous utterance likely reduced the chances of learners making cognitive comparisons, as they may have struggled to retain and process the information in working memory (WM). Supporting this, Lai et al. (2008) demonstrated that contingent recasts (i.e., occurring immediately after an error) are better noticed than non-contingent ones (i.e., occurring after a delay). Sauro's results suggest that contingent recasts may be more effective than non-contingent ones in SCMC environments (Ziegler, 2017) and when using full recasts, clarifying their corrective nature could increase their effectiveness.

Although smaller in number, some studies have also investigated the effectiveness of recasts during videoconferencing (Monteiro, 2014; Rassaei, 2017, 2019). Monteiro (2014) was among the first to compare the benefits of metalinguistic feedback and recasts in video-based focused tasks. Both feedback types proved equally useful for developing learners' explicit and implicit knowledge of the English past tense. In another study, Rassaei (2017) compared the impact of recasts on the use of definite versus indefinite articles during face-to-face and video-based interactions. Results from an oral production test, error correction test, and stimulated recall interviews found no significant differences between the two modalities. More recently, Rassaei (2019) examined the relative effects of recasts during mobile-mediated audio-based and video-based SCMC, focusing on the same target features. Learners' performance on the picture description and writing tests revealed a positive impact of recasts on accuracy, with the audio recast group producing more target-like modified utterances and showing more accurate perceptions of recasts.

In sum, despite the growing body of SCMC research on L2 grammatical development (e.g., Cerezo, 2021; Ziegler & Mackey, 2017), most existing studies have examined the effects of recasts in text-based SCMC. Furthermore, the limited research on video-based recasts has only considered oral recasts, overlooking other modalities of SCMC, such as multimodal provision of recasts (see however Bryfonski & Ma, 2020; Martin, 2018).

Multimodal Recasts as Pedagogical Intervention in L2 Development

Investigating multimodal recasts, defined here as captioned oral recasts that integrate both oral and written reformulations, appears promising due to their potential advantages over using either oral or written recasts alone. According to Mayer's (2014) Cognitive Theory of Multimedia Learning, processing information through two different channels enhances learning and understanding. In Mayer's view, there are two approaches to conceptualizing these two channels: the presentation-mode approach (Paivio, 2007), which distinguishes verbal and nonverbal channels, and the sensory-modality approach (Baddeley, 1999), which differentiates auditory and visual channels. The sensory-modality approach is particularly relevant for multimodal recasts, as the addition of visual information to auditory input activates both channels, facilitating mental connections between the two representation systems and improving recall. Moreover, the burden on WM in each channel is alleviated, fostering input processing and deeper learning (Mayer, 2014).

Given that this study implemented multimodal recasts in the form of captioned oral recasts, their potential for promoting L2 learning is also supported by previous research demonstrating the positive impact of captions in L2 acquisition. Meta-analyses (Montero Perez et al., 2013; Reynolds et al., 2022) have shown the benefits of captions for L2 learning. While most studies have focused on vocabulary acquisition, several studies have also confirmed the advantages of captioning for grammar acquisition (e.g., Cintrón-Valentín et al., 2019; Lee & Révész, 2018, 2020; Pattemore & Muñoz, 2020). These positive findings are attributed to the capacity of captions to help learners segment and parse spoken input into linguistic units

(Gass et al., 2019). Similarly, captioned oral recasts may facilitate the segmentation and parsing of oral input, directing learners' attention to target features and supporting grammar learning.

Multimodal recasts may also make the targeted feature more salient due to their hybrid nature, leveraging the distinct strengths of written and oral feedback. Written feedback offers longer processing time, visual/perceptual salience, and re-readability (e.g., González-Lloret, 2014; Lai et al., 2008; Sauro, 2009; Yilmaz, 2012), making the corrective intent of feedback more explicit to learners than that of oral feedback. Although on-screen written feedback in multimodal recasts may not fully replicate all the advantages of text-based CF, it still grants visual salience and extended processing time. Oral feedback, on the other hand, provides verbal and nonverbal cues, also enhancing the salience of recasts (Gass et al., 2017). Moreover, oral feedback has the advantage of being contingent upon the ill-formed utterance, which is often not feasible during text chat due to the absence of strict turn adjacency and a "split negotiation routine" (Smith, 2003, p. 48).

The Present Study

The literature review indicates that the majority of SCMC studies on L2 grammatical development have concentrated on text-based recasts, while research on video-based recasts and other modalities, including multimodal recasts, remains understudied. To address these gaps, we formulated the following research questions:

1. To what extent do recasts affect the development of *wh*-question formation during videoconferencing activities?
2. To what extent does modality of recasts impact the development of *wh*-question formation during videoconferencing activities?

Methodology

Design

The study employed a pretest–posttest–delayed posttest design with three treatment sessions. Sixty participants were randomly assigned into four groups: oral recast ($n = 15$), written recast ($n = 15$), multimodal recast ($n = 15$), and no recast ($n = 15$). All groups were administered a background questionnaire, the Oxford Placement Test, a pretest, a series of treatment tasks, an immediate posttest, a delayed posttest, and an exit questionnaire. Each testing session entailed a spontaneous oral production test, a more controlled oral production test, and a written production test.

Linguistic target

We selected the English *wh*-questions as the linguistic target for several reasons. First, *wh*-questions are physically salient, which makes them more likely to be susceptible to recasts (e.g., Mackey, 2006). Also, our choice of participants and assessment was informed by Pienemann et al.'s (1988) six-stage developmental sequence for question formation. Our study focused on developing learners' ability to produce questions at Stage 3 (fronting of the *wh*-word, *do*, or another element; e.g., *what you ate?*), Stage 4 (inversion in *wh*-questions with a copula or *yes-no* questions without *do*; e.g., *where were you?*), and Stage 5 (inversion in all *wh*-question with an auxiliary verb in the second position; e.g., *how did you lose her?*).

Participants

Of the 60 participants, 48 were female and 12 were male. They were all native Korean speakers learning English as a second language in the United Kingdom. The mean age was 27.55 ($SD = 2.33$). On average, participants had received 10.20 years of English instruction ($SD = 0.97$). Their mean length of residence in the United Kingdom was 15.12 months ($SD = 4.03$). None of the participants had previously lived in an

English-speaking country prior to their studies in the United Kingdom. One-way ANOVAs with age, length of English study, and residence as dependent variables found no significant differences among the four groups, age: $F(3, 56) = .49, p = .69, \eta^2 = .03$; length of study: $F(3, 56) = .33, p = .81, \eta^2 = .02$; and length of residence: $F(3, 56) = .13, p = .94, \eta^2 = .01$. Participants' proficiency levels fell within the A2–B1 bands according to the Common European Framework for Reference, as determined by their total scores on the Oxford Placement Test. One-way ANOVAs confirmed that all groups achieved comparable scores on both the listening: $F(3, 56) = .40, p = .75, \eta^2 = .02$, and grammar: $F(3, 56) = .86, p = .47, \eta^2 = .04$, sections of the test. We recruited participants who showed partial procedural knowledge of question formation by being able to produce two distinctive types of question forms from stages 3, 4, and 5. For example, they had to demonstrate two different uses of stage 5 questions, such as “Why did you follow me?” and “Where were you standing at the time?” in at least two pretest tasks. This selection method was based on previous research on question development (e.g., Philp, 2003; Kim et al., 2015).

Treatment Tasks

During videoconferencing, all four groups performed three multimodal communicative activities designed to elicit the use of *wh*-questions. Each version was set in an imaginary scenario where participants played different ‘interviewing’ roles, such as an FBI agent, police officer, or animal shelter volunteer, in a videoconferencing interrogation or interview. The first author acted as the interviewee, taking on roles such as a witness to a crime or incident, or as the owner of a lost dog. In each activity, participants were first presented with an incident form (activities 1 and 2) or a lost pet poster (activity 3) (Figure 1). Their task was to ask their interviewees questions in order to complete the report form or the lost dog poster. Subsequently, participants were given a set of question prompts (Figure 2), designed to create contexts for each of the nine types of *wh*-questions: subject *wh*-extraction (who, what), object *wh*-extraction (who, what), and adjunct *wh*-extraction (where, when, why, how/manner, how/degree). The treatment materials included only high-frequency words selected from the New General Service List (Brezina & Gablasova, 2015). Piloting confirmed that the activities successfully elicited the target *wh*-questions.

