Investigating the Effects of Multimodal Digital Recasts on Second Language Development Through Video-conferencing Tasks

by

Yeonwoo Jung

A thesis submitted to the UCL Institute of Education

in fulfilment of the requirements for the degree of

Doctor of Philosophy

UCL Institute of Education

2023

I, Yeonwoo Jung confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Word count (exclusive of front matter, reference list and appendices):

91,729

ABSTRACT

Previous second language (L2) acquisition research has provided compelling evidence that recasts facilitate L2 development (Goo & Mackey, 2013; Li, 2010). Building upon this foundation, an increasing number of studies have explored recasts' potential to promote L2 grammatical proficiency in synchronous computer-mediated communication (SCMC) settings. However, most SCMC studies have focused on text-based mode, with limited exploration of video-based recasts, which primarily involved oral recasts. Consequently, the feedback affordances of video-conferencing, which can offer a variety of feedback modalities, have been understudied.

This study addressed this gap by comparing the effects of oral, written, and multimodal (combined oral and written) recasts on L2 grammatical development in video-conferencing tasks. It also examined how feedback modalities influence learners' visual attention allocation and their L2 development using eye-tracking technology. Additionally, the research explored the impact of learners' working memory (WM) on the relationship between feedback modalities and L2 learning, as well as attention allocation.

This thesis reports on two empirical studies. Study 1, involving 60 Korean ESL learners, compared the effects of recast modalities on *wh*-question formation. Study 2, with 120 Korean EFL learners, further explored how feedback modalities affected learners' attention to target structures (present progressive and third person singular) and their subsequent learning outcomes. The study also scrutinized the role of different WM functions in the impact of feedback modalities on attention allocation and L2 acquisition.

The results demonstrated the superiority of multimodal recasts over unimodal recasts in facilitating the use of target constructions in oral and written production, as well as in

spontaneous oral production tests in both studies. Multimodal recasts also proved more effective in directing learners' attention to target forms. Furthermore, the study revealed the prominent role of WM in attention allocation and learning gains, with this relationship varying depending on the feedback modality and WM functions.

IMPACT STATEMENT

The motivation of this research arises from the extensive use of video-conferencing in various language education contexts and the limited exploration of its potential advantages for L2 acquisition. This thesis endeavors to elucidate the relative effectiveness of different modalities (oral, written, multimodal) of a corrective feedback (CF) technique, recasts, during video-conferencing tasks in facilitating L2 grammatical development. It also provides insights into the relationship between learners' visual attention to target language features and their learning outcomes, as well as the influence of working memory (WM) capacity on attention allocation and L2 learning. This impact statement underscores the key contributions of the thesis for researchers, educators, learners, curriculum developers, and policymakers.

For researchers, multimodal recasts enable the investigation of the simultaneous use and comparison of various synchronous computer-mediated communication (SCMC) modalities during video-based interactions. This innovative CF approach differs from prior research that employed feedback in single modalities corresponding to the mode of interaction. The study highlights the benefits of incorporating both early and late eye-tracking measures to explore the relationship between visual attention, feedback modality, and learning gains. The study's focus on the impact of WM on initial morphosyntactic achievement among elementary-level learners also ensured a focused investigation free from confounding variables associated with advanced language skills, laying the groundwork for exploring the link between WM components and initial L2 acquisition in video-based SCMC.

While this project was conducted in laboratory settings, its findings hold implications for educators and curriculum developers, given the prevalent use of video-conferencing in classroom language education. The research demonstrates the superiority of multimodal

recasts compared to unimodal recasts, offering an additional CF tool for educators and encouraging curriculum developers to integrate multimodal CF into language-learning programs using video-conferencing. The study also sheds light on the relationship between WM and linguistic outcomes, emphasizing variations dependent on the feedback modalities and specific grammatical structures. For educators, understanding this interplay underscores the need for personalized approaches in language instruction, allowing them to tailor their teaching methods and strategies to ensure a more inclusive and effective learning environment. Curriculum developers can leverage this insight to design flexible and adaptive language-learning programs, creating curricula that cater to diverse needs and preferences.

Learners can also benefit from the inclusion of multimodal recasts in their language-learning process. The combination of oral and written feedback provides learners with a holistic learning experience, fostering a better understanding of grammatical features. Moreover, the research highlights that paying more attention to multimodal recasts leads to greater learning gains; learners are encouraged to actively engage with and pay close attention to the target features presented in multimodal feedback, which can contribute to a deeper understanding and retention of grammatical constructions.

For policymakers, the research encourages the incorporation of video-conferencing and multimodal CF into language-learning initiatives. Understanding the impact of different feedback modalities can inform policies related to language education and technology integration, ultimately leading to more effective and inclusive language-learning programs.

ACKNOWLEDGEMENT

The completion of this thesis would not have been attainable without the support of many people. I extend my sincere gratitude to all who contributed to my Ph.D. journey.

First and foremost, I am deeply grateful to my principal supervisor, Prof. Andrea Révész, for her expert guidance, encouragement, and inspiration throughout the entire process. She was an exceptional mentor and a perfect role model to me in every aspect. I consider myself genuinely fortunate and privileged to have completed my master's dissertation and Ph.D. thesis under Andrea's supervision. She has demonstrated the qualities of an outstanding supervisor, proving the significance of this role in nurturing not only academic excellence but also personal growth of a Ph.D. student. Andrea's expertise across a wide range of SLA fields and incisive comments were pivotal in shaping and enhancing the conceptual and methodological quality of this thesis. Her generosity in sharing her knowledge and experience and her endless passion for motivating and supporting her students also made her a remarkable supervisor. Furthermore, her unwavering support has motivated me to overcome every challenge I faced during this journey. She guided me to achieve progress beyond my expectations, and her patience and belief in me dispelled my self-doubt, making me realize my potential and bring out the best in me. Indeed, she has been such a constant source of motivation throughout my entire academic life. Thank you so much, Andrea, I am forever grateful for being your student.

I would also like to express my sincere appreciation to my secondary supervisor, Dr. Ana Pellicer-Sánchez. Her insightful answers and thoughtful suggestions to my inquiry were essential in improving the quality of this thesis. In particular, her extensive expertise and wealth of experience in eye-tracking research significantly influenced how I conducted my

experiment and interpreted my findings. I am grateful to Dr. Effrosyni Argyri and Dr. Kazuya Saito for their thorough review of my upgrade document. Their constructive feedback was truly instrumental in enriching my work.

My sincere appreciation must go to the Viva examiners, Prof. Bryan Smith and Dr. Jim McKinley for their perceptive comments and inspiring suggestions during the viva. Their thought-provoking questions and stimulating discussion were truly enlightening and motivating. Their constructive feedback and suggestions assisted me in refining my arguments and engaging in an extended discussion, enabling me to gain diverse perspectives on my research and contemplate my future research directions more deeply. I also appreciate their generosity and kindness, which made the entire process of my viva such an enjoyable and memorable experience.

I also want to thank all members of the London Second Language Acquisition Research Forum (L-SLARF). Over the past few years, I have been lucky to receive constructive and insightful feedback on my ongoing research through my presentations at the conferences and seminars of this organization. Their feedback has significantly improved and enhanced my research design and approaches. I am also thankful for their interest in my work and the members' generosity in sharing their expertise.

I would also like to extend my special thanks to Dr. Jookyoung Jung and Dr. John Rogers for their thoughtful comments and invaluable assistance with different parts of my thesis, which greatly contributed to improving my research design. My appreciation also goes to Tobii Pro Korea manager Jongpil Joo for his technical assistance and valuable advice, which significantly contributed to the successful completion of the eye-tracking experiments. I am also grateful to all my dear participants, who generously shared their time and showed their genuine interest in my research. I am also truly grateful to the faculty members of

Institute of Education, University College London, for providing me with immense academic and professional support, which greatly contributed to the smooth completion of my study.

My sincere appreciation also must go to Language Learning and The International Research Foundation for English Language Education (TIRF) for their generous financial support of the research projects reported in this thesis. These grants allowed me to focus on my study and continue my academic endeavors, inspiring me to help others and give back to the community. I am also appreciative for the grant provided by the Institute of Education, University College London, which enabled my attendance at the 2022 annual conference of the American Association for Applied Linguistics (AAAL).

My heartfelt gratitude goes to my dearest grandmother, Myoung Ok Lee, my wonderful parents, Jinwoo Jung, Jongai Lee, my brother Sang Yeol Jung and his family. My cousins Monica Lee, Jiwon Park, Christina, and Kyoungja Park's family for their unwavering support throughout the process of writing my thesis. There are no words to express my appreciation for what my family has done for me during my Ph.D. journey. This thesis is dedicated to my family. I am truly blessed to have such a loving family who believes in me unconditionally. With their endless support, prayers, and love, I was able to pursue my research endeavors. I would also like to thank my Ph.D. peers, who has been there for me emotionally and intellectually as I worked toward my Ph.D. Last but not least, I thank God for everything.

DISSEMINATION

Conference Presentations

1. *Investigating the effects of multimodal digital feedback on second language development through video-conferencing tasks: An eye-tracking study* [Paper presentation]. London Second Language Acquisition Research Forum (L-SLARF), London, UK, January 2020.
2. *Exploring the effects of multimodal digital feedback on second language development through video-conferencing tasks* [Paper presentation]. the Annual International Conference on Education, Waseda University, Japan, December 2020.
3. *Investigating the effects of multimodal digital feedback on second language development through video-conferencing tasks* [Paper presentation]. American Association for Applied Linguistics (AAAL) Virtual Conference, March 2021.
4. *Exploring the effects of digital recasts on second language development through video-conferencing tasks* [Paper presentation]. London Second Language Acquisition Research Forum (L-SLARF), London, UK, January 2021.
5. *Investigating the effects of multimodal digital recasts on second language development through video-conferencing tasks* [Paper presentation]. American Association for Applied Linguistics (AAAL), Pittsburgh, US, March 2022.

Table of Contents

ABSTRACT....3
IMPACT STATEMENT....5
ACKNOWLEDGEMENT....7
DISSEMINATION....10
Table of Contents....11
Llst of Tables....17
List of Figures....23
List of Abbreviations....25
CHAPTER 1. INTRODUCTION....26
1.1. Rationale for the Thesis....26
1.2. Aims of the Thesis....39
1.3. Structure of the Thesis....41
CHAPTER 2. LITERATURE REVIEW....44
2.1. Role of Input in Second Language Acquisition....44
2.2. Attention and Awareness in SLA....46
2.2.1. Schmidt's Noticing Hypothesis....46
2.2.2. Tomlin and Villa's Model of Attention....48
2.2.3. Robinson's Model of Attention and Memory....49
2.2.4. Leow's Model of the L2 Learning Process in Instructed SLA....51
2.2.5. Summary....53
2.3. The Role of Interaction, Output and Corrective Feedback in SLA....54
2.3.1. Interaction Hypothesis....54
2.3.2. Output....54
2.3.3. Corrective Feedback....56
2.4. Recasts....62
2.4.1. Recasts in L2 Acquisition....62
2.4.2. Recasts in SCMC Research....68
2.4.3. Summary....78
2.5. Multimodal Recasts....84

2.5.1. Theoretical Support for Multimodal Recasts 84
2.5.1.1. Dual Coding Theory 84
2.5.1.2. Cognitive Theory of Multimedia Learning 86
2.5.1.3. Cognitive Load Theory 88
2.5.2. Benefits of Hybrid Characteristics of Multimodal Recasts 90
2.5.3. Empirical Research on Captioned Audiovisual Input in L2 Grammar 92
2.5.4. Possibility of Redundancy Effects of Audiovisual Input 97
2.5.5. Summary 99
2.6. Eye-Tracking 100
2.6.1. Eye-tracking Measures and Eye-tracking Research on L2 Grammar in SLA 101
2.6.2. SCMC Eye-tracking Research on L2 Grammar 107
2.6.3. Summary 116
2.7. Working Memory 117
2.7.1. Baddeley's Working Memory Model 118
2.7.2. Wickens' Multiple Resource Model 122
2.7.3. Working Memory Measures 125
2.7.4. Working Memory, Recasts and Learning of L2 Grammar in SLA 128
2.7.5. Working Memory and Recasts in SCMC research 147
2.7.6. Working Memory in Eye-tracking Research 164
2.7.7. Summary 170
CHAPTER 3. STUDY 1 172
3.1. Methodology 173
3.1.1. Overall Design 173
3.1.2. Research Ethics 173
3.1.3. Participants 174
3.1.4. Target Construction 175
3.1.5. Materials 176
3.1.5.1 General English Proficiency Test 176
3.1.5.2. Experimental Treatment Materials 176
3.1.5.3. Assessment Tasks 180
3.1.5.4. Background Questionnaire 182
3.1.5.5. Exit Questionnaire 183

3.1.6. Data Collection Procedure 183
3.1.7. Coding and Scoring 185
3.1.8. Statistical Analyses 186
3.2. Analyses and Results 189
3.2.1. Preliminary Results 189
3.2.1.1. Comparability of the Group 189
3.2.2. Main Analyses 192
3.2.2.1. OPT 192
3.2.2.2. WPT 195
3.2.2.3. SOPT 198
3.2.3. Summary of Results 200
3.3. Interim Discussion 201
3.3.1. Effects of Recasts on L2 Grammatical Development (RQ1) 201
3.3.2. Effects of Feedback Modality on L2 Grammatical Development (RQ2) 203
3.3.3. Limitations of Study 1 and Modifications in Study 2 207
CHAPTER 4. STUDY 2 210
4.1. Research Questions 211
4.2. Methodology 211
4.2.1. Overall Design 211
4.2.2. Participants 212
4.2.3. Research Ethics 213
4.2.4. Target Constructions 214
4.2.4.1 English Third Person Singular 214
4.2.4.2. English Present Progressive 217
4.2.5. Experimental Treatment Tasks 220
4.2.6. Type of Recasts 222
4.2.7. Materials 224
4.2.7.1. General Proficiency Test 224
4.2.7.2. Assessment Tasks 224
4.2.7.2.1. OPT 225
4.2.7.2.2. WPT 228
4.2.7.2.3. SOPT 229

4.2.7.3. Working Memory Measures 231
4.2.7.3.1. NWS Task 232
4.2.7.3.2. Forward Corsi Block Task 232
4.2.7.3.3. Color Shape Task 233
4.2.7.3.4. Stop Signal Task 235
4.2.7.3.5. AOSPAN Task 236
4.2.7.4. Background Questionnaire 238
4.2.7.5. Exit Questionnaire 238
4.2.8. Eye-tracking Procedure 239
4.2.9. Data Collection Procedure 239
4.2.10. Scoring and Data Analysis 241
4.2.10.1. Scoring 241
4.2.10.1.1. Oxford Placement Test 241
4.2.10.1.2. OPT, WPT, and SOPT 242
4.2.10.2. Eye-movement Data Analysis 243
4.2.10.3. Working Memory Measures 246
4.2.10.3.1. NWS Task 246
4.2.10.3.2. Forward Corsi Block Task 246
4.2.10.3.3. Color Shape Task 246
4.2.10.3.4. Stop Signal Task 247
4.2.10.3.5. AOSPAN Task 247
4.2.11. Statistical Analyses 248
4.2.11.1. Mixed-Effects Models in Study 2 251
4.3. Analyses and Results 256
4.3.1. Preliminary Analyses 256
4.3.1.1. Internal Consistency Reliability 256
4.3.1.2. Inter-coder Agreement 256
4.3.1.3. Number of Obligatory Contexts and Recasts 257
4.3.1.4. Comparability of the Groups 260
4.3.2. Effects of Recasts and Feedback Modality on Development in L2 Grammatical Knowledge (RQs 1–2) 265
4.3.2.1. OPT 265

4.3.2.1.1. Present Progressive 265
4.3.2.1.2. Third Person Singular 267
4.3.2.2. WPT 271
4.3.2.2.1. Present Progressive 271
4.3.2.2.2. Third Person Singular 273
4.3.2.3. SOPT 276
4.3.2.3.1. Present Progressive 276
4.3.2.3.2. Third Person Singular 277
4.3.3. Effects of Multimodal Recasts and Written Recasts on Learners' Attention Allocation (RQ3) 281
4.3.3.1. Present Progressive 285
4.3.3.2. Third Person Singular 293
4.3.4. Relationship Between Attention and L2 Grammatical Development (RQ4) 302
4.3.4.1. OPT 303
4.3.4.1.1. Present Progressive 303
4.3.4.1.2. Third Person Singular 305
4.3.4.2. WPT 308
4.3.4.2.1. Present Progressive 308
4.3.4.2.2. Third Person Singular 310
4.3.4.3. SOPT 313
4.3.4.3.1. Present Progressive 313
4.3.4.3.2. Third Person Singular 314
4.3.5. Effects of WM on Attention Allocation and L2 Grammatical Knowledge Development (RQ5) 317
4.3.5.1. Present Progressive 320
4.3.5.2. Third Person Singular 323
4.3.6. Relationship Between Attention and WM (RQ6) 326
4.3.6.1. Present Progressive 326
4.3.6.2. Third Person Singular 327
4.3.7. Summary of Results 329
4.3.8. Interim Discussion 335
4.3.8.1. Effects of Recasts on L2 Grammatical Development (RQ1) 335

4.3.8.2. Effects of Feedback Modality on L2 Grammatical Development (RQ2) 338
4.3.8.3. Effects of Modality of Recasts on Allocation of Visual Attention (RQ3) 339
4.3.8.4. Relationship Between Attention and L2 Development (RQ4) 344
4.3.8.5. Moderating Effects of WM on L2 Grammatical Knowledge Development (RQ5).....348
4.3.8.6. Relationship Between WM and Attention (RQ6).....357
CHAPTER 5. GENERAL DISCUSSION AND CONCLUSION.....364
5.1. Effects of Recasts (RQ1) and Recast Modalities (RQ2) on L2 Grammatical Development – Study 1 and Study 2.....365
5.2. Effects of Feedback Modality on Learners' Visual Attention Allocation (RQ3) and Relationship Between Attention and L2 Grammatical Development (RQ4) – Study 2 366
5.3. Moderating Effects of WM on L2 Grammatical Development (RQ5) – Study 2 368
5.4. Moderating Effects of WM on Attention Allocation (RQ6) – Study 2 371
5.5. Implications of the Study.....373
5.5.1. Theoretical Implications.....373
5.5.2. Methodological Implications.....378
5.5.3. Pedagogical Implications 380
5.6. Limitations and Future Research Directions.....382
REFERENCES 385
APPENDICES 445

List of Tables

Table 1. Taxonomy of CF Strategies (adapted from Lyster & Ranta, 1997)..... 58

Table 2. Classification of CF Strategies (Ellis, 2012) 59

Table 3. Definitions of Recasts..... 63

Table 4. Studies on the Role of CALL/SCMC Feedback Modality in L2 Learning 80

Table 5. Definitions of Eye-tracking Measures Commonly Used in Video-based Studies (Adapted from Conklin & Pellicer-Sánchez, 2016; Conklin, Pellicer-Sánchez, & Carroll, 2018; Godfroid, 2020)

..... 104

Table 6. Studies on the Role of WM in the Effect of Interactional Interventions 141

Table 7. CALL/SCMC Studies on the Role of WM in L2 Learning..... 158

Table 8. Scoring System for *Wh*-Question Formation..... 186

Table 9. Descriptive Statistics for Obligatory Contexts and Number of Recasts per Group

..... 189

Table 10. Descriptive Statistics for the Pretest 190

Table 11. Results for the Ordinal Logistic Mixed-Effects Model Examining Performance on the OPT Pretest..... 190

Table 12. Results for the Ordinal Logistic Mixed-Effects Model Examining Performance on the WPT Pretest..... 191

Table 13. Results for the Linear Mixed-Effects Models Examining Performance on the SOPT Pretest 191

Table 14. Descriptive Statistics for the OPT Scores..... 193

Table 15. Results of Post Hoc Contrasts for Significant Interactions Identified by Ordinal Logistic Mixed−Effects Models Examining Participants′ Scores on the OPT 194

Table 16. Descriptive Statistics for the WPT Scores..... 196

Table 17. Results of Post Hoc Contrasts for Significant Interactions Identified by Ordinal Logistic Mixed-effects Models Examining Participants' Scores on the WPT 197

Table 18. Descriptive Statistics for the SOPT 199

Table 19. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Participants' Scores on the SOPT 199

Table 20. Descriptive Statistics for Participants' Performance on the Oxford Placement Test 213

Table 21. Scoring System for Present Progressive 243

Table 22. Scoring System for Third Person Singular 243

Table 23. The Definitions of Five Eye-Tracking Measures for Study 2 244

Table 24. Average AOI Sizes of Each Target Construction in Each Treatment Session 245

Table 25. Internal Consistency Reliability 256

Table 26. Inter-coder Agreement on Three Assessment Tasks 257

Table 27. Descriptive Statistics for the Number of Obligatory Contexts Across Treatment Sessions 258

Table 28. Descriptive Statistics for the Number of Recasts Across Treatment Sessions 259

Table 29. Results of One-way ANOVAs on the Number of Obligatory Contexts and Recasts Across Treatment Sessions 260

Table 30. Descriptive Statistics for the Pretest – Present Progressive 261

Table 31. Descriptive Statistics for the Pretest – Third Person Singular 261

Table 32. Results for the Linear Mixed-effects Model Examining Performance on the OPT Pretest – Present Progressive 262

Table 33. Results for the Linear Mixed-effects Model Examining Performance on the WPT Pretest – Present Progressive 263

Table 34. Results for the Linear Mixed-effects Model Examining Performance on the SOPT Pretest – Present Progressive 263

Table 35. Results for the Linear Mixed-effects Model Examining Performance on the OPT Pretest – Third Person Singular 264

Table 36. Results for the Linear Mixed-effects Model Examining Performance on the WPT Pretest – Third Person Singular 264

Table 37. Results for the Linear Mixed-effects Model Examining Performance on the SOPT Pretest – Third Person Singular 265

Table 38. Descriptive Statistics for the OPT Scores – Present Progressive 266

Table 39. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed−Effects Models Examining Participants′ Scores on the OPT – Present Progressive 267

Table 40. Descriptive Statistics for the OPT Scores – Third Person Singular 268

Table 41. Results of Post Hoc Contrasts for Significant Interactions Identified by Logistic Mixed−Effects Models Examining Participants′ Scores on the OPT – Third Person Singular 270

Table 42. Descriptive Statistics for the WPT Scores – Present Progressive 271

Table 43. Results of Post Hoc Contrast for Significant Interactions Identified by Linear Mixed−Effects Models Examining Participants′ Scores on the WPT – Present Progressive 272

Table 44. Descriptive Statistics for the WPT Scores – Third Person Singular 273

Table 45. Results of Post Hoc Contrast for Significant Interactions Identified by Logistic Mixed-effects Models Examining Participants' Scores on the WPT – Third Person Singular 275

Table 46. Descriptive Statistics for the SOPT Scores – Present Progressive 276

Table 47. Results of Post Hoc Contrast for Significant Interactions Identified by Linear Mixed−Effects Models Examining Participants' Scores on the SOPT – Present Progressive 277

Table 48. Descriptive Statistics for the SOPT Scores – Third Person Singular 278

Table 49. Significant Interactions Identified by Logistic Mixed-effects Models Examining Participants' Scores on the SOPT – Third Person Singular

.....280

Table 50. Correlation Analysis of Five Attention Measurements – Present Progressive.....283

Table 51. Correlation Analysis of Five Attention Measurements – Third Person Singular 284

Table 52. Descriptive Statistics for Attention Measurements – Present Progressive.....286

Table 53. Results for the Linear Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Present Progressive 288

Table 54. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Present Progressive.....291

Table 55. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Present Progressive.....293

Table 56. Descriptive Statistics for Attention Measurements – Third Person Singular 295

Table 57. Results for the Linear Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Third Person Singular 297

Table 58. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction (Visit Count) – Third Person Singular 300

Table 59. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction (Skip Rate) – Third Person Singular 301

Table 60. Results for the Linear Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the OPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Present Progressive.....304

Table 61. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the OPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Third Person Singular – Third Person Singular.....306

Table 62. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the OPT (Visit count and Skip rate) – Third Person Singular 308

Table 63. Results for the Linear Mixed-Effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the Written Production Test – Present Progressive 310

Table 64. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the WPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Third Person Singular .. 311

Table 65. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the WPT (Visit Count and Skip Rate) – Third Person Singular 312

Table 66. Results for the Linear Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the SOPT – Present Progressive 314

Table 67. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the SOPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Third Person Singular 315

Table 68. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the SOPT (Visit count and Skip rate) – Third Person Singular 317

Table 69. Descriptive Statistics for WM Measures 318

Table 70. Results of One-way ANOVAs on Five WM Measures 318

Table 71. Spearman Correlation Matrix of Five WM Measures 319

Table 72. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of SOPT (NWS) – Present Progressive 321

Table 73. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (AOSPAN) – Present Progressive 322

Table 74. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of WPT (AOSPAN) – Present Progressive 322

Table 75. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of SOPT (AOSPAN) – Present Progressive 323

Table 76. Significant Interactions Identified by Logistic Mixed-effects Models Examining Participants' Scores on the Spontaneous Oral Production Test – Third Person Singular (Stop Signal) 325

Table 77. Correlation Analyses Between Working Memory and Eye-tracking Measures – Present Progressive 327

Table 78. Correlation Analyses Between Working Memory and Eye-tracking Measures – Third Person Singular 329

List of Figures

Figure 1. Robinson's Model of Memory and Attention (adapted from Robinson, 2003) 50
Figure 2. Dual Coding Theory (Paivio, 1986, p. 67) 85
Figure 3. Cognitive Theory of Multimedia Learning (Mayer, 2014, p. 52) 87
Figure 4. Eye-movement Measurements (Roberts & Siyanova-Chanturia, 2013, p. 218) 103
Figure 5. Multicomponent Model of WM (Baddeley, 2010, p. 138) 119
Figure 6. The Multiple Resource Model (Wickens, 2002) 122
Figure 7. An Example of a Lost Dog Poster 178
Figure 8. An Example of Question Prompts 178
Figure 9. An Example of a Written Recasts 180
Figure 10. Examples of Assessment Tasks (SOPT, OPT, WPT in this order) 182
Figure 11. Data Collection Procedures for Study 1 184
Figure 12. An Example of Experimental Treatment Tasks 222
Figure 13. Examples of Written Recasts for Each Target Constructions 222
Figure 14. An Example of the OPT 227
Figure 15. An Example of Video Footage and a Prompt for the Present Progressive 227
Figure 16. An Example of the WPT 229
Figure 17. An Example of the SOPT 231
Figure 18. Forward Corsi Block Task 233
Figure 19. Color Shape Task – Non-switching Block 234
Figure 20. Color Shape Task – Switching Block 235
Figure 21. Stop Signal Task 236
Figure 22. AOSPAN Task 237
Figure 23. Data Collection Procedures for Study 2 241
Figure 24. Examples of AOIs of Each Target Construction 245

Figure 25. Interaction Plot with Logged Means of Total Fixation Duration for Each Group (Present Progressive) 289
Figure 26. Interaction Plot with Logged Means of First-pass Reading Time for Each Group (Present Progressive) 289
Figure 27. Interaction Plot with Logged Means of Re-reading Time for Each Group (Present Progressive) 290
Figure 28. Interaction Plot with Logged Means of Visit Count for Each Group (Present Progressive) 292
Figure 29. Interaction Plot with Logged Means of Skip Rate for Each Group (Present Progressive) 293
Figure 30. Interaction Plot with Means of Total Fixation Duration for Each Group (Third Person Singular) 298
Figure 31. Interaction Plot with Means of First-pass Reading Time for Each Group (Third Person Singular) 298
Figure 32. Interaction Plot with Means of Re-reading Time for Each Group (Third Person Singular) 299
Figure 33. Interaction Plot with Mean of Visit Count for Each Group (Third Person Singular) 300
Figure 34. Interaction Plot with Means of Skip Rate for Each Group (Third Person Singular) 302

List of Abbreviations

AOI	Area of Interest
AOSPAN	Automated Operation Span
CALL	Computer-Assisted Language Learning
CEFR	Common European Framework of Reference for Languages
CF	Corrective Feedback
CMC	Computer-Mediated Communication
EFL	English as a Foreign Language
ESL	English as a Second Language
GJT	Grammaticality Judgement Test
IP	Interest Period
L1	First Language
L2	Second Language
NWS	Nonword Span
OPT	Oral Production Test
OR	Odds Ratio
PSTM	Phonological Short-Term Memory
SCMC	Synchronous Computer-Mediated Communication
SLA	Second Language Acquisition
SOPT	Spontaneous Oral Production Test
TOI	Time of Interest
VSTM	Visuospatial Short-Term Memory
WM	Working Memory
WPT	Written Production Test

CHAPTER 1. INTRODUCTION

1.1. Rationale for the Thesis

Interactionist perspectives on second language acquisition (SLA) (Gass, 1997; Gass & Mackey, 2007, 2015; Long, 1996; Mackey, 2007, 2012; Pica, 1994) suggest that drawing learners' attention to linguistic properties while engaging in meaningful interaction benefits second language (L2) development. One effective way to direct learners' attention to L2 linguistic constructions during interaction is to provide them with corrective feedback (CF), which refers to "any indication to the learners that their use of the target language is incorrect" (Lyster et al., 1999, p. 171). SLA scholars have claimed that CF promotes L2 acquisition as it can draw learners' attention to their erroneous production (Gass, Mackey, & Pica, 1998; Lightbown & Spada, 1990; Long, 1990, 1996; Schachter, 1986, 1991; Schmidt & Frota, 1986), thereby increasing the likelihood that they notice the gap between the target and their interlanguage form (Schmidt & Frota, 1986). This view is supported by a number of meta-analyses and narrative reviews, which concluded that CF facilitates L2 development (Brown, 2016; Goo et al., 2015; Li, 2010; Lyster & Saito, 2010; Lyster, Saito, & Sato, 2013; Mackey & Goo, 2007; Nassaji, 2016; Russell & Spada, 2006; Sheen, 2010).

To achieve a more detailed understanding of the role of CF in instructed SLA, considerable research effort has been devoted to investigating the effects of different types of CF on L2 development (see Plonsky & Brown, 2015 for meta-analysis). Of various feedback techniques, recasts have been identified as the most frequent feedback strategy used in a wide range of instructional settings (e.g., Bower & Kawaguchi, 2011; Lee, 2013; Lee & Tan, 2022; Li, 2010; Loewen & Philp, 2006; Lyster & Mori, 2006; Lyster & Ranta, 1997; Mackey, Gass, & McDonough, 2000; Sheen, 2004; Li & Vuono, 2019), leading to a keen interest among

SLA researchers to explore their potential benefits. Recasts constitute a partial or full reformulation of an erroneous piece of learner output into a target-like utterance or written form while retaining its central meaning (Long, 1996, 2007; Long, Inagaki, & Ortega, 1998), as in the example below taken from the current study:

Learner: Now, she dancing.

Researcher: She is dancing.

A number of theoretical reasons have been proposed to explain why receiving recasts may be beneficial for L2 learners. One of the main advantages is that recasts have the capacity to provide learners with both positive evidence (i.e., the provision of the correct form) and negative evidence (i.e., a signal indicating the occurrence of an error) about the target features (Leeman, 2003). The facilitative role of positive evidence and negative evidence in L2 acquisition is supported by the Direct Contrast Hypothesis and the Universal Grammar approach for language acquisition respectively. According to the direct contrast hypothesis (Saxton, 1997, 2000), one essential function of recasts is to serve as negative evidence, since it raises learners' awareness of their error through the immediate contrast of their ill-formed production with the correct alternative. The universal grammar approach, on the other hand, suggests that the effectiveness of recasts can be attributed to their role as positive evidence in that the appropriate correct form is presented in the input in an accessible way (Nicholas et al., 2001). Although these two frameworks exhibit different views regarding the explanation for the utility of recasts in language learning, they highlight two key characteristics of recasts. First, they can promote cognitive comparisons between the target form and a problematic utterance, which may ultimately lead to the rejection of the erroneous form (i.e., in the case of

negative evidence). Second, they provide a correct alternative, thereby enhancing the likelihood that learners notice the target-like norm and potentially incorporate it into their developing grammar (i.e., in the case of positive evidence) (Leeman, 2003).

Another noteworthy benefit of recasts lies in their contingency (e.g., Long, 2007), which means that recasts immediately follow the erroneous learner output. The contingency of recasts has been claimed to assist learners in noticing (Schmidt, 1990, 2001) discrepancies between their non-target-like production and target-like models by juxtaposing the erroneous learner output and the correct form (Doughty, 2001; Gass & Mackey, 2006; Long, 1996, 2007, 2015). This, in turn, may prompt learners to refine and restructure their interlanguage representation of the linguistic target, eventually leading to the acquisition of the target-like form-meaning connection. Moreover, recasts provided immediately after the learner's problematic utterances can create propitious conditions for L2 acquisition, in that the meaning of the feedback (at least part of the message) has already been comprehended by the learners, and this may allow more attentional resources to be directed to the linguistic target in recasts, thereby fostering form-function mapping (e.g., Long, 1996, 2007, 2015; Doughty, 2001). Empirical evidence has also demonstrated that contingent recasts are better noticed than non-contingent ones (Lai et al., 2008), providing further justification that contingency may play a significant role in facilitating the noticing of recasts. It is worth noting, however, that non-contingent recasts which are not presented immediately after learners have committed an error are less likely to trigger a cognitive comparison, as learners may not be able to store and process the information long enough in their working memory (WM). Considering that "verbal short-term memory is limited in the duration of storage as well as the number of items" (Cowan, 1988, p. 166), recasts provided in a non-contingent fashion are more likely to remain unnoticed (Lai et al., 2008).

Last but not least, recasts can promote learner attention to L2 features, allowing learners to focus on meaning and communication without interrupting the natural flow of conversation, unlike explicit corrections such as metalinguistic feedback or feedback techniques requiring learners to correct their erroneous production before continuing with the conversation. Thus, recasts are likely to create optimal conditions for establishing and developing L2 form-meaning mappings (e.g., Doughty, 2001; Gass & Mackey, 2006; Long, 2007).

In line with the aforementioned theoretical expectations, there is extensive evidence in the SLA literature supporting the effectiveness of recasts in fostering L2 development. This is well documented in a number of narrative reviews and meta-analyses (e.g., Goo & Mackey, 2013; Li, 2010; Mackey & Goo, 2007; Nassaji, 2015, 2016; Rassaei, 2022; Russell & Spada, 2006). Of particular relevance to the present study is the research that has found a positive relationship between recasts and the development of L2 grammatical features, including question formation (e.g., Ellis & Sheen, 2006; Li & Iwashita, 2019, 2021; Loewen & Nabei, 2007; Mackey et al., 2002; Mackey & Philp, 1998; Mackey & Sachs, 2012; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003; Sheen, 2004; Zhao & Wang, 2016); third person singular (e.g., Zhao, 2015; Zhao & Ellis, 2020), and present progressive tense (Lyster & Mori, 2006; Mackey & Gass, 2005).

Inspired by the positive findings for recasts in the broader field of SLA as well as pedagogical and technological advances using synchronous computer-mediated communication (SCMC) applications, an increasing number of scholars have sought to explore the potential of recasts in facilitating L2 development in different types of SCMC contexts. So far, this line of research has primarily been concerned with investigating recasts during text-based mode (Cerezo, 2021; Ziegler, 2016; Ziegler & Mackey, 2017). Researchers

have compared the developmental benefits of orally or textually enhanced versus non-enhanced recasts (Sachs & Suh, 2007; Sagarra & Abbuhl, 2013), contingent versus non-contingent recasts in learner noticing (Lai et al., 2008), recasts versus more explicit feedback techniques (Loewen & Erlam, 2006; Sauro, 2009; Yilmaz, 2012), levels of awareness in relation to type of recast and linguistic item (Gurzynski-Weiss et al., 2015), recasts provided during tasks with varying cognitive complexity (e.g., Baralt, 2010, 2015), feedback timing (e.g., Arroyo & Yilmaz, 2018; Kourtali & Borges, 2023), and recasts supplied in response to more and less salient linguistic features (Yilmaz & Yuksel, 2011; Yilmaz, 2012). Some studies have also contrasted the effectiveness of recasts in SCMC versus face-to-face contexts (e.g., Baralt, 2010, 2015; Kourtali, 2022; Yilmaz & Yuksel, 2011). Yilmaz and Yuksel (2011), for instance, found that participants showed greater gains in the acquisition of morphological features when they received recasts during text-chat interaction compared to face-to-face mode. Similarly, Yilmaz (2012) investigated the relative effects of recasts and explicit correction on L2 grammar learning in text-based and face-to-face environments, observing that both feedback types led to greater gains in the SCMC mode.

A small amount of research has also begun to explore the extent to which oral recasts may promote development in L2 grammatical accuracy during video-conferencing. In Monteiro's (2014) experiment, learners received recasts, metalinguistic feedback or no feedback while completing focused tasks during video-based interactions. The study yielded no difference in terms of grammatical development under the three conditions. In a more recent study, Rassaei (2017) has looked into whether recasts supplied in video-based or face-to-face interactions would result in a greater learning outcome of the target structure. The results suggested that modality of recasts did not influence the extent of L2 gains displayed by the learners. Rassaei (2019) also examined the effects of recasts on English as a foreign

language (EFL) learners' grammatical development in the use of English articles during mobile-mediated audio and video interactions. The results revealed that recasts had a positive impact on participants' accuracy development both in audio-based and video-based interactions, and that participants in the audio recast group produced more target-like modified utterances compared to those of the video recast group.

While the majority of SCMC research has focused on text-based feedback, investigating other modes of computer-mediated CF, particularly video-based recasts, has received less attention. Also, existing studies of video-based feedback including recasts have employed oral mode only, yielding mixed results. This highlights the need for further research examining whether and to what extent other SCMC modalities would mediate the effects of recasts on L2 learning in the context of multimodal video-conferencing SCMC.

One of the potential strengths of video-conferencing is that video-based SCMC can promote interaction and facilitate the processing of L2 acquisition with the availability of other modalities, such as text-chat, visual aids (e.g., pictures, video clips), as well as aural and paralinguistic cues including intonation and facial expressions (Hampel & Stickler, 2012). Thus, to fully investigate the value of video-based feedback for L2 acquisition, it appears important to explore whether multimodal recasts, involving the supply of oral and written recasts together, may offer advantages over exposure to oral or written recasts alone.

There are several reasons why multimodal recasts are likely to be more effective in promoting L2 development than the provision of recasts in a single modality. First, it is worthwhile to consider the potential benefits of multimodal recasts from a theoretical perspective. These advantages of multimodal recasts are grounded in two information processing models: the Dual Coding Theory (Paivio, 1986, 2007) and the Cognitive Theory of Multimedia Learning (Mayer, 2005, 2014). According to Paivio's Dual Coding Theory,

humans process information through two channels; one for verbal and the other for nonverbal information. These two channels are believed to be separate but interconnected, so information presented in one channel can activate the other. Therefore, providing information through multiple modes such as visual, auditory and textual may stimulate both channels, which is assumed to facilitate information processing, resulting in better learning (Paivio, 1986, 2007).

