APPLING
International Applied Linguistics Research Conference
https://doi.org/10.29140/9781763711662-9

Modality Matters: Eye-Tracking Evidence of Attentional Engagement in Second Language Feedback Processing

Yeonwoo Jung
Sogang University

Background

While recasts are widely recognized as beneficial for the development of second language (L2) (Li, 2010), their effectiveness in synchronous computer-mediated communication (SCMC), particularly videoconferencing, remains underexplored. A key affordance of videoconferencing is its capacity to deliver multimodal recasts, operationalized in this study as oral reformulations accompanied by synchronized on-screen captions (i.e., captioned oral recasts). This format draws on the sensory–modality framework of working memory (Baddeley et al., 2009), which highlights the benefits of distributing input across the auditory and visual channels. When aligned precisely, this dual input can enhance processing efficiency and reduce cognitive load (Mayer, 2014; Sweller, 2005). Empirical studies further support these claims. Reading-while-listening research shows that audio–text combinations foster consistent eye movements and deeper processing (Hui & Godfroid, 2025). Similarly, captioning research shows that textual support can enhance learners' attention to grammatical forms, aid speech segmentation, and improve comprehension and retention (e.g., Gass et al., 2019; Lee & Révész, 2020; Reynolds et al., 2022).

Multimodal recasts also combine the strengths of written and oral feedback. Written input offers visual salience and persistence—a record learners can reread (e.g., Smith, 2004; Yilmaz, 2012)—whereas oral feedback provides immediacy, prosodic cues, and turn adjacency, facilitating real-time comparison between the learner's output and the target form (Leeman, 2003; Long, 2007). Each modality, however, has limitations: oral recasts may be misinterpreted as

E-mail: yeonwooj@sogang.ac.kr

In Remy Decorte, Serge Bibauw & Glenn Stockwell (Eds.)
Proceedings of the Inaugural International Applied Linguistics Conference, 2025 (pp. 1–6). Castledown.

confirmation, and text-based feedback often lacks precise temporal alignment with the utterance it addresses. The hybrid nature of multimodal recasts mitigates these weaknesses, producing feedback that is both perceptually salient and interactionally contingent.

Despite these theoretical and empirical foundations, direct comparisons of multimodal versus unimodal recasts in video-based interaction are scarce (Jung & Révész, 2024). Consequently, little is known about how feedback modality shapes learners' visual attention to the target forms, and how such attentional engagement relates to subsequent grammatical gains during real-time video-based tasks. This study addresses these gaps by investigating whether multimodal recasts promote more focused visual attention and facilitate greater grammatical development than oral-only or text-only recasts, and whether the link between attention and learning varies depending on the feedback modality.

Research Questions

1. How does the recast modality affect L2 grammatical development during videoconferencing tasks?
2. How does the recast modality affect visual attention, and does the relationship between attention and grammar learning vary across modalities?

Methodology

This study employed a pretest–post-test–delayed post-test design with three treatment sessions. A total of 120 Korean learners of English as a foreign language were randomly assigned to one of four groups: oral recasts, written recasts, multimodal recasts (captioned oral), or no recasts (control). During the treatments, participants completed three videoconferencing tasks designed to elicit two target constructions: the third-person singular and the present progressive forms. Recasts were provided immediately following the learners' errors according to group conditions. Oral recasts were delivered live; written recasts appeared as real-time on-screen text; and multimodal recasts integrated both, with the written recast simultaneously displayed as a caption at the bottom of the screen during oral delivery.

Eye-tracking data were collected using a Tobii 60Hz tracker for the written and multimodal groups to examine visual attention to feedback. Five indices were analyzed: two reflecting early-stage processing (first-pass reading time and skipping rate) and three reflecting later and more effortful processing

(rereading time, total fixation duration, and number of visits) (Conklin et al., 2018; Godfroid, 2020).