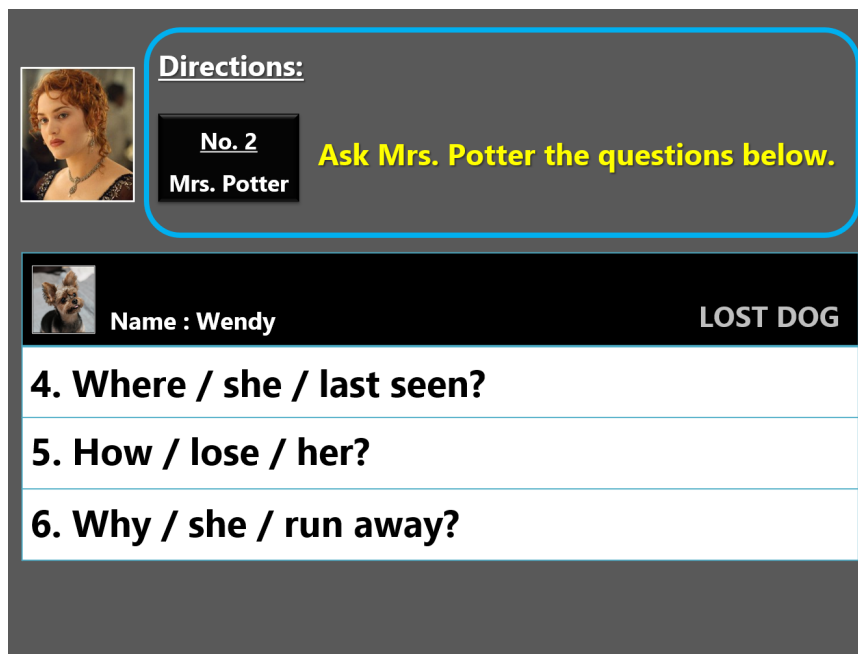
Figure 1

An Example of a Lost Dog Poster

LOST DOG!  Name : Wendy If you see her, please call 079-4321-7654!	1. How old?	2. What Features?	3. When?
	4. Where?	5. How?	6. Why?
	7. Home Address?	8. Who?	9. Contact Details?

Figure 2

An Example of Question Prompts



Directions:

No. 2
Mrs. Potter

Ask Mrs. Potter the questions below.

Name : Wendy **LOST DOG**

4. Where / she / last seen?

5. How / lose / her?

6. Why / she / run away?

During the treatment tasks, the three experimental groups received recasts for their errors in using *wh*-questions. The recasts were implemented as full, isolated recasts, focusing solely on reformulating the incorrect parts of learners' production (Sheen, 2006), as shown below:

Learner: Why she run away?

Researcher: Why did she run away?

This recast type was applied consistently across all experimental groups. The oral recast group received recasts orally, the written recast group was given written recasts, and the multimodal recast group received both oral and written recasts simultaneously. The no recast group received no feedback. For the multimodal and written recast groups, the first author provided written feedback as real-time on-screen text using Open Broadcaster Software (OBS) (Figure 3). This text was displayed at the bottom of the computer screen for 10 seconds immediately after participants made errors in their use of *wh*-questions. The duration of the written recasts was determined through piloting. Prior to the experiment, written recasts for each target question were prepared and presented following participants' errors. These pre-prepared recasts enabled immediate written feedback for the written recast group and simultaneous oral and written feedback for the multimodal recast group, ensuring prompt, accurate and contingent presentation of feedback without the need for real-time typing. It should be noted that, unlike pre-programmed recasts in Intelligent CALL (ICALL) tools, these recasts occurred during genuine exchanges between the learners the researcher. All treatment sessions were conducted and recorded via Zoom.

Figure 3*An Example of a Written Recast*

Directions:

No. 2 Ask Mrs. Potter the questions below.

Mrs. Potter

Name : Wendy LOST DOG

4. Where / she / last seen?

5. How / lose / her?

6. Why / she / run away?

How did you lose her?

A written recast in the form of real-time captions/on-screen text

Assessment Tasks

We developed three assessment tasks: a controlled oral production test (OPT), a controlled written production test (WPT), and a spontaneous oral production test (SOPT) (Figure 4). Three different but comparable versions of each test were designed and counterbalanced across testing sessions. All tests were administered and recorded using Zoom. The tests only included high-frequency words selected from the New General Service List (Brezina & Gablasova, 2015). Piloting confirmed that the tests successfully created obligatory contexts for the target constructions.

The OPT, similar to the treatment tasks, asked participants to act as interviewers investigating a particular incident (e.g., a police officer interviewing witnesses to a bank robbery). Each test version was designed to elicit 18 target *wh*-questions and 12 distractors ($n = 30$). Nine types of *wh*-question were included, with two items for each sub-type: subject *wh*-extraction (who, what), object *wh*-extraction (who, what), and adjacent *wh*-extraction (where, when, why, how/degree, how/manner). The distractors required the formation of four types of *yes/no* questions, with three items per type: *be* + bare infinitive, *do*-fronting, *can* + bare infinitive, and *have/has* + past participle. Except for modality, the WPT had the same format as the OPT, with the only difference being that a single item targeted each *wh*-question type ($n = 9$). The number of distractors remained the same, resulting in a total of 21 items. The internal consistency reliability for the three versions of the OPT (version A: $\alpha = .894$, version B: $\alpha = .813$, version C: $\alpha = .838$) and WPT (version A: $\alpha = .825$, version B: $\alpha = .838$, version C: $\alpha = .878$) fell within the good range.

In the SOPT, learners were provided with a report form (e.g., an accident report or a passenger report) containing less specific *wh*-question prompts (e.g., occupation) compared to the OPT (e.g., what / your job). Their task was to complete the form by asking questions to the first author. Given the less specific prompts, this task required learners to rely more on their own linguistic resources than in the OPT.

By incorporating varied assessments, we could evaluate different aspects of knowledge (e.g., Dekeyser, 2007). The OPT imposed time pressure, facilitating the use of procedural knowledge, whereas the WPT allowed more time for reflection, enabling participants to rely more on their declarative knowledge while also deploying procedural knowledge. Through utilizing different test modalities, we also assessed whether participants could transfer any gains from the varied recast types to different modalities. Moreover, the SOPT enabled us to measure learners' ability to use the target constructions in a more authentic scenario, thereby enhancing ecological validity.

Figure 4

Examples of Assessment Tasks (SOPT, OPT, WPT in this order)

The figure shows three examples of assessment tasks arranged horizontally. Each task includes a character image, a name, and a list of questions.

- Task 1 (SOPT):** A blue-bordered box titled "PASSENGER REPORT" for "Passenger No. 1 Mr. Stark" (with a passport icon). Questions: 1. Occupation, 2. Nationality, 3. Purpose of visit, 4. Length of stay, 5. Amount of money you have.
- Task 2 (OPT):** A green-bordered box titled "Directions: Ask Mr. Hunter the questions below." for "No. 1 Mr. Hunter" (with a clipboard icon). Questions: 1. What / your real job?, 2. How much / earn / a year?, 3. When / start / spying?, 4. Can / describe / current mission?, 5. kill / anyone / here?.
- Task 3 (WPT):** A black-bordered box titled "Directions: TYPE the following questions in the chat box." for "No. 1 Mr. Kingsman" (with a clipboard icon). Questions: 1. What / your previous job?, 2. How / get paid?, 3. When / start / this job?, 4. Your family / know / your job?, 5. Who / plan / this operation?, 6. Satisfied with / your current job?, 7. Can / speak / any foreign languages?.