Influenced by the Dual Coding Theory and Baddeley's model of WM (Baddeley, 1999; Baddeley, Eysenck, & Anderson, 2009), Mayer's Cognitive Theory of Multimedia Learning also provides support for the potential benefits of multimodal presentation of feedback. Mayer suggests that learning and understanding can be enhanced when learners construct information from two different channels. In Mayer's view, there are two approaches to conceptualizing these two channels: the presentation-mode approach (Paivio, 1986, 2006) distinguishes a verbal channel (e.g., spoken or printed words) and a nonverbal channel (e.g., pictures, graphs), whereas the sensory-modality approach (Baddeley, 1999; Baddeley, Eysenck, & Anderson, 2009) differentiates between an auditory channel (processing the information through the ears such as spoken words) and a visual channel (processing the information through the eyes, as with printed words). For multimodal recasts, Mayer's prediction based on the sensory-modality approach appears particularly relevant; it states that adding visual information to aural information (e.g., on-screen text to accompany spoken text) can prime both aural and visual channels, thus facilitating the establishment of mental connections between the two representation systems, which may lead to more effective learning and better recall (Mayer, 2014; Mayer & Fiorella, 2014).

Sweller's (2005) Cognitive Load Theory provides further support for the potential value of multimodal recasts. This theory suggests that the simultaneous presentation of

corresponding auditory and visual input (e.g., captioned video) can alleviate the cognitive burden on WM, resulting in improved short-term memory storage and retrieval, compared to relying on a single channel (Sweller, 1988, 1994). This can potentially assist learners in taking advantage of their full capacity for processing information (Mayer, 2014), thereby enhancing encoding and decoding of the input.

Considering the theoretical support for the advantages of multimodal input discussed above, multimodal recasts appear to have the capacity to promote the formation of mental connections between auditory and visual representation systems. In addition, the use of multimodal recasts may mitigate cognitive overload on WM by distributing linguistic input across multiple modalities. This, in turn, may reduce the risk of exceeding the learner's cognitive capacity in each channel (Mayer & Fiorella, 2014), potentially facilitating the processing and retrieval of linguistic input and subsequent L2 learning.

Another potential advantage of multimodal recasts is that, due to their hybrid nature, learners may benefit from the distinct strengths of both oral and written modalities of feedback, resulting in greater salience and noticeability of the target constructions. For example, features of written CF, such as the extended processing time, visual/perceptual salience, and re-readability of messages, have been suggested as factors that may enhance the noticeability of the feedback and target forms entailed (e.g., Gonzalez-Lloret, 2014; Long, 2007; Smith, 2004, 2010; Yilmaz, 2012). Thus, the corrective intention of text-based CF is more likely to be explicit to learners than that of oral feedback (e.g., Blake, 2008; Chapelle, 2001; Smith, 2004, 2010; Yilmaz, 2012). This may lead to increased opportunities for learners to focus on form, providing greater chances for L2 development (Salaberry, 2000). Indeed, these distinctive features of text-based feedback have been claimed to work as a

cognitive amplifier (Lai & Zhao, 2006; Warschauer, 1997), creating optimal conditions for learners to notice L2 linguistic forms and thus facilitate L2 acquisition (Schmidt, 2001).

Meanwhile, oral recasts provide verbal and nonverbal cues such as word or syllable stress, intonation, and gestures, which have been proposed to make recasts more salient and noticeable to learners (Loewen & Philp, 2006). More importantly, oral recasts can be provided in a contingent manner, that is, immediately following a learner's erroneous utterance. However, this is frequently unfeasible during text-based communication due to the absence of strict turn adjacency and "split negotiation routine" (Smith, 2003, p. 48). As previously explained, contingency of recasts (i.e., the immediate correction of errors) is regarded as a critical element that contributes to their effectiveness, as the contrast between the ill-formed learner output and the target-like model renders the negative evidence included in recasts more pronounced to learners, thereby promoting cognitive comparisons during speech processing (Doughty, 2001; Leeman, 2003; Long, 2007).

Taken together, multimodal recasts might have the capacity to mitigate the inherent weaknesses of oral as well as written recasts, enabling more effective focus on form during meaningful video-based interactions, in which recasts are often misinterpreted as confirmation checks or mere repetitions. In addition, supplying multimodal recasts might be particularly helpful when the aim is to draw learners' attention to grammatical errors, as in the current thesis, given that morphosyntactic recasts have been shown to be less salient for learners than recasts targeting lexis or phonology (e.g., Carpenter et al., 2006; Kourtali, 2022; Lyster, 1998; Lyster et al., 2013; Lyster, Saito, & Sato, 2013; Mackey et al., 2000; Mackey, Gass, & McDonough, 2000; Smith, 2010; Trofimovich et al., 2007).

Given that multimodal recasts in the present study were operationalized as captioned oral recasts that simultaneously present oral and on-screen written recasts, it seems

worthwhile to briefly overview prior research using captions in L2 grammar learning. Captions are defined as a simultaneous, on-screen written text that matches spoken audio signals (Pattemore & Muñoz, 2022; Vandergrift, 2007). An increasing amount of research is demonstrating the benefits of caption-supported audiovisual input (e.g., captioned video) for L2 learning, particularly vocabulary and comprehension (e.g. Montero Perez et al., 2013a; Peters & Webb, 2018; Rodgers & Webb, 2017). As reflected in the results of meta-analyses (Montero Perez et al., 2013b; Reynolds et al., 2022), this group of research, overall, has yielded positive effects for captioning. The utility of captions in L2 development has been attributed to the additional source of input they provide for learners. That is, captions facilitate learners' overall comprehension of speech by serving as a visual cue that highlights pertinent linguistic units such as words and phrases (Bird & Williams, 2002; Gass et al., 2019; Vanderplank, 1988). This, in turn, enables learners to recognize novel linguistic items in the input and create connections between the written and spoken forms of the features, thereby promoting their attention to and acquisition of the items (Montero Perez, 2020; Winke, Gass, & Sydorenko, 2010).

Of more relevance to the present thesis is research which has specifically explored the learning of grammar through audiovisual input supported by captions (e.g., Cintrón-Valentín et al., 2019; Cintrón-Valentín & García-Amaya, 2021; Lee & Révész, 2018, 2020, 2021; Muñoz et al, 2023; Pattemore & Muñoz, 2020, 2021, 2022). Captioning studies investigating the processing and acquisition of grammar have mostly focused on comparing different types of captioning (e.g., full captions, key-word captions, textually enhanced captions, and no captions). For instance, Lee and Révész (2020) examined the relative effects of textually enhanced and non-enhanced captions on learners' attention to and acquisition of the English present perfect tense versus the simple past construction from 24 short video clips. Compared

to a non-captions condition, both captioning modes resulted in greater learning gains for the present perfect, suggesting an advantage for captioned input.

Considering that multimodal recasts are the combination of aural and written input, it seems reasonable to assume that the concurrence of oral and written recasts in the form of captioned oral recasts may assist learners in segmenting and parsing the oral input while drawing their attention to morphosyntactic features. This may promote more in-depth processing and acquisition of the target grammatical constructions. While researchers have begun to explore the effects of audiovisual input using captioning on L2 grammar learning, previous research in this strand, as mentioned earlier, has mainly focused on development in L2 vocabulary and listening comprehension. To date, no research has investigated whether captioned oral feedback can promote development in L2 grammatical knowledge in the context of multimodal video-conferencing. Therefore, an exploration of the effects of this novel type of recasts on L2 grammar learning could contribute to expanding this area of research.

In addition to examining the impact of multimodal recasts on L2 outcomes, it would also be worthwhile to gain information about learners' real-time processes, particularly how they use and allocate their visual attention to feedback and target forms depending on the feedback mode. An effective technique that can shed light on learners' online cognitive processing of captioned recasts is eye-tracking (Conklin, Pellicer-Sánchez, & Carrol, 2018; Godfroid, 2020). Eye-tracking provides a means to directly record an individual's eye movements while processing information on a computer screen (Conklin, Pellicer-Sánchez, & Carroll, 2018; Godfroid, 2020). This methodology is versatile, unobtrusive, and sensitive, and has been widely utilized in various applied linguistics domains (Conklin et al., 2018; Godfroid, 2020). It allows researchers to identify the specific area of visual input that

participants focus on during a specific time period, along with the frequency and duration of these focal points (Pellicer-Sánchez & Siyanova-Chanturia, 2018). As aptly posited by Robinson (2003), "attention is the process that encodes language input, keeps it active in working and short-term memory, and retrieves it from long-term memory" (p. 361). When processing information, attention can be seen as a multifaceted set of cognitive functions that regulate our actions, such as selecting information, maintaining focus on relevant information, and inhibiting distractions. All of these functions collectively contribute to facilitating our learning process (Robinson et al., 2011).

Given that eye movements offer direct measurements of overt attention allocation and are closely linked to covert attentional processes (Leow, 2015; Rayner, 2009), eye-tracking emerges as a robust physiological tool for studying attention (Leow, 2015; Robinson et al., 2011). This technique has been employed in operationalizing attention in L2 studies (Conklin et al., 2018; Godfroid, 2020; Rayner, 2009). Therefore, it holds significant promise for unraveling how learners distribute their visual attention when exposed to multimodal recasts in video-based SCMC environments. To date, however, only a limited number of SCMC studies (e.g., Smith, 2010; Smith & Renauld, 2015) and research on captioning (e.g., Lee & Révész, 2020, 2021) have delved into the extent learners' attention is directed to target grammatical forms and how the amount of attention they allocate to these structures relates to L2 grammatical development. Consequently, there is a need for further research to explore the effects of this novel form of recasts on learners' allocation of visual attention and their subsequent progress in L2 grammar learning within video-based SCMC contexts.

Another way to make inferences about the cognitive processes in which learners engage during feedback processing is to examine how the effects of the treatment interact with pertinent cognitive individual difference factors (Dekeyser, 2012). When it comes to

multimodal recasts, a key individual difference worth considering is learners' WM capacity. WM is generally defined as an individual's ability to temporarily store and manipulate the information needed to perform complex tasks (Baddeley, 1992; Juffs, & Harrington, 2011).

According to Baddeley's (Baddeley, 1986, 2000; Baddeley & Hitch, 1974) influential model, WM consists of four components: the central executive; the phonological loop or the phonological short-term memory (PSTM); the visuospatial sketchpad or visuospatial short-term memory (VSTM); and the episodic buffer. The central executive is responsible for several functions, such as (a) task switching: the ability to switch between tasks; (b) inhibition: to inhibit responses when required; and (c) updating: to monitor, revise, and update incoming information (e.g., Baddeley, 1996; Miyake et al., 2000). The phonological loop performs the functions of holding, manipulating, and retaining auditory and verbal information for a brief duration, whereas the visuospatial sketchpad is the storage system which temporarily retains and processes visual images and spatial relations. The role of the episodic buffer (Baddeley, 2000) is to store and integrate information from various sources, including verbal, visual, and spatial, along with long-term memory, to construct episodes or multimodal units.

Given that an individual's WM capacity is limited, and its function is to regulate attentional resources while performing complex tasks, it seems logical to assume that individual differences in WM may affect the extent to which learners pay attention to input, depending on the feedback modalities (i.e., oral, written, or multimodal recasts). While there is abundance of research suggesting that WM influences the efficacy of oral and/or written feedback (e.g., Goo, 2012; Mackey et al., 2002, 2010; Révész, 2012; Sagarra, 2007, 2008, Sagarra & Abbuhl, 2013; Trofimovich et al., 2007; Yilmaz, 2013a; Zalbidea & Sanz, 2020), there has been no investigation so far into whether WM capacity mediates the effects of

multimodal feedback and how it relates to L2 grammatical development in the context of video-conferencing. This highlights the need for further research into whether and to what extent learners' WM capacity would affect the effects of multimodal recasts on L2 acquisition while performing video-conferencing tasks.

1.2. Aims of the Thesis

To fill the aforementioned research gaps in the literature, two studies were conducted as part of this thesis. The primary objective of the two studies was to investigate the pedagogical value of multimodal recasts provided in the context of multimodal video-conferencing tasks. In particular, the studies aimed to examine the effectiveness of multimodal recasts in drawing learners' attention to target grammatical constructions and promoting development in their use, as measured by a spontaneous oral production test, an oral production test, and a written production test. Study 1 was guided by the following research questions:

1. To what extent do recasts, as compared to no recasts, affect the development of L2 grammatical knowledge during video-conferencing tasks?
2. To what extent does modality of recasts affect the development in L2 grammatical knowledge during video-conferencing tasks?

That is, the first research question explored the effects of recasts on the development in L2 grammatical knowledge, compared to no recasts. The second research question aimed to examine the relative effectiveness of multimodal recasts in promoting L2 grammatical knowledge, compared to unimodal recasts, i.e., oral recasts only and written recasts only. As noted, the outcome measures included a spontaneous oral production test (SOPT), a more

controlled oral production test (OPT), and a written production test (WPT) to assess development in various aspects of L2 knowledge.

In addition to the two questions that were investigated in Study 1 (research questions 1–2 below), Study 2 addressed four additional research questions (research questions 3–6 below) as follows:

1. To what extent do recasts, as compared to no recasts, affect the development of L2 grammatical knowledge during video-conferencing tasks?
2. To what extent does modality of recasts affect the development in L2 grammatical knowledge during video-conferencing tasks?
3. To what extent does modality of recasts (in the case of written versus multimodal recasts) affect learners' visual attentional allocation to the target linguistic constructions during video-conferencing tasks, as reflected in participants' eye movements?
4. To what extent is learners' visual attention to the target constructions included in recasts, related to the development in L2 grammatical knowledge? Is this relationship influenced by whether learners are provided with written or multimodal recasts?
5. To what extent do individual differences in WM capacity moderate the effects of oral, written, or multimodal recasts on the development in L2 grammatical knowledge during video-conferencing tasks?
6. To what extent is learners' visual attention to the target constructions related to their WM capacity? Is this relationship influenced by whether learners are provided with written or multimodal recasts?

The additional questions were included to examine (a) the effects of feedback modality on learners' visual attention allocation to the target linguistic constructions and its association with their development in L2 grammatical knowledge; (b) the moderating role of WM in the relationship between the modality of recasts and the development of L2 grammatical knowledge; and (c) the relationship between learners' WM capacity and their visual attentional allocation to the target linguistic constructions.

In order to address these questions, various data collection tools were employed. Production-based outcome measures were used to assess development in grammatical knowledge. An eye-tracker was employed to capture learners' visual attention allocated to the recasts (in the case of written recasts and multimodal recasts) and the target structures while performing the treatment tasks. To assess learners' WM capacity, a battery of WM tests was administered: a nonword span (NWS) task (Jung, 2017), a color shape task (Miyake, Emerson, Padila, & Ahn, 2004), a stop signal task (Logan, 1994; Verbruggen, Logan, & Stevens, 2008), a forward Corsi block task, and an automated operation span (AOSPAN) task (Turner & Engle, 1989).

1.3. Structure of the Thesis

The present thesis consists of five chapters. The current chapter introduces the rationale and purpose of the thesis. Chapter 2 explains the theoretical framework and gives an overview of the relevant literature. First, it provides a brief introduction to the interactionist approach and theoretical accounts that address the benefits of interaction and the importance of attention and awareness in L2 acquisition. That is followed by a discussion of different models that explain the role and significance of attention, awareness, and noticing in input processing. The next section focuses on theoretical accounts that provide support for the advantages of

CF, particularly recasts, along with a review of empirical studies investigating the utility of recasts in L2 learning. The section which follows that one introduces the notion of computer-assisted language learning (CALL) and SCMC, with an overview of previous research exploring the effects of CF including recasts on L2 development in various CALL and SCMC contexts such as computerized feedback as well as text-based and video-based CF. This is followed by a discussion of the results as related to empirical studies. As the primary interest of the present thesis lies in the effectiveness of multimodal recasts, the next section elaborates theoretical perspectives that elucidate the benefits and pedagogical potential of a multimodal provision of input, followed by a discussion of different theories and approaches relating to processing information presented multimodally. Also, as multimodal recasts are operationalized as captioned oral recasts in the present thesis, the notion of captions is introduced, and their effectiveness in L2 acquisition is reviewed. The review comprises both theoretical underpinnings and empirical evidence, with a particular emphasis on how captions can facilitate L2 grammatical development when used in conjunction with audiovisual input. Then, the application of eye-tracking methods in L2 learning, including knowledge development in L2 grammar, is discussed followed by a review of pertinent empirical research. The next section explains the construct of WM, Baddeley's WM model, Wickens' Multiple Resource Model, and various measures of WM capacity. This chapter ends with a review of previous studies examining the potential impact of individual differences in WM on L2 acquisition.

Chapter 3 reports on Study 1, which examined the effects of recasts and feedback modality on the development in the use of L2 grammatical knowledge by comparing the provision of recasts to the absence of recasts, as well as oral-only r and written-only recasts, and multimodal recasts conditions. This chapter starts by describing the methodology,

including the participants, target linguistic constructions, materials employed (e.g., treatment tasks, outcome measures, and questionnaires), the type of feedback, data collection, coding and scoring procedures, and statistical analyses. After the methodology, the results are presented, followed by a discussion of the findings for each research question. This chapter concludes with a discussion of the perceived limitations of the study and an explanation of the modifications made for Study 2.

Chapter 4 focuses on Study 2, which investigated the effectiveness of recasts and the impact of feedback modality on L2 grammatical development as well as the effects of feedback modality on learners' visual attentional allocation to and development in the use of the target constructions. In Study 2, an eye-tracker was employed to capture learners' attentional allocation to the target structures. Study 2 also examined one role of WM in the effects of feedback modality on learners' allocation of attentional resources to the target forms and how it related to their knowledge development in the use of L2 grammar. Chapter 4 ends with a presentation of the results and discussion of the study's findings.

Chapter 5 synthesizes the overall findings from Study 1 and Study 2 and discusses possible theoretical, methodological, and pedagogical implications, followed by a consideration of the limitations of the studies and suggestions for future research.

CHAPTER 2. LITERATURE REVIEW

2.1. Role of Input in Second Language Acquisition

The significance of input in SLA has been acknowledged by many scholars (e.g., Ellis, 1994; Gass, 1997; Krashen, 1982, 1983, 1985; Long, 1983, 1985, 1996), and there is a general agreement on the necessity of input in language learning. Indeed, input is an essential component of all theories of language acquisition, including Krashen's (1983, 1985) Comprehensible Input Hypothesis. Krashen proposed this hypothesis as part of the monitor model, with a particular emphasis on the importance of input in L2 acquisition. According to this hypothesis, language acquisition occurs only by receiving *comprehensible input*, defined as language input that is slightly beyond the learner's current level of linguistic competence (*i+1*), arguing that exposure to comprehensible input is a sufficient condition for L2 acquisition to occur (Krashen, 1985, p. 23).

Although the comprehensible input hypothesis has been influential in the field of SLA as one of the earliest models to underscore the significance of input in L2 acquisition, there was also criticism mainly in terms of: (a) what *i+1* really means, and (b) its claim that comprehensible input is the only necessary and sufficient condition for successful L2 acquisition. The first criticism points out the vagueness and inconsistency in the notion of comprehensible input (Gass, 1988; White, 1987) in that it is not possible to define the concept of *i+1* or measure it objectively by means of a scientific method. Another major criticism is of its assumption that all input is automatically processed by the learner. However, many researchers have noted that comprehension of input itself does not guarantee intake of the input. Corder (1967, 1982), for example, addressed this issue by defining input as "what is available for going in" and intake as "what goes in" (1967, p. 165).

Similarly, Gass (1988) proposed a model for L2 acquisition that incorporated two types of input: *apperceived input* and *comprehended input*. She defined apperceived input as what learners notice in the ambient speech, highlighting that not all input is noticed by learners. According to this model, apperceived input precedes comprehended input, defined as the amount of apperceived input that learners comprehend. Gass also distinguished comprehended input from comprehensible input, explaining that the former means that the learner is in charge of comprehending, as opposed to the latter that disregards learner control over input. Another noticeable difference between these two types of input is that, unlike comprehensible input, comprehended input is not a dichotomous term. For example, input can be comprehended in various ways, as either in understanding overall meaning of message or at a metalinguistic level such as phonological or morphological features.

As can be seen, the comprehensible input hypothesis has been challenged by many researchers (e.g., Gass, 1988; Larsen-Freeman & Long, 1991; Lightbown, 1998; Robinson, 1995, 2003; Schmidt, 1990, 1993b, 2001; Schmidt & Frota, 1986; Swain, 1985, 1995; VanPatten, 1996; White, 1987). In particular, a general consensus has emerged over the fact that not everything presented in input is processed by the learner (Corder, 1967), and that attention is a principle cognitive mechanism affecting the process of selecting input for further processing (Robinson, 2003; Schmidt, 2001). Sharwood Smith (1986) also contends that the comprehensibility of input itself may not elicit learners' attention to form, resulting in the failure of complete intake of the input, highlighting the pivotal role of attention to the linguistic target. As such, learners' attention to input has been acknowledged as being crucial for subsequent processing and successful intake of the input (e.g., Schmidt, 1990, 1993b, 2001).

2.2. Attention and Awareness in SLA

A facilitative role for attention in the L2 learning process has been advocated by a number of scholars in the field of SLA and cognitive psychology (e.g., Gass, 1997; Izumi, 2002, Robinson, 1995, 2003; Schmidt 1990, 1993b, 1995, 2001 among others). Schmidt (1990, 1993b, 1995, 2001), for example, supports the critical role of attention and awareness in language learning, suggesting that "people learn about the things they attend to and do not learn much about the things they do not attend to" (2001, p. 30). While many SLA researchers agree about the significance of attention in language learning, there are differences in their views in terms of the type and level of attention and the necessity of awareness involved in noticing (e.g., Robinson, 1995, 2003; Schmidt, 1990, 1993b, 1995, 2001; Leow, 2015; Tomlin & Villa, 1994), which will now be discussed in turn.

2.2.1. Schmidt's Noticing Hypothesis

In the Noticing Hypothesis (Schmidt, 1990, 1993b, 1995, 2001), Schmidt suggests that learners need to focus their attention on specific aspects of the target language and be aware of the differences between their own language production and the correct way of using the language. Only then can they effectively understand and use new language input in their interlanguage systems (Schmidt, 1990, 2012; Schmidt & Frota, 1986). Thus, Schmidt defines noticing as a combination of focal attention and a low level of awareness.

The Noticing Hypothesis is rooted in Schmidt and Frota's (1986) early work, which reported the findings of a case study on Schmidt's own language-learning experience in Brazil over a period of five months. The authors found that input, interaction or classroom drills were not sufficient for Schmidt to learn particular verb forms in Portuguese, and he had to notice these forms in the input to learn them. Although the authors did not make a strong

assertion about whether noticing occurred consciously or not, in his later works, Schmidt (1995) argued for a significant role for consciousness as well as awareness in language learning, defining noticing as attention entailing a certain level of awareness.

More specifically, Schmidt (1995) identified two levels of awareness relevant to the context of L2 learning: awareness *at the level of noticing* and awareness *at the level of understanding*. Awareness at the level of noticing includes focal attention, and noticing occurs with conscious awareness, which enables subsequent learning. Noticing is characterized as awareness at a low level of abstraction, which means that, rather than underlying grammatical rules or patterns, only surface features of the input are noticed by learners (Schmidt, 2001). Awareness at the level of understanding, on the other hand, "implies recognition of a general principle, rule or pattern" (Schmidt, 1995, p. 29), in which learners are consciously engaged in deeper processing of the input. Thus, both problem-solving and all forms of metacognition are indicative of awareness at this level (Schmidt, 1990). Furthermore, if noticing is related to the transfer of information to long-term memory and item learning, then awareness at the level of understanding is connected to "the organization of material in long term memory, to restructuring, and to system learning" (Schmidt, 1993b, p. 213). Schmidt (1993, p. 26) clearly distinguishes these two levels of awareness as follows:

> I use noticing to mean registering the simple occurrence of some event, whereas understanding implies recognition of a general principle, rule, or pattern. For example, a second-language learner might simply notice that a native speaker used a particular form of address on a particular occasion, or at a deeper level the learner might understand the significance of such a form.

Schmidt (1990, 1993) originally proposed a strong version of the noticing hypothesis, considering noticing as "the necessary and sufficient condition for the conversion of input to intake" (Schmidt, 1990, p. 129). After a decade, drawing upon theories of learning discussed in the field of psychology, this version was modified into a weaker version, that of viewing the act of noticing as a facilitator of L2 development (Schmidt, 1994, 1995, 2001). In this revised version, Schmidt claimed that learning without noticing (i.e., implicit learning) may occur but, when learners notice more, they learn more. In both versions, however, the roles of attention and awareness during the early stage of language processing were highly underscored. That is, L2 acquisition is premised on attention as an essential factor for "the conversion of input to intake for learning" (Schmidt, 1994, p. 17), as "attended learning is far superior, and for all practical purposes, attention is necessary for all aspects of L2 learning" (p. 3).

2.2.2. Tomlin and Villa's Model of Attention

An alternative view regarding the role of attention in L2 acquisition was Tomlin and Villa's (1994) model of attention. They proposed a fine-grained analysis of attention in which two important claims were made: (a) that attention involves three theoretically and empirically separable functions; *alertness, orientation, and detection*, and (b) that awareness can be dissociated from attention (p. 198). First, Tomlin and Villa identified three main processes involved in attention, which are alertness, orientation, and detection. Alertness refers to learners' overall readiness for dealing with incoming data or stimuli that is directly related to the rate at which information is selected for further processing. Hence, this process is important for L2 acquisition as it affects learners' processing or performing a task.

Orientation refers to allocating attentional resources to a specific type of stimuli at the exclusion of others. The authors claimed that this component of attention is of greater importance to SLA than alertness, in that when attention is drawn to a particular piece of information, detection is generally facilitated. Orientation can have either a facilitative or inhibitory consequence for further processing, depending on whether the information that attentional resources are focused on is expected (positive effect) or not (negative effect).

Detection refers to the actual cognitive registration of stimuli that was attended to, which enables learners to select and process particular input. The researchers noted that although alertness and orientation may together or separately enhance detection, neither of these attentional processes is a prerequisite for detection to occur, arguing that only detection is essential for further processing of the information and that language learning must occur at this level. While detection is similar to noticing in Schmidt's model, a major difference between them is that Schmidt (2001) argues that noticing and awareness are directly linked, whereas for Tomlin and Villa (1994), detection "does not require awareness" (p. 197). In other words, they argue for a dissociation between attention and awareness, and awareness is not a necessary condition for input to be available for language acquisition.

2.2.3. Robinson's Model of Attention and Memory

In his model of attention and memory (1995, 2003), Robinson incorporated both detection (Tomlin & Villa, 1994) and awareness (Schmidt, 1990) in his account of noticing, expanding its relationship with working memory. Drawing on Cowan's (1988) model of memory and attention, Robinson defined noticing as "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (Robinson, 1995, p. 296). That is, Robinson (1995)

claimed that information in WM must be kept activated to be retained, and this is possible by rehearsing the information mentally (see Figure 1).

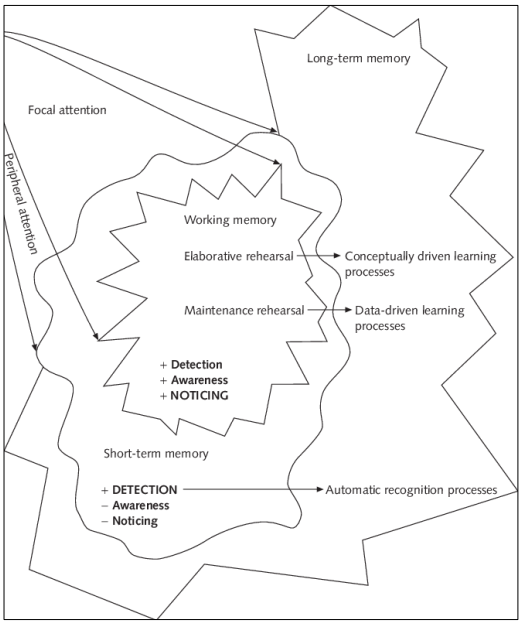


Figure 1. Robinson's Model of Memory and Attention (adapted from Robinson, 2003)

Robinson explains that detection precedes rehearsal, which is comprised of two types: *maintenance rehearsal* and *elaborative rehearsal*. Maintenance rehearsal is data-driven processing, involving repetition of the information in one's mind. Elaborative rehearsal, on the other hand, is conceptually driven processing, which is a more elaborated form of rehearsal involving effortful integration of the information detected with the information from long-term memory (i.e., prior knowledge) (Leow, 2015; Woolfolk, 1998). In other words, information processing (i.e., rehearsing detected information) entails conscious attention to form in the input to complete task demands, and awareness is equivalent to conceptually driven processing, which is a deeper processing of noticed input. In summary, Robinson explains that although detection plays an important role in the input processing, it

is at the level of noticing (i.e., attention with a low level of awareness) that linguistic input is available for further processing.

2.2.4. Leow's Model of the L2 Learning Process in Instructed SLA

Leow's (2015) model views the notion of learning as a combination of both processes and products. The model proposes three processing stages: *input processing stage*, *intake processing stage*, and *knowledge processing stage*. The input processing stage is the initial stage, that is "between the input and the intake of specific linguistic information and what is taken in (intake) is initially stored in working memory" (Leow, 2015, p. 241–242). At this stage, input transforms to intake. Initially, intake is stored in WM and the success of this stage is largely dependent upon the level of attention, which can be peripheral, selective, or focal, and may be accompanied by depth of processing, cognitive registration, and level of awareness.

Based on these variables, intake lodged into WM can be categorized as three types: *attended intake, detected intake,* and *noticed intake*. Attended intake, which is comparable to the initial stages of perception of input proposed by Chaudron (1985), results from learners' high peripheral attention, low cognitive registration, and low depth of processing, and thus is not further processed or stored in WM. Detected intake takes place with learners' selective attention accompanied by minimal depth of processing paid to input. The theoretical assumption underlying this is that detected intake occurs without learners' awareness. That is, learners may detect target linguistic features in the input without being aware of them (Leow, 2015). Hence, the detected intake is similar to "detection" in Tomlin and Villa's (1994) model. Noticed intake entails high focal attention, high cognitive registration, a low level of awareness and a low depth of processing, which is comparable to Schmidt's (1990)

"noticing". Hence, noticed intake which requires high focal attention entailing a certain level of awareness is more likely to be stored and processed in WM and, further, attributes to development of the L2 grammar system. Leow notes that learning is likely to occur when the learner pays minimal attention to new information in L2 instruction and "[c]rucially, unless further processed in the intake processing stage, any type of intake will be discarded from working memory" (Leow & Driver, 2021, p. 78).

During the second stage, the intake processing stage, preliminary intake (i.e., attended, detected, and noticed) is further processed in one of two ways depending on lower or higher depth of processing, which may also be accompanied by the activation of prior knowledge and/or lower or higher cognitive effort: (1) minimal data-driven processing is associated with a lower depth of processing that allows preliminary intake to be entered into learners' L2 developing system, encoded as "a non-systemized of language" (Leow & Driver, 2021, p. 79); or (2) conceptually driven processing or consciously encoding and decoding of the linguistic input occurs with higher levels of processing or amount of cognitive effort and levels of awareness. Higher depth of processing is typically accompanied by higher levels of awareness, which helps maintain L2 data in WM and promotes its incorporation into the developing system. Depending on how the L2 data are processed, they will eventually be encoded in the developing system via item learning or system learning. Activation of old or recent prior knowledge also plays an important role in the intake processing.

The third and final stage, the knowledge processing stage, takes place between the L2 developing system and their production of output (e.g., monitoring production in relation to learned grammar). In line with Swain (2005), Leow views learner output as a product (that represents the outcomes of learning) as well as part of the L2 learning process, given that learners' output may also serve as additional input; learners may monitor and modify their

output based on the feedback they receive to confirm or disconfirm their L2 production or prior knowledge. A depth of processing and a higher level of awareness may also be needed for learners at this stage, with the ability to activate prior knowledge to develop their current L2 grammatical system.

2.2.5. Summary

As discussed, Schmidt (1990, 1993b, 1995, 2001), Tomlin and Villa (1994), Robinson (1995, 2003), and Leow (2015) have different stances with regards to the type of attention and level of awareness involved in noticing input. The main difference in their theoretical perspectives is whether the learner's awareness is required or not for further processing of the input (Hama & Leow, 2010; Leow, 2001). However, they seem to agree that selective attention (i.e., with or without awareness) is necessary for successful SLA. What needs to be answered regarding learners' attention is, then, how their attention can be drawn to target linguistic features during the learning process to foster the intake of the target items. Among various suggestions, learners' engagement in conversational interaction has been claimed to effectively direct learners' attention to the input and facilitate L2 development. With much empirical evidence regarding a facilitative role of interaction in L2 acquisition, researchers have begun to investigate the specific aspects of interaction, such as negotiation for meaning, CF, and modified output, which most likely influence the extent to which learners notice and allocate their attention to the target linguistic input (e.g., Mackey, 2012; Mackey, Abbuhl, & Gass, 2012; Mackey & Goo, 2013; Gass,& Mackey, 2015; Loewen & Sato, 2018).

2.3. The Role of Interaction, Output and Corrective Feedback in SLA

2.3.1. Interaction Hypothesis

The interaction hypothesis (Gass, 1997, 2003; Gass & Mackey, 2007; Long, 1983, 1996, 2006; Pica, 1994) proposes that negotiated interaction is fundamental to SLA in that it "connects input, internal learner capacities, particularly selective attention, and output in productive ways" (Long, 1996, p. 451–452). Learners' engagement in a negotiated interaction likely enables learners to notice the gaps between their non-target-like output and the target-like form when provided with CF, and modify their erroneous output accordingly (e.g., Gass & Mackey, 2006).

Taken together, to take advantage of negotiated interaction, learners need to: (a) notice the target linguistic features in the linguistic input; (b) have opportunities to produce output; (c) notice errors in their own output; and (d) correct their erroneous output for comprehension or accuracy by virtue of the CF they may receive.

2.3.2. Output

The benefits of negotiated interaction for L2 learning discussed in the previous section provoked an examination of the impact of learners' output on directing learners' attention to and enhancing their noticing and awareness of the target linguistic features (Gass, 1997; Long, 1981, 1991, 1996; Pica, 1994; Swain, 1985, 1995). For example, upon studying Canadian immersion classrooms, Swain (1985, 1995) found that despite learners' prolonged and ample exposure to comprehensible input, the input alone did not result in the students' successful L2 acquisition (Swain, 1985; Swain & Lapkin, 1995). Specifically, while the immersion students exhibited native-like proficiency in their reading and listening comprehension abilities (i.e., receptive skills), they failed to acquire the same level of

language competence in their oral and written production (i.e., productive skills) in terms of grammatical accuracy. Based on this observation, Swain concluded that the lack of opportunities to produce output may hinder learners from developing holistic L2 proficiency and grammatical accuracy. She emphasized that learners' engagement in producing output may push or stretch them to produce syntactically more complex and accurate target language forms. Thus, in contrast to Krashen's (1985) assertation that "output is a result of acquisition, not its cause" (p. 80), Swain (1985, 1995) viewed output as a causal factor for L2 acquisition, not a mere result of it (Adams, 2003).

The premise that output plays a beneficial role in language-learning processes raised a question of exactly how output functions to facilitate L2 acquisition. According to Swain's (1985, 1993, 1995, 1998, 2001) output hypothesis, a learner's output has four functions: (a) through producing output, learners "notice a gap between what they *can* say and what they *wish* to convey" (emphasis added, Swain, 2005, p. 474); (b) output serves as a ground for hypothesis testing about the way the target language works (Swain, 1995); and (c) learners' output induces learners to reflect on their metalinguistic knowledge. A number of empirical studies have provided evidence supporting these three functions (e.g., Adams, 2003; Izumi, 2002; Izumi & Bigelow, 2000; Mackey, 2002; Philp & Iwashita, 2013; Swain & Lapkin, 1998). In her later work, Swain (2005) provided evidence for the fourth function of learner output, which states that (d) as output necessarily involves language use, it provides more opportunities for production practice to develop learner's fluency and automaticity of the target language (e.g., DeKeyser 2001; Lyster & Sato 2013; DeKeyser 2017).

With regard to the benefits of output, Pica et al. (1989) argued that what is "especially helpful to the acquisition process is the interaction in which learners are 'pushed' to make their output comprehensible" (p. 64). In this context, pushed output refers to learners' self-

modification of an utterance in response to negative feedback from their interlocutors signaling that their output has not been fully comprehended or is grammatically incorrect. In her later work, Mackey (2012) refers to pushed output as modified output, defining it as reformulation of "one's original utterance in response to feedback or self-monitoring" (p. 16). In this sense, a number of scholars have emphasized the importance of CF in negotiation of meaning. This is because solely aiming for mutual comprehension may not lead learners to identify and address the discrepancy between their interlanguage forms and the L2 forms, which is essential for acquiring specific language items. (Schmidt, 1995, 1998). Considering that CF may "[encourage] self-repair involving accuracy and precision and not merely comprehensibility" (Lyster & Ranta, 1997, p. 42), CF can come into play to prompt learners to focus on language-specific problems and stimulate them to produce modified output (Long, 1996).

2.3.3. Corrective Feedback

As discussed, the interactionist approach to SLA suggests that directing learners' attention to linguistic properties while engaging in L2 interaction facilitates L2 development (e.g., Gass, 1997; Gass & Mackey, 2007, 2015; Long, 1996; Mackey, 2007, 2012; Pica, 1994). Of various aspects of interaction, CF has been claimed to be an effective way to draw learners' attention to L2 linguistic constructions. It may be provided by teachers or more competent students in classroom contexts, by native speakers or more competent L2 speakers during laboratory research or in naturalistic environments.

SLA scholars have argued that CF promotes L2 acquisition, as it can draw learners' attention to their erroneous output, which is necessary for successful L2 acquisition (Gass, Mackey, & Pica, 1998; Lightbown & Spada, 1990; Long, 1990, 1996; Schachter, 1986, 1991;

Schmidt & Frota, 1986), thereby increasing the likelihood that they notice the gap between the target and their interlanguage forms (Schmidt & Frota, 1986). The rationale behind this argument is that attention and noticing are two key elements that bridge the gap between learner-internal and learner-external factors to promote L2 development (Schmidt, 2001), and that CF indicating that an error has occurred in the learner's production may potentially trigger these two fundamental cognitive functions. This view is supported by a number of meta-analyses and narrative reviews, which evidenced the benefits of CF in L2 development (e.g., Brown, 2016; Goo, Granena, Yilmaz, & Novella, 2015; Li, 2010; Lyster & Saito, 2010; Lyster, Saito, & Sato, 2013; Mackey & Goo, 2007; Nassaji, 2016; Russell & Spada, 2006; Sheen, 2010; Yilmaz, 2016).