Grammar development was assessed using three tasks: an oral production test, which, although untimed, induced greater perceived time pressure and primarily tapped procedural knowledge; a written production test, which allowed more planning time, thus engaging both declarative and procedural knowledge; and a spontaneous oral production test, which required the learners to use the target forms in a more communicative scenarios, thereby enhancing ecological validity. By utilizing different test modes, the design aimed to capture not only accuracy gains but also the depth and transferability of learning across modalities (DeKeyser, 2007).

Results and Discussion

Effects of Recast Modality on L2 Grammatical Development (Research Question 1)

Mixed-effects analyses revealed that the multimodal group significantly outperformed the unimodal and control groups across all three assessments in both the immediate and delayed post-tests. No significant differences emerged between the oral and written groups. These findings suggest that distributing feedback across the auditory and visual channels enhances processing and retention. The multimodal recasts may have amplified the feedback's salience by combining the clarity and durability of written input with the immediacy and prosodic cues of oral delivery. Moreover, the captioned format likely facilitated speech segmentation and form–meaning mapping by reinforcing phonological input with orthographic representations (Bisson et al., 2014).

Effects of Recast Modality on Visual Attention and the Attention–Learning Link (Research Question 2)

Eye-tracking data showed that the multimodal group exhibited significantly greater visual attention to the feedback than the written group, as evidenced by longer total fixation durations, greater first-pass and rereading times, more frequent visits, and lower skipping rates. Moreover, the attention–learning relationship was significantly stronger in the multimodal group than in the written group: participants who engaged in longer fixations, extended initial reading, and repeated reanalysis of the target forms achieved greater immediate and delayed post-test gains across all outcome measures. These findings suggest that multimodal recasts heightened the feedback's salience

and reduced the processing load, enabling learners to sustain attention for both initial detection and deeper reanalysis. Importantly, the emergence of a stronger attention–learning relationship in the multimodal condition indicates that multimodal input more effectively supported the consolidation and transfer of grammatical knowledge across modalities than written feedback alone.

Conclusion

Multimodal recasts (i.e., captioned oral reformulations) proved more effective than unimodal recasts in video-based synchronous computer-mediated communication, eliciting greater visual attention and facilitating more substantial grammatical development. Delivering the same correction synchronously through speech and text enhances salience, reduces cognitive load, and enables learners to align auditory and visual cues, thereby promoting deeper cognitive engagement with the target forms. The robust associations between visual attention and learning gains in the multimodal condition indicate that effective feedback should support both perceptual processing and cognitive integration. In line with the conference's theme, these findings suggest that multimodal recasts can leverage technological affordances while maintaining human-centered pedagogy: they sustain meaningful interaction and broaden access to interactive feedback. By engaging dual input channels, multimodal recasts offer flexible access to feedback, allowing learners to process input through the mode—auditory or visual—that best suits their individual abilities and learning conditions. This adaptability helps address challenges such as limited auditory proficiency, poor audio quality, or differences in cognitive style. In this way, multimodal feedback supports linguistically effective and socially inclusive learning in real-time, technology-mediated L2 classrooms.

References

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Psychology Press.

Bisson, M. J., van Heuven, W. J., Conklin, K., & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, *35*(2), 399–418. https://doi.org/10.1017/s0142716412000434

Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge University Press. https://doi.org/10.1017/9781108235636

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., 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. https://doi.org/10.64152/10125/44684

Godfroid, A. (2020). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. Routledge. https://doi.org/10.4324/9780429261631

Hui, B. and Godfroid, A. (2025), Listening, reading, or both? Rethinking the comprehension benefits of reading-while-listening. *Language Learning* (in press). https://doi.org/10.1111/lang.12721

Jung, Y., & Révész, A. (2024). Promoting grammatical development through multimodal digital recasts in videoconferencing tasks. *Language Learning & Technology*, *28*(1), 1–29. https://doi.org/10.64152/10125/73601

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

Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.

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

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

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

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press.

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