Data Collection Procedure

Each participant engaged in four individual sessions over two weeks. In session 1, the background questionnaire (10 minutes), the Oxford Placement Test (40 minutes), and the pretest (45–50 minutes) were administered in this order. During session 2, participants performed the first and second treatment activities. Session 3 included the third treatment activity, immediately followed by the posttest. In Session 4, which took place one week later, participants completed a delayed posttest and an exit questionnaire. In each testing session, participants first carried out the SOPT (14–16 minutes), followed by the OPT and WPT (13–15 minutes each). The WPT was administered last to minimize the likelihood of participants relying on their declarative knowledge during subsequent oral production tests. Likewise, the OPT was conducted before the SOPT, as it constituted a more controlled oral production test. Each treatment task lasted approximately 14–16 minutes. There was no time limit during the treatments or tests. The first author administered all sessions. Figure 5 illustrates the experimental schedule.

Figure 5

Experimental Schedule

Background Questionnaire / Oxford Placement Test / Pretest			
Oral recast (n = 15)	Written recast (n = 15)	Multimodal recast (n = 15)	No recast (n = 15)
Treatment session 1			
Treatment session 2			
Treatment session 3			
Immediate Posttest			
Delayed Posttest (1 week after) / Exit Questionnaire			

Coding and scoring

Participants' development of *wh*-questions was evaluated using a categorical scoring system based on Pienemann et al.'s (1988) question development scale ([Appendix A](#)). This scale has been widely employed in previous research (e.g., Kim, 2012; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006) and has proven to be a sensitive and accurate reference to measure progress in question formation. In all tests, we assigned scores of 3, 4 or 5 points to Stage 3, 4, and 5 questions, respectively. Only non-subject questions were included in the total scores, resulting in a maximum of 70 points (14 items) for the OPT and 35 points (7 items) for the WPT. The SOPT did not have a fixed number of *wh*-question contexts; instead, we counted the total number of obligatory contexts for *wh*-questions produced by each learner. The total score for the SOPT was calculated by dividing the sum of raw scores by the maximum possible score (i.e., the number of obligatory contexts $\times$ 5). Pronunciation and spelling errors were disregarded. Each test consisted of three comparable versions, which were counterbalanced across participants, modalities, and testing sessions. All assessment tests were coded and scored by the researcher, with a second coder scoring 25 percent of the data for each measure. Inter-coder agreement was determined by calculating Cohen's kappa, which showed a high level of agreement for each task (OPT: .98; WPT: .98; SOPT: .97).

Statistical analyses

We used SPSS Version 27 to assess internal consistency reliability for the OPT and WPT, calculate descriptive statistics, and perform one-way ANOVAs to examine the comparability of the four groups regarding the number of obligatory contexts elicited, the number of recasts provided, and pretest scores.

To analyze advancements in *wh*-question stages on the OPT and WPT, we employed ordinal logistic mixed-effects models utilizing the *clmm* function from the *ordinal* package in the R statistical environment (R Core Team, 2022), given the ordinal nature of the dependent variables. For the SOPT data, linear mixed-effects models were formulated with the *lm* function from the *lme4* package (Bates et al., 2015), due to the continuous nature of the dependent variables. In each analysis, the dependent variables were participants' pretest, posttest, and delayed posttest scores. The fixed effects were group, time, and their interactions, with the predictors of interest being the group-by-time interactions. For the OPT and WPT, participant and item were included as random effects. In the SOPT analyses, only participant served as a random effect, as participants generated different *wh*-questions. Each analysis used a different group as the reference to derive results for all group-time interactions of interest.

To estimate effect sizes, we calculated η^2 values for ANOVAs, odds ratios (ORs) for the logistic mixed-effects regression models, and R^2 values for the linear mixed-effects regression models. For the mixed-effects models, effect size estimates were computed with the command "r.squared GLMM" from the "MuMin" package (Barton, 2020). Following Plonsky and Oswald (2014), η^2 values of .01, .06, and .14 were interpreted as small, medium, and large, respectively. R^2 values up to .20 were labeled as small, while those exceeding .50 were classified as large (Plonsky & Ghanbar, 2018). ORs greater than 3 or less than 0.33 were considered strong (Haddock et al., 1998). We used an alpha level of $p < .05$ for significance.

To ensure the validity of our analyses, we conducted diagnostic procedures, which included Shapiro-Wilk tests for normality and variance inflation factors (VIF) for collinearity assessment. Homoscedasticity was evaluated by examining residual plots and conducting Levene's or Breusch-Pagan tests.

Results

Preliminary analyses

The Shapiro-Wilk tests indicated that all data were normally distributed ([Appendix B](#)). [Table 1](#) presents the descriptive statistics for the number of obligatory contexts created by the treatment tasks for the linguistic target during the treatment, as well as the number of recasts participants received in response to

errors in *wh*-question formation across the three treatment activities. A one-way ANOVA found no significant difference in the number of obligatory contexts among the four groups: $F(3, 56) = .290$, $p = .831$, $\eta^2 = .02$. Another one-way ANOVA confirmed that there was no significant difference in the number of recasts given to the three experimental groups: $F(2, 42) = 2.56$, $p = .091$, $\eta^2 = .11$.

Table 1

Descriptive Statistics for the Number of Obligatory Contexts and Recasts Per Group

Group	N	Mean	SD	95% CI	
				Lower	Upper
Obligatory contexts					
Oral	15	26.13	.83	25.67	26.60
Written	15	26.27	.80	25.82	26.71
Multimodal	15	26.33	.82	25.88	26.79
Control	15	26.40	.83	25.94	26.86
Number of recasts					
Oral	15	13.53	1.64	12.62	14.44
Written	15	12.87	2.10	11.70	14.03
Multimodal	15	12.00	1.81	11.00	13.00

Note. CI = confidence interval

The descriptive statistics for the pretest scores are presented in [Tables 3, 5, and 7](#). As shown in [Table 2](#), a series of one-way ANOVAs found no significant differences in performance among the four groups for any of the three assessment tasks at the pretest.

Table 2

Results of One-Way ANOVAs Comparing Pretest Scores

Test	df	F	p	η_p^2
Oral Production	3, 56	.20	.90	.01
Written Production	3, 56	.47	.70	.03
Spontaneous Oral Production	3, 56	1.39	.26	.07

Main Analyses

OPT

[Table 3](#) summarizes the descriptive statistics for the pretest, posttest, and delayed posttest scores on the OPT. Ordinal logistic mixed-effects model analyses, conducted to examine the effect of group on participants' performance, revealed eight significant group-by-time interactions. As shown in [Table 4](#) (see [Appendix C](#) for full model results), all recast groups showed greater pretest–posttest gains than the control group. As indicated by the odds ratios, the multimodal recast group achieved the largest gains compared to the control group, followed by the written and oral recast groups, in this order. The multimodal recast group retained its advantage over the control group on the delayed posttest, but no significant differences were found in pretest–delayed posttest gains between the unimodal (i.e., oral and

written) and control groups. Regarding modality differences, the multimodal recast group demonstrated greater pretest–posttest gains than both the oral and written recast groups, and the written recast group outperformed the oral recast group. On the delayed posttest, however, we detected a difference in gains only between the multimodal and oral recast groups. In general, the effect sizes were larger for pretest–posttest than for pretest–delayed posttest gains. Overall, these results suggest that recasts had a positive impact on participants' gains on the OPT, with multimodal recasts providing the greatest benefit followed by written and oral recasts.