Descriptive studies have provided useful taxonomies of CF strategies (Chaudron, 1977; Lyster & Ranta, 1997; Ranta & Lyster, 2007). Lyster and Ranta (1997), for example, proposed a six-part taxonomy of CF strategies (Table 1).

Table 1. Taxonomy of CF Strategies (adapted from Lyster & Ranta, 1997)

Strategy	Definition	Example
Explicit correction	The teacher points out an error directly and provides the correct form.	S: I play basketball yesterday. T: You should say "I played basketball yesterday."
Recast	The teacher reformulates part or all of the initial incorrect utterance.	S: I play basketball yesterday. T: I played.
Clarification request	The teacher indicates that there has been a mistake or misunderstanding.	S: I play basketball yesterday. T: Pardon?
Elicitation	The teacher endeavors to elicit correct forms by pausing before the initial erroneous word or asking students to reformulate their utterances.	S: I play basketball yesterday. T: I... (pause) basketball yesterday.
Repetition	The teacher repeats the erroneous part.	S: I play basketball yesterday. T: Play?
Metalinguistic feedback	The teacher indicates there is an error in learner's utterance using comments, questions, or metalinguistic knowledge of the error but does not provide the correct form.	S: I play basketball yesterday. T: Present tense or past tense? (Or) It should be past tense.

In their later work, Ranta and Lyster (2007) integrated recasts and explicit correction into a group known as "reformulation". This group constitutes input-providing strategies. The other strategies, namely repetition, elicitation, clarification requests, and metalinguistic feedback, were grouped together as "prompts". These strategies are also known as output-prompting strategies. In addition to the distinction between input-providing and output-prompting functions, Ellis (2012) classified CF strategies into explicit and implicit according to how apparent they appear as a form of correction. Given that input-providing and output-

prompting CF strategies exhibit varying degree of explicitness, the two dimensions of CF intertwine with each other (Table 2).

Table 2. Classification of CF Strategies (Ellis, 2012)

	Explicit	Implicit
Output-prompting	Metalinguistic clue Elicitation	Repetition Clarification requests
Input-providing	Explicit correction	Recasts

As such, several important distinctions have been made concerning different characteristics of CF, including: (a) negative evidence versus positive evidence (Leeman, 2003); (b) input-providing versus output-prompting feedback (Lyster, 2004; Goo & Mackey 2013; Lyster & Ranta, 2013); and (c) explicit versus implicit feedback (Ellis, 2007, 2008; Leeman, 2007; Loewen & Erlam, 2006; Lyster et al., 2013; Mackey, 2012; Sheen & Ellis, 2011). In terms of the first distinction, Long (1996) states that positive evidence is "models of what is grammatical and acceptable (not necessarily the same)" in the target language (p. 413). That is, positive evidence refers to exemplars of well-formed target language forms produced by native or competent non-native speakers. Negative evidence, on the other hand, is information that signals learners what is not grammatical or acceptable in the target language. In terms of the utility of each type of evidence, some have argued that negative evidence is an important component of CF (e.g., Schachter, 1991), whereas others have suggested that the positive evidence can also be important (e.g., Leeman, 2003).

Input-providing feedback offers the correct linguistic form for the learner. For example, a recast provides the correct alternative immediately after the learner's ill-formed utterance. In contrast, output-prompting feedback withholds the correct form; rather it prompts learners to self-correct their errors, thereby eliciting the target-like form from them.

The effectiveness of these two types of CF has been widely debated. Theoretical arguments and empirical investigation have been used to support the superiority of both input-providing feedback (e.g., Long, 2007; Goo & Mackey, 2013) and output-prompting feedback (e.g., Lyster, 2004; Lyster & Ranta, 2013), although some studies have also found similar effects for the two (e.g., Loewen & Nabei, 2007).

Finally, explicit feedback draws learners' attention directly or openly to their non-target-like production (e.g., You should say "I played basketball yesterday."), as opposed to implicit feedback where an error is treated in a less obvious manner (e.g., "Pardon?"). The effectiveness of these two CFs is also debatable. Some researchers have claimed that explicit feedback is more likely to be noticed by learners, thus has the better possibility of promoting development in the learner's interlanguage system (Ellis et al., 2006; Lyster, 1998a, 2004; Loewen & Philp, 2006; Lyster & Ranta, 1997, 2013; Lyster & Saito, 2010; Sheen, 2007). In contrast, other scholars (e.g., Goo & Mackey, 2013; Long, 1996, 2007, 2015; Long, Inagaki, & Ortega, 1998) have argued that CF should be relatively implicit, particularly taking the form of a recast so as not to interrupt the natural flow of communication. Mackey (2012) also suggests that although the main function of implicit and explicit feedback is to indicate a problem in learners' erroneous utterances, "explicit feedback (i.e., indicating a problem with the utterance with a focus on the form) can interrupt the flow of the interaction, whereas implicit feedback (i.e., indicating there is a problem with primary focus on meaning) usually allows for the interaction to continue uninterrupted" (p. 116).

In this vein, many studies have compared the differential effects of various types of CF; output-prompting CF strategies, especially prompts (a combination of repetition, clarification requests, metalinguistic clues and elicitation), versus input-providing CF strategies, particularly recasts (Ammar, 2008; Ammar & Spada, 2006; Guichet et al., 2015;

Loewen & Nabei, 2007; Lyster, 2004; Lyster & Izquierdo, 2009; Nassaji, 2009; Sato & Lyster, 2012; Yang & Lyster, 2010); and explicit CF (e.g., metalinguistic feedback) versus implicit CF (e.g., recasts) (Ahmadian, 2020; Ellis, 2007; Ellis, Loewen, & Erlam, 2006; Goo, 2012, 2016; Li, 2013, 2014, 2015; Loewen & Nabei, 2007; Nassaji, 2009; Sheen, 2007; Yilmaz, 2013a; Yilmaz & Granena, 2019).

However, it is difficult to argue that one type of feedback has an advantage over the other since the effectiveness of a particular CF is largely dependent upon various factors. To this end, a considerable number of empirical studies have been devoted to exploring factors potentially influencing the efficacy of CF (e.g., Li, 2010; Nassaji, 2015; Sheen & Ellis, 2011), such as target language feature (Egi, 2007; Mackey et al., 2000; Saito & Lyster, 2012); feedback type (Liao & Zhang, 2022; Lyster & Izuierdo, 2009; McDonough, 2007; Yilmaz, 2013a, 2013b); outcome measures (Mifka-Profozic, 2013; Révész, 2012; Zhao, 2015); context of feedback (e.g., Lyster & Saito, 2010); task complexity (Kim et al., 2015; Révész et al., 2014); task type (Révész, Sachs, & Mackey, 2011); and learner factors such as proficiency level (Ammar & Spada, 2006), developmental readiness (Mackey & Philp, 1998; Mackey, 2012), WM capacity (Ahmadian, 2020; Goo, 2012, 2014a, 2014b, 2016; Kim et al., 2015; Li, 2013, 2015; Mackey et al., 2002, 2010; Mackey & Sachs, 2012; Mackey, 2013; Révész, 2012; Yilmaz & Granena, 2019), language analytic ability (Yilmaz, 2013a), and feedback timing (e.g., Li at al., 2016; Sanz et al., 2016; Quinn, 2014; Yilmaz & Sagdic, 2019).

In addition, due to a wide range of methodological variability in research on CF, care should be taken when making conclusions about which type of feedback may be better. Considering the existence of a multitude of factors that mediate the effectiveness of CF, making the interpretation of comparative studies difficult, Goo and Mackey (2013) suggest

that it would be more fruitful to devote research efforts toward investigating how and to what extent various types of CF would impact different target features in different contexts. In an effort to address this issue, a large body of research has examined the effects of different types of CF on different linguistic targets in various learning environments. Of the feedback strategies, recasts have received most attention from SLA scholars.

2.4. Recasts

2.4.1. Recasts in L2 Acquisition

As has been previously noted, recasts have been identified as the most frequent feedback strategy, used in a wide range of instructional settings (e.g., Bower & Kawaguchi, 2011; Lee, 2013; Li, 2010; Loewen & Philp, 2006; Lyster & Mori, 2006; Lyster & Ranta, 1997; Mackey, Gass, & McDonough, 2000; Panova & Lyster, 2002; Sheen, 2004, 2006), leading to a keen interest among SLA researchers to explore their potential benefits. Despite some mixed findings on the effectiveness of recasts in L2 acquisition (e.g., Ellis, 2007; Ellis et al., 2006; Lyster, 2004; Lyster & Ranta, 2013; Sheen, 2007; see Spada & Lightbown, 2009; Lyster & Saito, 2010; Lyster et al., 2013; Nassaji, 2016 for reviews), a facilitative role for recasts in SLA is well established (see Goo & Mackey, 2013; Mackey & Goo, 2007; Nassaji, 2015, 2016; Rassaei, 2022; Russell & Spada, 2006 for reviews). As the most extensively researched feedback, recasts have been defined by various scholars. Some definitions by scholars who are well-known for their work on CF are described in Table 3.

Table 3. Definitions of Recasts

Reference	Definitions
Long (2007, p. 77)	A *corrective recast* may be defined as a reformulation of all or part of a learner's immediately preceding utterance in which one or more non-target like (lexical, grammatical, etc.) items is/are replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on *meaning*, not language as an object. (emphasis in original)
Loewen & Reinders (2011, p.148)	A response to an error in learners' oral production that involves the reformulation of the incorrect linguistic element while maintaining the overall meaning of the utterance.
Lyster & Ranta (1997, p. 46)	[…] teacher's reformulation of all or part of a student's utterance, minus the error.
Mackey (2012, p.14)	[…] a form of feedback in which learners are provided with more target-like versions of their immediately preceding non-target-like utterance.
Sheen (2011, p. 2)	A reformulation of the learner's erroneous utterance that corrects all or part of the learner's utterance and is embedded in the continuing discourse.

As can be inferred from the definitions above, there are three primary characteristics of recasts: (a) they provide both positive and negative evidence, (b) they follow immediately after a learner error, and (c) they are meaning-oriented and unobtrusive. A number of theoretical accounts have been proposed to explain why these characteristics of recasts may be beneficial for L2 learners. First, recasts have the capacity to provide learners with both positive evidence (i.e., the provision of the correct form) and negative evidence (i.e., a signal indicating the occurrence of an error) about the target features (Leeman, 2003). The facilitative role of positive and negative evidence in L2 acquisition is supported respectively by the Direct Contrast Hypothesis and the Universal Grammar approach for language

acquisition. According to the direct contrast hypothesis (Saxton, 1997, 2000), one essential function of recasts is to serve as negative evidence, since it raises learners' awareness of their error through immediately contrasting the correct alternative with their ill-formed production. Within the first language (L1) acquisition, Saxton (1997) explains:

> When a child produces an utterance containing an erroneous form which is responded to immediately with an utterance containing the correct adult alternative to the erroneous form (i.e., negative evidence is supplied), the child may perceive the adult form as being in *contrast* with the equivalent child form. Cognizance of a relevant contrast can then form the basis for perceiving the adult form as a correct alternative to the child form. (emphasis in original, p.155)

In a similar vein, SLA scholars have suggested that the juxtaposition of the learner's non-target-like utterance with the feedback provider's target-like reformulation may prompt the learner to notice the gap and subsequently make a cognitive comparison between the grammatical norm and their interlanguage form (Doughty, 2001; Long, 1996, 2007).

The universal grammar approach, on the other hand, argues that the effectiveness of recasts can be attributed to their role as positive evidence, in that the appropriate correct form is presented in the input in an accessible way (Nicholas et al., 2001). In the context of L1 acquisition, White (2003) states:

> The child's linguistic experience includes what is known as *positive evidence*; that is, the primary linguistic data include utterances that in some sense reveal properties of the underlying grammar […]. *Negative evidence*, or information about ungrammaticality, is not (reliably) available. Nevertheless, children come to know that certain sentence types are disallowed; furthermore, they acquire knowledge that certain interpretations are permitted only in certain contexts […]. (emphasis added, p. 4)

Similarly, in the context of L2 acquisition recasts, with their capacity to model, can also assist learners in encoding new knowledge and creating initial mental representation of an L2 form. In contrast, output-prompting feedback strategies do not provide positive evidence that learners can refer to when constructing a target-like representation of the L2 form in their interlanguage system. Thus, output-prompting feedback may only be able to promote automatizing the retrieval of structure that has already been internalized in learners' interlanguage system regardless of whether it is grammatical or otherwise (Long, 2007). In Sharwood Smith's (1986, 1993) terminology, while recasts can advance the acquisition of new knowledge, other corrective moves may only increase learners' control over already-internalized forms.

Taken together, although these two frameworks have garnered competing views regarding the utility of recasts in language learning, both highlight two key characteristics of recasts. First, recasts can promote cognitive comparisons between the target form and a problematic utterance, which may ultimately lead to the rejection of the erroneous form (i.e., in the case of negative evidence). Second, they provide a correct alternative, thereby enhancing the likelihood that learners notice the target-like norm and potentially incorporate it into their developing grammar (i.e., in the case of positive evidence) (Leeman, 2003).

Second, in terms of the contingency of recasts (e.g., Long, 2007), as previously discussed as part of the direct contrast hypothesis (Saxton, 1997, 2000), it has been claimed that contingency assists learners in noticing discrepancies between their non-target-like production and the target-like model, as it allows learners to juxtapose their ill-formed output and the grammatical alternative, thereby allowing cognitive comparisons of two contrasting structures (Doughty, 2001; Gass & Mackey, 2006; Leeman, 2003; Long, 1996, 2007, 2015). This, in turn, may prompt learners to refine and restructure their interlanguage representation

of the linguistic target, eventually leading to the acquisition of the target-like form-meaning connection. Moreover, recasts provided immediately after the learner's problematic utterances can create optimal conditions for L2 acquisition in that they are reformulations of the learners' own production. By virtue of this, the meaning of the feedback (at least part of the message) has already been comprehended by the learners, and this may allow more attentional resources to be directed to the linguistic target in recasts, thereby fostering form-function mapping (e.g., Doughty, 2001; Goo, 2020; Long, 1996, 2007, 2015;). Empirical evidence has also demonstrated that contingent recasts are better noticed than non-contingent ones (Lai et al., 2008), providing further justification that contingency may play a beneficial role in facilitating the noticing of recasts.

It is worth noting that non-contingent recasts, those not immediately following an error, are less likely to trigger a cognitive comparison, as learners may not be able to store and process the information long enough in their WM, given that "verbal short-term memory is limited in the duration of storage as well as the number of items" (Cowan, 1988, p. 166). Thus, non-contingent recasts are more likely to remain unnoticed (Lai et al., 2008). Moreover, non-contingent recasts may cause learners to devote more attention to deciphering the meaning of the feedback, resulting in less cognitive resources that can be allocated to form and form-function mapping.

Finally, the meaning-oriented and unobtrusive nature of recasts can promote learner attention to L2 features, helping learners focus on meaning and communication without either interrupting the natural flow of conversation or diverting their focal attention from meaning (Doughty, 2001). This is an important advantage of recasts over other types of CF that inevitably disrupt communicative interactions, such as an explicit correction or feedback techniques requiring learners to correct their erroneous production before continuing with the

conversation. Thus, recasts likely provide ideal conditions for establishing and developing L2 form-meaning mappings (e.g., Doughty, 2001; Gass & Mackey, 2006; Long, 2007) with "roving attention to linguistic forms while the remainder of selective attention remains engaged in processing meaning" (Doughy, 2001, p. 251).

Consistent with these theoretical predictions, the effectiveness of recasts in promoting L2 development is well documented in SLA research. Of particular relevance here is an extensive number of studies that found a positive relationship between recasts and development in the knowledge of L2 grammatical features, such as question formation (e.g., Ellis & Sheen, 2006; Li & Iwashita, 2019, 2021; Loewen & Nabei, 2007; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003; Sheen, 2006; Zhao & Wang, 2016); third person singular (Kourtail & Révész, 2020; Zhao & Ellis, 2020); articles (Sheen, 2008, 2010; Nassaji, 2017); simple past tense (Doughty & Valera, 1998; Han, 2002; McDonough, 2007; Sue & Jiang, 2020); present progressive tense (Lyster & Mori, 2006; Mackey & Gass, 2005); past progressive tense (Révész, 2012; Révész & Han, 2006); possessive determiners (Ammar & Spada, 2006); *that*-trace filter (Goo, 2012); grammatical gender (Lyster, 2004; Lyster & Izquierdo, 2009); adjective ordering (Long et al., 1998); relative clauses (Kim, 2019); locative construction (Long et al., 1998); locative prepositions (Nanatsukasa, 2016); and noun–adjective agreement (Leeman, 2003) among many others.

Given that the effectiveness of recasts is influenced by a number of variables, researchers have also examined various potential factors that may mediate the effects of recasts on L2 development in relation to: recast features/types (e.g., Al-Surmi, 2012; Erlam & Loewen, 2010; Nassaji, 2017; Sagarra & Abbuhl, 2013; Wacha & Liu, 2017); task features (e.g., Kim, Payant, & Pearson, 2015; Kourtali & Révész, 2020; Révész, Sachs, & Hama, 2014); individual differences in cognitive capacities (e.g., Goo, 2012, 2014a, 2014b, 2016;

Granena & Yilmaz, 2019; Kim, 2019; Kim et al., 2015; Kourtali & Révész, 2020; Li, 2013; Li, Ellis, & Zhu, 2019; Mackey et al., 2002, 2010; Mackey & Sachs, 2012; Révész, 2012; Yilmaz & Granena, 2016, 2019); proficiency level (e.g., Ammar & Spada, 2006; Mackey & Philp, 1998); and target structures (e.g., Kim, 2019; Li, 2014; Li & Iwashita, 2019, 2020; Long et al., 1998; Mackey, 2006; van de Guchte et al., 2015; Varnosfadrani & Basturkmen, 2009).

Another potential mediating factor that has attracted growing interest from SLA researchers is SCMC modality (e.g., text chat, audio-chat, and video-conferencing), which will be discussed in turn.

2.4.2. Recasts in SCMC Research

In light of the favorable findings for recasts in the broader realm of SLA as well as pedagogical and technological advances using SCMC applications, an increasing number of scholars have embarked on research endeavors aimed at exploring the potential utility of recasts for L2 grammatical development in diverse SCMC contexts. Two main groups of research have examined the effectiveness of recasts in this area. One group has focused on text-based SCMC, while the other has been conducted in video-based SCMC. So far, this line of research has primarily been concerned with investigating recasts during text-based interaction (see Ziegler, 2016; Ziegler & Mackey, 2017 for a comprehensive review). Research on text-based recasts has been largely concerned with two interactional contexts: SCMC (e.g., Lai et al., 2008; Baralt, 2010, 2015; Gurzynski-Weiss et al., 2016; Kourtali, 2022; Kourtali & Borges, 2023; Loewen & Erlam, 2006; Sachs & Suh, 2007; Sauro, 2009; Yilmaz, 2012; Yilmaz & Sagdic, 2019; Yilmaz & Yuksel, 2011); and non-communicative CALL environments (e.g., Sachs, 2010; Sagarra, 2007, Sagarra & Abbuhl, 2013;

Trofimovich et al., 2007; Zalbidea & Sanz, 2020). Along with other types of feedback, researchers have examined the developmental benefits of recasts in a wide range of topics: contingent versus non-contingent recasts and learner noticing (Lai et al., 2008); textually enhanced versus non-enhanced recasts and awareness (Sachs & Suh, 2007); computerized oral and/or written feedback/recasts (Sachs, 2011; Trofimovich et al., 2007; Zalbidea & Sanz, 2020); computer-based orally and/or textually enhanced recasts (Sagarra & Abbuhl, 2013); recasts versus more explicit feedback strategies (Loewen & Erlam, 2006; Sauro, 2009; Yilmaz, 2012); levels of awareness in relation to type of recast and type of linguistic item (Gurzynski-Weiss et al., 2016); recasts provided during tasks with varying cognitive complexity (Baralt, 2010, 2013, 2015); recasts supplied in response to more and less salient linguistic features (Yilmaz, 2012; Yilmaz & Yuksel, 2011); face-to-face versus text-based recasts (Baralt, 2010, 2015; Kourtali, 2022); timing of recasts (i.e., immediate versus delayed) (Kourtali & Borges, 2023; Yilmaz & Sagdic, 2019); recasts on learners' uptake (Kourtali, 2022); correlation between WM capacity and learners' ability to benefit from recasts (Baralt, 2010, 2015; Lai et al., 2008; Sachs, 2010; Saggara, 2007; Sagarra & Abbuhl, 2013; Trofimovich et al., 2007; Zalbidea & Sanz, 2020; see Section 2.7.5 for detailed reviews of these studies).

SCMC research exploring different types of CF has primarily compared recasts and explicit CF (e.g., recasts versus metalinguistic feedback or explicit correction) and has yielded mixed results, either showing no difference between them (Loewen & Erlam, 2006; Sauro, 2009), or an edge of explicit CF over recast (Yilmaz, 2012). Loewen and Erlam (2006), for example, investigated the efficacy of recasts and metalinguistic feedback on the acquisition of English regular past tense in text-based SCMC. This work is a partial replication of Ellis et al.'s (2006) study, which compared the effectiveness of these two types

of CF in the face-to-face classroom setting. Thirty-one elementary learners of L2 English completed two communicative tasks, during which they received either recasts or metalinguistic feedback in response to their erroneous utterances of the target structure. L2 development was measured by timed and untimed grammaticality judgement tests. In contrast to the original study, that showed a clear advantage for metalinguistic feedback over recasts for both the delayed imitation and grammaticality judgement posttests, Loewen and Erlam found no significant learning gains in response to either type of feedback. The authors speculated that one possible reason for this is the reduced contingency of the feedback. That is, unlike face-to-face interaction in Ellis et al.'s (2006) study, CF that was given in response to a particular error during text chat often appeared after multiple intervening turns unrelated to the initial error due to the overlap between interlocutor turns. Such non-contingent CF may not have been perceived as correction in relation to the given error and, consequently, the focus on the target linguistic form may not have been achieved.

Similarly, Sauro (2009) compared the effectiveness of text-based recasts and metalinguistic feedback in the acquisition of the zero article with abstract non-count nouns (e.g., employment, global warming, culture) with twenty-three intermediate and advanced learners of English at a Swedish university. Results of the immediate and delayed posttests (acceptability judgements) showed no significant advantage for either feedback type on immediate or delayed posttest gains. However, the metalinguistic feedback group descriptively outperformed the recast group and achieved statistically greater immediate posttest gains than the control group. Sauro interpreted the results as suggesting that the full recasts that reformulated a considerable length of initial utterance in her study might not have been perceived by learners as correction but confirmation checks or repetition of their prior messages, resulting in lower learning gains than those of metalinguistic feedback. Another

possible factor, she speculated, was the lack of adjacency of the recast to the initial utterance containing an error. Similar to Loewen and Erlam (2006), additional turns often separated the recast from the initial utterance. Hence, coupled with the length of recast, the lack of adjacency may have limited opportunities for the learners to perceive the recasts as correction, making comparisons between their own output and the more target-like reformulation difficult. Taken together, from the findings of the two aforementioned studies, it is suggested that in order to increase the effectiveness of recasts in SCMC-based interaction, it seems beneficial to make their corrective nature sufficiently apparent to learners and provide them in a contingent manner.

Some studies have contrasted recasts and explicit correction in SCMC versus face-to-face contexts. Yilmaz (2012) investigated the relative efficacy of the two feedback strategies and target form salience (salient versus non-salient) on the acquisition of two L2 Turkish morphemes in text-based and face-to-face modes. Forty-eight native speakers of English carried out two information gap tasks during which their errors on the target structures were treated according to their feedback conditions. Participants' learning gains were measured using oral production, comprehension, and recognition tests. Yilmaz found that scores were significantly higher in SCMC than face-to-face mode on the immediate oral production posttest and the immediate and delayed recognition posttest for the learning of two linguistic targets. A clear advantage was found for explicit correction over recasts in the oral production and comprehension tasks on both immediate and delayed posttests, but not on a recognition measure.

Results showed that both explicit correction and recasts were more effective in the SCMC mode on the oral production and recognition tests. Also, there was no interaction between feedback type and communication mode. In other words, explicit correction was

more effective than recasts, regardless of mode. With regard to the type of linguistic construction, the learners achieved higher scores on the salient construction in comparison to the non-salient one for all outcome measures. No interaction was found between the salience of the target forms and type of CF. That is, both explicit correction and recasts were more successful in facilitating L2 learning when they addressed salient construction.

The reason why explicit correction was more advantageous over recasts in text-based mode may be attributable to a differential salience of the corrective intent between these two types of feedback. Thus, considering that recasts are viewed as more implicit on the continuum of explicitness (Mackey, 2012), in order to take most advantage of the written modality, it seems helpful to provide them immediately following the error, thereby facilitating learners' noticing and cognitive comparisons (Doughty, 2001; Gass & Mackey, 2006; Leeman, 2003; Long, 1996, 2007, 2015).

Other research has compared the impact of text-based versus face-to-face recasts on the acquisition of two L2 Turkish morphemes. Yilmaz and Yuksel (2011) examined the impact of text-based versus face-to-face recasts on the acquisition of Turkish locative and plural morphemes. The study also investigated whether learners' scores on the target constructions differed depending on the communication mode. Twenty-four participants were paired up with either one of the authors to carry out the treatment tasks, and each participant received both face-to-face and computer-mediated recasts. Two oral production tasks were used as a posttest to measure learners' performance on the two target structures. Aligning with Yilmaz's (2012) findings, the researchers found that the participants' scores were significantly higher when they received recasts during text chat, with greater noticing of form than they did in face-to-face interactions. This suggest that communication modality may influence the extent to which learners benefit from recasts. However, learners' performance

did not vary depending on the target structures. Yilmaz and Yuksel concluded that re-readability of messages in text-based mode could have facilitated noticing of corrective force of recasts as well as making accurate form-meaning connections, However, the authors noted that care should be taken when generalizing their results, as the study was conducted in a lab where contingency between the error and recasts was tightly controlled, in contrast to naturalistic SCMC environments in which recasts may not be contingent on errors (Lai et al., 2008). This seems to be a logical conclusion, as it can be seen in SCMC research mentioned above, the common "split turns" in text-based interactions (Smith, 2003), in which the reformulation of an error appears many lines after the relevant error, likely hinders learners' cognitive comparison between their erroneous production and the well-formed model, as well as focus on form (e.g., Lai & Zhao, 2006; Lai et al., 2008; Loewen & Erlam, 2006; Sauro, 2009). Taken together, the findings may indicate that if the text-based recasts are controlled so as to be presented immediately after the relevant error during SCMC interactions, they are likely to be more effective than oral recasts, which do not provide re-readable, written records of the feedback. Yilmaz and Yuksel also suggested combining additional measures of noticing with performance measures to shed light on the question of whether the difference in performance between the two communication modes is due to differences in quantity or quality of noticing.

By contrast, Kourtali (2022) found that face-to-face recasts were more effective in promoting successful uptake and generating L2 gains than the recasts delivered via text chat. Her study examined whether the mode of interaction impacts L2 learners' successful uptake and learning gains when receiving recasts. 60 young Greek English as foreign language learners were assigned to either face-face or text-based modes, completing information transmission tasks that required them to describe the habits of fictional characters. Both

groups received interrogative, partial recasts addressing errors in the present third person singular, the target construction. L2 development was measured through oral and written production tests. The study found that recasts delivered in face-to-face mode generated significantly more modified output than those supplied in text-based mode. Moreover, the face-to-face group achieved significantly greater L2 gains than the SCMC group on both oral and written production tests. Modified output was associated with learners' improvement only in the face-to-face condition.

Kourtali suggested that when learners receive text-based recasts, they may prioritize meaning over form, thus allocating less attention to low salient and communicatively redundant linguistic features; in her study, many learners responded to the recasts by providing additional information related to the pictures of the task without correcting their ill-formed responses, indicating that the corrective intention of recasts remained unattended. This finding is in part consistent with previous research that has shown that written SCMC recasts are less effective when addressing low salient features (e.g., the past tense in Loewen & Erlam, 2006; the English zero article in Sauro, 2009; Turkish locative case morphemes in Yilmaz, 2012). Nonetheless, it is still unclear whether this pattern would also arise in video-based SCMC, where low salient features are corrected by recasts delivered via various feedback modalities. This issue seems to merit further research, particularly given contradictory results observed in previous studies regarding the effect of target feature salience on learners' learning gains (e.g., Yilmaz & Yuksel, 2011).

As of now, scant research has explored the potential impact of recasts on L2 grammatical development during video-conferencing. Of the few existing studies, Monteiro (2014) investigated the effects of recasts and metalinguistic feedback in the development of implicit and explicit knowledge of English regular past tense in relation to focused tasks (i.e.,

designed to create a context for a pre-determined linguistic feature; see Ellis, 2003, 2016, 2019 for the characteristics of unfocused and focused tasks). Forty-two Brazilian learners of various levels received recasts, metalinguistic feedback, or no feedback on their errors in the target form while completing focused tasks during video-based interaction. Learning gains were measured using untimed and timed grammaticality judgement tests, a metalinguistic knowledge test, and an oral imitation test. Results revealed that all three groups were equally effective in developing learners' implicit and explicit knowledge of the target structure. The effectiveness of recasts and metalinguistic feedback with focused tasks in SCMC context is aligned with laboratory-based studies, which suggest that both feedback strategies tend to be effective when they are intensive, individualized, and controlled (e.g., Loewen & Nabei, 2007; Lyster & Izquierdo, 2009).

In a more recent inquiry, Rassaei (2017) explored whether the provision of recasts in video-based or face-to-face interactions would lead to greater learning outcomes in the use of the English definite article "the". Fifty-seven Iranian EFL learners were assigned to one of three conditions (face-to-face recasts, video-based recasts, and no recasts) and involved in three dyadic storytelling treatment tasks during which the experimental groups' errors were corrected by the interlocutor's recasts. L2 gains were measured by an oral production task and an error correction test, and the learners' interpretations of recasts in the two modes was further examined in a stimulated recall interview. Results indicated that computer-mediated and face-to-face recasts were equally effective, and both experimental groups outperformed the control group. This suggests that modality of recasts did not have impact on the extent of L2 gains exhibited by the learners. Furthermore, stimulated recall interviews revealed no statistical difference in learners' interpretations of recasts in the two modalities.

In his later work, Rassaei (2019) delved into the effects of recasts on the development of EFL learners' grammatical knowledge in using English articles "a" and "the" during mobile-mediated audio and video interactions. Fifty-two Iranian intermediate EFL learners were randomly assigned into four groups: audio recasts and video recasts groups as well as audio control and video control groups. The experimental groups received either audio-based or video-based recasts for their errors while performing communicative treatment tasks with their interlocutors via their smartphones. The control groups completed the same tasks in the audio and video modes without receiving any recast. Learners' knowledge development was measured using a picture-description task and a writing task. Results showed that recasts had a positive impact on participants' accuracy development both in audio-based and video-based interactions and that participants in the audio recast group produced more target-like modified output compared to those of the video recast group. With respect to learners' perceptions of mobile-mediated recasts, stimulated recall interviews revealed that both experimental groups were mostly successful in correctly interpreting mobile-mediated recasts as CF, but the audio recast group were more accurate in interpreting recasts than the video recast group. The results thus suggest that learners' accurate interpretation of recasts and the effectiveness of recasts may be associated. Previous research has also provided evidence for a positive association between the rate of modified output and learners' noticing of feedback during face-to-face and text-based interaction (e.g., Gurzynski-Weiss & Baralt, 2015; Gurzynski-Weiss et al., 2017).

Similarly, Aeen, Rassaei, Riasati, and Zamanian (2021) examined the relationship between learners' perception of recasts and their modified output in video-based mobile-mediated interaction. Thirty intermediate Iranian EFL learners were asked to retell three short stories and received recasts for their errors in the use of English definite and indefinite

articles. During a stimulated recall interview, learners were asked if they had noticed what the teacher was saying at the time it was said (noticing) and why (perception). Then the video-recorded mobile interactions were analyzed to determine whether the learners' perception of the recast triggered them to modify their erroneous output. Results indicated that a great majority of the reformulated errors occurred after recast that had been perceived by the learners. The results also found a high rate of modified output where the participants perceived the mismatch between their erroneous utterances and the recast. On the contrary, where the learners simply noticed the errors without being able to perceive the mismatch between them and the recast, the rate of modified output decreased significantly. Further analysis also confirmed that there was a significant relationship between learners' perception of mobile-mediated recast and their modified output.

Martin's (2018) study is the only one that addressed the differential effects of various SCMC modalities on L2 grammar learning. The study involved 65 intermediate learners of Spanish who completed a dialogic story retell task, where metalinguistic feedback was used to correct errors in their use of Spanish past subjunctive. Participants were randomly assigned to one of the three groups: video-based, voice-based, and text-based. To prevent learners from repeatedly reading the written feedback, the text-based group was required to close the chat after receiving feedback and reopen it to continue the interaction. L2 gains were measured using an oral production test (a controlled, monologic retell story task) and a written production test (a semi-spontaneous picture-description task). The results showed that all three groups exhibited improvements from the pretest to the immediate posttest, with many retaining these gains one week after the treatment. The learning gains did not significantly differ across groups. This suggests that feedback modality may not impact the extent to which learners benefit from the metalinguistic feedback in SCMC interactions.

However, it remains uncertain whether these findings would be replicated when employing an implicit type of feedback, such as recasts. Therefore, further research is warranted.

Another important finding of the study was that L2 learning in one mode (oral/written) was transferable to the other. That is, participants who carried out a task in the written modality exhibited improvements in the oral production test. Likewise, those who completed a task in the oral mode demonstrated learning gains in the written production test. Thus, it is worth examining whether this transferability of L2 learning is also applicable to the present study, in which learners complete tasks via different modalities and have to produce the target forms in both oral and written production tasks.

In order to provide a concise overview of the research discussed, Table 4 below summarizes the studies in the same order as they were presented in this section.

2.4.3. Summary

As the above review suggests, compared to the studies that investigated the effects of recasts on L2 development during face-to-face interactions, how computer-delivered recasts affect L2 learning remains underexplored. More importantly, most SCMC research has focused on recasts during text-based interaction, while little attention has been given to other modes of SCMC, particularly video-based recasts. In addition, previous studies of video-based CF including recasts have primarily employed oral mode only, yielding inconsistent results, only exception being Martin's (2018) study. This highlights a clear necessity for further investigations into whether and to what extent other SCMC modalities can enhance the effectiveness of recasts in facilitating L2 acquisition in the context of multimodal video-conferencing SCMC.

One of the key benefits of video-conferencing is the availability of multiple modalities such as oral-based input, text-based input, visual aids (e.g., pictures, video clips), as well as aural and paralinguistic cues (e.g., intonation and facial expressions). These features have the potential to achieve contingency and enhance the salience of recasts, leading to greater corrective force, compared to face-to-face and unimodal recasts (Hampel & Stickler, 2012; Wang, 2006). As the preceding review of SCMC research suggests, the effectiveness of recasts is significantly affected by the contingency and salience of their corrective intent. Making suggestions for future SCMC research, Monteiro (2014, p. 71) also rightly suggests:

> Future research in video-conferencing is promising, since it allows not only for audio and video interactions, but also for the use of different modes (e.g., text chat, pictures) simultaneously; therefore, future studies should explore how those modes can be combined in order to enhance corrective feedback.

Therefore, to fully assess the potential benefits of video-based recasts for L2 acquisition, it seems crucial to explore whether multimodal recasts that combine oral and written recasts may offer advantages over receiving oral or written recasts alone.

Table 4. Studies on the role of CMC/SCMC feedback modality in L2 learning

Study (year)	Modality	Participants	Target Structure	CF Type	Treatment	Assessment	Results
Loewen & Erlam (2006)	Text-based	31 elementary learners of L2 English	English regular past tense	1) Recasts 2) Metalinguistic feedback	Communicative tasks	Timed and untimed GJTs	No significant learning gains in response to either type of feedback
Sauro (2009)	Text-based	23 intermediate and advanced learners of English	Zero article with abstract non-count nouns	1) Recasts 2) Metalinguistic feedback	Two collaborative writing activities	Acceptability judgments	1) No significant advantage for either feedback type on immediate or delayed posttest gains. 2) Descriptively, metalinguistic feedback group outperformed the recast group and achieved statistically greater immediate posttest gains than the control group.
Yilmaz & Yukcel (2011)	Text-based	24 Turkish learners	Turkish plural and the locative case morphemes	Recasts	Two picture-description tasks	Two picture-description tasks	Scores were significantly higher when they received SCMC recasts than when they received FTF recasts.
Yilmaz (2012)	1) Text-based 2) FTF	48 Turkish learners	Turkish plural morpheme, and the locative case morpheme	1) Recasts 2) Explicit correction	Information gap tasks	1) OPT 2) Comprehension task 3)Recognition task	1) Explicit feedback was more effective than recasts. 2) Effect for communication mode (in favor of SCMC) on the immediate OPT and both recognition posttests, but not on the comprehension task.

Study (year)	Modality	Participants	Target Structure	CF Type	Treatment	Assessment	Results
							3) No interaction between feedback type and communication mode
Monteiro (2014)	Video-based	42 Brazilian learners	Regular past tense	1) Recasts 2) Metalinguistic feedback 3) Control	Focused tasks	1) Untimed and timed GJTs 2) Metalinguistic knowledge test 3) Oral imitation test	All three groups were equally effective in developing learners' implicit and explicit knowledge of the target structure.
Rassaei (2017)	1) Video-based 2) Face-to-Face	57 Iranian EFL learners	English definite article 'the'	1) Face-to-Face Recasts 2) Video-based recasts 3) Control	Dyadic story telling	1) Oral production task 2) Error correction test 3) Stimulated recall interview	1) Both experimental groups outperformed the control group. 2) No modality impact on the extent of the learning gains.

Study (year)	Modality	Participants	Target Structure	CF Type	Treatment	Assessment	Results
Rassaei (2019)	Mobile-mediated 1) Audio-based 2) video-based	52 Iranian EFL learners	English articles “a” and “the”	1) Audio-based recasts 2) Video-based recasts 3) Audio control 4) Video control	Communicative tasks	1) Oral picture-description task 2) Writing task 3) Stimulated recall interview	1) Recasts had a positive impact on participants’ accuracy development both in audio-based and video-based interactions 2) The audio recast group produced more target-like modified output compared to the video recast group.
Aeen et al. (2021)	Video-recorded mobile interaction	30 intermediate Iranian EFL learners	English definite and indefinite articles	Mobile-mediated video-based recasts	Video-recorded mobile interaction	1) Stimulated recall interview 2) Analysis of video-recorded mobile interactions	There was a significant relationship between learners’ perception of mobile-mediated recast and their modified output.
Martin (2018)	1) Video-based 2) Voice-based 3) Text-based	65 intermediate learners of Spanish	Spanish past subjunctive	Metalinguistic feedback	Interactive dialogic story retelling	1) OPT 2) WPT	1) All three groups successfully acquired the target structure, and many maintained the learning gains at the delayed posttest. 2) No modality impact on the learning gains.