Table 3

Descriptive Statistics for the OPT Scores

Group	Test	Mean	SD	95% CI	
				Lower	Upper
Oral (<i>n</i> = 15)	Pretest	57.80	1.57	56.93	58.67
	Posttest	60.87	1.13	60.24	61.49
	Delayed posttest	58.53	1.41	57.75	59.31
Written (<i>n</i> = 15)	Pretest	57.47	1.18	56.72	58.22
	Posttest	62.87	1.41	62.09	63.65
	Delayed posttest	60.20	1.74	59.24	61.16
Multimodal (<i>n</i> = 15)	Pretest	57.67	1.18	57.02	58.32
	Posttest	66.47	1.89	65.42	67.51
	Delayed posttest	61.60	.83	61.14	62.06
Control (<i>n</i> = 15)	Pretest	57.80	1.32	57.07	58.53
	Posttest	58.47	1.06	57.88	59.05
	Delayed posttest	58.40	0.99	57.85	58.95

Note. The maximum score was 70.

Table 4

Significant Interactions Identified by Ordinal Logistics Mixed-Effects Models Examining Participants' Scores on the OPT

	Fixed effects				OR	95% CI
	Estimate	SE	<i>z</i>	<i>p</i>		
ControlOral:Time2	0.58	0.28	2.04	.041	1.78	[1.02, 3.10]
ControlWritten:Time2	1.23	0.29	4.25	<.001	3.43	[1.94, 6.07]
ControlMulti:Time2	1.85	0.30	6.15	<.001	6.33	[3.52, 11.40]
ControlMulti:Time3	0.79	0.28	2.81	.005	2.21	[1.27, 3.84]
OralWritten:Time2	0.66	0.29	2.23	.026	1.93	[1.02, 3.07]
OralMulti:Time2	1.27	0.30	4.18	<.001	3.55	[2.18, 6.87]
OralMulti:Time3	0.70	0.28	2.49	.013	2.02	[1.16, 3.52]
WrittenMulti:Time2	0.61	0.31	1.98	.048	2.19	[1.22, 3.91]

Note. OR = Odds Ratio; Time 1 = pretest (reference level); Time 2 = immediate posttest; Time 3 = delayed posttest

WPT

Table 5 provides the descriptive statistics for the participants' pretest, posttest, and delayed posttest scores on the WPT. As summarized in Table 6, the ordinal logistic mixed-effects models, conducted to examine the impact of group on participants' test performance, yielded nine significant interactions between group and time (see Appendix D for full model results). Similar to the results from the OPT, all recast groups displayed larger pretest–posttest gains than the control group. The odds ratios indicated that, relative to the control group, the greatest pretest–posttest gains were achieved by the multimodal recast group, followed by the written and oral recast groups in this order. Both the multimodal and written recast groups also maintained their advantage over the control group on the delayed posttest. Turning to modality effects, the multimodal recast group generated greater pretest–posttest and pretest–delayed posttest gains than the oral and written recast groups, with no difference found between the oral and written recast groups. Overall, the presence of recasts, regardless of modality, improved participants' written production of *wh*-question formation. However, multimodal recasts were the most effective in promoting gains compared to their oral and written counterparts.

Table 5

Descriptive Statistics for the WPT Scores

Group	Test	Mean	SD	95% CI	
				Lower	Upper
Oral (<i>n</i> = 15)	Pretest	26.73	1.28	26.02	27.44
	Posttest	30.27	1.39	29.50	31.03
	Delayed posttest	28.93	1.75	27.96	29.90
Written (<i>n</i> = 15)	Pretest	26.53	1.41	25.75	27.31
	Posttest	31.27	0.89	30.78	31.76
	Delayed posttest	29.40	1.24	28.71	30.09
Multimodal (<i>n</i> = 15)	Pretest	26.93	1.28	26.22	27.64
	Posttest	33.60	0.74	33.19	34.01
	Delayed posttest	31.87	1.06	31.28	32.45
Control (<i>n</i> = 15)	Pretest	26.40	1.30	25.68	27.12
	Posttest	27.40	1.12	26.78	28.02
	Delayed posttest	27.20	0.94	26.68	27.72

Note. The maximum score was 35.

Table 6

Significant Interactions Identified by Ordinal Logistic Mixed–Effects Models Examining Participants’ Scores on the WPT

	Fixed effects				<i>OR</i>	95% CI
	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>		
ControlOral:Time2	0.90	0.38	2.38	.017	2.46	[1.18, 5.14]
ControlWritten:Time2	1.44	0.39	3.72	<.001	4.21	[1.98, 8.95]
ControlMulti:Time2	2.41	0.41	5.82	<.001	11.17	[4.98, 25.08]
ControlWritten:Time3	0.78	0.38	2.06	.039	2.17	[1.04, 4.53]
ControlMulti:Time3	1.62	0.39	4.20	<.001	5.06	[2.38, 10.75]
OralMulti:Time2	1.51	0.42	3.64	<.001	4.54	[2.02, 10.21]
OralMultimodal:Time3	1.11	0.39	2.87	.004	3.02	[1.42, 6.40]
WrittenMulti:Time2	0.98	0.42	2.31	.002	2.65	[1.16, 6.05]
WrittenMulti:Time3	0.85	0.39	2.19	.029	2.33	[1.09, 4.95]

SOPT

Table 7 shows the descriptive statistics for the pretest, posttest, and delayed posttest scores on the SOPT. The linear mixed-effects model analyses, performed to explore the influence of group on participants’ pretest–posttest and pretest–delayed posttest changes, detected ten significant group-by-time interactions. As presented in Table 8 (see Appendix E for full model results), all three recast groups outperformed the control group, with larger gains observed from pretest to posttest and from pretest to delayed posttest. Consistent with the findings from the other assessments, the multimodal recast group exhibited the most substantial gains over the oral and written recast groups, with the difference in gains being large. We found no significant difference between the gains of the oral and written recast groups. Overall, the SOPT results, in line with those of the OPT and WPT, confirmed the advantage of recasts, with the multimodal recasts leading to the greatest improvements.

Table 7*Descriptive Statistics for the SOPT Scores*

Group	Test	Mean	SD	95% CI	
				Lower	Upper
Oral (<i>n</i> = 15)	Pretest	72.85	1.20	72.18	73.51
	Posttest	78.90	1.18	78.25	79.55
	Delayed posttest	76.91	1.06	76.32	77.51
Written (<i>n</i> = 15)	Pretest	72.16	1.50	71.33	72.99
	Posttest	79.61	1.92	78.54	80.66
	Delayed posttest	77.08	1.54	76.23	77.94
Multimodal (<i>n</i> = 15)	Pretest	72.44	0.96	71.91	72.98
	Posttest	83.79	1.18	83.14	84.44
	Delayed posttest	80.83	1.16	80.19	81.47
Control (<i>n</i> = 15)	Pretest	73.05	1.51	72.21	73.88
	Posttest	74.88	1.41	78.39	80.19
	Delayed posttest	73.72	1.22	73.04	74.39

Table 8*Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-Effects Models Examining Participants' Scores on the SOPT*

	Fixed effects					Random effects		
	Estimate	SE	<i>t</i>	<i>p</i>	<i>R</i> ² <i>m</i>	<i>R</i> ² <i>c</i>	by participant variance	SD
ControlOral:Time2	0.03	0.01	5.72	<.001	.81	.92	<.01	<.01
ControlWritten:Time2	0.04	0.01	6.49	<.001	.77	.83	<.01	<.01
Control:Multi:Time2	0.09	0.01	15.08	<.001	.94	.97	<.01	<.01
ControlOral:Time3	0.03	0.01	4.29	<.001	.58	.70	<.01	<.01
ControlWritten:Time3	0.03	0.01	4.29	<.001	.63	.70	<.01	<.01
Control:Multi:Time3	0.08	0.01	12.76	<.001	.89	.94	<.01	<.01
OralMulti:Time2	0.07	0.01	9.35	<.001	.93	.97	<.01	<.01
OralMulti:Time3	0.05	0.01	8.47	<.001	.88	.93	<.01	<.01
WrittenMulti:Time2	0.05	0.01	8.58	<.001	.90	.93	<.01	<.01
WrittenMulti:Time3	0.05	0.01	8.47	<.001	.85	.90	<.01	<.01

Summary of Results

Table 9 gives a summary of significant results with relevant effect sizes.