Study (year)	Modality	Participants	Target Structure	CF Type	Treatment	Assessment	Results
Kourtali (2022)	1) Text-based 2) FTF	60 Greek EFL learners	English third person singular	1) Text-based Recasts 2) FTF recasts	1) Information transmission tasks 2) Describing fictional characters	1) OPT 2) WPT	1) FTF recasts generated significantly more modified output and achieved significantly greater L2 gains than text-based recasts. 2) Modified output was associated with improvement only in the FTF condition.

Note. OPT: Oral Production Test, WPT: Written Production Test, GJT: Grammaticality Judgment Test, FTF: Face-to-Face Communication.

2.5. Multimodal Recasts

2.5.1. Theoretical Support for Multimodal Recasts

There are several reasons why multimodal recasts are likely to be more effective in promoting L2 development than the provision of recasts in a single modality. First, it is worthwhile to consider the potential benefits of multimodal recasts from a theoretical perspective. The theoretical advantages of multimodal recasts are mainly grounded in two information processing models: Dual Coding Theory (Paivio, 1986, 2007) and the Cognitive Theory of Multimedia Learning (Mayer, 2005, 2009, 2014). Insights from Cognitive Load Theory (Chandler & Sweller, 1991; Sweller, 1988, 1994, 2005) can also help us gain a deeper understanding of how the use of multimodal input can facilitate L2 learning.

2.5.1.1. Dual Coding Theory

According to Paivio's (1986, 2007) Dual Coding Theory, humans process information through two systems; one is *a verbal system*, which handles verbal information (e.g., auditory, articulatory input, spoken/written words), and the other is *an imagery system*, which processes nonverbal objects and event (e.g., images and nonlinguistic input) (see Figure 2). These two systems are activated after external stimuli are received by the corresponding *logogens* (i.e., verbal representation) or *imagens* (i.e., nonverbal representations) respectively.

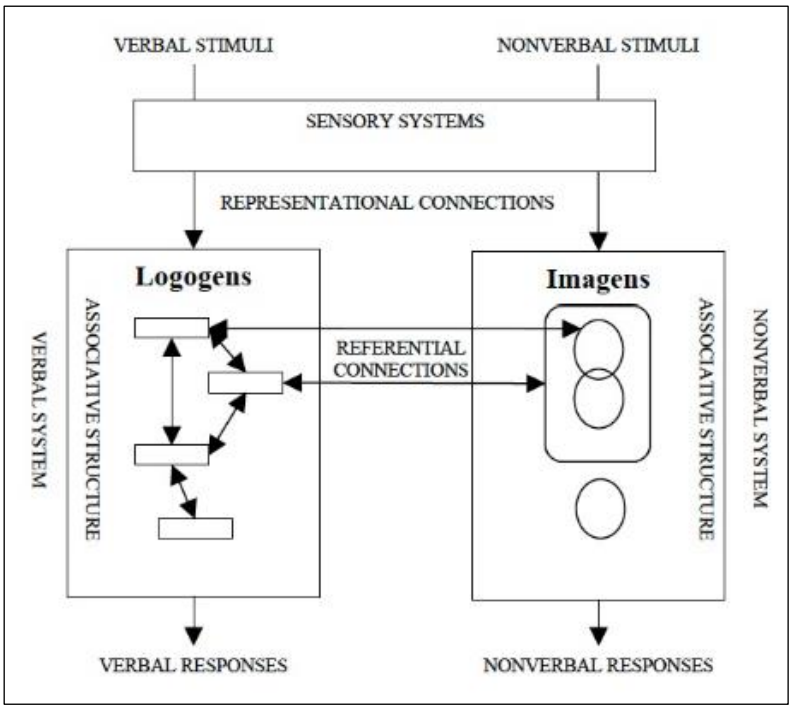


Figure 2. Dual Coding Theory (Paivio, 1986, p. 67)

Paivio hypothesizes that these two systems are functionally independent, namely, that either only one system is active or both systems are active in parallel, but that they are also interconnected and work cooperatively through *referential connections* (i.e., indirect activation between imagens and logogens). Thus, representations in one channel can activate those in the other. For instance, a linguistic stimulus (e.g., a particular word) in the verbal system can evoke related images and/or feelings in the nonverbal system. Another link, *associated connections*, links representations within each system. That is, a particular word received in the verbal system (e.g., winter) may elicit other related words (e.g., snow). Similarly, a particular imagen may evoke related imagen in the nonverbal system. Based on this assumption, Paivio (2007) hypothesizes that activation of both verbal and nonverbal systems can facilitate information processing, which may lead to better learning and recall. In other words, the provision of input through multimodalities (i.e., aural, visual, textual) may increase the opportunities to stimulate both the verbal and imagery systems, and this is

expected to facilitate information processing, resulting in better learning and recall of the input processed (Paivio, 1986, 2007).

2.5.1.2. Cognitive Theory of Multimedia Learning

Another theoretical support for the pedagogical potential of multimodal recasts is Mayer's (2001, 2005) Cognitive Theory of Multimedia Learning, which was influenced by the Dual Coding Theory and Baddeley's (Baddeley, 1999; Baddeley, Eysenck, & Anderson, 2009) Model of Working Memory. This theory makes three fundamental assumptions: (a) a dual-channel assumption, (b) a limited capacity assumption, and (c) an active processing assumption. These three assumptions incorporate elements from: Paivio's (1986, 2007) Dual Coding Theory; Sweller's (1988, 2005; Chandler & Sweller, 1991) Cognitive Load Theory, and Baddeley's (1986, 1999) Model of Working Memory with an emphasis on limited processing capacity; and Atkinson and Shiffrin's (1968) Memory Stores and Cognitive Processes.

First, similar to Dual Coding Theory, *the dual-channel assumption* is that there are two channels dealing with different information. There are two ways to conceptualize these two channels: the presentation-mode approach (Paivio, 1986, 2006), which distinguishes between a verbal channel (e.g., spoken or written words) and a nonverbal channel (e.g., pictures, graphs); and the sensory-modality approach (Baddeley, 1999; Baddeley, Eysenck, & Anderson, 2009), which differentiates between an auditory channel (processing the information through the ears such as spoken words) and a visual channel (processing the information through the eyes such as printed words). *The limited capacity assumption* is that due to a limited processing capacity, humans can only process a limited amount of information in each channel at a time. Therefore, instead of processing all input available,

humans are forced to make decisions about the allocation of attention and the degree to which the information is organized and integrated. *The active processing assumption* refers to the idea that instead of passively receiving information, humans are engaged in active processing: first, selecting relevant information, then organizing that information into auditory and visual mental representations, and finally integrating the selected information with prior knowledge in WM.

As depicted in Figure 3, there are five cognitive processes that learners must engage in for meaningful learning to occur in a multimedia environment: 1) selecting relevant words for processing in verbal WM; 2) selecting relevant images for processing visual WM; 3) organizing selected words into a verbal model; 4) organizing selected images into a pictorial model; and 5) integrating verbal and pictorial representations with each other and also with prior knowledge (Mayer, 2014). These processes do not necessarily occur linearly, but more likely occur segment by segment and frequently during a multimedia presentation.

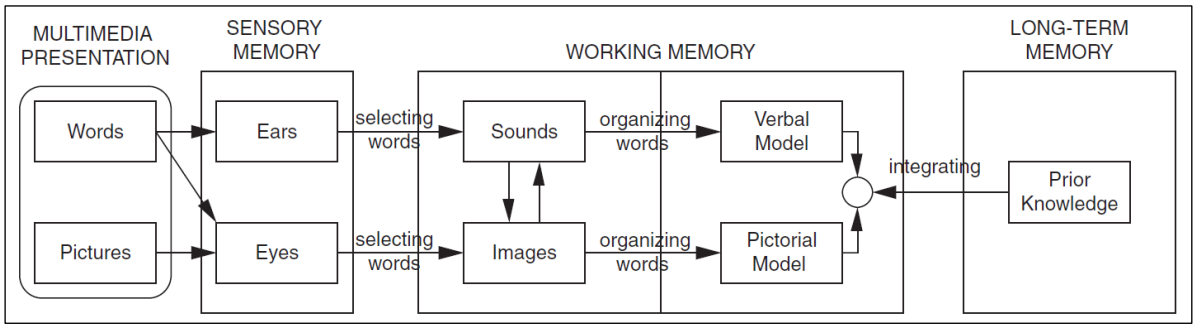


Figure 3. Cognitive Theory of Multimedia Learning (Mayer, 2014, p. 52)

As such, the use of both auditory and visual presentations can prime both aural and visual systems, thus facilitating the establishment of mental connections between the two representation systems, potentially leading to deeper and greater cognitive processing and better retrieval (Mayer, 2014; Mayer & Fiorella, 2014). Thus, similar to Paivio's view, Mayer

claims that learning and understanding can be enhanced when learners construct information from two different channels. It is worth noting, however, the difference between the Dual Coding Theory and the Cognitive Theory of Multimedia Learning is that, while the former focuses on the benefits of presenting information in verbal and nonverbal forms to promote learning process, the latter also emphasizes the importance of considering learners' limited cognitive capacity and their active role in multimedia learning.

In this regard, when discussing the potential value of multimodal recasts, Sweller's (2005) Cognitive Load Theory also seems relevant, which suggests that individuals have a limited capacity for information processing in a single mode. Thus, when processing linguistic input, learners must be selective in attending only to certain input in a particular modality, as not being selective may result in a cognitive-processing overload (Paas & Sweller, 2014; Sweller, 2005). Thus, when aiming to design effective instructional interventions with multimodal input, this theory suggests it is important to consider different types of cognitive load. We now turn to a discussion of how this may be achieved, drawing upon this framework.

2.5.1.3. Cognitive Load Theory

Cognitive Load Theory (Sweller, 2005) pertains to the cognitive effort necessary in the process of learning new information. The theory posits that acquiring new knowledge and expertise is made easier when the type of learning instruction employed maintains a low cognitive load, thus minimizing the demand on a learner's WM (Chandler & Sweller, 1991). To achieve this, there are three types of cognitive load that need to be either reduced or increased to promote learning. First, *intrinsic cognitive load* is determined by the inherent characteristics or difficulty of the information needed to be processed, which thus cannot be

altered; *extraneous cognitive load*, which is irrelevant for learning, results from inappropriate instructional designs that consume unnecessary cognitive resources, and should be reduced; *germane cognitive load*, on the other hand, is the cognitive load that promotes schema construction. A schema refers to "a cognitive construct that organizes the elements of information according to the manner with which they will be dealt" (Sweller, 1994, p. 296). That is, germane cognitive load relates the new information to prior knowledge, therefore, it should be maximized with instructional interventions. In summary, when designing instructional interventions using multimodal input, the aim should be to reduce the extraneous load and increase the germane load to facilitate effective learning.

For example, when learning a new linguistic target in L2, learners may struggle with processing auditory input as they often face challenges in distinguishing individual words and must use cognitive resources to consciously perceive each word in the spoken input (Mayer et al., 2014). This causes L2 learners to devote resources to decoding unfamiliar sounds, which places a considerable load on their WM (Lee & Mayer, 2018). As a result, only limited cognitive resources are available for comprehension and other more advanced processing tasks, which negatively impacts learners' L2 knowledge acquisition. Thus, to alleviate the intrinsic cognitive load associated with decoding L2 auditory input, providing corresponding text on the screen can be useful (Meskill, 1996). This is because on-screen text helps preserve word availability, enabling learners to encode spoken words more efficiently (Mayer et al., 2014). This, in turn, assists learners in segmenting and comprehending an otherwise incomprehensible stream of speech (Winke, Gass, & Sydorenko, 2010). This likely reduces the cognitive load and enables more resources to be available for learning (Lee & Mayer, 2018).

As such, Cognitive Load Theory predicts that, when auditory and visual information complement and correspond to each other, the burden on WM decreases compared to a single channel used alone (Sweller, 1988, 1994). This potentially assists learners in taking advantage of their full capacity for processing information (Mayer, 2014), thereby enhancing encoding and decoding of the input.

Taken together, based on the theoretical support for the benefits of multimodal input as discussed, it appears logical to expect that multimodal recasts may also have the potential to facilitate building mental connections between the auditory and visual representation systems and to reduce the cognitive load imposed in each modality by distributing the linguistic input through multiple modalities. Consequently, the amount of processing load required in each channel is less likely to exceed the learner's cognitive capacity (Mayer & Fiorella, 2014), which may then facilitate processing linguistic input and subsequent L2 learning.

2.5.2. Benefits of Hybrid Characteristics of Multimodal Recasts

Another potential advantage of multimodal recasts is their hybrid nature, which may enable learners to benefit from the distinct strengths of both oral and written SCMC feedback. Specifically, it has been suggested that text-based SCMC enhances the salience of feedback and target forms as well as allowing more cognitive resources through greater processing time, re-readability of messages, and visual/perceptual salience of the written input (e.g., Hummel & French, 2010; Long, 2007; Payne & Ross, 2005; Payne & Whitney, 2002; Sauro, 2009; Smith, 2004, 2010; Warschauer, 1996a, 1996b; Yilmaz, 2012; Yilmaz & Yuksel, 2011; see Gass et al., 2018 for definitions and types of salience in SLA). Thus, the corrective intention of written recasts is more likely to be explicit to learners than that of oral recasts

(e.g., Blake, 2008; Chapelle, 2001; Smith, 2004, 2010; Yilmaz, 2012). This increased salience may lead to improved opportunities for learners to notice and focus on form, providing better chances for L2 development (Salaberry, 2000). These unique characteristics of text-based SCMC are believe to serve as a cognitive amplifier (Lai & Zhao, 2006; Warschauer, 1997), creating optimal conditions for learners to identify L2 linguistic forms, thereby aiding L2 acquisition (Schmidt, 2001).

Meanwhile, oral recasts provide verbal and nonverbal cues such as stress, intonation, and gestures, which have been proposed as ways to make recasts more salient and noticeable (Gass et al., 2017; Loewen & Philp, 2006). Of greater significance, oral recasts can be provided in a contingent fashion, immediately after learners have made an error in their speech. This is often not possible during text-based interactions, as negotiation routine is often split due to the lack of strict turn adjacency (Smith, 2003). Indeed, contingency of recasts is a key factor in contributing to their effectiveness, as the contrast between the ill-formed learner output and the target-like exemplar renders the negative evidence included in the recasts more salient to learners, thereby prompting noticing and cognitive comparisons during speech processing (Doughty, 2001; Leeman, 2003; Long, 2007).

Taken together, multimodal recasts may have the potential to mitigate the inherent limitations of oral and written recasts by allowing for more effective focus on form during meaningful video-based interaction. This seems to be highly advantageous during video-based interactions, where recasts are frequently misinterpreted as confirmation checks or non-corrective repetitions. Moreover, utilizing multimodal recasts might be particularly helpful when aiming to direct learners' attention to grammatical errors, as morphosyntactic recasts are less apparent to learners compared to recasts targeting lexis or phonology (e.g., Al-Surmi,

2012; Carpenter et al., 2006; Gass & Lewis, 2007; Kim & Han, 2007; Lyster, 1998; Lyster et al., 2013; Mackey et al., 2000; Sheen, 2006; Trofimovich et al., 2007).

2.5.3. Empirical Research on Captioned Audiovisual Input in L2 Grammar

As has been noted, multimodal recasts can be operationalized as captioned oral recasts that present oral and on-screen written recasts simultaneously (see Section 3.1.5.2 for the detailed operationalization of multimodal recasts). A considerable body of research indicates the advantages of caption-supported audiovisual input for L2 learning, particularly in vocabulary acquisition and comprehension (Montero Perez et al., 2013a, 2013b; Peters & Webb, 2018; Rodgers & Webb, 2017). A meta-analysis by Montero Perez et al. (2013b) examining 18 studies on the effects of captions indicated that captions have a statistically significant influence on L2 listening comprehension and vocabulary learning. A comparison of the presence and absence of captions found that the use of captions tend to be more effective than no captions for vocabulary form recognition (e.g., Markham, 1999; Montero Perez et al., 2018; Neuman & Koskinen, 1992; Teng, 2018), meaning recognition (e.g., Neuman & Koshinen, 1992; Teng, 2018), and meaning recall (e.g., Jelani & Boers, 2018; Sydorenko, 2010; Winke et al., 2010).

Researchers have attributed these positive findings to the capacity of captions to aid speech comprehension by assisting learners in parsing and segmenting the speech stream into relevant linguistic input, such as word boundaries and forms (Bird & Williams, 2002; Gass et al., 2019; Vanderplank, 1988). Access to segmented speech, in turn, is assumed to facilitate word recognition, which enables learners to identify new lexical items in the input with greater ease, promoting attention to and learning of these items (Montero Perez, 2020; Winke, Gass, & Sydorenko, 2010).

So far, only a limited number of studies have explored L2 grammar learning from audiovisual input (e.g., d'Ydewalle & Van de Poel, 1999; Lee & Révész, 2018; Van Lommel et al., 2006) or captioning (e.g., Cintrón-Valentín et al., 2019, 2021; Lee & Révész, 2018, 2020, 2021; Muñoz et al., 2021, 2023; Pattemore & Muñoz, 2020, 2022). Of particular interest to the present study are those that have explored L2 grammar learning through captions (see Section 2.6.1 and 2.7.6 for detailed reviews of Lee & Révész, 2018 and 2021 respectively). This line of research has primarily focused on comparing different types of captioning, such as full captions, key-word captions, textually enhanced captions, and no captions. Lee and Révész (2020), for instance, examined the relative effects of textually enhanced versus non-enhanced captions on how learners pay attention to and acquire the English present perfect and simple past. Seventy-two Korean learners of English randomly assigned to one of three groups, an enhanced captions group, an unenhanced captions group, and a no-captions group, completed a series of treatment tasks during which they watched 24 short news clips under their respective captioning conditions. The oral and written productive tests and a fill-in-the-blank test were used to assess participants' gains. Results showed that compared to the non-caption condition, both captioning groups achieved greater gains, but the enhanced captions were found to be the most beneficial for the learning of the present perfect tense. The findings indicate that captioned input may have an advantage over non-captioned for the acquisition of L2 grammar.

Cintrón-Valentín et al. (2019) compared the effects of non-caption and textually enhanced captions on the acquisition of Spanish vocabulary and grammar from four animated videos designed specifically for the target features. A number of grammatical constructions were targeted, including the Spanish *preterite* and imperfect forms, copula verbs and *gustar*-type verbs, and the subjunctive. Participants divided into three groups: no captions, captions

with enhanced vocabulary, and captions with enhanced grammar. Learner gains were assessed by recognition and production tests. Results showed that, while textually enhanced captions had a positive impact on vocabulary learning, they only yielded an advantage for some of the targeted grammatical forms (*gustar*-type verbs, subjunctive) on the productive test. The authors interpreted this finding as suggesting that the differential salience of each grammatical form might have influenced the effectiveness of enhanced captions. However, due to some methodological limitations (i.e., a lack of a pretest and the inclusion of pre-teaching and pre-practice of the target grammar), it is difficult to determine whether the effectiveness of captioning was the single contributing factor to any learning gains.

In their follow-up study (Cintrón-Valentín & García-Amaya, 2021) the authors included a pretest and a no pre-teaching condition. Results indicated that there were significant immediate effects of captioning on target vocabulary, with advantageous effects of captioning on some target grammatical constructions (i.e., *gustar*-type verbs, the *preterite*/imperfect, and the conditional) over others (i.e., the subjunctive). These findings suggest that captions can not only facilitate the acquisition of L2 vocabulary (Montero Perez et al., 2013b) but also have the capacity to promote L2 grammatical knowledge development.

Pattemore and Muñoz (2020) conducted an 8-week study with 69 Spanish/Catalan bilingual undergraduate students, in which the participants watched 10 consecutive episodes of a TV series, each lasting 227 minutes, twice a week. To assess the impact of captions on grammar learning, the participants were divided into two groups: a captions group and a non-captions group. Sixteen grammar target structures were selected, and the pretest and posttests were comprised of productive grammar exercises as well as fill-in-the-gap tasks. The study found that while both groups demonstrated improvement in their knowledge of grammar, the captions group outperformed the no-captions group on the posttest. The researchers construed

their findings as indicating the learning potential of multimodal input. Their results, however, should be interpreted with caution since participants received all the testing materials in written form, which might have favored the captions group. Pattemore and Muñoz suggested that future research could address this limitation by incorporating both written and aural tasks.

Pattemore and Muñoz's findings tie in well with those of Cintrón-Valentín et al. (2019, 2021) and Lee and Révész (2020), who also observed improved grammar learning following captioned audiovisual exposure. The overall pattern of evidence indicates that captioned audiovisual input supports the learning of constructions better than non-captioned audiovisual input, probably owing to the presence of complementary written text and the balanced processing demands on the verbal and visual input channels (Vanderplanck, 2016). Likewise, other studies on grammar with captioned input (Lee & Révész, 2018; Plotnikova, 2017) demonstrated a significant improvement in grammar production after non-enhanced captioned exposure. While the number of studies is limited and drawing conclusions may be premature, the existing evidence shows that captions (regardless of textual enhancement) may be more conducive to grammar learning than non-captions.

A follow-up study by Muñoz et al. (2021) which replicated the identical input conditions yielded comparable results. Consistent with Pattemore and Muñoz (2020), one group viewed the episodes with captions, while the other group watched them without captions. The results suggested that the group with captions outperformed the non-captions group on the grammar posttest, mirroring the earlier findings. Interestingly, the findings also revealed that the group without captions derived greater benefits from the repetition of target constructions compared to the captions group. This phenomenon might be attributed to the fact that learners tend to show a frequency effect more evidently when they lack other forms

of more supportive resources, such as captions. The researchers concluded that the presence or absence of captions can significantly affect the way in which learners acquire grammar structures.

Pattemore and Muñoz (2022) also found that unenhanced captions were more effective for grammar learning than both textually enhanced captions and no captions. Their study involved 112 Catalan/Spanish bilingual undergraduate students who were shown 10 episodes of a TV series and asked to complete an immediate and a delayed posttest. The unenhanced captions group outperformed the other groups on the delayed posttest, leading the researchers to conclude that additional highlighting of the target constructions may not have been necessary and may have detracted from learning.

In a more recent inquiry, Muñoz, Pujadas, and Pattemore (2023) examined the impact of frequency, the role of subtitles and captions, and the moderating effect of learner proficiency on language gains across two longitudinal investigations. Study 1 examined vocabulary acquisition in two groups of adolescents with elementary L2 proficiency levels. They watched 24 episodes of a TV series over an academic year, with one group watching with subtitles in their L1 and the other with captions in their L2 during weekly viewing sessions. Study 2 investigated grammatical structure acquisition in two groups of university students with intermediate proficiency levels, who watched 10 episodes over five weeks, with one group viewing captions and the other group without captions. Both experiments revealed a significant correlation between language gains and input frequency, although the magnitude of the frequency effect varied depending on the type of support provided by the on-screen text. The analyses also demonstrated that captions or subtitles did not provide significant benefits for vocabulary learning at this proficiency level. However, there is a significant advantage of captions over no captions for learning grammatical constructions. Additionally,

learner proficiency level was found to be a significant factor in language acquisition. Overall, the study underscores the value of utilizing audiovisual materials in L2 learning. Taken together, previous research on captioning provides positive evidence for the benefits of captioned audiovisual input for L2 grammatical development, as compared to a no captioning condition. However, despite the increasing amount of research exploring the effects of audiovisual input using captions, most studies in this strand have mainly focused on L2 development in vocabulary and listening comprehension, with few studies looking into the effects of captions on L2 grammatical development. Furthermore, to date, no research has examined whether caption-supported oral feedback (i.e., multimodal feedback) can promote the acquisition of L2 grammar in the context of multimodal video-conferencing. While multimodal recasts were operationalized in a way that is somewhat different from how conventional captions have been used, they may also have the capacity to help learners segment and parse oral input, drawing their attention to morphosyntactic features. This may in turn assist learners to identify new grammatical items in the input with greater ease and ultimately acquire the target constructions. Thus, exploration of the effectiveness of this novel type of recasts on L2 grammar learning would expand our understanding about this area of research and help determine whether video-based SCMC interactions with multimodal recasts have, as expected, a positive impact on L2 grammatical development.

2.5.4. Possibility of Redundancy Effects of Audiovisual Input

In evaluating the effectiveness of multimodal input, what should also be considered is the *redundancy principle*, which states that redundant materials inhibit learning, rather facilitate it (Kalyuga & Sweller, 2014; see Mayer, 2009 for the explanation of twelve principles grounded in the Cognitive Theory of Multimedia Leaning). This principle, grounded in the

Cognitive Theory of Multimedia Learning (Moreno & Mayer, 2002), suggests that redundancy occurs when learners are presented with identical information in two different forms/media or presented with additional information in the same form as elaborate information (Sweller, 2014). For instance, adding captions to narrated audio/video might require students to coordinate the redundant information, which may increase extraneous cognitive load on their WM and thus be detrimental for learning (Chandler & Sweller, 1991; Kalyuga & Sweller, 2014; Mayer, Lee, & Peebles, 2014; Sweller, 2005).

Some research has provided support for the principle that the concurrent presentation of both the aural and written forms of the words hinders learning (Kalyuga & Sweller, 2014; Mayer, 2014). However, it has also been claimed that when applying this principle to L2 learning care should be taken, as previous studies in favor of the redundancy principle used academic materials with a high level of complexity (mainly from science and mathematics) to teach subject-specific content in an L1 learning context (e.g., Mayer et al., 2014; Montero Perez, 2020b). In other words, redundancy may be beneficial in the case of L2 learning, since it appears as reinforcement rather than as a redundant element by reducing the cognitive processing load needed for identifying and decoding transient auditory input (Clark & Mayer, 2011; Lee & Mayer, 2018; Mayer et al., 2014; Meskill, 1996; Winke, Gass, & Sydorenko, 2010). This contradicts the detrimental impact of text suggested in the redundancy principle.

In this regard, the Redundancy Facilitation Hypothesis (Mayer et al., 2014) has been proposed, which suggests "a reverse redundancy effect in scenarios where redundant material can support basic cognitive processing that is not yet automated in non-native speakers while minimizing extra cognitive load" (p. 654). For example, the provision of redundant written information, such as adding on-screen text to oral input, can support L2 learners' cognitive processing of the oral input if it is attended to. Moreover, on-screen text may also be helpful

for raising L2 learners' noticing of unfamiliar input and compensate for their less-advanced listening ability. The Redundancy Facilitative Hypothesis appears to be supported by a number of L2 studies which have provided evidence that, contrary to the redundancy principle, duplicating the same information did not impede comprehension or learning. For instance, viewing television programs with captions and subtitles (i.e., given aural, visual, and written forms of the words all together) did not hinder comprehension or learning (e.g., Pellicer-Sánchez et al., 2020; Peters, 2019; Teng, 2019; Webb & Chang, 2020). Also, previously reviewed captioning research is in favor of the utility of captioned audiovisual input in L2 acquisition. Therefore, as Sweller (2005) points out, the redundancy of information should always be determined depending on different learning contexts, as there is no universally applicable standard for redundancy.

2.5.5. Summary

As predicted by theoretical models, a review of previous research in this section provides support for a number of advantages of multimodal input, specifically captioned audiovisual input over unimodal input. However, while the role of captioning in L2 comprehension and vocabulary has been the subject of much recent research, its effects on L2 grammar acquisition have rarely been investigated. In addition, despite the supporting evidence for the effectiveness of captions, no attempt has been made to employ oral recasts accompanied by captions and explore whether this novel type of multimodal feedback may promote or hinder the acquisition of L2 grammatical features, compared to unimodal recasts during video-conferencing. Thus, it appears imperative to address this gap in instructed SLA research, particularly considering that a variety of video-conferencing platforms suitable for providing multimodal input are becoming increasingly accessible and used by L2 learners in a wide

range of language learning and teaching contexts. Moreover, this may help to determine whether the redundancy principle or the Redundancy Facilitation Hypothesis holds true when L2 learners are exposed to and process multimodal linguistic input during video-based SCMC. Also, given that recasts can be presented through different input modalities during video-conferencing (i.e., oral, written, and multimodal), it would be useful to explore which modalities impose an optimum cognitive load on L2 learners and create the best conditions for them to notice and learn grammatical constructions.

In addition to investigating the effects of multimodal recasts on L2 development, it would be valuable to acquire insights into learners' cognitive mechanisms, such as their attention allocation to different modes of feedback and target forms. To uncover how learners process input in real-time, eye-tracking would appear a useful tool (Conklin, Pellicer-Sánchez, & Carroll, 2018; Godfroid, 2020). Thus, the subsequent sections will delve into eye-tracking techniques and pertinent empirical studies.

2.6. Eye-Tracking

Eye-tracking allows researchers to track, record, and investigate an individual's spatio-temporal sequences of gaze points in real time while they are interacting with a visual stimulus, usually on a computer screen (Conklin, Pellicer-Sánchez, & Carroll, 2018; Godfroid et al., 2013, 2020; Rayner, 1998, 2009). The core concept behind eye-tracking is based on the "eye-mind link" principle (Reichle, Pollatsek, & Rayner, 2006. p. 4), which proposes a strong correlation between the focal points of the eyes and the cognitive processes of the mind. Godfroid et al. (2013, p. 489) describe the assumption underlying the notion of eye-mind link as follows:

> […] overt attention (as manifested by the exact eye location) and covert attention (mental focus) are tightly linked, and cognitive processing is considered a major determinant of when and where the eyes move during complex task performance, such as reading. (Just & Carpenter, 1980; Rayner, 2009).

That is, a person's overt eye gaze may effectively reflect their visual attention, supplying reliable and comprehensive information about an individual's online cognitive operations and covert attentional process occurring while interacting with a visual stimulus (Just & Carpenter, 1976; Rayner, 2009). In this vein, eye-tracking has been considered a robust means of revealing a precise trajectory of L2 learners' attentional focus during online activities and has led to a plethora of eye-tracking research (Duchowski, 2007; Leow, 2015, Rayner, 2009; Robinson et al., 2011).

2.6.1. Eye-tracking Measures and Eye-tracking Research on L2 Grammar in SLA

Eye-tracking data are captured according to different *area(s) of interest* (AOI, also known as *interest area(s)*, IA, or *region(s) of interest*, ROI), which refers to specific areas that researchers are interested in and eye-tracking data can be extracted for (Godfroid, 2020). Eye-tracking usually records the three main components of eye-movement data, *saccades*, *regressions*, and *fixations*, in relation to a selected AOI. Saccades refer to rapid movements of the eyes from one location to another (Rayner, 2009). In text-based reading, saccades generally move from left to right (e.g., English, Korean), but they can also be from right to left (Arabic, Hebrew) depending on the language. Approximately 10–15% of saccades are referred to as regressions (Rayner, 2009), which are backward eye movements that allow learners to reinspect an earlier part of the visual input (Winke, Godfroid, & Gass, 2013).

Fixations occur when the eyes remain relatively stationary at a particular position to process visual information (Rayner, 2009).

In the examination of reading, fixation durations are often discussed in terms of early or late measures, which reflect different stages of the reading process (e.g., Clifton, Staub, & Rayner, 2007; Conklin et al., 2018; Rayner, 2009). *Early measures* (e.g., first-pass reading) capture the initial phases of text processing, which reflect automatic word recognition, lexical access, and the spontaneous incorporation of incoming information (Clifton et al., 2007; Conklin et al., 2018; Roberts & Siyanova-Chanturia, 2013). *Late measures* (e.g., re-reading time), on the other hand, reflect subsequent stages of processing, during which learners engage in more conscious and strategic processes such as reanalysis, discourse integration, revisits, and recovery from processing difficulties. These measures may also indicate an interrupted reading process (Godfroid, 2020; Conklin et al., 2018). Both early and late measures are generally presented together in order to obtain a comprehensive understanding of the process being examined. The most commonly reported eye-tracking measures in SLA are summarized in Figure 4 (Roberts & Siyanova-Chanturia, 2013; see Conklin and Pellicer-Sánchez, 2016; Conklin et al., 2018; Godfroid, 2020 for a comprehensive review of eye-tracking measures).

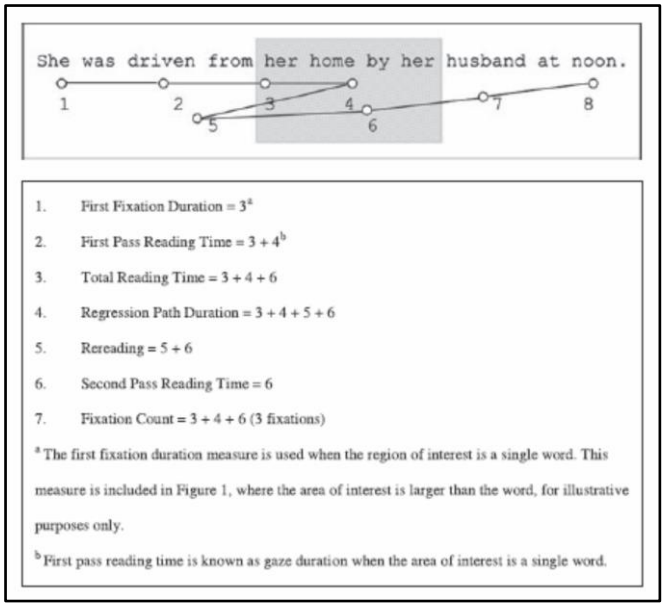


Figure 4. Eye-movement Measurements (Roberts & Siyanova-Chanturia, 2013, p. 218)

Of most relevance to the current study is video-based eye-tracking research using captions, which often uses eye-movement data as representational measures, which refer to "when and where people fixate as the utterance unfolds" (Tanenhaus, 2007, p. 318). Although the measures that have been reported in previous video-based studies are presented in Table 5, there is no standard in terms of which eye-tracking measures should be used in video-based studies. Rather, it has been emphasized that the selection of relevant measures should always be appropriate for the purpose of each study (Conklin et al., 2018; Godfroid, 2020). For example, if a researcher is only interested in the total length of time a learner spent on an AOI, the use of total reading time may be sufficient. However, if the researcher wants to further explore how many times the AOI was visited, run count would also be relevant.

Apart from AOIs, what is also vital in video-based research is to identify specific time points in the video to be investigated (Conklin et al., 2016; Godfroid, 2020; Montero & Perez, 2019). Specific time points are called interest periods (IPs) or time of interest (TOI). After identifying both AOIs and TOIs, learners' eye movements that are only within AOIs

during active TOIs are included in the analysis. In the present study, for example, only the eye-movement data that occurred within the AOIs (written recasts) during specifically defined TOIs were included in the analysis. In this case, TOIs were defined as the presentation times of the written recasts in the screen-based recordings (see Section 4.2.10.2 for detailed information on the AIOs and TOIs used in the current study).

Table 5. Definitions of Eye-tracking Measures Commonly Used in Video-based Studies (adapted from Conklin & Pellicer-Sánchez, 2016; Conklin, Pellicer-Sánchez, & Carroll, 2018; Godfroid, 2020)

Eye-tracking measure	Definition
Total reading time %	Proportion of all summed time spent on an AOI within the defined TOI
Average/Mean fixation duration	Average duration time of fixations on an AOI
Fixation count	Total number of fixations on an AOI within the defined TOI
Skip rate	Likelihood that an AOI is skipped (not fixated on at all) if the AOI is not fixated during first-pass reading
Total fixation duration	Duration of all summed fixations on an AOI within the defined TOI
Visit count	Total number of visits to an AOI within the defined TOI
First-pass reading time	Sum of all fixations on an AOI before exiting the AOI within the defined TOI
Re-reading time	Summed duration of all fixations on an AOI excluding those during the first pass within the defined TOI

To date, a limited number of studies have investigated the extent to which learners attend to target grammatical constructions and, further, the extent to which the amount of attention allocated to these structures relates to L2 acquisition (e.g., Alsadoon & Heift, 2015;

Godforid et al, 2013; Godfroid & Uggen, 2013; Indrarathne & Kormos, 2016; Issa, Morgan-Short, Villegas, & Raney, 2015; Loewen and Inceoglu, 2016; Montero Perez et al., 2015; Simard & Foucambert, 2013; Winke, 2013).

However, all of these studies were conducted in a reading context. More closely relevant to the present study is the research on captioned audiovisual input or captioned video using eye-tracking to measure learners' attentional allocation to grammatical items. So far, only three studies (Lee & Révész, 2018, 2020, 2021; see Section 2.7.6 for the review of Lee & Révész, 2021) have addressed this subject. Lee and Révész (2018) compared the effects of textually enhanced and non-enhanced captions on learner attention to the English pronominal anaphoric reference and the relationship between the amount of attention paid to and knowledge development in the target items. A pretest–posttest experimental design was employed, with three treatment sessions. Forty-eight intermediate Korean learners of L2 English were divided into a captions group and an enhanced captions group, and exposed to twenty-seven static pictures with audio-recordings. Learner attention to the target antecedents and pronouns were gauged with four eye-tracking indices: first-pass reading time, second-pass reading time, total fixation duration, and number of visits. Learning gains were assessed by oral and written grammaticality judgement tests. Results revealed that participants in the enhanced captions group paid greater attention to the anaphora antecedents and generated gains in receptive knowledge of pronominal anaphora. These findings suggest a positive correlation between the amount of attention paid to and L2 development of the target grammatical features.

As reviewed in the previous section, Lee and Révész (2020) examined the relationship between L2 learners' attentional processes during exposure to captioned (enhanced versus non-enhanced) video and L2 grammatical development in the use of the

English present perfect versus the past simple. Participants watched twenty-four short video clips (20 to 50 seconds each) in one session while their eye movements were recorded. Learner attention to the target constructions was analyzed with four eye-tracking measures: first-pass reading time, second-pass reading time, number of visits, and skip rate. The eye-tracking data revealed that enhanced captions drew learners' attention to the target constructions significantly more frequently than unenhanced captions. The authors suggested that enhanced captions increased visual salience of the target structures, and learners who attended to the target constructions longer and more frequently were likely to obtain higher gains in the production tasks.