Table 9

Summary of Significant Differences in Gains Across Groups

Pretest-posttest gains (effect size)	Pretest-delayed posttest gains (effect size)
<i>Oral Production Test (Odds Ratios)</i>	
control < multimodal (6.33)	control < multimodal (2.21)
control < oral (1.78)	
control < written (3.43)	oral < multimodal (2.02)
oral < multimodal (3.55)	
written < multimodal (2.19)	
oral < written (1.93)	
<i>Written Production Test (Odds Ratios)</i>	
control < multimodal (11.17)	control < multimodal (5.06)
control < oral (2.46)	
control < written (4.21)	control < written (2.17)
oral < multimodal (4.54)	oral < multimodal (3.02)
written < multimodal (2.65)	written < multimodal (2.33)
<i>Spontaneous Oral Production Test (R^2)</i>	
control < multimodal (.94)	control < multimodal (.89)
control < oral (.81)	control < oral (.58)
control < written (.77)	control < written (.63)
oral < multimodal (.93)	oral < multimodal (.88)
written < multimodal (.90)	written < multimodal (.85)

Discussion

Our first research question aimed to examine the extent to which recasts, irrespective of modality, affect L2 development in the context of videoconferencing. We defined development as stage advancement based on Pienemann et al.'s (1988) developmental sequence for question formation. We found that the provision of recasts, as compared to their absence, positively impacted participants' advancement in *wh*-question stages across all three tests. Therefore, the results suggest that the positive outcomes observed in previous research regarding the effectiveness of recasts in enhancing L2 grammatical development in face-to-face and text-chat settings are also applicable to recasts provided during videoconferencing. Moreover, our findings align with prior research that has shown the advantages of recasts in improving grammatical accuracy during video-based SMC (e.g., Monteiro, 2014; Rassaei, 2017, 2019).

Our second research question explored whether the modality of recasts—oral, written or multimodal— influences their effectiveness. The results indicated that the multimodal recast group produced the highest

pretest–posttest and pretest–delayed posttest gains across all assessments, with the written and oral recast groups following in that order. The only exception was the lack of a significant difference between the multimodal and written recast groups in the OPT pretest–delayed posttest gains.

The superior performance of the multimodal recast group is in line with our expectations and can be attributed to several factors. First, consistent with the Cognitive Theory of Multimedia Learning (Mayer, 2014), learners may have formed stronger and more elaborate mental representations of the target constructions, as multimodal recasts primed both visual and oral channels. Presenting information through both modalities could also have reduced the cognitive load imposed in each modality, leading to deeper input processing.

Second, the hybrid nature of multimodal recasts may have enabled learners to leverage the strengths of both oral and written CF. Factors such as longer processing time, increased visual/perceptual salience, contingency, and the ability to re-read the input likely heightened the noticeability of multimodal recasts, making their corrective force more explicit to learners. This seems particularly beneficial when targeting grammatical constructions, which tend to be less noticeable than lexical and phonological features (e.g., Lyster et al., 2013).

Third, when presented as captioned oral recasts, multimodal recasts may have helped learners segment and parse the spoken input (Gass et al., 2019). This likely facilitated word recognition, assisting learners in identifying linguistic items in the input with greater ease, promoting attention to and learning of these items (Montero Perez, 2020; Winke et al., 2010). This increased focus on linguistic information may have further contributed to the overall effectiveness of multimodal recasts.

It is also worthwhile to discuss how recasts in this study affected the development of different types of L2 knowledge. Multimodal recasts led to superior performance across both oral and written production tests. Furthermore, the written recast group outperformed the oral recast and control groups on both test modalities. Similarly, the oral recast group also achieved greater gains than the control group in both tests. These results suggest that participants' improvements extended beyond procedural knowledge to declarative knowledge, demonstrating the transferability of L2 knowledge across different modalities. According to Skill Acquisition Theory (e.g., DeKeyser, 2007), transferring procedural knowledge across modalities is challenging, and successful transfer between skills (e.g., from written to oral performance) often relies on declarative knowledge of the target structure. These findings are consistent with previous research (e.g., Révész, 2012), which has shown that recasts can facilitate both declarative and procedural knowledge. While this has been established for oral recasts, our study is among the first to provide evidence that multimodal recasts and written recasts have the capacity to promote the development of both types of knowledge (see, however, Baralt, 2013; Yilmaz & Yuksel, 2011).

Finally, it is important to consider why written recasts proved more advantageous than their oral equivalents. One explanation could be that the modality contrast between written recasts and the ongoing oral interactions captured learners' attention more effectively, due to the greater salience of the written format. Another possibility is that the consistent contingency of written recasts could have enhanced their corrective intent, aiding learners in making cognitive comparisons between their errors and the correct models (Long, 2007). This may have prompted learners to reflect more on the target forms, thereby facilitating the application of their declarative knowledge during tasks. As a result, they may have been better able to automatize their explicit knowledge of the use of *wh*-questions.

Limitations

Several limitations to the current study should be acknowledged. Firstly, we only included L2 production tests to assess changes in participants' use of *wh*-question formation. While this choice aligned with the

nature of the treatment, future research could investigate whether the positive effects observed for multimodal recasts extend to comprehension-based outcome measures. Another weakness concerns the short interval between the immediate and delayed posttests (one week). This timeframe was chosen to minimize the possibility of learners receiving external feedback or explicit instruction on the target construction outside the study context, which might have influenced their performance during the delayed posttest. Although a longer interval could have provided insights into the longer-term effects of the treatment, we opted for a more cautious approach to attribute the observed improvements in grammatical accuracy primarily to the effects of the treatment provided during the study. Additionally, the small number of participants in each group could make the study susceptible to outliers and other extraneous variables, highlighting the need for larger sample sizes in future research. Furthermore, given that previous research has suggested that WM plays a role in mediating the effectiveness of computer-mediated recasts (e.g., Sagarra & Abbuhl, 2013), future investigation of multimodal recasts could benefit from controlling for individual differences in WM capacity among learners. Finally, following previous video-based SCMC on different interactional features, such as types of CF (e.g., Canals et al., 2021; Martin, 2018), language-related episodes (LREs), and modified output (e.g., Bueno-Alastuey, 2013), as well as studies in other areas of language, including phonology (e.g., Bryfonski & Ma, 2020; Parlak & Ziegler, 2017), replication studies investigating different interaction-driven language learning opportunities and linguistic targets are warranted to explore how these variables might influence the effectiveness of multimodal feedback.

Conclusion and Pedagogical Implications

This study aimed to investigate the extent to which recasts provided during videoconferencing tasks promote L2 development in English *wh*-question formation, and whether the modality of recasts (oral, written, or multimodal) influences learning outcomes. Our findings showed that, regardless of modality, recasts had a positive impact on participants' knowledge of *wh*-questions. As expected, multimodal recasts, delivered as captioned oral recasts, resulted in significantly greater immediate and delayed gains in oral, written, and spontaneous oral production compared to unimodal recasts. This suggests that the integration of oral and written feedback may have helped learners create more elaborate mental representations, reduce cognitive load, and enhance the perceptual salience of the input. Also, multimodal recasts appeared to facilitate the segmentation of oral feedback, thereby enhancing learners' processing of linguistic items.

These findings have important implications. The superiority of multimodal recasts underscores the pedagogical advantage of supplying feedback through a combination of oral and written modalities, suggesting that this approach may be more effective for promoting L2 grammatical development than using a single modality in video-based SCMC. With the growing prevalence of online language learning platforms equipped with multimodal capacities, future research replicating this study could offer valuable insights for language educators seeking optimal feedback strategies in virtual environments.

Acknowledgements

This research was supported in part by the International Research Foundation for English Language Education (TIRF) Doctoral Dissertation Grant and the Language Learning Dissertation Grant.

References

- Baddeley, A. D. (1999). *Essentials of human memory*. Psychology Press.
- Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition*, 35, 689–725.
<https://doi.org/10.1017/S0272263113000429>

- Barton, K. (2020). *MuMIn: Multi-Model Inference*. R package version 1.43.17. <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, 36(1), 1–22. <https://doi.org/10.1093/applin/amt018>
- Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research*, 20(4), 436–458. <https://doi.org/10.1177/1362168814563200>
- Bryfonski, L., & Ma, X. (2020). Effects of implicit versus explicit corrective feedback on Mandarin tone acquisition in a SCMC learning environment. *Studies in Second Language Acquisition*, 42(1), 61–88. <https://doi.org/10.1017/S0272263119000317>
- Bueno-Alastuey, M. C. (2013). Interactional feedback in synchronous voice-based computer mediated communication: Effect of dyad. *System*, 41(3), 543–559. <https://doi.org/10.1016/j.system.2013.05.005>
- Canals, L., Granena, G., Yilmaz, Y., & Malicka, A. (2021). The relative effectiveness of immediate and delayed corrective feedback in video-based computer-mediated communication. *Language Teaching Research*, 0(0). <https://doi.org/10.1177/13621688211052793>
- Cerezo, L. (2021). Corrective feedback in computer-mediated versus face-to-face environments. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (pp. 494–519). Cambridge University Press. <https://doi.org/10.1017/9781108589789.024>
- 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>
- DeKeyser, R. (2007). Situating the concept of practice. In R. DeKeyser (Ed.), *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1–18). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667275.002>
- Gass, S. M., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition: An introduction. *Modern Language Journal*, 82, 299–307.
- Gass, S. M., Spinner, P., & Behney, J. (Eds.). (2017). *Salience in second language acquisition*. Routledge. <https://doi.org/10.4324/9781315399027>
- Gass, S. M., Winke, P., Isbell, D. R., & Ahn, J. (2019). How captions help people learn languages: A working-memory, eye-tracking study. *Language Learning & Technology*, 23(2), 84–104. <http://hdl.handle.net/10125/44684>
- González-Lloret, M. (2014). The need for needs analysis in technology-mediated TBLT. *Technology-mediated TBLT: Researching technology and tasks*, 6, 23–50. <https://doi.org/10.1075/tblt.6.02gon>
- Haddock, C. K., Rindskopf, D., & Shadish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods*, 3(3), 339–353. <https://doi.org/10.1037/1082-989X.3.3.339>
- Kim, Y. (2012). Task complexity, learning opportunities, and Korean EFL learners' question development. *Studies in Second Language Acquisition*, 34, 627–658. <https://doi.org/10.1017/S0272263112000368>

- Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory: L2 question development through recasts in a laboratory setting. *Studies in Second Language Acquisition*, 37(3), 549–581. <https://doi.org/10.1017/S0272263114000618>
- Kourтали, N. E. (2022). The effects of face-to-face and computer-mediated recasts on L2 development. *Language Learning & Technology*, 26(1), 1–20. <https://hdl.handle.net/10125/73457>
- Lai, C., Fei, F., & Roots, R. (2008). The contingency of recasts and noticing. *CALICO Journal*, 26(1), 70–90. Retrieved from <https://www.learntechlib.org/p/74206/>
- Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal*, 102(3), 557–577. <https://doi.org/10.1111/modl.12503>
- Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition*, 42, 625–651. <https://doi.org/10.1017/S0272263120000108>
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25(1), 37–63. <https://doi.org/10.1017/S0272263103000020>
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365. <https://doi.org/10.1111/j.1467-9922.2010.00561.x>
- Li, H., & Iwashita, N. (2021). The role of recasts and negotiated prompts in an FL learning context in China with non-English major university students. *Language Teaching Research*, 25(2), 209–233. <https://doi.org/10.1177/1362168819839727>
- Li, J., & Tan, D. (2022). Resurveying corrective feedback meta-analysis. *Mobile Information Systems*, 2022. <https://doi.org/10.1155/2022/3444160>
- Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in system. *System*, 84, 93–109. <https://doi.org/10.1016/j.system.2019.05.006>
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & B. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press. <https://doi.org/10.1016/B978-012589042-7/50015-3>
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26(3), 399–432. <https://doi.org/10.1017/S0272263104263021>
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37–66. <https://doi.org/10.1017/S0272263197001034>
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265–302. <https://doi.org/10.1017/S0272263109990520>
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40. <https://doi.org/10.1017/S0261444812000365>
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27(3), 405–430. <https://doi.org/10.1093/applin/ami051>
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82(3), 338–356. <https://doi.org/10.1111/j.1540-4781.1998.tb01211.x>

- Martin, A. (2018). *How to synchronize? A study of video-based, voice-based & text based synchronous computer-mediated communication, working memory, and second language learning*. [Doctoral dissertation, Georgetown University]. <http://hdl.handle.net/10822/1051974>
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 43–71). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition*, 27(1), 79–103. <https://doi.org/10.1017/S0272263105050047>
- McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning*, 56(4), 693–720. <https://doi.org/10.1111/j.1467-9922.2006.00393.x>
- Monteiro, K. (2014). An experimental study of corrective feedback during video-conferencing. *Language Learning & Technology*, 18(3), 56–79. <https://ilt.msu.edu/issues/october2014/monteiro.pdf>
- Montero Perez, M. (2020). Multimodal input in SLA research. *Studies in Second Language Acquisition*, 62, 653–663. <https://doi.org/10.1017/S0272263120000145>
- 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. <http://doi.org/10.1016/j.system.2013.07.013>
- Paivio, A. (2007). Basic principles of dual coding theory. In A. Paivio (Ed.), *Mind and its evolution: A dual coding theoretical approach* (pp. 25–57). Psychology Press.
- Parlak, Ö., & Ziegler, N. (2017). The impact of recasts on the development of primary stress in a synchronous computer-mediated environment. *Studies in Second Language Acquisition*, 39(2), 257–285. <https://doi.org/10.1017/S0272263116000310>
- Pattemore, A., & Muñoz, C. (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System*, 93, Article 102303. <https://doi.org/10.1016/j.system.2020.102303>
- Philp, J. (2003). Constraints on “noticing the gap”: Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25(1), 99–126. <https://doi.org/10.1017/S0272263103000044>
- Pienemann, M., Johnston, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for second language assessment. *Studies in Second Language Acquisition*, 10(2), 217–243. <https://doi.org/10.1017/S0272263100007324>
- Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R^2 values. *The Modern Language Journal*, 102(4), 713–731. <https://doi.org/10.1111/modl.12509>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912. <https://doi.org/10.1111/lang.12079>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rassaei, E. (2017). Video chat vs. face-to-face recasts, learners' interpretations and L2 development: a case of Persian EFL learners. *Computer Assisted Language Learning*, 30(1–2), 133–148. <https://doi.org/10.1080/09588221.2016.1275702>