Overall, Lee and Révész (2018, 2020) have shown that the amount of time spent on the target grammatical structures while viewing captioned listening activity or captioned video is positively related to subsequent L2 development. However, these are only two empirical studies that have attempted to investigate the relationship between learner attention and development in L2 grammatical knowledge. To the best of my knowledge, no research exists that has used eye-movement data to investigate the effects of captioned oral recasts on L2 grammar learning during video-conferencing. Considering the positive findings that captioned listening and video activities can be beneficial to achieving L2 grammatical development, it seems worthwhile to examine whether captioned oral feedback might influence development in the use of L2 grammatical knowledge. Also, the studies reviewed have shown that the increased salience of the enhanced captions was more successful in drawing learners' attention. In SCMC and CALL contexts, in order to enhance their corrective nature, some researchers have suggested that recasts be provided in the written mode (e.g., Sagarra & Abbuhl, 2013), and/or enhanced orally (e.g., through the use of prosodic highlighting) or typographically (e.g., using bolding or coloring) (e.g., Gurzynski-

Weiss et al., 2016; Sachs & Suh, 2007; Saggar & Abbuhl, 2013). However, no attempt has been made to combine different SCMC modalities as a means of enhancing the salience of recasts available in multimodal video-conferencing. Thus, it seems useful to explore whether multimodal recasts will be more successful in enhancing the salience as well as the corrective force of recasts and draw greater amount of attention to the target forms than unimodal recasts.

2.6.2. SCMC Eye-tracking Research on L2 Grammar

In recognition of its benefits, a growing body of L2 research has employed eye-tracking technology in SCMC contexts. However, a relatively small number of empirical studies have used eye-tracking to identify what learners attend to and how it relates to L2 development during synchronous online interaction (e.g., Michel & O'Rourke, 2019; Michel & Smith, 2014, 2018, 2019; Nezami, 2012; Örnberg Berglund, 2013; O'Rourke, 2008, 2012, 2013; Prendergast, 2011; Shintani & Ellis, 2013; Smith, 2010, 2012; Smith & Renaud; 2013; Stickler & Shi, 2015; see Michel & Smith, 2017 for an overview). Of most relevance to the present study is the empirical work that has used eye-tracking to examine the extent to which learners pay attention to CF, including recasts and, further, the extent to which the amount of attention allocated to recasts is related to the noticing of the feedback and L2 acquisition. As with other SCMC research, most of these studies have been conducted in text-based SCMC, and some studies have used eye-tracking in combination with other data collection sources, mainly stimulated recall interviews and/or keystroke logging.

Smith's (2010) work was the first empirical study using eye-tracking to investigate whether text-based recasts were attended to and noticed during task-based SCMC. Eight EFL learners participated in the text chat and were asked to retell a story of a video clip they

watched to the native-speaker researcher in English. While performing the task, participants received intensive recasts (i.e., focusing on lexical and grammatical items only) on their errors and their eye movements were recorded using an eye-tracker. Noticing was operationalized as eye fixation on the target item of 500 milliseconds (ms) or longer, while less than 500 ms or no evidence was classified as "no noticing". After the text chat, all participants completed a timed writing task individually to retell the same story. Results showed that learners noticed about 60 % of the recasts they received. Lexical recasts were found to be much easier for learners to notice, retain, and accurately produce in subsequent chat interaction and delayed written posttest than were grammatical recasts. Smith also found that learners fixated more on lexical recasts than grammatical recasts, suggesting the positive relationship between the amount of eye gaze and L2 development. The results are aligned with L2 studies in face-to-face contexts, which have shown that morphosyntactic recasts are less salient and less noticed than recasts targeting lexis or phonology (e.g., Carpenter et al., 2006; Lyster, 1998; Lyster et al., 2013; Mackey et al., 2000; Mackey, Gass, & McDonough, 2000; Lyster et al., 2013; Trofimovich et al., 2007).

Building on this study, Smith (2012) further explored the potential of employing eye-tracking to measure learners' noticing of corrective recasts and learning gains for the items targeted by the recasts in text-based SCMC. The study utilized a more rigorous research design, which included an increased number of participants (N = 18), a pretest, posttest, and delayed posttest design, and a stimulated recall protocol. Learners engaged in a short text chat with a native speaker who provided intensive explicit corrective recasts targeting morphological, semantic, and syntactic items. Participants were explicitly trained on the nature and intent of recasts, and informed that their interlocutor would ask questions and provide comments and CF, including recasts. Learners' eye movements on the feedback

given were tracked to examine the noticing of feedback and the targeted features. Borrowing the term *noticing event* from Godfroid et al. (2010), Smith operationalized a noticing event as increased visual attention. This increased visual attention was marked by the overall dwell time, which consists of the sum of all fixations on a word, including first run dwell time and all regressions. He used heat maps, in which the relative number of eye gaze fixations is reflected by the main colors on the color spectrum beginning with purple (the fewest number of eye fixations) to blue, green, yellow, orange, and red (the highest number of eye fixations). Where there was a heat map, it was coded as a noticing event, whereas no heatmap on a target item was coded as the absence of a noticing event. In the stimulated recall, in cases where learners indicated that a recast was targeting an earlier non-target-like item and were able to specifically identify verbally the noted difference, that item was coded as a noticing event occur. Smith compared participants' eye-gaze data to those of a stimulated recall to determine whether noticing events (i.e., increased visual attention) compiled from these two measures resulted in similar findings. Noticing events from each measure were also compared to results of immediate and delayed posttests of the target features. Overall, the findings from the study confirmed that both the eye-tracking and stimulated recall data predicted learners' noticing of recasts during text-based SCMC. The results were similar to those found in Smith's (2010) study. Learners engaged in similar amounts of viewing activity across recasts targeting three linguistic categories (i.e., morphological, semantic, and syntactic). However, in the stimulated recall interviews, learners reported that they noticed recasts targeting syntactic and semantic items better than morphological items. In other words, even though learners spent comparable amounts of time attending to recasts for each category, they were more aware of and focused on recasts that addressed syntactic and semantic aspects of the language than morphological aspects. Notably, the study found that

even when learners were aware that they would receive recasts prior to the treatment sessions, they noticed grammatical recasts least. This underscores the necessity for more effective feedback manipulations to draw learners' attention to grammatical recasts during SCMC interactions. However, when interpreting the results of Smith's work, caution should be taken, as learners noticing target items were reflected in a heat map without specific attention measures.

Adapting Smith's (2010) study design, Berglund (2012) examined the patterns indicative of noticing in relation to recasts and metalinguistic feedback during text-based interaction. Eight EFL students at a Swedish upper secondary school participated in a chat task with a teacher (the researcher). They were shown a short animated video clip without verbal content and were asked to retell the story during the text chat with the researcher, who provided recasts and metalinguistic feedback on language and content. The learners were then required to retell the same story once more in an MS Word document. Data from chat logs were complemented with screen recordings, keystroke logging, and eye-tracking data, enabling analysis of both verbalized and non-verbalized signs of noticing. Results revealed that while all of the teacher's feedback was seemingly noticed, at times the learners noticed the feedback long after it was posted, up to several minutes later. In other words, learners tended to focus on their own writing and read the teacher's feedback only once they had finished with their own composition and sent it. These findings are consistent with a previous study which also observed learners' patterns of "post-send monitoring" during text-based SCMC, rather than engaging in two other patterns, "simultaneous monitoring" or "pre-send monitoring" (O'Rourke's, 2012, p. 319). This suggests that learners may not consistently attend to the feedback immediately or in the way that the teachers intend when feedback is presented in text. Thus, in order to draw learners' attention to the written feedback

immediately after it has been given, it might be more effective to provide oral feedback, so the learners can be alerted to and notice the provision of feedback on their errors, thereby promoting their noticing and a cognitive comparison between them.

Smith and Renaud (2013) investigated the relationship between what L2 learners attend to and notice in recasts and their posttest scores on the target items during text-based SCMC tasks. Sixteen university students (L1 English) from low-intermediate level Spanish and low-intermediate level German classes participated in this study. As the treatment task, learners engaged in a 15-minute one-on-one writing conference with their interlocutor about a written assignment draft due the following week. To measure what learners noticed, the researchers used the total number of fixations and the total fixation duration, with fixations operationalized as those lasting over 200 ms. The researchers found a positive relationship between noticing of lexical and grammatical items and posttest gains. Specifically, learners attended to 75% of teacher recasts, with between 20% and 33% of these scored as correct on the posttest administered one week after. The strongest predictor of the posttest score was whether the learner fixated on a target item for at least 200 ms. A suggestive but not significant effect was established for the number of fixations and posttest success, while total fixation duration and posttest success were not related.

In the context of the current study, it is worth acknowledging the significant contributions made by Smith (2010, 2012) and Smith and Renaud (2013) in exploring the potential association between learners' visual attention to target features and L2 learning in SCMC. However, there are still questions in this domain. One area of uncertainty pertains to the specific nature of the relationship between attention measures and L2 learning. The aforementioned studies have relied on a limited set of attention metrics, primarily focusing on the total number of fixations and fixation duration. This narrow scope of measurements may

not offer a comprehensive understanding of learners' attention patterns. To draw more robust conclusions regarding the connection between visual attention and L2 learning outcomes in SCMC environments, further research is needed to address these methodological challenges. Also, investigating the impact of different target features, such as lexical versus grammatical features on capturing learners' attention, requires further investigation. Moreover, it is crucial to broaden the research focus beyond text-based recasts and consider other modalities, particularly video-based recasts within SCMC.

While their direct focus was not on CF, combining eye-tracking and stimulated recall interviews, Shi, Stickler and their colleagues have carried out multiple eye-tracking experiments within a sociocultural paradigm (Stickler & Shi, 2015, 2017, 2021; Shi, Stickler, & Lloyd, 2017). In these studies, they have explored the attention focus of language teachers and learners during synchronous multimodal online tutorials and their reflections on their own teaching practices and learning experiences, with an emphasis on sociocultural factors. Stickler and Shi (2015), for example, examined L2 learners' attention focus during synchronous online Chinese tutorials in a multimodal environment. Ten learners of Chinese as a foreign language at beginning or low-intermediate level took part in reading and interactive online activities via Elluminate, the online conferencing system used for tutorials at the Open University, UK. The reading task was presented as a short text in Chinese characters with Pinyin (i.e., the Chinese phonetic alphabet) transcription below, followed by three comprehension questions in English. The second activity was synchronous online speaking practice around the theme of "transportation" involving a tutor (one of the authors) and three or four other students, in addition to the participants being recorded in the eye-tracking laboratory. While performing the reading and speaking activities, participants' gaze focus was recorded in the different AOIs on the screen; the whiteboard containing

information for the task, vocabulary help, and images; the participants' window showing the names of all participants in the tutorial and the emoticons they displayed; and the technical areas in which learners use tools (e.g., the microphones or text chat). The eye movements were analyzed using two measures: fixation duration and fixation count. In subsequent stimulated recall sessions, participants were asked to reflect on their engagement with the screen and their intention while reading or speaking online. Results showed that during the reading tasks, all participants predominantly focused on the whiteboard where Pinyin, Chinese characters, and comprehension questions were displayed. There was little attention focus on the social areas, with very occasional glances at the technical areas. The stimulated recall revealed that participants' reasons for paying attention to Pinyin was for (a) convenience, (b) confirmation, and (c) comprehension. During interactive tasks, 70% of learners' eye gaze was drawn to content (e.g., the Pinyin section of the whiteboard), 20% on social areas of the interface (e.g., the names of other participants) and 10% on technical information (e.g., the microphone button) areas. Participants also reported their need to pay attention to peers' social presence. Aligning with the previously reviewed studies, both eye-tracking and stimulated recall data revealed that learners were able to notice semantic and syntactic targets more easily than morphological targets. Combining both eye-tracking and stimulated recall interviews for participants' views, Stickler and Shi found the important role of social presence in multimodal SCMC tutorials and the usefulness of Pinyin in both reading comprehension and speaking production.

Using the same research method, Shi, Stickler, and Lloyd (2017) uncovered the intricate dynamics between online teachers' attention focus, teaching skills, and online teaching experience. Three online tutorials were held using Blackboard Collaborate software (a German tutorial for 17 minutes, Chinese tutorial for 17 minutes, and French tutorial for 25

minutes). It was found that teachers' attention to the content, social, and technical areas was influenced by their familiarity with the online teaching tool as well as their prior online teaching experience. A highly experienced online teacher who was familiar with the tools could focus more on the teaching content and social areas without paying much attention to the technical area. If a teacher was unfamiliar with the specific online teaching tool or had little experience of online teaching, the attention focus on the technical area was significantly higher. Compared to the student study, this study found teachers' attention on social areas was slightly higher than the proportion of students' attention on the same areas. The reason was that teachers were clearly aware of the limitations of online communication and were trying to compensate for the loss of nonverbal cues by paying attention to students' social presence indicators (e.g., emoticons in the participant window). The stimulated reflection data revealed that more experienced tutors spent more time on attending to social and content areas and reflected more frequently about their online teaching skills at a higher level, based on a comparison with Hampel and Stickler's (2005) pyramid of online teaching skills.

However, the two studies mentioned above investigated the perspectives of learners and teachers separately, making comparisons only in retrospect. To delve into synchronous, joint attention between online teachers and students, a dual-point eye-tracking approach was adopted. In a recent exploration, Shi and Stickler (2021), utilizing dual-point eye-tracking study, sought to understand how students and teachers establish joint attention during online language tutorials. It also examined the various means teachers employed to direct students' attention and scaffold their meaning-making process. Data were collected from two online language tutorials where the eye movements of one teacher and one student were tracked simultaneously. Recordings of their stimulated reflection while watching the replay of their own eye-tracking visualization were also obtained. Two language teachers and two language

students participated in on-site eye-tracking at the university lab, while other students took part online. To maintain ethical restrictions, all participants were also language teachers learning an additional language. The screens used for the online tutorials and eye-tracking were identical in dimension and interface to ensure accuracy. Both the teacher and the learner underwent a nine-point calibration to ensure precise eye-tracking data, and the online tutorial and eye-tracking commenced simultaneously. There were two tutorials, one for beginner German students and the other for intermediate ESL students. The online tutorials were conducted via OULive, a synchronous online teaching tool. In the first tutorial, a German teacher taught four students for 15 minutes while one student's eye movements were tracked by an eye-tracker. The tutorial was conducted in German, and the session involved the teacher presenting key grammar points to the students and then engaging them in an interactive speaking task using the newly learned grammar structure. The interface was divided into two parts, with the microphone button, participant window, text chat, and editing tools on the left and the whiteboard on the right. During the interactive speaking task, the whiteboard was divided into top and bottom sections, with the top section displaying the teacher's sample structures and the bottom section displaying the student's own sentences. The teacher used "crosses" and "smiles" to express different degrees of likes and dislikes in German. After the tutorial, both the teacher and the student watched their eye-movement gaze-plot visualizations together on side-by-side screens while researchers recorded their dialogue. In the second tutorial, an English teacher instructed four adult ESL learners on the word order of adjectives, with the eye movements of both the teacher and one student tracked. The same research procedure as the German tutorial was followed.

The results showed that when the eye-tracking recordings of the teacher and the student were replayed side by side, it became apparent that their attention met during the task.

The teacher's attention focus shifted from technical aspects to the participant window and text-chat box during the interactive speaking phase, while the student's attention was mainly on the content, with a focus on the teacher's sample structures in the top half of the screen and the structures listed at the bottom of the whiteboard when his peers were speaking. This shift in attention and focus may indicate "learning from peers" and that "knowledge is co-constructed" in a live online session. The student also paid attention to the names in the participants window and the microphone icon, which are linked to social presence issues.

2.6.3. Summary

Findings from SCMC research using eye-tracking has shown that eye-tracking is a robust tool to scrutinize L2 learners' attention allocation to recasts and target items during synchronous online interaction. However, eye-tracking SCMC studies on recasts are limited to text-based SCMC. Although Shi and Stickler and their colleagues (2015, 2017, 2021) explored language teachers and L2 learners' attention allocation in synchronous multimodal language-learning environments, they only dealt with the participants' attention directed to various screen areas (i.e., linguistic, social, and technical) without an investigation into the effectiveness of CF and L2 grammatical development. So far, there has been no research that has used eye-tracking to investigate the effects of multimodal CF on learners' attentional allocation to the feedback and L2 grammar learning during video-conferencing. Thus, given the observed positive link between attention and L2 development in text-based SCMC, further research is warranted to explore whether the findings obtained from synchronous text-based interaction would transfer to synchronous multimodal video-based interaction. In addition, as in previous face-to-face L2 research, it was found that learners noticed better and attended more to recasts targeting syntactic and semantic items than those targeting morphological items,

achieving greater gains in syntactic and sematic items. Thus, to achieve a full picture of learner attention and L2 development in multimodal video-conferencing, it appears important to investigate which feedback modality (i.e., oral, written, or multimodal) would be more successful in enhancing the salience of the target grammatical features and thus attract more attention than unimodal recasts.

In examining the extent to which feedback modality affects the likelihood that learners will attend to and learn L2 grammatical constructions, a key individual difference worth considering is learners' WM capacity. This is particularly the case as increasing evidence suggests that recasts do not benefit all learners equally but may be more beneficial for some learners than for others. Hence, examining whether individual differences in WM capacity might influence the efficacy of recasts, feedback modality, and L2 grammatical development seems to be a worthwhile research endeavor.

2.7. Working Memory

WM is generally defined as the ability to simultaneously store and process incoming information during complex tasks for a short period of time (Alloway, 2009; Baddeley & Logie, 1999; Swanson & Alloway, 2012). Over the past two decades, the role of WM in L2 learning has been extensively researched in the field of SLA, and individual differences in WM capacity have been demonstrated to predict and account for the ability to perform and learn a wide range of cognitively complex tasks, including the development of various types of L2 knowledge and skills. Indeed, evidence from narrative reviews (Jackson, 2020; Jeon & Yamashita, 2014; Juffs & Harrington, 2011; Wen, 2016; Williams, 2012) and meta-analyses (Linck et al., 2014; Shin, 2020) has shown that, overall, WM has robust, favorable connections with L2 processing as well as learning outcomes. Given its prominence in L2

acquisition research, this section will overview the cognitive architecture of WM and empirical studies that have examined its relation to L2 acquisition in face-to-face and SCMC learning contexts.

2.7.1. Baddeley's Working Memory Model

As an important cognitive construct, WM has been studied extensively in the field of psychology (Baddeley, 2003a, 2003b, 2007; Conway, Jarrold, Kane, Miyake, & Towse, 2007a; Cowan, 2005; Jarrold & Towse, 2006; Miyake, 2001; Miyake & Shah, 1999), and various theories and models for WM capacity have been suggested (see Miyake & Shah, 1999 and Sagarra, 2013 for a review). Of the various models, Baddeley and colleagues' (Baddeley, 1986, 2000, 2007, 2010, 2017; Baddeley & Hitch, 1974; Baddely & Logie, 1999) multicomponent model has been the most influential in previous studies in SLA (Miyake, 2001; Wen, Mota, & McNeill, 2015). According to this model, the WM system consists of four subsidiary components: (1) the central executive, (2) the phonological loop or PSTM, (3) the visuospatial sketchpad or VSTM, and (4) the episodic buffer (Baddeley, 2000) (Figure 5).

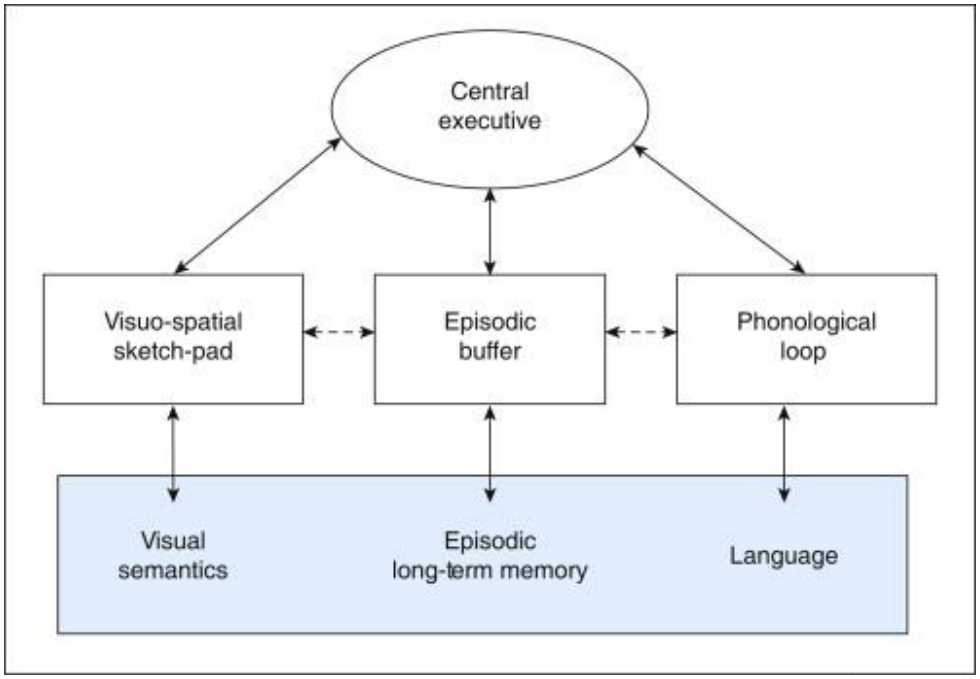


Figure 5. Multicomponent Model of WM (Baddeley, 2010, p. 138)

The central executive is responsible for controlling and regulating cognitive processes (Miyake et al., 2000), such as allocating attentional resources, task switching, inhibiting processing routines, and updating. Among these functions, three frequently referenced executive functions are task-switching, inhibition, and updating (e.g., Baddeley, 1996; Logan, 1985; Miyake et al., 2000; Smith & Jonides, 1999). Task-switching refers to the ability to shift attention from one task or activity to another, which involves disengaging from one task and engaging in another. Cognitive flexibility is required to switch between tasks, that is the ability to maintain task-specific information in memory while quickly adjusting to changing task demands. Inhibition is the ability to suppress or inhibit automatic or prepotent responses, selectively attending to pertinent information. This involves inhibiting irrelevant or distracting information that could interfere with the task at hand, as well as inhibiting impulse or habitual responses that are not consistent with the current task. Thus, inhibition requires the cognitive control to sustain attention and focus on the relevant information. Updating

concerns the ability to monitor, revise, and update incoming information (Baddeley, 1996; Miyake et al., 2000) It integrates new information into existing knowledge and modifies mental representations accordingly. This includes the process of adding, deleting, or modifying information in WM to keep it up-to-date and relevant to the current task. This is necessary for successful goal-oriented behavior, as learners need to flexibly adapt to changing demands, shift attention between relevant features of the task, and maintain multiple tasks in WM. While these three executive functions appear to have a common underlying construct, empirical evidence suggests that they are associated with different subprocesses and thus are separable (Miyaki et al., 2000).

Overall, the central executive WM is robustly implicated in intricate cognitive processes (Baddeley, 2003); it has been established through research that executive functions play a role in numerous L2 skills, such as reading comprehension (e.g., Peng et al., 2018), accuracy and complexity in language production (e.g., Zalbidea, 2017), and, of particular relevance to the present study, L2 grammar acquisition (e.g., Lado, 2017; Li et al., 2019; Serafini & Sanz, 2016).

The phonological loop or PSTM, is a sub-system for temporarily storing, manipulating, and retaining verbal and acoustic information over a short period of time (Baddeley, 1986, 2003; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). Specifically, PSTM subsumes two sub-systems: a temporary storage system which holds phonological information for a few seconds until it decays, and an articulatory rehearsal system which maintains and registers the information stored in short-term memory. Given its instrumental and supportive role in storing, chunking, combining, and recalling newly acquired phonological units, PSTM is considered to underlie the acquisitional and developmental aspects of language domains, such as lexis or vocabulary acquisition (e.g.,

Atkins & Baddeley, 1998), phrases or formulaic chunks (e.g., Martin & Ellis, 2012), as well as grammatical structures or morphosyntactic constructions in both L1 and L2 (see Wen, 2015, 2016 for reviews).

The visuospatial sketchpad is the storage system which temporarily retains and processes visual images and spatial relations. It contains two components: the visual component processes the characteristics of objects, such as size, color, and shape, whereas the spatial component addresses relational or spatial information and the control of movements. Although the visuospatial sketchpad does not comprise an independent sub-system for rehearsal, as the PSTM does, rehearsal in the sketchpad "appears to depend on a process sometimes known as 'refreshing', involving sustained attention to a selected item" (Baddeley, 2017, p. 303). Baddeley (2015) suggests that visuospatial WM contributes to the acquisition of new scripts or orthographic features. Despite the fact that this WM component may have a crucial role in L2 tasks that rely on visual memory (Kellogg, 1996, 2007), only a few studies have examined the relationship between VSTM and L2 development (e.g., Lee & Révész, 2021; Sachs, 2010; Zalbidea & Sanz, 2020).

Finally, the episodic buffer (Baddeley, 2000), which was later added as a memory sub-system under executive control, is considered a limited capacity storage system. Its role is to temporarily store and integrate information from the phonological loop, the visuospatial sketchpad, and long-term memory into episodes or multimodal units (Baddeley, 2003). That is, it serves as an interface between the sub-components and long-term memory. The idea of separate processing systems for verbal and visual information is consistent with Paivio's (1971, 1986) Dual Coding Theory which, as discussed earlier, hypothesizes how verbal and visual information is processed, integrated, stored, and recalled in the human mind.

2.7.2. Wickens' Multiple Resource Model

While WM is generally assumed to have a limited capacity, it has also been suggested that time-sharing across various activities will be differentially demanding depending on the type of information processing required. Wickens' (2007) Multiple Resource Model proposes that different types of cognitive resources can be used simultaneously. In other words, when different tasks require the same cognitive resource (e.g., auditory perception), information must be processed in sequence, whereas if the task requires different resources (e.g., auditory perception and visual perception), they can be processed simultaneously (see Figure 6 for the representations of these cognitive resources on the model).

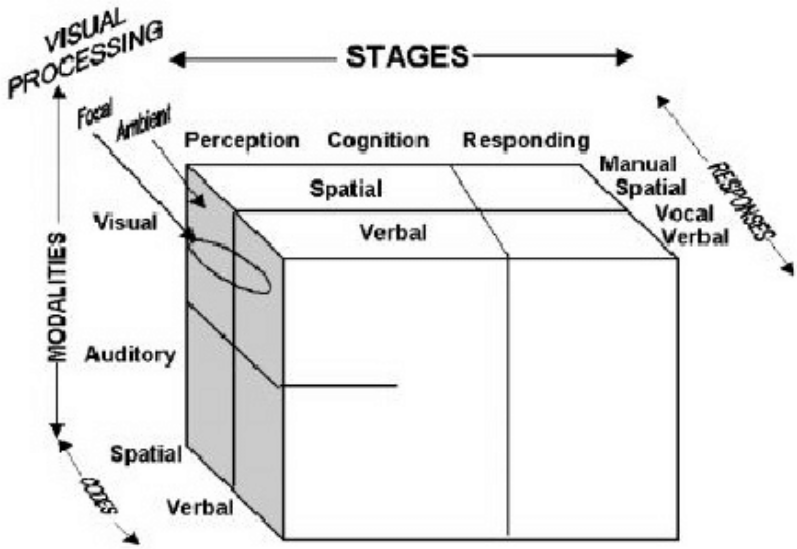


Figure 6. The Multiple Resource Model (Wickens, 2002)

There are four dimensions included in the model: *perceptual modalities* (visual versus auditory), *visual channel* (focal versus ambient), *processing codes* (verbal versus spatial), and *input processing stage* (cognitive/perceptual versus responding activities). Wickens' model explains that perceptual modalities, which are visual/manual, auditory/vocal, tactile, and olfactory modes used in time-sharing tasks, use different resources. In particular, there is a

dichotomy between tasks that utilize separate modalities or tasks that require one modality, referred to as cross-modal time-sharing (e.g., auditory and visual) or intra-modalities (e.g., auditory and auditory) respectively. The model predicts that there will be less interference when using cross-modalities rather than intra-modalities as a result of separate perceptual resources being used at the same time. For example, dividing attention between the auditory and visual channels (e.g., listening while reading) is better than dividing it between two auditory channels (e.g., listening to two messages simultaneously). Thus, in the case of multimodal input, interference is likely to be greater between on-screen text and pictorial information than between on-screen text and narrations, since both on-screen text and pictorial input are presented in the visual modality. This interference can hinder learning. However, if the on-screen text is presented with auditory input, they are processed separately in the visual and auditory channel respectively, making it easier for a learner to attend to and process both types of input simultaneously. This notion is consistent with dual channel and limited capacity assumptions in Mayer's Cognitive Theory of Multimedia Learning.

The visual channel, which was added in Wickens' (2002) updated theory, suggests two visual channels are used in visual processing, focal and ambient (Leibowitz& Post, 1982; Previc, 1998), a nested dimension within visual resources. Focal vision is linked to eye movement and supports object recognition, in particular, high acuity perception such as reading text and recognizing symbols. Ambient vision, on the other hand, involves the use of peripheral vision and is responsible for perception of orientation and movement for tasks such as moving toward targeted directions or keeping to a lane on the highway (Horrey, Wickens, & Consalus, 2006). The model infers that dual tasks involving one focal and one ambient processing will result in little interference. Examples of dual tasks which use both

channels include walking down a corridor (ambient) while reading a book (focal) or keeping a car in the center of a lane (ambient) while reading a road sign (focal).

The processing codes distinguish between analogue/spatial processes and categorical/symbolic (verbal or linguistic) processes. Wickens (1980) postulates that spatial codes, which involve understanding and processing spatial relationships, use different mental resources compared to verbal/linguistic codes. This distinction is evident across various cognitive activities, including perception, WM (Baddeley, 1986), and actions such as speech compared to manual control (Liu & Wickens, 1992; Wickens & Liu, 1988). Thus, manual and vocal responses can be time-shared very efficiently. Manual responses are assumed to be spatial in nature (tracking, steering, or mouse movement) and vocal responses are verbal (speaking). One useful implication of these distinction between processing codes is the ability to predict when it might or might not be beneficial to employ voice versus manual control (Liu & Wickens,1992; Wickens & Liu, 1988).

The processing stages indicate that perceptual and cognitive tasks use different resources from those underlying the selection and execution of action (Isreal, Chesney, Wickens, & Donchin, 1980), and that dual tasks can be affected depending on whether multiple tasks require cognitive/perceptual activities (e.g., listening) or response activities (e.g., writing). These activities represent stages of the model that are perceptual and response stages, and tasks requiring the same stages are more likely to have a greater interference during dual task performance than in those dual tasks where one task requires perceptual activity and the other a response.

The potentially most important application of this model for L2 learning appears to predict the performance of multiple time-sharing tasks. As was previously explained, task performance is negatively affected if multiple tasks share the same modalities. Wickens'

model defines attention as a process of allocating resources, whereby tasks using a single modality must be processed in a sequential manner, while the utilization of multiple sensory pathways enables learners to process information in parallel. Thus, tasks are more efficiently time-shared when multiple resources are used, resulting in less interference and reduced cognitive load. In his review of several studies, Wickens (1980) found greater positive impact of cross-modal (e.g., visual and auditory) over intra-modal presentations (e.g., visual and visual). This implies that the combination of visual and auditory information likely facilitates deeper understanding (Tabbers, et al., 2004; Cuevas et al., 2002) and improves retention of the input (Mayer, 2002). Therefore, it seems reasonable to assume that multimodal recasts have the potential to facilitate task performance, given that oral and written recasts are presented via auditory and visual channels respectively.

In this vein, this model seems to share a common assumption with Paivio's Dual Coding Theory and Mayer's Cognitive Theory of Multimedia Learning; they all suggest that different types of information require different cognitive resources and these resources are limited. Thus, presenting information in multiple modalities may help to reduce learners' cognitive load and promote the integration and organization of new information for more generative processing, ultimately leading to better learning outcomes and retention.

2.7.3. Working Memory Measures

In the area of instructed SLA, of particular interest to most previous WM studies have been the role of the central executive and/or PSTM in L2 development. Executive control has been usually measured using a single WM index (see, however, Indrarathne & Kormos, 2018; Lee & Révész, 2021) on the premise that the central executive functions are a unitary cognitive process (e.g., Duncan et al., 1996; Engle, Kane, & Tuholski, 1999). For example, a number of

studies have attempted to examine the extent to which individual differences in WM influence L2 acquisition using one type of task such as reading span task (Daneman & Carpenter, 1980), an operation span task (Turner & Engle, 1989), or a listening span task (Martin & Ellis, 2012). These complex WM tasks have been presumed to provide information about the overall functions of the central executive components of WM, in that these tasks require both storing and processing information (Gathercole, 1999). Against this view, however, a question has been raised as to whether the three subprocesses (switching, inhibition, and updating) can be measured using a single WM measure and whether they are separable, meaning that a specific measure should be used for each function of the central executive.

To avoid the possible danger of construct underrepresentation, it seems worthwhile for SLA researchers to identify the extent to which each function of the central executive affects L2 learning by employing carefully chosen indices of executive functions. For example, task-switching, inhibition, and updating all seem pertinent to learning from exposure to multimodal recasts, the pedagogical intervention investigated in the present study. To benefit from multimodal recasts, learners would likely need to be able to switch between a focus on meaning and forms, and keep updating information through replacing irrelevant with relevant pieces of incoming input. Inhibition is crucial for avoiding interference between multiple sources of information. For instance, when learners encounter recasts from multiple modalities simultaneously, they may need to inhibit extraneous information from each modality and concentrate only on the input that is most relevant to their learning. Inhibition may also be important in terms of preventing interference from prior knowledge, which may conflict with new information learners receive. Hence, to gain a holistic picture of the role of WM in learning through multimodal recasts, the present study

included measures of these various executive functions: an automated operation span (AOSPAN) task for updating; a color shape task for task-switching; and a stop signal task for inhibitory control (see Section 4.2.7.3.2 for the detailed description of these WM measures).

Meanwhile, as PSTM pertains to the retention of sequential information, its capacity has typically been measured by span tasks that require recall of a rising number of digits or words/nonwords of increasing length in the order of presentation (Baddeley, 2000). Visuospatial sketchpad capacity has rarely been assessed in SLA research, as this component is considered less relevant to the processing of verbal information than PSTM (see however, Lee & Révész, 2021; Michel et al., 2019; Révész, et al., 2017; Sachs, 2010; White, 2021; Zalbidea & Sanz, 2020). However, it appears beneficial to include VSTM in studies exploring the processing of multimodal linguistic input. For example, based on their prior research (Lee & Révész, 2020), Lee & Révész (2021) further investigated whether WM mediates the effects of captions on attentional allocation and L2 grammatical development, and whether L2 development is influenced by WM capacity in the absence of captions. To assess various aspects of WM, the authors employed measures of PSTM (nonword repetition) and VSTM (forward Corsi block) as well as the executive functions of updating (AOSPAN), task-switching (color shape), and inhibitory control (stop signal). They found that, in both captioning conditions (enhanced versus non-enhanced), higher PSTM was associated with higher oral production gains. Participants in the no-captions group showed a positive link between VSTM and oral production gains. Attentional allocation only correlated positively with updating ability and PSTM under the enhanced captions condition. These findings indicate that WM can moderate the effects of captions on attention and L2 development, and varied WM components may play a differential role under various captioning conditions.

Similarly, given that WM is limited in capacity and interference is likely to arise when learners need to perceive different input sources simultaneously (Wickens, 2007), VSTM capacity may influence the extent to which learners pay attention to verbal input (i.e., oral recasts) concurrently presented with visual input (i.e., written recasts). To this end, along with the measure for the executive functions listed above, indices of PSTM (a nonword span task) and VSTM (forward Corsi block task) were also included in the present study (see also Section 4.2.6.5 for the detailed description of these WM indices).

2.7.4. Working Memory, Recasts and Learning of L2 Grammar in SLA

As mentioned earlier, the fundamental benefit of CF is its ability to draw learners' attention to target linguistic elements while engaged in conversational interactions. Considering that the main function of the central executive is focusing, dividing, and switching attentional resources (Baddeley, 1986, 1996), it seems reasonable to predict an important role for WM in learners' noticing and attention allocation to CF.

In this vein, WM has received much attention from L2 scholars interested in the efficacy of recasts, as they require learners to process concurrent form-meaning relationships. That is, in order to take in the target form provided in a recast, learners need to hold both their incorrect form and the target-like model long enough in their short-term memory to compare them and identify the location and nature of the discrepancy during interaction (Doughty, 2001; Goo 2012; Mackey et al. 2002; Mackey & Sachs 2012; Sagarra, 2007; Sagarra & Abbuhl 2013). This cognitive comparison is directly related to WM, which simultaneously stores and processes incoming information during complex cognitive tasks (Baddeley, 2003). This leads to the prediction that learners with greater WM span may

benefit more from recasts than other students with less WM (Robinson, Mackey, Gass, & Schmidt, 2012).

In the last decades, a large number of studies in SLA have explored the extent to which WM would mediate the effects of CF (Granena & Yilmaz, 2018). Some studies have examined the potential link between WM and recasts (e.g., Dai, 2013; Egi et al., 2002; Kim et al., 2015; Kim & Cho, 2019; Li, 2015; Li et al., 2019; Mackey et al., 2002; Mackey & Sachs, 2012; Révész, 2012; Sanz et al., 2016; Zhao, 2015), while others compared multiple feedback moves (i.e., clarification requests, metalinguistic feedback, and explicit correction) with recasts (e.g., Ahmadian, 2020; Goo, 2012; Li, 2013, 2015; Mackey et al., 2010; Mifka-Profozic, 2013; Yilmaz, 2013a). WM was assessed by various measures across these studies. For example, measurements for PSTM includes nonword repetition (e.g., Egi et al., 2002; Kim & Cho, 2017; Mackey et al., 2002; Mackey & Sachs, 2012; Révész, 2012; Zhao, 2015) and digit span (e.g., Révész, 2012), while executive WM tests consisted of listening span (e.g., Kim & Cho, 2017; Li, 2013, 2015; Mackey et al., 2002, 2010; Mackey & Sachs, 2012; Zhao, 2015), aural running span (e.g., Kim et al., 2015), reading span (e.g., Goo, 2012; Révész, 2012), speaking span (e.g., Mifka-Profozic, 2013), and operation span tasks (e.g., Ahmadian, 2020; Goo, 2012, 2016; Li, Ellis, & Zhu, 2019; Li & Fu, 2018; Yilmaz, 2013a). VSTM, the least explored WM component was assessed via odd-one-out tasks (White, 2021). Despite the bulk of studies on the relationship between WM and recasts, their association remains unclear, as some studies have found a significant association between them, while others have not.

Mackey et al. (2002), for example, conducted a small yet pioneering study examining the relationship between WM and noticing of recasts and the learning of L2 English question formation. Thirty adult Japanese ESL learners had three 30-minute sessions of dyadic task-

based interaction, which were followed by stimulated recall interviews to evaluate noticing. Following Robinson's model, the authors defined noticing as "detection with awareness and rehearsal in short-term memory" (Robinson, 1995, p. 318). Learning of the target construction was determined by the production of target-like higher stage questions suggested by Pienemann and Johnston (1987). The noticing data were based on the learners' comments during a stimulated recall and the learners' responses to an exit questionnaire. WM was assessed by L1/L2 listening span tasks and a nonword recall test, which were used as a measure of executive WM and PSTM, respectively. Results showed that learners' phonological WM capacity was positively related to their noticing of recasts as well as in the use of L2 question formation. The researchers reported that learners with a high WM span showed more noticing of recasts than learners with low WM. Interestingly, learners with a low WM span showed short-term L2 development in immediate posttests, whereas those with a high WM span demonstrated greater long-term L2 development in delayed posttests. That is, WM had an impact on the duration of effectiveness of recasts.

Mackey et al.'s work is important in that it is the first attempt to address the relationship between WM and the effects of recasts in SLA research. However, due to its small sample size, the researchers cautioned against the generalizability of their findings. In particular, regarding the impact of WM on the delayed effects of recasts, Mackey and Sachs (2012) reported somewhat different results. They examined whether older learners with higher WM benefit from recasts on the acquisition of English question formation. Nine ESL learners with a mean age of 72 engaged in three communicative tasks (i.e., spot-the-difference, picture drawing, and picture sequencing), in which each of them interacted with a native speaker and received recasts on their problematic question forms. Learning gains were measure by spot-the-difference and picture-description tasks. WM was measured by nonword

repetition and listening span tests. The correlation analysis of the development and WM scores revealed that only the learners with the highest listening span scores demonstrated development on the immediate posttest, whereas none of the learners with lower listening span scores did so. Also, it seems worth noting that it was only the two learners with the highest listening spans who maintained their stage progression at the delayed posttest. As noted, this specific pattern of results differs from Mackey et al.'s (2002) study, in which lower WM participants showed initial improvement but did not maintain it, whereas higher WM participants demonstrated development only after a delay. While the results of both studies are limited in generalizability due to the small sample size, both studies revealed a potential relationship between WM and the duration of recasts effectiveness. Further research is warranted to understand the nature of this relationship in greater depth and to determine whether this association may transfer to other communicative L2 learning contexts such as video-based SCMC interaction.

Mackey et al. (2010) also examined the association between WM capacity and modified output. Forty-two English-speaking learners of Spanish as a foreign language participated in four dyadic interactional tasks (a map task, a picture-drawing task, a spot-the-difference task, and a story completion task) with a Spanish native speaker who provided feedback in the form of clarification requests, repetitions with rising intonation, and recasts. An L1 listening span task administered to measure learners' WM found that WM capacity positively predicted the production of modified output; high-span learners modified their output following interactional feedback more often than low-span learners. This suggests that WM mediates the extent to which learners notice recasts and incorporate them into their own production, which is necessary for restructuring and improving their L2 knowledge.

A positive association between WM and noticing of recasts was observed in Kim et al.'s (2015) study as well. The researchers investigated the impact of task complexity and WM on the noticeability of recasts as well as the development of English question formation. Regardless of the task condition (simple or complex), the participants received recasts during task-based interaction with a native speaker. They were instructed to repeat the last-heard utterance when the native speaker knocked on the desk twice. Noticing was measured by the type of response after the knocks (e.g., full repetition of the recast). WM was measured using an aural running span test. The analysis of oral production tests administered three times indicated that WM capacity was the sole predictor of the level of noticing as well as L2 development. Thus, this study also suggests that high WM capacity plays a beneficial role in the noticing of recasts and learning of morphosyntactic features during interactional tasks involving recasts.

In relation to noticing and modified output, Zhao (2015) examined how WM predicts the learners' repair following recasts. The learners' repair was operationalized as "uptake-with-repair", namely, the learners' attempts to repair their erroneous utterances by repeating or integrating the recasts. The study involved one-way information gap tasks in a classroom setting, where the teacher provided either corrective recasts (a combination of repetitions and recasts) or implicit recasts (without repetitions) to correct errors in the third person *-s* and the formation of embedded questions. In the corrective recast condition, the teacher explicitly pointed out the error by repeating it and provided a recast, while in the implicit recast group, the teacher only provided a recast without drawing learners' attention to the error. The effects of these two types of recasts were examined both in terms of learners' acquisition of implicit knowledge (measured by elicited imitation and oral production tests) and their acquisition of

explicit knowledge (measured by a written production test and an untimed GJT). The WM test was comprised of listening span (executive WM) and nonword span (PSTM) tests.

Results showed that corrective recasts facilitated the development of both implicit and explicit knowledge, but implicit recasts only led to significant gains in explicit knowledge. Specifically, corrective recasts were more effective than implicit recasts in promoting implicit knowledge of embedded questions and explicit knowledge of third person *-s*. These results suggest that explicitness in recasts makes a difference to their efficacy, and the superior effect of corrective recasts over implicit recasts on the development of implicit or explicit knowledge depends on the target structure. The repair rate was also higher after corrective recasts than implicit recasts. With regard to the relationship between WM and learners' repair rate, the correlation analysis revealed that learners' repair rate was not constrained by their PSTM or executive WM capacity in the corrective recast condition. However, one exception was a significant correlation between PSTM and the ability to repair the third person *-s* errors after receiving implicit recasts. The researcher attributed the results to the classroom context where implicit recasts were not salient. As Lyster (1998b) argued, implicit recasts are indistinguishable from non-corrective repetitions, and their corrective function is easily overridden by their discourse function in communicative classroom settings. In addition, the redundant non-salient structure may increase the difficulty of attending to form. Thus, learners with higher PSTM are more likely to perceive the corrective intention of implicit recasts and modify their own production accordingly. To sum up, Zhao's study suggests that it is important to consider how recasts are operationalized, as this can affect the extent to which learners notice the recasts and thus impact their effectiveness.

Kim and Cho (2017) suggested that the learning context might have a potential impact on the influence of WM. They examined how WM affected Korean EFL learners' acquisition

of regular and irregular English past tense in one-on-one conversational interaction. Forty-two participants divided into a corrective recast group and a control group completed two story sequence tasks and one information gap task during the treatment session. Only the corrective recast group received feedback on their erroneous responses. Learners' WM and PSTM were measured by a listening span test and a nonword span test respectively. Results showed that neither learners' PSTM nor WM had a significant correlation with their accuracy improvement, rejecting previous findings that learners with higher WM span may gain more from recasts. Kim and Cho suggested that, along with the characteristics of recasts in their study (i.e., shortened length, one change, and interrogative intonation), the intensive dyadic interactional tasks conducted in a controlled laboratory setting may have made the recasts more salient and noticeable to the learners, rendering the high involvement of their WM unnecessary. Hence, expanding upon the work of Zhao (2015) and Kim and Cho (2017), it would be informative to investigate whether the impact of WM varies depending on the types of recasts that have different degrees of corrective salience, as well as the nature of L2 learning contexts.

Besides exploring the connection between WM and recasts, other researchers have also empirically assessed this relationship by comparing recasts with explicit types of feedback, yielding inconsistent results (e.g., Ahmadian, 2020; Goo, 2012; Yilmaz, 2013a). Goo (2012) explored the effects of recasts and metalinguistic feedback on the acquisition of English *that*-trace filter, and how WM capacity was related to the extent to which learners benefitted from these two different types of feedback. Learners' WM was measured using two executive WM measures: a reading span task and an operation span task. Results showed that both metalinguistic feedback and recast groups outperformed the control group and there was no statistically significant difference between the two experimental groups in their

performance. In terms of the role of WM in the effectiveness of recasts and metalinguistic feedback, WM was found to be a significant predictor of the beneficial effects of recasts on the development of the *that*-trace filter. Specifically, both reading span and operation span scores predicted the effect of recasts, indicating that WM, especially executive attention, is involved in the noticing recasts. However, for metalinguistic feedback, WM had no significant influence.

Goo speculated that this difference may have been attributed to a function of externally driven processing for metalinguistic feedback and internally driven processing for recasts. According to this interpretation, while the explicitness of metalinguistic feedback may be helpful in perceiving the feedback easily, it may not be effective in triggering the cognitive comparison necessary for processing the error-target form relationship (Doughty, 2001). Recasts, on the other hand, may have been effective in inducing the learners' cognitive control of attention to the target forms. Thus, the learners with higher WM capacity may have been able to capitalize on this advantage. Goo's study supports the previously reviewed studies, which also found that high-span learners are more inclined to notice recasts and produce a greater degree of modified output, compared to those with low-span WM capacity (Mackey et al., 2002, 2010, Mackey & Sachs, 2012; Kim et al., 2015).

Similarly, Ahmadian (2020) found a significant association between WM and recasts. The researcher examined the effectiveness of recasts and explicit instruction in teaching refusal strategies to upper-intermediate EFL learners, along with the interactions between WM and each feedback type. Seventy-eight participants engaged in peer interaction via role-play to learn pragmatic feature of refusals, which required learners to spontaneously respond to their interlocutor's request or suggestions without much preparation. Participants in the recast group were given an input-enhanced text where the target structure was underlined and

received recasts on their inappropriate use of refusal expressions during the role-play. The explicit instruction group was provided with descriptions and examples of the target structure along with explicit corrective feedback. The control group did not engage in any task that contained an element of refusal. To assess the learning gains, a discourse completion test (productive knowledge) and a comprehension questionnaire (receptive knowledge) were used. Participants' WM was measured by an operation span test. Results revealed that both recast and explicit feedback groups outperformed the control group, and explicit feedback was found to have superior efficacy over recasts in both outcome measures. Also, the learners' WM had a strong positive correlation with the development in the immediate and delayed posttests for both productive and receptive knowledge. As in Goo (2012), the correlation was only significant in the recast group. Ahmadian explained these findings as suggesting that the learners in the recast group may have had a heavier cognitive and attentional load due to the implicit nature of the recasts than those in the explicit instruction group. Although the primary focus of this study was not on grammar learning, its findings provide some insights into the role of WM in the effectiveness of recasts for both productive and receptive knowledge development, as well as their long-term effects. Furthermore, this research demonstrates the impact of WM on learners' spontaneous oral production, which was one of the assessments in the present study.

By contrast, Yilmaz (2013a) reported contradictory results in his research investigating the role of executive WM (as measured by an operation span task) in the effects of explicit correction and recasts. Forty-eight beginners of Turkish performed one-way information gap tasks and received explicit correction or recasts on their incorrect use of the target constructions (Turkish plural and locative morphemes), which have different levels of salience. Learning gains were assessed by comprehension, oral production, and written

recognition tests. Contrary to Goo (2012) and Ahmadian (2020), results revealed that WM had a positive and significant correlation with the performance of the explicit correction group. Yilmaz speculated that the effects of WM were observable only when learners noticed the corrective intent of the feedback. In this study, the use of interactional tasks requiring learners to answer meaning-oriented questions may have obscured the corrective intent of the recasts, possibly making it difficult for learners to recognize them as corrections. As a result, this may have led to a low level of WM involvement. Yilmaz's study also revealed that WM was linked to L2 development only for the more prominent target form, suggesting that the impact of WM may be influenced by the salience of the target structure. Taken together, these findings suggest that the predictive power of WM may be contingent on the explicitness of feedback and the salience of the target structure.

As has been noted, visuospatial WM has rarely been explored in SLA. In a recent work, White (2021) investigated the relationship between WM and various dimensions of L2 learning, including syntax, pragmatics, and semantics during a year-long interaction-based instruction. The study involved 27 young ESL learners (aged 5–6) in their first year of formal immersive education. They were taught different subjects and interacted with their teacher and peers in the L2. The participants' L2 proficiency was tested at three points: at the beginning (T1), in the middle (T2), and at the end (T3) of the year. The outcome measures consisted of L2 development in vocabulary, syntactic structures, semantic interpretation (e.g., verb contrast, preposition contrast, and quantifiers), and pragmatics (e.g., turn-taking, short narratives, and question-asking). Nonword repetition and odd-one-out tests assessed PSTM and VSTM respectively. The results revealed significant positive correlations between all WM measures and all linguistic performance measures across the year. Notably, VSTM was significantly related to syntax at each time point throughout the year, indicating that learners

with higher VSTM were better at acquiring the grammatical structures of the L2 language over time. From this finding, it can be assumed that the ability to retain and manipulate visual information helps learners to recognize and understand the grammatical structures of a language, leading to better learning outcomes. Thus, it is advisable to include measures for VSTM along with other WM assessments when the aim is to examine L2 learners' knowledge development in grammar.

It is also worth noting that the few studies which have included separate measures of PSTM and executive functions found distinct results for the different indices (e.g., Révész, 2012; Yilmaz & Granena, 2019). For example, Révész (2012) investigated the link between recasts and WM, postulating that different aspects of WM are associated with different types of L2 development, as assessed by various outcome measures. In this study, 90 EFL learners in Hungary were divided into three groups, recast, non-recast, and control, and took part in contextualized treatment tasks. Three different outcome measures (i.e., an oral production test, a written production test, and GJT), were administered as a pretest, an immediate posttest, and a delayed test to determine learners' knowledge development in the use of the English past progressive form. Three different WM measures were used: digit span and nonword span tasks for PSTM, and a reading span task for executive WM capacity. Results showed that the WM capacity of the participants who received recasts had significant correlation with developmental measures. Learners with a high executive WM span achieved greater improvement on the written production test and GJT, whereas those with a greater PSTM achieved greater improvement on the oral production test. These findings indicate that different types of WM may be related to different aspects of L2 learning. Given the results of this study, it can be interpreted that PSTM is more closely involved in enhancing the benefits

of recasts for oral production but not for written production, while the opposite may be true for the complex WM.

Révész interpreted her findings as suggesting that high PSTM capacity may have enabled learners to retain the grammatical input entailed in recasts longer in short-term memory, leading to increased likelihood that long-term memory traces were encoded. These long-term memory traces could then be subjected to data-driven processes, promoting development in procedural knowledge (N. Ellis, 2005). On the other hand, high executive WM capacity might have made learners more capable of paying attention to the recasts they received, leading to greater improvement in declarative knowledge. As a result, increased procedural and declarative knowledge is likely to have helped learners to perform better on the oral and written tests respectively. The findings also suggest that the mediating effect of executive WM and PSTM may vary according to the outcome measures, which clearly points to a need to incorporate multiple measures to precisely capture the relationship between WM and different aspects of L2 knowledge.

Yilmaz and Granena (2019) investigated whether attention control and PSTM predicted the effects of explicit correction and recasts on learners' development of Spanish gender agreement (N = 112). Attention control was measured through a Task Switching Numbers test, which evaluated the learners' ability to shift attention among different cognitive tasks. PSTM was assessed by a letter span test. The learners were randomly assigned to one of the three groups (explicit correction, recast, or control) and interacted with a native speaker during spot-the-difference tasks. Participants' improvement in the target structure was assessed using an oral production task and a GJT. The results showed that in the recast group, neither PSTM nor attentional control had an impact on the learning gains in the oral production test and GJT. In the explicit correction group, on the other hand, PSTM

was predictive of development on the GJT. That is, high PSTM learners benefitted from explicit correction in terms of developing of declarative knowledge. These findings contrast with those of Révész (2012), who found a positive correlation between PSTM and the development of procedural knowledge in the recast group condition.

Yilmaz and Granena attributed the differences in the results to the target feature salience and task design. The target structure used in their study (i.e., Spanish gender agreement) was less physically salient than the English past progressive form used in Révész's investigation. Therefore, even learners with high PSTM might have had difficulty detecting the structure in the recasts, given its low salience. In addition, the oral production test utilized in their study was more likely to measure declarative knowledge than procedural knowledge, in that it required the participants to identify differences between two pictures, a discrete item primarily assessing explicit knowledge. In contrast, the task employed in Révész's study asked learners to describe what eight people were doing for ten minutes, which was relatively more spontaneous and thus likely a better measure of procedural knowledge. Taken together, these two studies underscore the importance of incorporating multiple measures of WM and learning outcome to precisely evaluate the relationship between WM and different aspects of L2 knowledge. The studies reviewed are summarized in Table 6 below (the studies are presented in the order in which they appeared in this section).

Table 6. Studies on the Role of WM in the Effects of Interactional Interventions

Study (year)	Study Setting	Participants	Target Structure	Interventions	Treatment Tasks	Assessments	WM Measurements	Results
Mackey et al., (2002)	Lab	30 adult Japanese ESL learners	English question formation	Recasts	1) Picture drawing 2) Picture difference 3) Story completion	1) Noticing: Stimulated recall, questionnaire 2) Development: Interactional tasks	1) EWM: LSPAN 2) PSTM: NWR	1) High WM predicted more noticing. 2) Learners with low WM showed more development in immediate posttest while high WM learners outperformed in delayed posttest.
Mackey & Sachs (2012)	Lab	9 older ESL learners	English question formation	Recasts	1) Spot-the-difference 2) Picture-drawing 3) Picture-sequencing tasks	1) Spot-the-difference 2) Picture-sequencing tasks	1) EWM: LSPAN 2) PSTM: NWR	Older learners with higher WM benefited from interaction in L2 learning.

Study (year)	Study Setting	Participants	Target Structure	Interventions	Treatment Tasks	Assessments	WM Measurements	Results
Mackey et al., (2010)	Lab	42 English L1 learners of Spanish	Not specified	1) Prompt (i.e., mainly clarification request and repetition) 3) Recasts	1) A map task 2) Picture-drawing task 3) Spot-the-difference task 4) A story completion task	1) Response to feedback 2) Interactional tasks	EWM: LSPAN	Learners with higher WM modified their output more.
Kim et al. (2015)	Lab	81 ESL learners	English question formation	Recasts	Two-way collaborative information gap (simple vs. complex tasks)	1) Noticing: immediate cued recall 2) Development: 3 oral production tests	1) EWM: aural running span	WM had a significant positive association with noticing and L2 development regardless of task complexity.

Study (year)	Study Setting	Participants	Target Structure	Interventions	Treatment Tasks	Assessments	WM Measurements	Results
Zhao (2015)	Classroom	109 L2 Chinese learners of English	English third person *-s*	1) Corrective recasts 2) Implicit recasts	One-way information gap tasks (focused task)	1) Uptake: recording of the treatment; exit questionnaire 2) Implicit knowledge: EIT and oral production tests 3) explicit knowledge: the written production and untimed GJT	1) EWM: LSPAN 2) PSTM: NWR	The only significant association was found between PSTM and the repairs after implicit recasts.
Kim & Cho (2017)	Lab	42 Korean EFL learners	Regular and irregular English past tense	1) Corrective recast 2) Control	1) Two story sequence tasks 2) one information gap task		EWM: LSPAN PSTM: NWS	Neither learners' PSTM nor WM had a significant correlation with their accuracy improvement.

Study (year)	Study Setting	Participants	Target Structure	Interventions	Treatment Tasks	Assessments	WM Measurements	Results
Goo (2012)	Classroom	60 Korean EFL learners	*That*-trace filter	1) Recast 2) Metalinguistic feedback 3) Control	One-way information gap activities	1) GJT 2) WPT	EWM: RSPAN, OSPAN	WM significantly predicted the effects of recasts, not metalinguistic feedback.
Ahmadian (2020)	Classroom	78 upper-intermediate EFL learners	Pragmatic feature of refusal	1) Implicit instruction with recasts 2) Explicit instruction with explicit correction	Role-play	Pragmatics development: 1) Discourse Completion Test 2) Comprehension questionnaire	EWM: OSPAN	WM had a strong positive relationship with learning gains in the implicit instruction group.
Yilmaz (2013a)	Lab	48 beginners of Turkish	Turkish plural and locative morphemes	1) Explicit correction 2) Recasts 3) Control	One-way information gap task	Learning: oral production, comprehension, and recognition tests	EWM: OSPAN	Positive correlation between WM and explicit correction group; WM linked to L2 development only for the more prominent target form

Study (year)	Study Setting	Participants	Target Structure	Interventions	Treatment Tasks	Assessments	WM Measurements	Results
White (2021)	Immersive language learning program	27 young ESL learners, (aged 5–6)	Vocabulary, Syntactic structures, Semantic interpretation, Pragmatics	Immersive language learning	Immersive language learning	1) Receptive skills: Picture vocabulary test 2) Comprehension, production skills: (Developmental Evaluation of Language Variation-Criterion Referenced Edition)	1) PSTM: NWR 2) VSM: Odd-one-out	1) PSTM consistently predicted learning gains on syntax and pragmatics for 1 year. 2) PSTM predicted further outcomes on semantics and pragmatics in later time points of the year.
Révész (2012)	Lab	90 EFL learners in Hungary	Past progressive	1) Recasts 2) No recasts 3) Control	Picture description	1) GJT 2) Written picture description	PSTM: DS, NWS EWM: RSPAN, OSPAN	1) Complex verbal WM associated with improvement on written production test and GJT. 2) PSTM associated with improvement

Study (year)	Study Setting	Participants	Target Structure	Interventions	Treatment Tasks	Assessments	WM Measurements	Results
								on oral production test.
Yilmaz & Granena (2019)	Lab	112 learners of Spanish	Spanish gender agreement	1) Explicit correction 2) Recasts 3) Control	One-way information gap task	1) OPT 2) GJT	1) PSTM: Letter Span test 2) Attention Control: Task Switching Numbers Test	PSTM predicted the improvement on the GJT, only in the explicit correction group.

Notes. EIT: Elicited Imitation Test, OPT: Oral Production Test, WPT: Written Production Test, GJT: Grammaticality Judgement Test, EWM: Executive WM, PSTM: Phonological Short-Term Memory, VSTM: Visuospatial Short-Term Memory, LSPAN: Listening Span Task, NWR: Nonword Repetition Task, RSPAN: Reading Span Task, WMC: WM Capacity.

2.7.5. Working Memory and Recasts in SCMC research

There has been an increasing interest in the variable WM as an important factor in L2 learning not only in the CMC (Sagarra, 2007; Saggara & Abbuhl, 2013; Trofimovich et al., 2007; Zalbidea & Sanz, 2020) but also in SCMC mode (Baralt, 2010, 2015; Lai et al., 2008; Payne & Ross, 2005; Payne & Whitney, 2002). These studies have examined the role of WM capacity in learners' noticing of contingent versus non-contingent recasts (Lai et al., 2008); face-to-face versus text-based recasts and task complexity (Baralt, 2010, 2015); oral versus written tasks and feedback modalities (Zalbidea & Sanz, 2020); text-based SCMC versus face-to-face interactions and general proficiency development (Payne & Whitney, 2002); SCMC interactions and repetition and relexicalization (Payne & Ross, 2005); synchronous versus asynchronous written feedback (Shintani & Aubrey, 2016); computerized pre-recorded oral recasts (Sagarra, 2007; Trofimovich et al., 2007); automated computer-delivered oral and written recasts with various input manipulation (e.g., input enhancement) (Sagarra & Abbuhl, 2013); and various SCMC modalities on L2 grammar acquisition (Martin, 2018). Overall, as with previous research in face-to-face contexts, findings from the SCMC research in this strand remain inconclusive.

First, Lai et al. (2008) examined whether WM predicted the noticing of contingent versus non-contingent recasts among seventeen ESL learners in text-based SCMC. The participants completed two spot-the-difference tasks while chatting individually with one of the researchers. Learners' noticing was measured through think-aloud protocols during their task performance, and through stimulated recall while viewing the video-recordings of their performance of the tasks. A reverse digit span test was used to test WM capacity. Results indicated that contingent sequences of recasts were better noticed than non-contingent ones. Moreover, WM capacity was significantly correlated to the noticing of recasts, and higher

WM capacity was associated with greater noticing of non-contingent recasts in text-based SCMC. The study suggests that non-contingent recasts in text-based interaction consume more WM, compared to contingent ones. Adding to the benefit of being more noticeable (e.g., Leeman, 2003; Long, 2007), this study provides further evidence that contingent recasts are advantageous for reducing WM load in text-based interaction. More research is needed to investigate whether this advantage applies to various types of recasts (i.e., oral, written, and multimodal) provided in a contingent fashion during video-based SCMC. Although their research did not directly address synchronous communication, both Sagarra (2007) and Trofimovich et al. (2007) employed pre-recorded, computer-delivered oral recasts to examine the impact of WM on recasts. Interestingly, these two studies reported contradictory results. In Sagarra's work, 36 native speakers of English college students in first-semester Spanish courses at a US university participated in the study. They were asked to complete computerized fill-in-the-blank exercises on Spanish noun–adjective agreement. After entering incorrect responses, participants in the recast group received computer-delivered oral recasts of their erroneous responses, whereas the control group did not receive any feedback. The effects of recasts were assessed by oral production and written production tests. The WM test was adapted from Waters and Caplan's sentence span tests (1996). Sagarra found that computer-delivered recasts improved L2 accuracy both in the oral and written production tests. In addition, the impact of WM on the recast group's L2 learning was evident in not only immediate posttests but also delayed posttests, suggesting that the effectiveness of recasts positively correlated with the learners' WM capacity. Trofimovich, Ammar, and Gatbonton (2007), on the other hand, found that WM did not influence learners' noticing of recasts nor their subsequent learning. The researchers examined the correlation between different aptitude components (attention control, PSTM, WM, and analytic ability) and learners'

ability to notice and benefit from computerized oral recasts. Thirty-two adult Francophone learners of English were presented with pictures on a one-on-one basis and required to describe them using the target structures, which were morphosyntactic (English possessive determiners), lexical (intransitive verbs), or both (transitive verbs followed by a possessive determiner/noun combination). Regardless of its accuracy, the learner's description of each picture was followed by the pre-recorded native speaker's response that served as a recast. Recasts took the form of morphosyntactic recasts, lexical recasts, or both. After each recast, learners were asked whether they perceived any differences between their own production and the recast to capture their noticing. The same type of task that learners performed during the treatment was also used for the pretest, posttest, and delayed posttest. Two WM measures were used: an L2 nonword recognition test measuring PSTM, and a L1 Letter-Number Sequencing subtest of the Wechsler Adult Intelligence test (Psychological Corporation, 1997) for executive WM. The Words in Sentences subtest of the MLAT was used to measure analytic ability, and attention control was tested using the Trail Making Test of the US Army Individual Test Battery.

Results revealed that participants were significantly better at detecting their lexical errors compared to the morphosyntactic errors following recasts, which aligns with previous research demonstrating lower noticing rates for morphosyntactic recasts (e.g., Carpenter et al., 2006; Lyster, 1998; Lyster et al., 2013; Lyster, Saito, & Sato, 2013; Mackey et al., 2000). The learners' accuracy in producing linguistic targets improved overall after receiving recasts. The study also found that learners' attention control, analytic ability, and PSTM predicted their accuracy development, whereas executive WM did not have a significant impact. However, it should be noted that executive WM was assessed through the Letter-Number Sequencing subtest of the Wechsler Adult Intelligence test, which is not commonly

used in the field of SLA. There was no significant relationship between learners' noticing and the four factors examined (attention control, WM, PSTM, and analytic ability). The researchers speculated that, unlike Sagarra's (2007) study, in which learners only received recasts on their errors, recasts in their study were given after each response regardless of accuracy. This may have made recasts frequent and prominent, making them highly noticeable and predictable to the learners. As a result, learners may have been able to pay attention to and detect the feedback without heavily relying on cognitive resources.

Sagarra and Abbuhl (2013) further examined the effects of automated oral versus written recasts on L2 learning and explored whether executive WM capacity (as measured by an L1 reading span test) was related to the efficacy of each feedback type. Two hundred and eighteen beginners of Spanish participated in a computerized fill-in-the-blank grammar practice and received one of the following types of feedback on their incorrect use of Spanish noun–adjective gender agreement: (1) no feedback (control group), (2) written utterance rejection (i.e., 'Wrong. Move on to the next sentence'), (3) oral utterance rejection (i.e., same as the written utterance rejection but delivered orally), (4) written recasts, (5) oral recasts, (6) textually enhanced written recasts (i.e., 'La silla es blancA'), and (7) orally enhanced recasts with stress. Learning gains were assessed with written fill-in-the-blank and face-to-face oral interactive production tests. Learners' ability to modify their output was also examined. Results indicated that recasts, regardless of whether they were oral or written, were more effective than oral or written utterance rejections in eliciting modified output and promoting the development of noun–adjective agreement. It was also found that orally enhanced recasts led to a significantly greater degree of modified output and learning compared to both orally unenhanced and textually enhanced recasts. WM capacity was positively associated with the oral and written gains, but only in the oral recast condition (both enhanced and unenhanced).

WM was also significantly associated with modified output production, but, again, only for those who received oral recasts. The findings indicate that the cognitive demands of processing written recasts may be lower than those of oral recasts, leading to reduced reliance on WM. As a result, learners with high executive WM span are more likely to benefit from oral recasts.

Compared to executive WM and PSTM, little research has dealt with the role of visuospatial WM in the effectiveness of CF in a computerized learning environment (Sachs, 2011; Zalbidea & Sanz, 2020). Of the few studies, Sachs (2011) investigated the relationship between VSTM and the effects of visual metalinguistic feedback. The visual metalinguistic feedback differed from the typical form of metalinguistic feedback in its use of tree diagrams providing visual metalinguistic information on the syntax of the target structures in sentences rather than verbal comments on learners' errors. VSTM was measured using the Visual Patterns test, which assesses the ability to store visuospatial information. During the treatment, participants engaged in a computerized practice on interpreting sentences containing the target construction (Japanese reflexive *zibun*). They received one of the three types of feedback: no feedback, right/wrong feedback, or visual metalinguistic feedback. Knowledge development in the target form was evaluated by a truth–value judgement test, which was similar to the treatment task. Results showed that learners who received the visual metalinguistic feedback demonstrated clear evidence of improvement, and only those with higher VSTM benefitted significantly from the feedback. This suggests that VSTM plays a critical role when feedback is presented visually, as visual feedback may impose a taxing processing load on learners, making VSTM essential to their ability to extract complicated syntactic rules from the visual input.

Sachs' study highlights the significance of the feedback type used in computer-mediated learning environments. It shows that the presence of visual input in the feedback has a substantial impact on learning outcomes, particularly for learners with varying levels of visuospatial WM. Therefore, it is advisable that studies examining the effectiveness of different types of feedback involving visual input need to account for VSTM and its interaction with the feedback types. Sachs' work also shows the pedagogical potential of explicit feedback visually presented in computer-mediated communication (CMC). Hence, it appears to be worth exploring whether this benefit extends to implicit types of feedback containing visual input of the target features, such as written and multimodal recasts in video-based SCMC, and whether they also facilitate L2 acquisition in this setting.

In a more recent study, Zalbidea and Sanz (2020) examined how learners' WM capacity related to their morphosyntactic knowledge development and perceived mental and task difficulty in computerized feedback conditions. Fifty-eight beginners of Spanish were assigned either to a speaking group or a writing group which differed only in terms of the modality in which learners had to produce the target constructions and received feedback. During the treatment, both groups completed two computerized, focused sentence production tasks targeting two Spanish grammatical structures carrying different salience (the simple future tense and the indirect clitic). Following the production of each sentence, both groups were presented with exemplars that demonstrated the correct usage of each grammatical structure, irrespective of the accuracy of their own responses. The speaking group received auditory feedback in the form of spoken sentences, while the writing group received the same feedback in written form. After each task, participants assessed their perceived level of mental effort and task difficulty. To measure immediate and delayed gains, production along with written and aural acceptability judgement tasks were conducted. Automated operation

span, nonword recognition, and forward Corsi block-tapping tasks were used to assess the participants' executive, phonological, and visuospatial WM abilities, respectively.

The results showed that each WM component only predicted L2 outcomes and task perception ratings in the speaking group. Specifically, PSTM and VSTM were linked to production and acceptability judgement performance accuracy, while executive WM was associated with learners' perceived mental effort. Regarding the two grammatical constructions, PSTM was positively linked to immediate aural acceptability judgement accuracy for the more salient future form. VSTM, on the other hand, was positively associated with immediate production and delayed written acceptability judgement accuracy. For the less salient clitic form, PSTM had a positive correlation with immediate production and aural acceptability judgement accuracy, and VSTM was a beneficial factor for immediate written and aural acceptability judgement accuracy. Overall, these findings suggest that each WM component concerns different aspects of L2 development and influences the effect duration of recasts in a computerized L2 learning practice.

Zalbidea and Sanz's results align with Sagarra and Abbuhl's (2013) research, which also found a strong correlation between WM and L2 gains in the oral feedback group. In addition, expanding upon the previous findings, Zalbidea and Sanz's study presents further evidence for the relevance of WM abilities in learning gains using oral and written focused tasks. Although the primary focus of this study was not the relationship between WM and recasts, it can still benefit the current study in that it demonstrated how different components of WM (i.e., executive, phonological, and visuospatial WM skills) contribute to different domains of L2 learning and affect the effects of different feedback modalities in computer-mediated environments. The results are particularly significant in that they shed light on the

under-researched visuospatial WM function in computer-mediated L2 acquisition settings, which has received little attention in previous research on L2 learning and recasts.

Payne and Ross (2005) also found a positive link between WM and L2 development in an online course with a focus on interactive communication. The study involved 24 college students taking a Spanish course who participated in 20 chat sessions involving group interactions such as role-plays, personal theme discussions, and jigsaw reading activities. The participants' oral proficiency was evaluated by oral production tasks at the beginning and end of the semester, and their phonological WM and executive WM were assessed using nonword repetition and reading span tests, respectively. The findings revealed that although executive WM did not contribute to the development of oral proficiency, participants with higher PSTM demonstrated significantly more learning gains than those with lower PSTM over the 15 weeks of the course, providing further evidence that PSTM is concerned with learners' oral performance.

Some studies empirically attested to the role of WM in the effectiveness of recasts in face-to-face versus text-based SCMC (Baralt, 2010, 2015). Baralt (2010) examined how task complexity and WM affected the development of the Spanish past subjunctive through recasts by comparing face-to-face and text-based interactions. Task complexity was operationalized as [+/− intentional reasoning]. The simple group was asked to retell a story (−intentional reasoning), whereas the complex group was also required to hypothesize the character's reason for a certain action (+intentional reasoning) while retelling the stories. Learners in all groups received recasts during their interaction. Improvement in productive knowledge and receptive knowledge were measured by story retell and multiple-choice tests respectively. Three different memory span tasks were used: an operation span task, a counting span task, and a reading span task. Baralt found a significant correlation between

WM and face-to-face oral recasts in a simple task group in terms of the improvement of receptive knowledge. WM did not play any role in the text-based SCMC, regardless of task complexity.

Based on this study, Baralt (2015) further explored the relationship between WM and the efficacy of recasts provided during text-based SCMC. Results also showed that WM was not related to learning outcomes. Also, WM did not moderate learning for either high- or low-WM span learners. Specifically, while participants did learn the Spanish past subjunctive, the construct of WM capacity did not affect the extent to which they benefitted from recasts in the text-based mode.

Baralt's findings suggest that, as in Sagarra and Abbuhl's (2013) study, the role of WM in learning from recasts is not significant when recasts are delivered in written form. The rationale behind this is that the executive function of WM is crucial for filtering out irrelevant stimuli and focusing on pertinent cues, primarily in processing the incoming auditory input of L2. However, when CF is provided in a written form via text-based communication, the processing and comparison of incorrect and corrected forms in WM is not required, since the written feedback is durable and readily noticeable without putting much burden on WM resources. In contrast, oral feedback, particularly recasts, necessitate the use of WM in comparing erroneous and target-like forms, as feedback given in oral interactions is transient and fleeting, thus making learners more likely to depend on the availability of WM to a greater extent.

With regard to the role of executive WM in text-based recasts, Yilmaz and Sagdic (2019) reached the same conclusion. They explored the relationship between inhibitory control, one of the functions of the central executive, and the effects of immediate and delayed provision of feedback in text-based SCMC. The learners in the immediate feedback

group received recasts during one-way information gap interaction, while the delayed feedback group received recasts after the task. Learners' inhibitory control was measured using an Eriksen Flanker test, which tapped the ability to suppress extraneous information to make a quick decision. Results suggested that learning gains, measured by an oral production test and a grammaticality judgement test, were not affected by inhibitory control on either outcome measure. Yilmaz and Sagdic attributed the results to the characteristics of text-based SCMC. The slower pace of text-based interactions provides learners with extra time to process the input in the recasts, and the permanent nature of written feedback enables learners to reread it during the chat. This can help alleviate learners' cognitive load on WM and level out the role of inhibitory control. However, to date, scant L2 research has attested empirically to the diversity of central executive functions (i.e., switching, inhibiting, and updating) and their distinctive role in second-language acquisition. Thus, further studies based on the theoretical assumption that the central executive functions may be separable are expected to shed more light on the construct and role of WM in L2 acquisition.

Finally, as previously reviewed, Martin's (2018) study focused on the role of executive WM (as measured by an operation letter span task) in L2 learning during text-based, video-based, and voice-based SCMC. Martin expected WM to have a greater impact on video or voice groups due to the cognitively more challenging level of oral communication. However, the results showed that all groups successfully acquired the target structure regardless of their WM capacity. This may have been possible due to the explicit nature of the feedback and the manageable level of task difficulty, even for the video and voice groups.

Nonetheless, it is unclear whether the results are the same when an implicit type of feedback, such as recasts, are used. As explained, recasts require more cognitive resources, as

learners must hold their own utterance and the correct form in WM to deduce the target-like form from the reformulation. Thus, the explicit nature of metalinguistic feedback may be less demanding on WM than recasts, making it easier for learners to acquire the target structure. Thus, it is plausible that the role of WM in the feedback modality may be more pronounced when recasts are provided. However, this is a speculative assumption that requires further research to examine the effect of feedback type on WM's role in L2 learning using different SCMC modalities. Overall, these findings suggest that WM plays an important role in terms of mediating the benefits of recasts for learners' L2 grammatical development. However, little is known about how WM might moderate the effectiveness of recasts as well as the effects of feedback modality in the context of multimodal video-conferencing tasks. The reviewed studies in this section are summarized in Table 7.

Table 7. CALL/SCMC Studies on the Role of WM in L2 Learning

Study (year)	Modality /Experimental Settings	Participants	Target Structure	Interventions	Treatment tasks	Measures	WM Measures	Results
Lai et al. (2008)	Text-based	17 ESL learners	N/A	Text-based 1) contingent recasts 2) non-contingent recasts	Spot-the - difference	Noticing: 1) Think-aloud protocols 2) Stimulated recalls	EWM: Reverse digit span	WM had a strong, significant relationship with noticing of non-contingent recasts.
Sagarra (2007)	Computerized grammar practice	36 L1 English learners of Spanish	Spanish noun-adjectives agreement	1) Computerized pre-recorded oral recasts 2) Control	Computerized fill-in-the blank exercises	1) OPT 2) WPT	WMC: WM test adapted from Waters and Caplan's sentence-span tests	1) Recasts improved L2 accuracy in oral and written production tests. 2) Effectiveness of recasts positively correlated with learners' WM capacity, evident in immediate and delayed posttests.