- Rassaei, E. (2019). Computer-mediated text-based and audio-based corrective feedback, perceptual style and L2 development, *System*, 82, 97–110. <https://doi.org/10.1016/j.system.2019.03.004>
- Rassaei, E. (2024). The effects of recasts on L2 grammar: a meta-analysis. *The Language Learning Journal*, 52(1), 16–36. <https://doi.org/10.1080/09571736.2022.2097298>
- Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*, 62, 93–132. <https://doi.org/10.1111/j.1467-9922.2011.00690.x>
- Reynolds, B. L., Cui, Y., Kao, C. W., & Thomas, N. (2022). Vocabulary acquisition through viewing captioned and subtitled video: A scoping review and meta-analysis. *Systems*, 10(5), 133. <https://doi.org/10.3390/systems10050133>
- Sagarra, N., & Abuhl, R. (2013). Optimizing the noticing of recasts via computer-delivered feedback: Evidence that oral input enhancement and working memory help second language learning. *Modern Language Journal*, 97(2), 196–216. <https://doi.org/10.1111/j.1540-4781.2013.01427.x>
- Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, 13(1), 96–120. <http://hdl.handle.net/10125/44170>
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139524780.003>
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361–392. <https://doi.org/10.1191/1362168806lr203oa>
- Smith, B. (2003). The use of communication strategies in computer-mediated communication. *System*, 31(1), 29–53. [https://doi.org/10.1016/S0346-251X\(02\)00072-6](https://doi.org/10.1016/S0346-251X(02)00072-6)
- Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition*, 26(3), 365–398. <https://doi.org/10.1017/S027226310426301X>
- Smith, B., & Sauro, S. (2009). Interruptions in chat. *Computer Assisted Language Learning*, 22(3), 229–247. <https://doi.org/10.1080/09588220902920219>
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14, 65–86. <http://hdl.handle.net/10125/44203>
- Yilmaz, Y. (2012). The relative effects of explicit correction and recasts on two target structures via two communication modes. *Language Learning*, 62(4), 1134–1169. <https://doi.org/10.1111/j.1467-9922.2012.00726.x>
- Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research*, 15(4), 457–477. <https://doi.org/10.1177/1362168811412873>
- Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition*, 38(3), 553–586. <https://doi.org/10.1017/S027226311500025X>
- Ziegler, N. (2017). Task modality, noticing, and the contingency of recasts: Insights on salience from multiple modalities. In *Salience in second language acquisition* (pp. 269–290). Routledge.
- Ziegler, N. & Mackey, A. (2017). Interactional feedback in computer-mediated communication: A review and state of the art. In H. Nassaji and E. Kartchava (Eds.), *Corrective feedback in second language teaching and learning*. Routledge. <https://doi.org/10.4324/9781315621432>

Appendix A. Developmental stages in L2 English Question Formation with Data from the Present Study (stages adapted from Pienemann et al., 1988)

Stages	Constructions	Examples
Stage 1: Rising intonation	Single words	Why? This?
	Single units	My passport? What color?
Stage 2: Rising intonation	Declarative word order, no inversion, no fronting	This is your car?
Stage 3: Fronting	<i>Wh</i> -fronting	*Where your head office? *How long they stayed?
	Do-fronting	Do you have any questions? *Do your family came here?
	Fronting other followed by uninverted sentence	*Is the lab is open to everyone?
Stage 4: Inversion	<i>Yes/no</i> questions with auxiliary or copula	Is the lab recorded in real time? Have you experienced a similar incident?
	<i>Yes/no</i> questions with modal	Can you recall the situation?
	<i>Wh</i> -questions with copula (not auxiliary)	Where is your head office? *What are your position?
Stage 5: Inversion	Auxiliary (e.g., <i>is</i>) in 2 nd position	*What are you discover at the scene? Where were you trained?
	<i>Do</i> operator (e.g., <i>does/do</i>) in 2 nd position	When did you discover the scene? Who do you work for?
	Modal (e.g., <i>may</i>) in 2 nd position	Who will help you?
Stage 6: Complex question	Negative questions with <i>Do</i> operator	Why didn't you call the police?
	Tag questions	It's better, isn't it?
	Cancelled inversion?	Can you tell me where you come from?

Appendix B. Results of the Shapiro-Wilk Test of Normality

	OPT			WPT			SOPT		
	Statistics	df	Sig.	statistics	df	sig.	Statistics	df	Sig.
Control	.916	15	.165	.908	15	.128	.911	15	.141
Oral	.915	15	.159	.912	15	.148	.903	15	.105
Written	.904	15	.109	.893	15	.072	.900	15	.094
Multimodal	.896	15	.082	.897	15	.083	.887	15	.061

Appendix C. Results for the Ordinal Logistic Mixed-Effects Models Examining Participants' Scores on the Oral Production Test

Fixed effects	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>OR</i>	95% CI for OR
Control group as reference level						
3 4	-1.08	0.34	-3.15	.002**	0.33	[0.16, 0.66]
4 5	0.25	0.34	0.74	.461	1.63	[0.81, 3.27]
Oral	0.02	0.20	0.09	.926	1.02	[0.70, 1.48]
Written	-0.06	0.20	-0.31	.755	0.94	[0.65, 1.37]
Multimodal	-0.04	0.20	-0.20	.843	0.96	[0.66, 1.40]
Time2	0.16	0.20	0.74	.461	1.14	[0.78, 1.64]
Time3	0.10	0.20	0.53	.598	1.08	[0.74, 1.56]
ControlOral:Time2	0.58	0.28	2.04	.041*	1.78	[1.02, 3.10]
ControlWritten:Time2	1.23	0.29	4.25	<.001***	3.43	[1.94, 6.07]
ControlMulti:Time2	1.85	0.30	6.15	<.001***	6.33	[3.52, 11.40]
ControlOral:Time3	0.08	0.28	0.31	.755	1.12	[0.66, 1.88]
ControlWritten:Time3	0.50	0.28	1.77	.076	1.71	[1.01, 2.91]
ControlMulti:Time3	0.79	0.28	2.81	.005**	2.21	[1.27, 3.84]
Oral recast group as reference level						
3 4	-1.11	0.34	-3.21	.001**	0.32	[0.16, 0.64]
4 5	0.23	0.34	0.68	.494	1.60	[0.80, 3.22]
Control	-0.02	0.20	-0.09	.926	0.98	[0.68, 1.43]
Written	-0.08	0.20	-0.40	.686	0.93	[0.64, 1.34]
Multimodal	-0.06	0.20	-0.29	.771	0.95	[0.65, 1.37]
Time2	0.77	0.20	3.55	<.001***	2.04	[1.39, 2.99]
Time3	0.16	0.20	0.96	.335	1.20	[0.83, 1.74]
OralControl:Time2	-0.58	0.28	-2.04	.041*	0.56	[0.33, 0.95]
OralWritten:Time2	0.66	0.29	2.23	.026*	1.93	[1.02, 3.07]
OralMulti:Time2	1.27	0.30	4.18	<.001***	3.55	[2.18, 6.87]
OralControl:Time3	-0.09	0.28	-0.31	.755	0.90	[0.53, 1.51]
OralWritten:Time3	0.41	0.28	1.46	.144	1.53	[0.90, 2.61]
OralMulti:Time3	0.70	0.28	2.49	.013*	2.02	[1.16, 3.52]
Written recast group as reference level						
3 4	-1.02	0.34	-2.97	.001**	0.35	[0.17, 0.70]
4 5	0.31	0.34	0.92	.360	1.73	[0.86, 3.48]
Control	0.06	0.20	0.31	.755	1.06	[0.73, 1.54]
Oral	0.08	0.20	0.40	.686	1.08	[0.74, 1.57]
Multimodal	0.02	0.20	0.11	.910	1.02	[0.70, 1.49]
Time2	1.38	0.21	6.44	<.001***	3.61	[2.42, 5.38]

Time3	0.60	0.20	3.00	.003**	1.84	[1.26, 2.70]
WrittenControl:Time2	-1.23	0.29	-4.25	<.001***	0.31	[0.18, 0.54]
WrittenOral:Time2	-0.66	0.29	-2.23	.026*	0.56	[0.33, 0.98]
WrittenMulti:Time2	0.61	0.31	1.98	.048*	2.19	[1.22, 3.91]
WrittenControl:Time3	-0.50	0.28	-1.77	.076	0.58	[0.34, 0.99]
WrittenOral:Time3	-0.41	0.28	-1.46	.144	0.65	[0.38, 1.11]
WrittenMulti:Time3	0.29	0.28	1.03	1.03	1.31	[0.76, 2.24]
Random effects	<i>variance</i>		<i>SD</i>			
Participant (Intercept)	<.01		<.01			
Item (Intercept)	1.64		1.17			