Study (year)	Modality /Experimental Settings	Participants	Target Structure	Interventions	Treatment tasks	Measures	WM Measures	Results
Trofimo v-ich et al., 2007	Pre-recorded morphosyntact ic and lexical recasts	32 adult Francophon e learners of English	1) Morphosynta ctic: English possessive determiners 2) lexical: intransitive verbs 3) both	Computerized oral recasts targeting 1) morphosyntacti c item 2) lexical item 3) both	Picture-description tasks	Pretest, posttest, and delayed posttest using the same task as during treatment	PSTM: L2 nonword recognition test WM: L1 Letter-Number Sequencing subtest of the Wechsler Adult Intelligence test	1) Attention control, analytic ability, and PSTM were predictive of interlanguage accuracy development. 2) WM had no impact. 3) Learners better detected lexical errors than morphosyntacti c errors 4) No substantial relationship between learners' noticing and four indices examined.

Study (year)	Modality /Experimental Settings	Participants	Target Structure	Interventions	Treatment tasks	Measures	WM Measures	Results
Saggara & Abbuhl (2013)	Computerized grammar practice	218 beginners of Spanish	Spanish noun–adjective gender agreement	1) No feedback 2) Written utterance rejection 3) Oral utterance rejection 4) Written recast; 5) Oral recast; 6) Typographically enhanced recast; 7) Orally enhanced recast	Computer-based grammar practice	1) Written fill-in-the-blank test, 2) Oral FTF interactive tasks 3) Modifying output pattern: one of the interactive tasks	EWM: RSPAN	–WM was significantly correlated with the effects of oral recasts, both enhanced and unenhanced.
Sachs (2011)	Text-based	80 English-speaking learners of Japanese as a foreign language (JFL)	Japanese reflexive *zibun*	1) Right/Wrong Feedback 2) Tree-Diagram Metalinguistic feedback 3) No Feedback	Computerized sentence interpretation practice	Truth–value judgment tasks	VSTM: Visual Patterns Test	VSM was significantly correlated with learning gains only for the Tree-diagram metalinguistic feedback group.

Study (year)	Modality /Experimental Settings	Participants	Target Structure	Interventions	Treatment tasks	Measures	WM Measures	Results
Zalbidea & Sanz (2020)	Computerized production task	58 beginners of Spanish	Spanish 1) Simple future tense 2) Indirect clitic	1) Oral target-like model 2) Written target-like model	Two computerized focused sentence production tasks	1) Production test 2) Written and aural acceptability judgment tasks 3) Perceived mental effort and task difficulty ratings	1) EWM: OSPAN 2) PSTM: NWR 3) VSTM: Forward Corsi block test	1) PSTM was linked to immediate aural acceptability judgment accuracy for the more salient future form 2)VSTM was associated with immediate production and delayed written acceptability judgement accuracy for both grammatical constructions. 3) Executive WM was related to learners' perceived mental effort.

Study (year)	Modality /Experimental Settings	Participants	Target Structure	Interventions	Treatment tasks	Measures	WM Measures	Results
Baralt (2010)	1) Text-based 2) FTF	70 intermediate-level learners of Spanish	Spanish past subjunctive	1) FTF, complex 2) FTF, simple 3) CMC, complex 4) CMC, simple (+/– intentional reasoning)	Two-way information gap	1) OPT 2) Multiple-choice test	–EWM: OSPAN, CSPAN, RSPAN	–WM had a significant effect on receptive learning for FTF, simple task group.
Baralt (2015)	1) Text-based 2) FTF	34 intermediate-level learners of Spanish	Spanish past subjunctive	1) FTF, complex 2) FTF, simple 3) CMC, complex 4) CMC, simple (+/– intentional reasoning)	Two-way interactive story retells	1) Productive tasks 2) Multiple-choice test	-EWM: OSPAN, CSPAN, RSPAN	–WM did not play any role in SCMC regardless of task complexity.
Yilmaz & Sagdic (2019)	Text-based	43 learners of Spanish	Spanish noun–adjective gender agreement	1) Immediate feedback 2) Delayed feedback	One-way information gap picture description activity	OPT (oral production test) and GJT	EWM (Inhibitory control): Eriksen Flanker Task	No significant relationship was found between inhibitory control and learning for both groups.

Study (year)	Modality /Experimental Settings	Participants	Target Structure	Interventions	Treatment tasks	Measures	WM Measures	Results
Payne & Ross (2005)	SCMC (classroom)	24 learners of Spanish	Blended instruction	Blended instruction	Blended instruction	OPT	–EWM: RSPAN –PSTM: NWR	-Only PSTM predicted learning gains over the 15 weeks of the course.
Martin (2018)	1) Video-based 2) Voice-based 3) Text-based	65 intermediate learners of Spanish	Spanish past subjunctive	Metalinguistic feedback in three SCMC modes	Interactive dialogic story retell	1) OPT 2) WPT	EWM: Operation-Letter Span Task	WM did not predict learning gains in any modes.

Notes. OPT: Oral Production Test, WPT: Written Production Test, GJT: Grammaticality Judgement Test, EWM: Executive Working Memory, PSTM: Phonological Short-Term Memory, VSTM: Visuospatial Short-Term Memory, LSPAN: Listening Span Task, NWR: Nonword Repetition Task, RSPAN: Reading Span Task, WMC: Working Memory Capacity, FTF: Face-to-Face Communication, CMC: Computer-Mediated Communication, SCMC: Synchronous Computer-Mediated Communication.

2.7.6. Working Memory in Eye-tracking Research

Given that WM appears to be involved in L2 learning in a wide range of instructed contexts, it has been of interest to further examine how WM is related to attention allocation as manipulated by various instructional interventions. Recently, there has been a growing interest in the association between WM and learners' attention allocation to visual input. To date, only a few eye-tracking studies have investigated whether WM affects the extent to which learners pay attention to L2 target constructions entailed in captions and learn from them (e.g., Gass et al., 2019; Indrarathne & Kormos, 2018; Issa, 2019; Lee & Révész, 2021; Pattemore & Muñoz, 2020).

Most closely relevant to the present study is the research of Lee and Révész (2021), who investigated whether WM mediates the effects of captions on attentional allocation and L2 grammatical development, and whether L2 development is influenced by WM capacity in the absence of captions. Seventy-two Korean learners of English were randomly assigned to one of three groups. The groups differed as to whether they were exposed to news clips without captions, with textually enhanced captions, or with unenhanced captions during the treatment. Learners' attentional allocation was measured with eye-tracking methodology, and learning gains were assessed through oral production, written production, and fill-in-the-blank tests. To assess various aspects of WM, the researchers employed measures of PSTM (L1 nonword span task), VSTM (forward Corsi block task), and the executive functions of updating (automated operation span task), task-switching (color shape task), and inhibitory control (stop signal task). Results showed that in both caption groups, higher PSTM was associated with higher oral production gains. For the enhanced captions group, PSTM was also positively related to gains on the written production test. Participants in the no-captions group, however, showed a positive link between VSTM and oral production gains.

Attentional allocation only correlated positively with updating ability and PSTM under the enhanced captions condition. These results indicate that WM can moderate the effects of captions on attention and L2 development, and various WM components may play a differential role under various captioning conditions.

Gass et al. (2019) explored the effects of participants' WM on L2 comprehension and global caption-reading behavior while watching a short video clip (4 minutes 37 seconds) twice (Time 1 and Time 2) in two groups of participants whose L2 was either English or Spanish. Specifically, the authors conducted two parallel experiments with participants sampled from two distinct populations: Spanish foreign-language learners (i.e., Spanish L2 learners) (Experiment 1) and English second-language learners (i.e., ESL learners) (Experiment 2). A reading span test was used to measure WM. In both experiments, comprehension was assessed by asking participants to recall the story and type a detailed summary. Three eye-movement measures were used: the first-pass reading time and re-reading time and total reading time.

Results showed that captions promoted L2 video comprehension in both groups and that the recall scores of the ESL learners were higher than those of the Spanish L2 learners. While the WM capacity did not impact Spanish L2 learners' comprehension, ESL students with higher WM tended to perform better in a comprehension task. WM also influenced caption-reading behavior. The patterns were similar in both experimental groups. High WM learners appeared to spend less time reading captions during Time 2 and spend less caption-reading time in general. Learners with low WM capacities, on the other hand, increased their reading time during Time 2. The researchers speculated that, contrary to those with low WM span, high WM learners might have been more able to hold key information effectively in the episodic buffer during Time 1 and, therefore, might have gleaned more of the information

needed from the captions the first time around. With regard to the relationship between comprehension and the caption-reading behaviors, the two groups demonstrated different reading behaviors. In the Spanish L2 group, high-comprehension learners spent less time overall reading captions during the first-pass reading time and less time re-reading than the low-comprehension group. For the ESL group, there were fewer differences between the high- and low-comprehension learners.

Taken together, these two studies suggest that WM does impact learners' caption-reading behaviors and L2 development. Thus, it would be worthwhile to explore whether these findings were transferable to other audiovisual input conditions, such as caption-supported oral recasts (i.e., multimodal recasts), and how WM may relate to the effectiveness of this novel type of recasts in facilitating attention to and development in the knowledge of an L2 grammatical feature in the context of video-based SCMC.

Some researchers explored the role of WM in L2 grammar learning and processing using computer-mediated tasks (Indrarathne & Kormos, 2018; Issa, 2019). Indrarathne and Kormos (2018), for example, investigated how attention allocation and textually enhanced target grammatical construction was affected by WM capacity under various instructional conditions. In this study, eighty Sri Lankan intermediate learners of English received instruction on L2 grammar (causative *had*) under implicit and explicit instruction conditions. In the implicit conditions, learners read a story with either the target feature textually enhanced in bold font (Implicit 1) or not (Implicit 2); neither group received information about the target form. In the explicit conditions, learners read the textually enhanced story; one group additionally received explicit a metalinguistic explanation (Explicit 1), whereas the other group was further instructed to pay attention to the textually enhanced parts (Explicit 2). A sentence reconstruction task and a GJT were administered as pretest and posttest to

assess the production and comprehension of the target construction respectively. Particularly relevant to the present study is that, similar to Lee and Révész (2021), the authors employed a broad range of WM measures to assess both the storage capacity and attention regulation functions of WM, including an index of PSTM (forward digit span test) and measures of the updating (Keep Track task), task-switching (Plus Minus task) and inhibitory control (Stroop task) functions of the central executive. A composite score was then calculated in a regression analysis to explore how WM interacted differently with learning under the four different instructional conditions (i.e., aptitude-treatment interactions). Amount of attention paid to the target feature was assessed by the means of eye-tracking methodology.

Results showed that WM was associated with learning in all four conditions, with a significantly stronger association in the more explicit conditions. Furthermore, learners with higher WM abilities performed better overall than did learners with less WM capacity. In terms of the relationship between eye-tracking measures and WM, results revealed that in the most explicit learning condition (i.e., Explicit 2), individual differences in WM abilities are highly predictive of how much attention L2 learners pay to the target input. This suggests that learners with high WM storage capacity and efficient attention regulation abilities engage in more attentional processing of the input if their awareness of the existence of a target syntactic construction in the input is experimentally manipulated. In other words, these learners might respond better to the instruction that they should pay specific attention to the target construction embedded in the text because they are more efficient in directing their attentional resources.

In contrast, Issa's (2019) study using a computerized structured input practice found that the lower WM group allocated more attention to the target structure (Spanish direct object pronouns). In this study, Issa investigated how attention to form and meaning was

related to WM capacity during L2 structured input practice. Thirty-four beginners of Spanish were asked to complete tasks according to the treatment conditions they were assigned to: an attentional condition group (structured input practice) and a control group. The experimental group was exposed to the target constructions through picture-sentence interpretation practice, which was manipulated to draw learners' attention toward the target form. The control group performed a similar activity without such manipulation. WM capacity was measured by an automated Symmetry Span Task (Unsworth, Heitz, Schrock, & Engle, 2005). Learners' eye-movement was examined using four measures (first-pass reading time, second-pass reading time, total reading time, and fixation count). Learning gains in L2 grammatical knowledge were assessed with a sentence interpretation task.

Results indicated that while the learners who received structured input practice showed large, significant gains in L2 development of the target structure, the control group did not evidence such gains. Interestingly, there was no difference in learning outcomes for the high- and low WM groups. Issa suggested that it could be because the structured input treatment may have compensated for lower WM by allowing learners to process input in depth. Additionally, correlation analyses revealed that for the control group, no significant correlations were found between any of the eye-tracking measures examined and the WM scores, whereas the experimental group showed significant, medium-sized, negative correlations between WM scores and second-pass reading time as well as total reading time. These results suggest that, aligning with Gass et al.'s (2019) findings, learners with lower WM spent longer attending to the target form on their second-pass and overall reading times. Isaa interpreted this finding as suggesting that structured practice engages WM and that attentional allocation may be affected by a learners' WM capacity. Also, learners with lower WM may need to attend to the target form for longer overall and return to the form after

attending to it on their first-pass reading time. These negative relationships may be explained by the type of practice employed in the study, which asked participants to examine a visual scene, hold that information in WM, and then process a sentence to make a decision about it. Thus, the structured input practice may have been more taxing for low WM learners' cognitive resources, causing them to fixate on the pronoun and verb longer to accurately respond to the practice items, whereas higher WM learners may not have needed to attend to the pronoun and verb for as long to respond correctly.

To sum up, while Issa's findings indicate that lower WM prompts more attention to the target constructions, Indrarathne and Kormos found a positive relationship between WM and attentional allocation (i.e., higher WM learners allocated more attention to the target forms). The difference in the direction of the relationship may be related to the demands of the tasks completed by the participants. In Issa's study, the instantiation of the structured input practice may be more cognitively demanding for learners, whereas textual enhancement, used in Indrarathne and Kormos (2018), may be less so.

Overall, the findings of these two studies (Indrarathne &Kormos, 2018; Issa, 2019) suggest that WM plays a significant role in not only in learners' attention allocation to the target features but also in online instructional conditions. In particular, WM appears to have a facilitative role in instructional conditions whereby learners' attention is drawn to linguistic form, that is, form-focused instruction, in comparison to instructional conditions whereby the emphasis is on communication, that is, solely meaning-focused instruction (Ellis, 2012). This warrants further investigation into whether WM also has a positive impact on facilitating attention to and development in the knowledge of an L2 grammatical feature in instructional interventions that are designed to manipulate learner attention, such as the provision of CF,

one of the subsets of form-focused instructions, and how the impact of WM may vary under different feedback modality conditions.

2.7.7. Summary

Given that an individual's WM capacity is limited, and its function is to regulate attentional resources while performing complex tasks, it seems logical to assume that individual differences in WM may affect the extent to which learners pay attention to and process input, depending on feedback modality, as a different modality may consume a different amount of cognitive resources. Previous literature indicates that WM abilities may play a determining role in L2 learning through oral and written feedback, including recasts. However, it is yet to be investigated whether WM capacity mediates the effects of multimodal recasts and how it relates to L2 grammatical development in the context of video-conferencing. This warrants further research, examining whether and to what extent learners' WM capacity would affect the efficacy of multimodal recasts on L2 acquisition while performing video-conferencing tasks.

Moreover, the dominance of cognitive-oriented L2 research on feedback has primarily focused on complex WM and PSTM, with less empirical attention given to visuospatial WM. However, some studies suggest that VSTM plays a crucial role when feedback is presented visually (e.g., Sachs, 2010; Zalbidea & Sanz, 2020). Therefore, it appears to be important to include an index of VSTM in studies exploring the impact of feedback that includes visual input.

Last but not least, given the findings from previous research suggesting that each component of WM influences L2 grammar learning to a varying degree in a different domain, more empirical research that considers a range of WM measures tapping into executive,

phonological, and visuospatial components is needed to gain a more detailed understanding of the contributions of WM in conjunction with feedback modality. It is also desirable to administer assessment tests in both modalities (i.e., oral and written), as the involvement of WM resources can vary with the type of assessment, to reach a more fine-grained understanding of the contributions of WM alongside the feedback modality. Also, it appears desirable to administer assessment tests in both modalities (i.e., oral and written) as the involvement of WM resources can vary with the type of assessment (e.g., Révész, 2012).

CHAPTER 3. STUDY 1

The purpose of Study 1 was to investigate the impact of recasts and feedback modality on L2 acquisition in the context of video-based SCMC. To this end, this study involved examining the effects of recasts compared to the absence of recasts on L2 grammatical development and how the modality of recasts influenced their effectiveness during video-conferencing tasks. The study introduced multimodal recasts, presented as captioned oral recasts, which concurrently provided both oral and on-screen written recasts. Another noteworthy aspect was the comparison of various feedback modalities within video-based interactions. The research questions guiding Study 1 were as follows:

1. To what extent do recasts, as compared to no recasts, affect the development in L2 grammatical knowledge during video-conferencing tasks?
2. To what extent does modality of recasts affect the development in L2 grammatical knowledge during video-conferencing tasks?

This chapter comprises four sections: methodology, results, discussion, and limitations and modifications for Study 2. The methodology section elaborates the overall design, participants, target linguistic construction, materials, data collection procedure, and coding and scoring system, as well as statistical analyses used in the study. Subsequently, the results are presented for each research question. The discussion section interprets and discusses the findings in light of research questions as well as previous literature and empirical studies. The chapter finishes with a discussion of the study's limitations and modifications for Study 2, taking into account the limitations identified in Study 1.

3.1. Methodology

3.1.1. Overall Design

The study employed a pretest–immediate posttest–delayed posttest design with three treatment sessions. Sixty participants were randomly assigned to one of four groups: an oral recast group (N = 15), a written recast (N = 15) group, a multimodal recast group (N = 15), and a no recast group (N = 15). All four groups were administered a background questionnaire, a proficiency 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 in this sequence. Detailed descriptions of the research instruments and procedures employed in this study are presented in subsequent sections.

3.1.2. Research Ethics

Prior to initiating data collection, this study received research ethics approval from the Research Ethics Committee at the UCL Institute of Education, University College London. At the commencement of the study, participants were provided with a comprehensive introduction, with an information sheet that outlined: (a) the objective of the study, (b) the overall procedures and duration of the study, (c) their right to withdraw from the research at any point, and (d) the measures in place to ensure the preservation of their privacy and confidentiality (Appendix A1). Each participant's agreement to take part in the study was formally documented through their signature on a consent form (Appendix A2). It was imperative that participants remained unaware of the specific target structure under investigation and the type of recasts they would receive prior to the treatment sessions, as the primary aim of this study was to investigate how recasts and their modality influenced

participants' noticing and knowledge development in the target construction. Therefore, participants were informed that the research aimed to analyze the online communication and interactional patterns of Korean ESL learners during one-on-one multimodal video-conferencing tasks. This approach was adopted to ensure that participants' responses and behaviors remained impartial and authentic, thereby upholding the integrity and validity of the study.

3.1.3. Participants

Of the 60 participants, 48 were female and 12 were male. They were all native Korean speakers learning English as a second language (ESL) in the UK. The mean age was 27.55 (*SD* = 2.33). On average, participants had received 10.20 years of English instruction (*SD* = 97). Their mean length of residence in the UK was 15.12 months (*SD* = 4.03). None of the participants had lived in an English-speaking country prior to their studies in the UK. 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: $F(3, 56) = .33$, $p = .81$, $\eta^2 = .02$; length of residence: $F(3, 56) = .13$, $p = .94$, $\eta^2 = 01$. Participants' proficiency levels fell into the A2–B1 bands according to the Common European Framework for Reference (CEFR), 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.

3.1.4. Target Construction

English *wh*-questions were selected as the linguistic target for several reasons. Firstly, *wh*-questions are physically salient. This makes them likely to be susceptible to recasts, given that recasts have been found to be more effective when addressing errors provided in more salient linguistic constructions (e.g., Mackey, 2006). Also, the selection of both participants and assessment was guided by Pienemann et al.'s (1988) developmental sequence for question formation. According to this sequence, language learners gradually advance from less to more target-like question forms through an established six-stage developmental sequence (Appendix A3). In that the empirical support for the sequence is robust (e.g., Mackey, 1999; Mackey & Philp, 1998; Pienemann & Mackey, 1993), and developmental stages enable researchers to evaluate and control learners' readiness to acquire specific forms (e.g., Mackey, 1999), the progress through the various stages in this sequence has been widely employed to assess L2 question development in numerous instructed SLA studies (e.g., Adams, 2007; Kim, 2012; Kim et al., 2015; Loewen & Nabei, 2007; Mackey, 1999, 2006; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003).

This study primarily focused on examining learners' ability to produce *wh*-questions in stage 3 (fronting of *wh*-word, *do*, or another element without inversion; 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*-questions with an auxiliary verb in the second position; e.g., "how did you lose her?"). To ensure that learners had prior knowledge of producing questions at these stages and to control for their proficiency in generating these question forms, a conservative emergence criterion was employed: learners who demonstrated any evidence of producing two distinctive uses of question types of stages 3, 4,

and 5 in at least two production tasks out of the three assessment tasks during the pretest were recruited for the present study. For example, learners would be considered capable of producing stage 5 questions if they could generate different forms of the same stage 5 question types, such as "Why did you follow me yesterday?", "Where were you standing at the time?". This screening criterion aligns with methodologies used in other SLA studies involving question development (Mackey, 1999; Mackey & Philp, 1998; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003; Kim et al., 2015; Lightbown & Spada, 1997; Spada & Lightbown, 1993, 1999).

3.1.5. Materials

3.1.5.1 General English Proficiency Test

As previously explained, the participants' overall English proficiency was assessed using the Oxford Placement Test, comprising 200 questions: 100 in the listening section and 100 in the grammar section. During the test, participants were required to select the correct answers from the options provided. In accordance with the test scoring guideline, one point was awarded for each correct answer, resulting in a maximum score of 100 for each section.

3.1.5.2. Experimental Treatment Materials

Over the course of the three treatment sessions, all four groups performed three video-conferencing focused tasks which were specifically designed to elicit the use of *wh*-questions (Appendix A4) (see Ellis, 2003 for the distinction between unfocused and focused tasks). These tasks also adhered to the criteria established for tasks as defined by Ellis (2009) and Ellis and Shintani (2014), which included: (a) a primary focus on meaning, (b) the presence of some kind of gap (e.g., missing information or the need for opinion), (c) the primary

utilization of participants' linguistic and non-linguistic resources, and (d) the requirement of a communicative outcome. Each task version was situated in a hypothetical scenario, where participants assumed different "interviewing" roles during video-conferencing interrogations or interviews. These roles included that of an FBI agent, a police officer, and an animal shelter volunteer, with the researcher taking on the role of interviewee, portraying a witness to a crime, an incident, or the owner of a lost dog. In the beginning of each treatment task, participants were presented with either an incident report form (treatment task 1 and 2) or a lost dog poster (treatment task 3) (Figure 7). Their responsibility was to ask their interviewees questions which were crucial for completing the report form or the lost dog poster they were provided.

Subsequently, participants were presented with a set of questions in the form of prompts (Figure 8). These prompts were meticulously designed to create specific 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 consisted exclusively of high-frequency words selected from the New General Service List (Brezina & Gablasova, 2015). The tasks underwent pilot testing with a population similar to the study's participants, and the results confirmed that the tasks succeeded in eliciting the targeted *wh*-questions.



Figure 7. An Example of a Lost Dog Poster

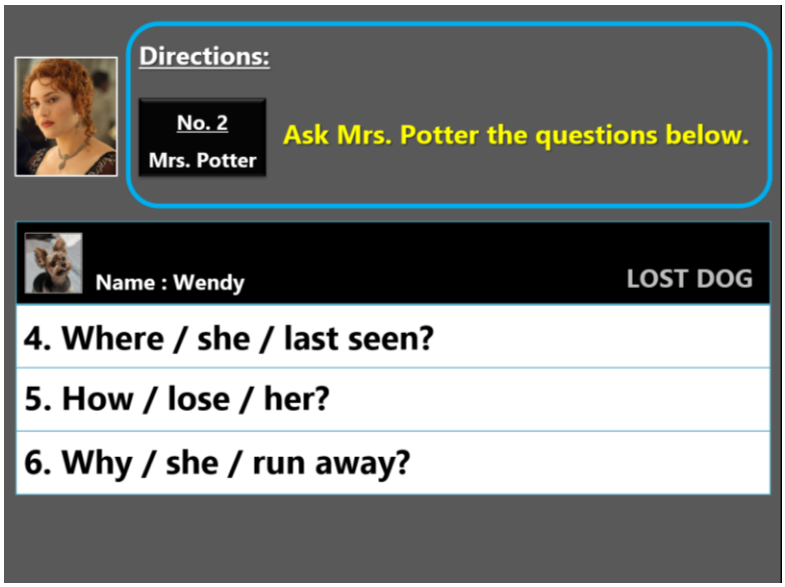


Figure 8. An Example of Treatment Tasks Serving as Prompts

During the treatment tasks, the three experimental groups received recasts from the researcher for their errors in using *wh*-questions. The recasts were implemented as full interrogative isolated recasts (Lyster, 1998; Sheen, 2006), which reformulated only the

incorrect parts of the learners' production without adding new information. An example is shown below:

Learner: Why she run away?

Researcher: Why did she run away?

This uniform recast type was consistently applied across all experimental groups. The oral recast group exclusively received oral recasts, while the written recast group was only provided written recasts. The multimodal recast group were presented with both oral and written recasts simultaneously. The no recast group, which served as a control group, did not receive any feedback. The written recasts for the multimodal recast and written recast groups were displayed as real-time, on-screen text using Open Broadcaster software, as depicted in Figure 9. These caption-type recasts appeared at the bottom of the computer screen for 10 seconds immediately following participants' errors in using *wh*-questions. The duration of the written recasts was established through piloting procedures and informed by participant feedback. Prior to the experiment, corresponding written recasts to each target question had been prepared and were presented via OBS immediately after a participant's erroneous utterance. This approach enabled the provision of real-time written recasts to the written recast group and simultaneous oral and written recasts to the multimodal group during their communicative video-based interactions. It is important to emphasize that these pre-prepared recasts differed from the pre-programmed recasts typically used in CALL tools, which usually involve interactions between learners and computers within a computerized language learning practice. In this study, on the other hand, recasts were presented during genuine exchanges between learners and the researcher. All treatment sessions were administered and

recorded using Zoom, and the recorded sessions were subsequently transcribed for further analysis.

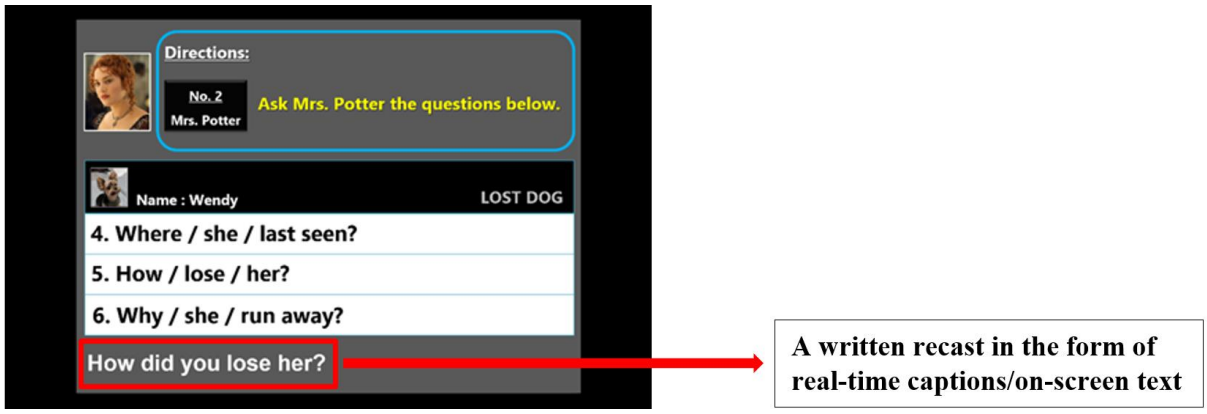


Figure 9. An Example of a Written Recast

3.1.5.3. Assessment Tasks

To gauge learners' development in the use of the target construction, three assessment tasks were employed: a controlled OPT, a controlled WPT, and an SOPT (Appendix A5–A7). For each test, three different yet comparable versions were created, which were counterbalanced across three testing sessions. All assessments were conducted and recorded via Zoom. Similar to the treatment tasks, these tests included only 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 and WPT were designed to evaluate participants' ability to use *wh*-questions in a controlled context. The OPT, akin to the treatment tasks, required participants to act as an interviewer investigating incidents, such as a police officer conducting interviews with witnesses regarding a bank robbery. Each test version was designed to elicit 18 target *wh*-questions and included 12 distractors, resulting in a total of 30 items. Nine types of *wh*-question were included, with two items for each sub-type: subject *wh*-extraction (who, what),

object *wh*-extraction (who, what), adjacent *wh*-extraction (where, when, why, how/degree, how/manner). The distractors involved the formation of four types of yes–no questions, with three items per type: be + bare infinitive, do-fronting, can + bare infinitive, have/has + past participle. Except for modality, the WPT shared the same format as the OPT. However, in the WPT, a single item targeted each type of *wh*-question, totalling nine items. The number of distractors remained consistent, resulting in a test containing 21 items. The internal consistency reliability for the three version of the OPT (version A: α =.894, version B: α =.813, version C: α =.838) and WPT (version A: α =.825, version B: α =.838, version C: α =.878) fell within the good range.

The SOPT was designed to elicit the use of *wh*-questions in a more spontaneous context, aiming to ascertain whether participants could transfer acquired knowledge to a less controlled environment. As part of the SOPT, learners were provided with a report form (e.g., an accident report or a passenger report), with 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 playing various roles, such as a police officer interviewing a driver involved in a car accident or a customs officer questioning a passenger under suspicion. Given the less specific prompts, this task required learners to rely more on their own resources than the OPT.

Incorporating varied assessments aimed to evaluate different aspects of knowledge. The OPT imposed greater time pressure compared to the WPT, thereby facilitating the evaluation of learners' procedural knowledge and the measurement of the process of automatization (e.g., Dekeyser, 2007). The written production test, on the other hand, allowed more time for reflection, enabling the assessment of participants' declarative knowledge and grammatical accuracy (McDonough & Mackey, 2006), along with their procedural knowledge. Moreover, the utilization of various assessments in different modalities enabled

an investigation into whether the assessment task modality had any bearing on the relationship between the feedback modality and the learning gains of the target constructions (e.g., Révész, 2012). Also, different test modalities aimed to ascertain whether participants could transfer acquired gains from the varied recast modes to different modalities. Furthermore, the SOPT provided a measure of learners' ability to use target constructions in a more authentic scenario, enhancing ecological validity. Examples for each assessment task are provided in Figure 10.



Figure 10. Examples of Assessment Tasks (SOPT, OPT, WPT in this order).

3.1.5.4. Background Questionnaire

A background questionnaire was administered to obtain general information about the participants. It covered various aspects such as gender, age, duration of English study, prior experiences living in English-speaking countries or outside of South Korea. Additionally, participants provided details on their social networking sites and SCMC usage habits, including average time spent, preferred platforms, and communication types (e.g., text-based, audio-based, or video-based). They also shared previous experience with learning languages through these methods and their level of comfort in using these tools. Given the online nature of the experiment using the SCMC platform (i.e., Zoom), this questionnaire aimed to ensure that participants were comparably familiar with communication in SCMC modes. (Appendix

A8). To assist participants in providing their responses, the questionnaire was administered in Korean, their native language.

3.1.5.5. Exit Questionnaire

After completing the experiment, participants completed an exit questionnaire (Appendix A9). This questionnaire aimed to collect information on: (a) their understanding of the purpose of the study, (b) whether they noticed and paid attention to any particular linguistic features or grammar rules, (c) whether they referred to external sources outside the study, and (d) their perception of task difficulty. Additionally, open-ended questions sought their opinions regarding the effectiveness of recasts and the influence of modality (for experimental groups), as well as their overall experience during the video-conferencing treatment and assessment tasks. To ensure clearer responses, the exit questionnaire was also conducted in Korean.

3.1.6. Data Collection Procedure

The total duration of data collection was approximately four months, with each session conducted individually. Each participant took part 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 sequence. In Session 2, participants completed the first and second treatment tasks. Session 3 included the third treatment task, immediately followed by the posttest. Session 4 occurred a week later, with participants completing a delayed posttest and filling in an exit questionnaire. In each testing session, participants first completed the SOPT (14–16 minutes), followed by the OPT and WPT (13–15 minutes each). The decision to administer the WPT last was driven by the

intention to mitigate any potential priming effect. This precaution aimed to prevent participants from being influenced to depend excessively on their explicit, declarative knowledge in the subsequent oral tests. Likewise, the decision to administer the OPT prior to the SOPT took into account the controlled nature of the OPT, as the OPT constituted a more structured oral production test featuring more specific prompts compared to the SOPT. Each treatment task lasted approximately 14–16 minutes. There was no time limit for the treatment or assessment tasks. All sessions were administered by the researcher, which were conducted and recorded by Zoom for subsequent analyses. Figure 11 shows the data collection procedures for Study 1.

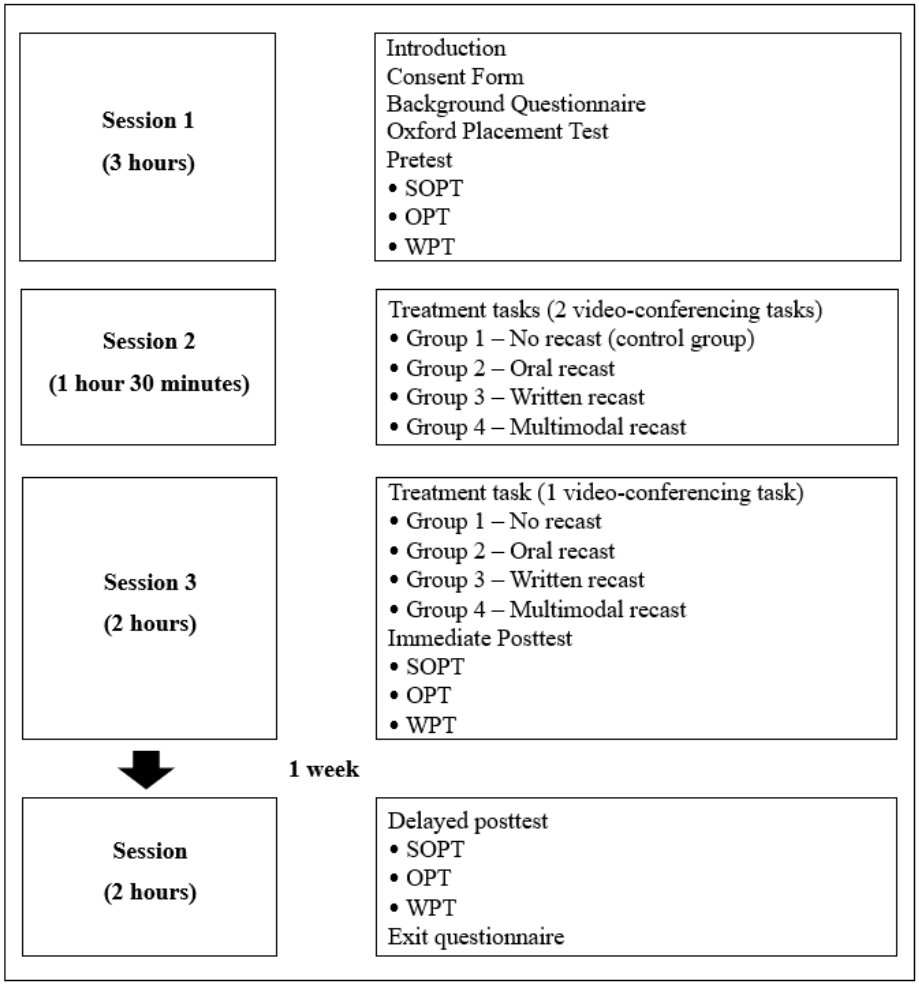


Figure 11. Data Collection Procedures for Study 1

3.1.7. Coding and Scoring

As previously noted, participants' *wh*-question development was assessed and coded using a categorical scoring system derived from Pienemann et al.'s (1988) question development scale. This scale typically comprises six stages, each corresponding to a specific developmental phase, where higher stages (e.g., Stage 5) indicate more advanced learners compared to lower stages (e.g., Stage 3). Based on this scale, the scoring system of the present study awarded a score of 3, 4, or 5 points to the participants' production of target items in both the OPT and WPT, reflecting the participants' differing developmental stages. For example, if a learner's response was coded as Stage 3, it received 3 points (Table 8). Total scores were calculated by considering only non-subject questions (cf., Li & Iwashita, 2019, 2021), with a maximum score of 70 for the OPT (14 items) and 35 points for the WPT (7 items).

As the SOPT did not have a fixed number of obligatory contexts for *wh*-questions, it underwent coding and scoring through the following three steps. Firstly, the total number of obligatory contexts for *wh*-questions produced by each learner was determined. Secondly, as with the OPT and WPT, each question form was awarded a score of 3–5 based on its developmental stage. Finally, the total score for each participant was calculated by dividing the sum of raw scores by the maximum possible score (i.e., the number of obligatory contexts × 5). In all tests, pronunciation and spelling errors were disregarded, as the primary focus was on development in question formation. Each test (pretest, posttest, delayed posttest) had three versions, which were counterbalanced across participants, modalities, and testing sessions to mitigate the test order effect. The assessment tasks were coded and scored by the researcher and a second coder, with the second coder scoring 25 per cent of the three outcome measures. Inter-coder agreement was determined by calculating Cohen's kappa. The inter-coder

agreement for the OPT was.98 with a kappa of.98, and 1.00 for the WPT with a kappa of 1.00 and.98, with a kappa of.97 for the SOPT, indicating a high level of agreement between the two raters.

Table 8. Scoring System for *Wh*-Question Formation

Learners Production	Coded Stages in Pienemann et al.'s (1988) scale	Points Awarded
"What you ate?"	Stage 3 (fronting of *wh*-word)	3
"Where were you?"	Stage 4 (inversion in *wh*-questions with a copula)	4
"How did you lose her?"	Stage 5 (inversion in all *wh*-questions with an auxiliary verb in the second position)	5

3.1.8. Statistical Analyses

In Study 1, the preliminary statistical analyses were performed using IBM SPSS 27.0 for Windows. The aim of these analyses was to assess the reliability of the tests, compute descriptive statistics and carry out one-way ANOVAs to examine whether the four groups were comparable in terms of the number of obligatory contexts elicited and number of recasts provided during the treatment tasks. Cronbach's alpha was employed to determine the internal consistency reliability of the different tests, while Cohen's kappa was used to assess inter-coder reliability.