Appendix D. Results for the Ordinal Logistic Mixed-Effects Models Examining Participants' Scores on the Written Production Test

Fixed effects	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>	<i>OR</i>	95% CI for OR
Control group as reference level						
3 4	−0.34	0.37	−0.92	.357	0.71	[0.35, 1.46]
4 5	1.25	0.37	3.35	<.001***	3.48	[1.68, 7.19]
Oral	0.12	0.27	0.46	.643	1.13	[0.73, 2.04]
Written	0.03	0.27	0.12	.901	1.03	[0.67, 1.90]
Multimodal	0.20	0.26	0.76	.450	1.22	[0.61, 1.74]
Time2	0.35	0.27	1.30	.193	1.41	[0.84, 2.37]
Time3	0.28	0.26	1.06	.289	1.32	[0.79, 2.22]
ControlOral:Time2	0.90	0.38	2.38	.017*	2.46	[1.18, 5.14]
ControlWritten:Time2	1.44	0.39	3.72	<.001***	4.21	[1.98, 8.95]
ControlMulti:Time2	2.41	0.41	5.82	<.001***	11.17	[4.98, 25.08]
ControlOral:Time3	0.52	0.37	1.38	.168	1.68	[0.81, 3.48]
ControlWritten:Time3	0.78	0.38	2.06	.039*	2.17	[1.04, 4.53]
ControlMulti:Time3	1.62	0.39	4.20	<.001***	5.06	[2.38, 10.75]
Oral recast group as reference level						
3 4	−0.46	0.37	−1.26	.207	0.63	[0.31, 0.64]
4 5	1.12	0.37	3.04	.002**	3.08	[1.49, 3.22]
Control	−0.12	0.27	−0.46	.643	0.88	[0.53, 1.48]
Written	−0.09	0.26	−0.34	.734	0.91	[0.55, 1.53]
Multimodal	0.08	0.26	0.29	.770	1.08	[0.65, 1.80]
Time2	1.25	0.27	4.61	<.001***	3.47	[2.05, 5.88]
Time3	0.80	0.27	3.00	.002**	2.22	[1.32, 3.73]
OralControl:Time2	−0.90	0.38	−2.38	.017*	0.41	[0.19, 0.85]
OralWritten:Time2	0.54	0.39	1.38	.167*	1.71	[0.80, 3.65]
OralMulti:Time2	1.51	0.42	3.64	<.001***	4.54	[2.02, 10.21]
OralControl:Time3	−0.52	0.37	−1.38	.168	0.60	[0.29, 1.24]
OralWritten:Time3	0.26	0.38	0.69	.490	1.30	[0.62, 2.70]
OralMulti:Time3	1.11	0.39	2.87	.004**	3.02	[1.42, 6.40]
Written recast group as reference level						
3 4	−0.37	0.37	−1.01	.310	0.69	[0.33, 1.41]
4 5	1.21	0.37	3.27	.001**	3.36	[1.63, 6.94]
Control	−0.03	0.27	−0.12	.901	0.97	[0.57, 1.63]
Oral	0.09	0.26	0.34	.734	1.09	[0.65, 1.83]
Multimodal	0.17	0.26	0.63	.527	1.18	[0.71, 1.97]
Time2	1.78	0.28	6.30	<.001***	5.94	[3.42, 10.32]

Time3	1.06	0.27	3.93	<.001***	2.88	[1.70, 4.86]
WrittenControl:Time2	−1.44	0.39	−3.72	<.001***	0.24	[0.11, 0.51]
WrittenOral:Time2	−0.54	0.39	−1.38	.167	0.58	[0.27, 1.25]
WrittenMulti:Time2	0.98	0.42	2.31	.002*	2.65	[1.16, 6.05]
WrittenControl:Time3	−0.78	0.38	−2.06	.039*	0.46	[0.22, 0.96]
WrittenOral:Time3	−0.26	0.38	−0.69	.490	0.77	[0.37, 1.61]
WrittenMulti:Time3	0.85	0.39	2.19	.029*	2.33	[1.09, 4.95]
Random effects	<i>variance</i>		<i>SD</i>			
Participant (Intercept)	<.01		<.01			
Item (Intercept)	.70		.84			

Appendix E. Results for the Linear Mixed-Effects Models Examining Participants' Scores on the Spontaneous Oral Production Test

	Fixed effects				Random effects	
	<i>Estimate</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>by participant</i> <i>variance</i>	<i>SD</i>
Control group as reference level						
(Intercept)	0.72	0.00	206.15	<.001***	<.01	<.01
Oral	0.00	0.00	0.80	.423		
Written	0.01	0.00	1.34	.423		
Multimodal	−0.00	0.00	−0.80	.182		
Time2	0.02	0.00	5.29	<.001***		
Time3	0.01	0.00	2.49	.010*		
ControlOral:Time2	0.03	0.01	5.72	<.001***		
ControlWritten:Time2	0.04	0.01	6.49	<.001***		
ControlMulti:Time2	0.09	0.01	15.08	<.001***		
GroupOral:Time3	0.03	0.01	4.29	<.001***		
GroupWritten:Time3	0.03	0.01	4.29	<.001***		
GroupMulti:Time3	0.08	0.01	12.76	<.001***		
Oral recast group as reference level						
(Intercept)	0.73	0.00	207.29	<.001***	<.01	<.01
Control	−0.00	0.00	−0.80	.423		
Written	0.00	0.00	0.54	.593		
Multimodal	−0.01	0.00	−1.61	.110		
Time2	0.06	0.01	13.38	<.001***		
Time3	0.04	0.01	8.56	<.001***		
OralControl:Time2	−0.03	0.01	−5.72	<.001***		
OralWritten:Time2	0.00	0.01	0.77	.443		
OralMulti:Time2	0.07	0.01	9.35	<.001***		
Oral:Time3	−0.03	0.01	−4.29	<.001***		
OralWritten:Time3	−0.00	0.01	0.00	1.00		
OralMulti:Time3	0.05	0.01	8.47	<.001***		
Written recast group as reference level						
(Intercept)	0.73	0.00	208.05	<.001***	<.01	<.01
Control	−0.01	0.00	−1.34	.182		
Oral	−0.00	0.00	−0.54	.593		
Multimodal	−0.01	0.00	−2.14	.034*		
Time2	0.06	0.00	14.47	<.001***		
Time3	0.04	0.00	8.56	<.001***		

WrittenControl:Time2	−0.04	0.01	−6.49	<.001***
WrittenOral:Time2	−0.01	0.01	−0.77	.443
WrittenMulti:Time2	0.05	0.01	8.58	<.001***
WrittenControl:Time3	−0.03	0.01	−4.29	<.001***
WrittenOral:Time3	−0.00	0.01	0.00	1.00
WrittenMulti:Time3	0.05	0.01	8.47	<.001***
<i>R</i> ² <i>m</i> : 0.87 <i>R</i> ² <i>c</i> : 0.91				

About the Authors

Yeonwoo Jung is an Honorary Senior Research Fellow at University College London. She holds a PhD in Applied Linguistics from University College London. Her research interests include second language acquisition, instruction, and testing with a focus on corrective feedback and feedback modalities in task-based synchronous computer-mediated communication (SCMC) environments.

E-mail: yeonwoo.jung.15@ucl.ac.uk

ORCID: <https://orcid.org/0000-0001-8708-5868>

Andrea Révész is a Professor of Second Language Acquisition at University College London. Her research interests lie in the areas of second language acquisition, instruction, and assessment. Her current research investigates the cognitive mechanisms underlying instructed second language performance and development, with a focus on individual differences and task-related factors.

E-mail: a.revesz@ucl.ac.uk

ORCID: <https://orcid.org/0000-0003-1093-4336>