The primary statistical analyses involved constructing a series of logistic mixed-effects models using the *clmm* function from the ordinal package and linear mixed-effects models (*lmer* functions) from the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R Version 4.2.1 (R Core Team, 2022), a statistical software. To assess the advancement in *wh*-question stage increase on the OPT and WPT, a series of ordinal logistic mixed-effects models were constructed, given the ordinal nature of the dependent variables. In each

analysis, the fixed effects were Group, Time, and their interactions, while participant and item were included as the random effects. The dependent variables were participants' pretest, posttest, and delayed posttest scores, with the predictors of interest being the Group-by-Time interactions. For the SOPT data, a series of linear mixed-effects models were formulated based on the continuous nature of the dependent variables. Again, the fixed effects were Group, Time and their interactions, and the dependent variables were participants' pretest, posttest, and delayed posttest scores, with the predictors of interest being the Group–Time interactions. It should be noted that while participants were presented with the same set of target items, they produced *wh*-questions for different items, leading to variations in the items addressed. Consequently, the variable "item" was not included as a random effect in the models.

To estimate the magnitude of the effects, η^2 values were computed for ANOVAs, odds ratios (*ORs*) were determined for the logistic mixed-effects regression models, and R^2 values were calculated for the linear mixed-effects model analyses. In logistic regression models, an *OR* larger than one indicated a positive relationship and an *OR* less than one indicated a negative relationship. *ORs* greater than 3 or less than 0.33 were considered strong (Haddock, Rindskopf, & Shadish, 1998). The LmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) was employed to acquired *p*-values for linear mixed-effects models. Marginal and conditional R^2 values, which measure the variance explained by fixed and random effects, were computed using the r.squared generalized linear mixed models (GLMM) function in the MuMin package (Barton, 2020). R^2 values of.06,.16, and.36 were interpreted as small, medium, and large, respectively (Plonsky & Oswald, 2014).

Assumption tests for logistic mixed-effects models and linear mixed-effects models involved a range of diagnostic procedures. Collinearity was evaluated by inspecting the

correlation matrix or variance inflation factors for predictor variables. The normal distribution assumption was scrutinized through diagnostic plots specific to the type of residuals, such as Q-Q plots or histograms, and statistical tests like the Shapiro-Wilk test or Kolmogorov-Smirnov test. Homoscedasticity, indicating constant variance of residuals, was assessed using residual plots and statistical tests like Levene's test or the Breusch-Pagan test. The sjPlot package (v 2.8.4; Lüdecke, 2020) was employed to conduct these assumption tests in logistic mixed-effects models. These tests provided insights into the validity of the model assumptions, ensuring the robustness of the statistical analyses.

To detect potential outliers in the data, various outlier detection procedures were implemented. Residual analysis, which included plotting the Pearson residuals against predicted values in linear mixed-effects models, was undertaken to assess the presence of outliers. In logistic mixed-effects models, deviance residuals or standardized residuals were examined. Data points deviating significantly from the expected pattern or exhibiting unusual patterns were flagged as potential outliers. Cook's distance, a measure of influence, was computed to identify potentially influential outliers. Furthermore, a model criticism approach using the romr.fnc() function from the LMERConvenienceFunctions package (Tremblay & Ransijn, 2020) was utilized to pinpoint outliers. Data points with absolute standardized residuals exceeding 2.5 standard deviations ($|z| > 2.5$) were categorized as outliers. Sensitivity analyses were performed by re-running the models with and without the identified outliers to assess their potential impact on the statistical significance of fixed effects. Outliers that significantly influenced the fixed effects were excluded from the analyses. These procedures for checking outliers ensured the validity and reliability of the statistical models, providing confidence in the interpretation of the results.

3.2. Analyses and Results

3.2.1. Preliminary Results

3.2.1.1. Comparability of the Group

Table 9 presents the descriptive statistics pertaining to the number of obligatory contexts that the treatment tasks created for the linguistic target during the treatment and the number of recasts participants received in response to their errors in *wh*-question formation. A one-way ANOVA indicated 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 the absence of a significant difference in the number of recasts provided to the experimental groups, $F(2, 42) = 2.56$, $p = .091$, $\eta^2 = .11$.

Table 9. Descriptive Statistics for Obligatory Contexts and Number of Recasts per Group

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

Note. Control = Control group; Oral = Oral recast group; Written = Written recast group; Multi = Multimodal recast group.

Table 10 provides descriptive statistics for the pretest scores. As depicted in Tables 11–13, a series of ordinal logistic mixed-effects models (OPT, WPT) and linear mixed-effects model (SOPT) analyses revealed no significant difference in the performance of the four groups across the three assessment tasks at the pretest.

Table 10. Descriptive Statistics for the Pretest

					95% CI	
	Group	*N*	*M*	*SD*	*Lower*	*Upper*
OPT	Control	15	57.80	1.32	57.07	58.53
	Oral	15	57.80	1.57	56.93	58.67
	Written	15	57.47	1.18	56.72	58.22
	Multi	15	57.67	1.18	57.02	58.32
WPT	Control	15	26.40	1.30	25.68	27.12
	Oral	15	26.73	1.28	26.02	27.44
	Written	15	26.53	1.41	25.75	27.31
	Multimodal	15	26.93	1.28	26.22	27.64
			M (%)	*SD (%)*	*Lower*	*Upper*
SOPT	Control	15	73.05	1.51	72.21	73.88
	Oral	15	72.85	1.20	72.18	73.51
	Written	15	72.16	1.50	71.33	72.99
	Multi	15	72.44	0.96	71.91	72.98

Note. OPT maximum score = 18; WPT maximum score = 18; SOPT maximum score = 100%.

Table 11. Results for the Ordinal Logistic Mixed-effects Model Examining Performance on the OPT Pretest

	Fixed effects					Random effects			
						by participant		*by item*	
	Estimate	*SE*	*z*	*p*	*OR*	*variance*	*SD*	*variance*	*SD*
Control group as reference level									
3|4	–1.01	0.35	–2.92	.004**	0.36	<.01	<.01	1.48	1.21
4|5	0.47	0.34	1.35	.176	1.59				
Oral	0.02	0.19	0.09	.931	1.02				
Written	–0.06	0.19	–0.30	.761	0.94				
Multi	–0.03	0.19	–0.18	.856	0.97				
Oral group as reference level									
3|4	–1.03	0.35	–2.97	.003**	0.36	<.01	<.01	1.48	1.21
4|5	0.45	0.34	1.31	.192	1.57				
Control	–0.02	0.19	–0.09	.931	0.98				
Written	–0.07	0.19	–0.39	.670	0.93				
Multi	–0.05	0.19	−0.27	.790	0.95				
Written group as reference level									
3|4	–0.95	0.35	–2.75	.006**	0.39	<.01	<.01	1.48	1.21
4|5	0.52	0.34	1.52	.129	1.69				
Control	0.06	0.19	0.30	.761	1.06				
Oral	0.07	0.19	0.39	.670	1.08				
Multi	0.02	0.19	0.12	.901	1.02				
Formula: clmm(OPT_Pretest ~ 1 + Group + (1|Participant) + (Group|Item))									

Table 12. Results for the Ordinal Logistic Mixed-effects Model Examining Performance on the WPT Pretest

	Fixed effects					Random effects			
						by participant		*by item*	
	Estimate	*SE*	*z*	*p*	*OR*	*variance*	*SD*	*variance*	*SD*
Control group as reference level									
3|4	–0.35	0.42	–0.83	.407	0.70	<.01	<.01	.96	.98
4|5	1.39	0.43	3.25	.001**	4.01				
Oral	0.14	0.27	0.51	.607	1.15				
Written	0.05	0.27	0.18	.902	1.04				
Multi	0.27	0.27	0.84	.399	1.25				
Oral group as reference level									
3|4	–0.49	0.42	–1.16	.245	0.61	<.01	<.01	.96	.98
4|5	1.25	0.42	2.94	.003**	3.49				
Control	–0.14	0.27	−0.51	.607	0.87				
Written	–0.09	0.27	–0.34	.802	0.91				
Multi	0.09	0.26	0.33	.900	1.09				
Written group as reference level									
3|4	–0.40	0.42	–0.94	.345	0.67	<.01	<.01	.96	.98
4|5	1.34	0.46	3.15	.002**	3.83				
Control	–0.05	0.27	−0.18	.902	0.95				
Oral	0.09	0.27	0.34	.802	1.09				
Multi	0.18	0.27	0.67	.901	1.20				
Formula: clmm(WPT_Pretest ~ 1 + Group + (1|Participant) + (Group|Item))									

Table 13. Results for Linear Mixed-effects Models Examining Performance on SOPT Pretest

	Fixed effects						Random effect	
							By participant	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*Variance*	*SD*
Control group as reference level								
(Intercept)	0.73	0.00	223.12	<.001***	0.09	0.09	<.01	<.01
Oral	0.00	0.00	0.87	.388				
Written	0.01	0.00	1.45	.153				
Multi	–0.00	0.00	–0.87	.388				
Oral group as reference level								
(Intercept)	0.73	0.00	224.35	<.001***	0.09	0.09	<.01	<.01
Control	–0.00	0.00	–0.87	.388				
Written	0.00	0.00	0.58	.566				
Multi	–0.01	0.00	–1.74	.				
Written group as reference level								
(Intercept)	0.73	0.00	225.17	<.001***	0.09	0.09	<.01	<.01
Control	–0.01	0.00	–1.45	.153				
Oral	–0.00	0.00	–0.58	.566				
Multi	–0.01	0.00	–0.88	.389				
Formula: lmer(SOPT_Pretest ~ 1 + Group + (1|Participant))								

3.2.2. Main Analyses

3.2.2.1. OPT

Table 14 summarizes the descriptive statistics for the participants' pretest, posttest, and delayed posttest scores on the OPT. The ordinal logistic mixed-effects regressions carried out to examine the effect of Group and Time on participants' performance on the test identified eight significant Time-by-Group interactions (see Appendix A10 for full model results). As shown in Table 15, all recast groups showed greater pretest–posttest gains than the control group. As reflected in odds ratios, the multimodal recast group achieved the largest gains as compared to the control group (*OR* = 6.33), followed by the written recast (*OR* = 3.43) and oral recast groups (*OR* = 1.78) in that order. The multimodal recast group retained its advantage over the control group on the delayed posttest with a large effect size (*OR* = 2.21), but no significant difference in pretest–delayed posttest gains were found between the control and unimodal (i.e., oral and written) groups.

Moving on to the modality differences, the multimodal recast group showed greater pretest–posttest gains than the oral (*OR* = 3.55) and written recast groups (*OR* = 2.19), and the written recast group outperformed the oral recast group (*OR* = 1.93). On the delayed posttest, however, gain differences were only detected between the multimodal recast and oral recast groups (*OR* = 2.02). In general, the effect sizes were larger for pretest–posttest than pretest–delayed posttest gains. Overall, these results indicate that recasts provided during video-conferencing had a positive impact on participants' gains on the oral production test, with multimodal recasts yielding the greatest advantage followed by written recasts and oral recasts in that order.

Table 14. Descriptive Statistics for the OPT Scores

Group	Test	*M*	*SD*	95% CI *Lower*	95% CI *Upper*
Oral (*N* = 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 (*N* = 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
Multi (*N* = 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 (*N* = 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 15. Results of Post Hoc Contrasts for Significant Interactions Identified by Ordinal Logistic Mixed−Effects Models Examining Participants' Scores on the OPT

	Fixed effects						Random effects			
							by participant (intercept)		by item (intercept)	
	Estimate	*SE*	*z*	*p*	*OR*	95% CI	*Variance*	*SD*	*Variance*	*SD*
ControlOral:Time2	0.58	0.28	2.04	.041	1.78	[1.02, 3.10]	<.01	<.01	1.86	1.37
ControlWritten:Time2	1.23	0.29	4.25	<.001	3.43	[1.94, 6.07]	<.01	<.01	1.73	1.32
ControlMulti:Time2	1.85	0.30	6.15	<.001	6.33	[3.52, 11.40]	<.01	<.01	1.42	1.19
ControlMulti:Time3	0.79	0.28	2.81	.005	2.21	[1.27, 3.84]	<.01	<.01	1.64	1.28
OralWritten:Time2	0.66	0.29	2.23	.026	1.93	[1.02, 3.07]	<.01	<.01	1.54	1.24
OralMulti:Time2	1.27	0.30	4.18	<.001	3.55	[2.18, 6.87]	<.01	<.01	1.25	1.12
OralMulti:Time3	0.70	0.28	2.49	.013	2.02	[1.16, 3.52]	<.01	<.01	1.68	1.30
WrittenMulti:Time2	0.61	0.31	1.98	.048	2.10	[1.22, 3.91]	<.01	<.01	1.25	1.12

Note. *OR* = Odds Ratio; Time1 = pretest (reference level); Time2 = immediate posttest; Time3 = delayed posttest.

3.2.2.2. WPT

Table 16 presents the descriptive statistics for the participants' pretest, posttest, and delayed posttest scores on the written production test. As summarized in Table 17, the logistic mixed-effects regression analyses conducted to examine the impact of Group and Time on participants' test performance yielded nine significant interactions between Group and Time (see Appendix 11 for full model results). As for the OPT, all recast groups displayed larger pretest–posttest gains than the control group. As reflected in the odds ratio, compared to the control group, the greatest pretest–posttest gains were achieved by the multimodal group (*OR* = 11.17), followed by the written recast group (*OR* = 4.21) and oral recast group (*OR* = 2.46) in that order. Moreover, the multimodal recast and written recast groups also maintained their advantages over the control group on the delayed posttest. The odds ratios indicated that participants in the multimodal recast group were 5.06 times more likely to experience development in the delayed gains, while those in the written recast group were 2.17 times more likely to exhibit greater progress than the control group in the WPT. When examining the effects of modality, the multimodal recast group outperformed the oral recast and written recast groups, showing greater gains from the pretest to the posttest and the pretest to the delayed posttest. The odds of experiencing improvements in WPT gains were 4.54 times higher for the multimodal recast group than the oral recast group and 2.65 times higher than the written recast group for the pretest–posttest gains. These patterns persisted for the pretest–delayed posttest gains, with the multimodal recast group being 3.02 times more likely to improve than the oral recast group, and 2.33 times more likely than the written recast group. However, no difference was found between the gains of the oral and written recast groups. Taken together, the presence of recasts, regardless of modality, promoted participants' development in *wh*-question formation, as reflected in their performance on the

WPT. Multimodal recasts, in particular, proved to be more effective in engendering gains than their oral and written counterparts.

Table 16. Descriptive Statistics for the WPT Scores

Group	Test	*M*	*SD*	95% CI *Lower*	95% CI *Upper*
Oral (*N* = 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 (*N* = 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 (*N* = 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 (*N* = 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 17. Results of Post Hoc Contrasts for Significant Interactions Identified by Ordinal Logistic Mixed-effects Models Examining Participants' Scores on the WPT

	Fixed effects						Random effects			
							by participant (intercept)		*by item* (intercept)	
	Estimate	*SE*	*z*	*p*	*OR*	95% CI	*Variance*	*SD*	*Variance*	*SD*
ControlOral:Time2	0.90	0.38	2.38	.017	2.46	[1.18, 5.14]	<.01	<.01	.78	.88
ControlWritten:Time2	1.44	0.39	3.72	<.001	4.21	[1.98, 8.95]	<.01	<.01	1.01	1.01
ControlMulti:Time2	2.41	0.41	5.82	<.001	11.17	[4.98, 25.08]	<.01	<.01	.90	.95
ControlWritten:Time3	0.78	0.38	2.06	.039	2.17	[1.04, 4.53]	<.01	<.01	.99	.99
ControlMulti:Time3	1.62	0.39	4.20	<.001	5.06	[2.38, 10.75]	<.01	<.01	.69	.83
OralMulti:Time2	1.51	0.42	3.64	<.001	4.54	[2.02, 10.21]	<.01	<.01	.41	.64
WrittenMulti:Time2	0.98	0.42	2.31	.002	2.65	[1.16, 6.08]	<.01	<.01	.51	.72
OralMultimodal:Time3	1.11	0.39	2.87	.004	3.02	[1.42, 6.40]	<.01	<.01	.41	.64
WrittenMulti:Time3	0.85	0.39	2.19	.029	2.33	[1.09, 4.95]	<.01	<.01	.51	.72

3.2.2.3. SOPT

Table 18 provides the descriptive statistics for the pretest, posttest, and delayed posttest scores on the SOPT. The linear mixed-effects model analyses, which were carried out to examine the effect of Group and Time on participants' pretest–posttest and pretest–delayed posttest changes, detected ten significant Group-by-Time interactions (see Appendix A12 for full model results). As presented in Table 19, all three recast groups achieved greater pretest–posttest and pretest–delayed posttest gains than the control group, exhibiting large effect sizes. Parallel to the results obtained for the other assessments, the multimodal recast group also achieved greater pretest–posttest and pretest–delayed posttest gains than the oral and written recast groups, with the difference in gains being large. No significant difference was found between the pretest–posttest and pretest–delayed posttest gains of the oral recast and written recast groups. In summary, the results of the SOPT, similar to those of the OPT and WPT, demonstrated an advantage for recasts, with multimodal recasts generating the greatest gains.

Table 18. Descriptive Statistics for the SOPT

Group	Test	*M (%)*	*SD (%)*	95% CI *Lower*	*Upper*
Oral (*N* = 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 (*N* = 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 (*N* = 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 (*N* = 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 19. 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*	*t*	*p*	*R2m*	*R2c*	*by participant variance*	*SD*
ControlOral:Time2	0.04	0.01	5.72	<.001	.81	.92	<.01	<.01
ControlWritten:Time 2	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:Time 3	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	0.06	0.01	9.35	<.001	.93	.97	<.01	.01
WrittenMulti:Time2	0.05	0.01	8.58	<.001	.90	.93	<.01	<.01
OralMulti:Time3	0.05	0.01	8.47	<.001	.88	.93	<.01	<.01
WrittenMulti:Time3	0.05	0.01	8.47	<.001	.85	.90	<.01	<.01

*Formula: lmer(SOPT ~ 1 + Group*Time + (1|Participant))*

3.2.3. Summary of Results

RQ1. To what extent do recasts affect development in L2 grammatical knowledge during video-conferencing tasks?

The presence of recasts, in contrast to their absence, had a beneficial influence on the development of *wh*-question formation. In all three outcome measures (OPT, WPT, SOPT), participants who received recasts exhibited more significant improvements in pretest–posttest gains, showing large effect sizes. Concerning pretest–delayed posttest gains, the results varied across the experimental groups. While all experimental groups outperformed the control group on the SOPT, only the multimodal recast group surpassed the control group on the OPT. On the WPT, both the multimodal recast group and the written recast group outperformed the control group, also with large effect sizes.

RQ2. To what extent does modality of recasts affect development in L2 grammatical knowledge during video-conferencing tasks?

The results pointed strongly to the greater effectiveness of multimodal recasts compared to both oral and written recasts in promoting participants' knowledge development in using *wh*-questions. Across all three assessments, the group exposed to multimodal recasts consistently outperformed those who received unimodal recasts in terms of their progress from pretest to posttest and pretest to delayed posttest, with effect sizes ranging from small to large. The only exception was the lack of a significant difference in gains from pretest to delayed posttest between the multimodal recast and the written recast groups in the OPT. The differences in gains between the oral recast and the written recast groups were significant only in the pretest–posttest improvements in the OPT, showing a large effect. In summary,

the data unequivocally established the superiority of multimodal recasts over unimodal recasts in facilitating the development of learners' competence to form *wh*-questions.

3.3. Interim Discussion

3.3.1. Effects of Recasts on L2 Grammatical Development (RQ1)

The first research question aimed to explore the extent to which recasts, irrespective of modality, affect the development of L2 grammatical knowledge in the context of video-conferencing tasks. Specifically, the progress of learners in forming *wh*-questions was operationalized as their stage advancement through Pienemann et al.'s (1988) development sequence for question formation. The results revealed that the provision of recasts significantly facilitated participants' *wh*-question stage advancement in all three assessments when compared to not receiving recasts. These findings align with the theoretical claim made in favor of recasts (e.g., Doughty, 2001; Gass & Mackey, 2006; Goo & Mackey, 2013; Li, 2013; Long, 2007; Mackey, 1999; Mackey & Goo, 2007; Mackey & Philp, 1998). The results also suggest that the positive outcomes observed in prior research regarding the effectiveness of recasts in promoting L2 grammatical development, particularly in question formation (e.g., Ellis & Sheen, 2006; Li & Iwashita, 2019, 2021; Loewen & Nabei, 2007; Mackey et al., 2002; Mackey & Philp, 1998; Mackey & Sachs, 2012; McDonough, 2005; McDonough & Mackey, 2006; Philp, 2003; Sheen, 2004; Zhao & Wang, 2016) can also be extended to recasts provided during video-conferencing tasks. Moreover, these findings corroborate previous studies that have demonstrated the benefits of recasts for L2 grammar learning in

text-based (e.g., Baralt, 2013; Kourtail & Révész, 2020; Yilmaz, 2012; Yilmaz & Yuksel, 2011) and in video-based SCMC (Monteiro, 2014; Rassaei, 2017, 2019).

It is important to note that the positive impact of recasts was observed not only in the posttest but also in the delayed posttest. The experimental groups outperformed the control group in both the posttest and delayed posttest in the SOPT. Additionally, the written recast group exhibited greater improvements from the pretest to the posttest and from the pretest to the delayed posttest compared to the control group in the WPT. Furthermore, the multimodal recast group displayed greater pretest–posttest and pretest–delayed posttest gains in all assessments compared to the control group. These findings highlight the pedagogical value of incorporating recasts into L2 learning environments within video-conferencing. Moreover, the enduring benefits of recasts imply their role in maintaining and retaining the acquired grammatical knowledge, although the extent of sustained benefits may differ depending on different recast modalities. This underscores the significance of further exploring the influence of feedback modality on L2 grammatical development in video-based SCMC.

It is also worthwhile discussing how recasts in the present study affected the development of different types of L2 knowledge. That is, multimodal recasts led to greater gains in both the oral and written production than were observed in the control group. Also, the oral recast and written recast groups outperformed the control group in both types of production tests. These results imply that learners were able to transfer the L2 knowledge they had acquired to different modalities. According to Skill Acquisition Theory (e.g., DeKeyser, 2007), transferring procedural knowledge across modalities is a challenging process, and successful transfer between skills (e.g., from written to oral performance) often relies on declarative knowledge of the target construction. These results are also in line with previous empirical research (Ellis et al., 2006; Révész, 2012) showing that recasts can

facilitate gains in both declarative and procedural knowledge. While this has been established for oral recasts in several studies, the present study is among the first to provide evidence for the capacity of multimodal recasts and written recasts to promote both declarative and procedural knowledge (see, however, Baralt, 2013; Yilmaz & Yuksel, 2011). To further validate the transferability of learning gains across various modalities and assessment methods, additional research is needed that incorporates a diverse range of SCMC modalities and outcome measures using different modes.

3.3.2. Effects of Feedback Modality on L2 Grammatical Development (RQ2)

The second research question was concerned with whether the modality of recasts – oral, written or multimodal – affects their effectiveness. The results indicated that the multimodal recast group achieved the highest pretest–posttest and pretest–delayed posttest gains on all tests, followed by the written recast and oral recast groups in that order. The only exception was a lack of significant gain difference between the multimodal recast and written recast groups in pretest–delayed posttest gains on the OPT. The differences in gain between the oral recast and the written recast groups were significant only in the pretest–posttest improvements in the OPT, showing a large effect size. Various factors, individually or in combination, may have contributed to the success of multimodal recasts. First, the multimodal presentation of recasts could have decreased the cognitive load imposed by each modality (Sweller, 1988, 1994), facilitating improved input processing and successful learning of the target constructions. According to Sweller's (Sweller, 1988, 2005; Chandler & Sweller, 1991) Cognitive Load Theory, the combination of auditory and visual information diminishes cognitive load on WM compared to the use of a single channel (Sweller, 1988, 1994). Also, the concurrent presentation of pertinent phonological and orthographic forms

might have enabled learners to create more robust mental representations of the linguistic features (Perfetti & Hart, 2002) targeted by the feedback.

These predicted advantages of multimodal recasts also align with the Dual Coding Theory (Paivio, 1986, 2007) and the Cognitive Theory of Multimedia Learning (Mayer, 2005a, 2014). These theories posit that when aural and visual input complement each other, learning is enhanced, providing multiple input sources for comprehension and reducing the burden of relying on a single source. Processing information in two channels activates both systems, facilitating a mental connection between auditory and visual channels, potentially leading to better learning and recall. Considering these potential benefits of multimodal input, multimodal recasts likely not only balanced cognitive load but also primed both channels, resulting in superior learning outcomes and retention compared to unimodal recasts.

The hybrid nature of multimodal recasts may also have contributed to the greater learning gains observed in the multimodal recast group. As was previously mentioned, multimodal recasts possess the advantages of both oral and written CF, such as longer processing time, visual/perceptual salience, contingency, and the re-readability of input (e.g., Doughty, 2001; Gass, 2003; Lai et al., 2008; Leeman, 2003; Long, 2007; Payne & Whitney, 2002; Sauro, 2009; Saxton, 1997; Smith, 2004, 2010; Yilmaz, 2012). These factors collectively heightened the noticeability of multimodal recasts, rendering their corrective force more explicit to learners. This appears particularly beneficial when targeting grammatical features, which tend to be less noticeable than semantic, syntactic, and phonological recasts (e.g., Lyster et al.,, 2013; Mackey et al., 2000; Kourtali & Borges, 2023; Smith, 2010), resulting in varying degrees of noticeability of grammatical constructions (e.g., Lyster et al., 2013).

In particular, the contingency of recasts, where target-like alternatives are immediately presented, is a crucial factor contributing to their salience and noticeability (Arroyo & Yilmaz, 2018; Lai et al., 2008; Long, 2007). In multimodal recasts, this contingency extended to both oral and written recasts, as delivered in the form of captioned oral recasts. This immediacy of feedback affords learners greater attentional resources to focus on the target forms, having already comprehended the meaning of the feedback, either fully or partially, thereby facilitating the process of form-function mapping (e.g., Long, 1996, 2007, 2015; Doughty, 2001). In line with theoretical propositions from the Focus on Form literature (e.g., Doughty, 2001), the short time lapse in immediate CF between learners' errors and feedback helps learners compare communicative intent, their own output, and the target forms received through feedback, all while the information is still held in WM. Consequently, as the temporal gap between the error and the feedback increases, the efficacy of the feedback is likely to decrease (Doughty & Long, 2003). Therefore, the contingency of multimodal recasts is likely to be instrumental in directing learners' attention more effectively toward the linguistic target, fostering a stronger connection between the structural aspects of the linguistic target and its intended communicative or semantic purpose.

In summary, the combination of reduced cognitive load, increased salience, and contingency of multimodal recasts successfully heightened their noticeability and corrective intent. Consequently, this engaged learners' attention toward the target forms and facilitated cognitive comparisons between their errors and the target-like model, ultimately leading to improved learning and retention. Further research utilizing eye-tracking methodology and working memory measures is essential to validate whether multimodal recasts genuinely alleviate learners' cognitive burden, draw more attention from learners, and if the reduced WM load and increased visual attention correlate with their learning gains.

It is worthwhile discussing how recasts in the present study affected the development of different types of L2 knowledge. The findings underscore the superiority of multimodal recasts evident across all assessment tasks. Also, the written recast group outperformed the control group on OPT, SOPT, and WPT. Similarly, the oral recast group achieved higher gains than the control group in all outcome measures. This suggests that participants, through recasts, not only improved procedural knowledge but also exhibited gains in declarative knowledge, demonstrating the transferability of knowledge across different modalities. According to Skill Acquisition Theory (e.g., DeKeyser, 2007), the transfer of procedural knowledge across modalities is challenging, and the shift between skills, such as moving from written to oral performance, is likely to happen through declarative knowledge of the target construction. Thus, the observed results align with prior empirical research (Ellis et al., 2006; Révész, 2012), supporting the idea that recasts can facilitate gains in both declarative and procedural knowledge. While this is relatively well established for oral recasts, there has been limited exploration of this aspect concerning other types of recasts (e.g., Baralt, 2013; Yilmaz & Yuksel, 2011). This research is among the first to provide evidence for the capacity of multimodal and written recasts to promote both types of knowledge.

Finally, it is also worth discussing why written recasts proved more beneficial than their oral counterparts. One possible explanation is that the written recasts, differing in modality from the ongoing oral interaction, attracted learners' attention more effectively, given the salience created by the modality difference. Another possibility is that, as noted in prior text-based studies (e.g., Lai et al., 2006; Loewen & Erlam, 2006; Yilmaz & Yuksel, 2012; Sauro, 2009), the contingency of written recasts might have assisted learners in making cognitive comparisons between their erroneous production and the target-like models (Long, 2007). This may have rendered the corrective nature of recasts more explicit to learners

(Blake, 2008), prompting them to reflect more on the target forms (Salaberry, 2000) and apply their declarative knowledge more throughout task performance. Consequently, they might have been able to automatize their explicit knowledge about the use of *wh*-questions to a greater extent.

3.3.3. Limitations of Study 1 and Modifications in Study 2

The empirical findings of this study should be interpreted in the light of several limitations, which can be categorized into six primary areas: 1) research design, 2) sample size and generalizability of findings, 3) learners' proficiency level, 4) linguistic target, 5) the absence of investigation into learners' visual attention allocation, and 6) the need to address individual differences in WM capacity.

Firstly, in terms of research design, the adoption of a pretest–posttest–delayed posttest design to evaluate participants' learning gains was methodologically sound. However, one drawback lies in the relatively short interval between the immediate and delayed posttests (one week). This decision aimed to minimize the potential external influences (e.g., external feedback or explicit instruction on the target constructions) on participants' performance during the delayed posttest. While this approach strengthens the attribution of improvements to the treatment, in order to strike a balance between maintaining study context integrity and tracking the persistence of grammatical enhancements, the interval between the immediate and delayed posttests was extended from one week to two weeks.

Secondly, the study had 60 participants, a relatively small number. This limitation raises concerns about the robustness and generalizability of the study's findings to a broader population. To address this concern, Study 2 expanded the sample size to 120 participants.

Thirdly, this research was confined to participants within A2–B1 proficiency band according to CEFR. Therefore, it is important to explore the influence of feedback modality on L2 learning across various proficiency levels. Study 2 tackled this issue by including less proficient learners, to investigate whether different feedback modalities have similar or differential effects on L2 grammatical knowledge development in comparison to their impact on the more proficient English users in Study 1.

Fourth, this study focused on a single grammatical feature. However, previous research suggested that the effects of feedback modality may differ when addressing different linguistic constructions, especially those varying in salience and redundancy. In alignment with the primary linguistic focus of the thesis (L2 grammatical development) and the proficiency levels of the participants, Study 2 selected the present progressive and the third person singular forms as target linguistic features, as they differ in salience and communicative redundancy.

A fifth limitation is the lack of insight into how learners allocate their visual attention to the target structures and whether this allocation is related to feedback modality and learning outcomes. Study 2 addressed this limitation by incorporating eye-tracking methodology and various attention measurements.

The final shortcoming of this research pertains to the absence of an inquiry into individual differences among the participants, particularly regarding their WM abilities. WM is widely recognized for its pivotal role in attention control and noticing of L2 forms during input processing (Wen et al., 2017) and in establishing connection between form and meaning (VanPatten, 2004). Moreover, considering that task modality is assumed to impose different cognitive demands on L2 learners (e.g., Gilbert et al., 2016; Kormos, 2014; Williams, 2012; Zalbidea, 2017), WM is likely to have influenced the varied learning

outcomes associated with various feedback modalities in Study 1. Consequently, Study 2 delved into how individual differences in WM impacted the effects of SCMC feedback modality on the enhancement of L2 grammatical knowledge and the direction of learners' visual attention toward the target constructions. This investigation was also grounded in the well-established association between WM and CF (see Granena & Yilmaz, 2018; Li, 2017 for a meta-analytic reviews), as well as the potential relationship between WM and attention (Robinson et al., 2011).

CHAPTER 4. STUDY 2

This chapter reports on Study 2, the main study of this thesis. Study 2 further examined the impact of different recast modalities on learners' visual attention allocation to target linguistic features and their subsequent development in L2 grammatical knowledge, particularly with the provision of multimodal and written recasts. In addition, the study considered individual differences in WM among the participants by exploring the extent to which WM capacity might moderate the effects of recast modalities on L2 grammatical development and visual attentional allocation. Five distinct WM measures were employed to assess different constructs of WM capacity: (a) a NWS task, (b) a forward Corsi block task, (c) a color shape task, (d) a stop signal task, and (e) an AOSPAN task. To address the limitations identified in Study 1, several modifications were implemented. One involved recruiting a larger participant pool to increase the generalizability of the findings. Also, participants at a lower proficiency level were recruited to test whether the observed superiority of multimodal recasts in Study 1 would transfer to different proficiency levels. As well, the interval between the immediate and delayed posttests was extended from one to two weeks, allowing for a more extended observation of the learning outcomes.

This chapter begins by presenting the research questions that Study 2 aimed to address. The subsequent sections outline the methodology employed in the study, encompassing details of the participants, research ethics, target constructions, materials, data collection procedures, scoring system, and statistical analyses. Following this, the results of the study are provided for each research question. Finally, this chapter concludes with a comprehensive discussion of the findings, offering insights into the implications and significance of the research outcomes.

4.1. Research Questions

The following research questions were addressed in Study 2:

1. To what extent do recasts, as compared to no recasts, affect the development in L2 grammatical knowledge during video-conferencing tasks?
2. To what extent does modality of recasts affect the development in L2 grammatical knowledge during video-conferencing tasks?
3. To what extent does modality of recasts (in the case of written versus multimodal recasts) affect learners' attentional allocation to the target linguistic constructions during video-conferencing tasks, as reflected in participants' eye movements?
4. To what extent is learners' visual attention to the target constructions included in recasts, related to the development in L2 grammatical knowledge? Is this relationship influenced by whether learners are provided with written or multimodal recasts?
5. To what extent do individual differences in WM capacity moderate the effects of oral, written, or multimodal recasts on development in L2 grammatical knowledge during video-conferencing tasks?
6. To what extent is learners' visual attention to the target constructions related to their WM capacity? Is this relationship influenced by whether learners are provided with written or multimodal recasts?

4.2. Methodology

4.2.1. Overall Design

This study employed a pretest–immediate posttest–delayed posttest design with three treatment sessions. A total of 120 participants learning English as a foreign language (EFL) in Seoul, South Korea were recruited for the research. They were randomly assigned to four

groups: an oral recast group (N = 30), a written recast group (N = 30), a multimodal recast group (N = 30), and a no recast group (N = 30). All four groups were administered a background questionnaire, a general proficiency test, a pretest, a series of treatment tasks, an immediate posttest, WM measures, a delayed posttest, and an exit questionnaire. Each testing session involved completing a spontaneous oral production test (SOPT), followed by an oral production test (OPT), and a written production test (WPT) in that order. The groups differed only in what type of recasts they received during the treatment sessions. That is, the treatment tasks were administered either without recasts, with oral recasts only, with written recasts only, or with oral and written recasts together (i.e., multimodal recasts). The research instruments and procedures employed in the study are delineated in greater detail in the following sections.

4.2.2. Participants

Of the total number of 120 participants, 90 were female and 30 were male. They were recruited from two different language institutes located in Seoul, South Korea. All participants reported normal or corrected-to-normal vision. They were all native speakers of Korean. Their ages ranged from 26 to 39, M = 29.48, SD = 3.88, 95% CI [28.81, 30.21]. On average, they had received between 10 and 14 years of English instruction prior to the study, M = 12.44, SD = 1.06, 95% CI [12.25, 12.63]. One-way ANOVAs with age and length of English study as dependent variables found no significant differences among the four groups, age: $F(3, 116) = .50$, $p = .681$, $\eta^2 = .02$; length of study: $F(3, 116) = .12$, $p = .952$, $\eta^2 = .01$. All participants reported that they had not lived in a country where English was the main spoken language prior to the present study. To identify the general proficiency level of the participants, an Oxford Placement Test was administered at the outset of the study.

Participants' average test scores ranged from 107 to 118, M = 111.49, SD = 2.56, 95% CI [111.03, 111.95], indicating that their proficiency level was in the A2 (elementary) band in terms of the CEFR. Table 20 presents the descriptive statistics for the participants' performance on the Oxford Placement Test. One-way ANOVAs confirmed that all groups achieved comparable scores on both the listening, $F(3, 116) = .1.25$, $p = .294$, $\eta^2 = .03$, and grammar, $F(3, 116) = .994$, $p = .398$, $\eta^2 = .03$, sections of the placement test.

Table 20. Descriptive Statistics for Participants' Performance on the Oxford Placement Test

	Listening Section			Grammar Section		
	M	*SD*	95% CI	*M*	*SD*	95% CI
Control	61.30	2.18	[60.48, 62.12]	50.17	2.67	[49.17, 51.16]
Oral	60.37	2.25	[59.53, 61.21]	51.00	2.78	[49.96, 52.04]
Written	60.77	1.79	[60.10, 61.44]	49.93	2.23	[49.10, 50.77]
Multimodal	60.53	1.68	[59.91, 61.16]	50.23	2.46	[49.32, 51.15]

4.2.3. Research Ethics

Prior to commencing data collection, the study obtained research ethics approval from the Research Ethics Committee at the UCL Institute of Education, University College London. At the beginning of the study, the participants were provided with a general introduction in a form of an information sheet explaining: (a) the purpose of the study, (b) the overall procedures and duration of the study, (c) the participants' right to withdraw from the research at any time, and (d) the measures taken to ensure the maintenance of their privacy and confidentiality (Appendix A13). Each participant's agreement to participate in the study was obtained through a signed consent form (Appendix A14). As in Study 1, it was crucial for participants not to be aware of the target structures and the type of recasts they would receive before the treatment sessions, given the primary aim of this study was to examine the effects

of recasts and their modality on participants' noticing and development in the use of the target constructions. Thus, participants were informed that the purpose of the research was to examine Korean EFL learners' online interactional patterns and on-screen eye movements during video-conferencing tasks. This approach ensured that participants' responses and behaviors remained unbiased and genuine, contributing to the integrity and validity of the study.

4.2.4. Target Constructions

Two linguistic targets were selected for the present study: the present third person singular and the English present progressive.

4.2.4.1 English Third Person Singular

The present third person singular verb form is a grammatical construction which denotes tense and agreement. According to Swan (2016, p. 30–31), there are four typical contexts in which the present simple is used as in Example 1:

(Example 1)

a. To talk about timeless truths, permanent situations, or things that happen regularly, repeatedly or all the time. (e.g., I **play** tennis every Wednesday.)
b. For the present meaning with verbs that do not usually have progressive forms such as like, believe. (e.g., I **like** this wine.)
c. To talk about the future:
    a) For timetabled future events. (e.g., His train **arrives** at 11.46.)
    b) In subordinate clauses. (e.g., I'll phone you when I **get** home.)

c) In suggestions with *Why don't you…?* (e.g., Why don't you **take** a day off tomorrow?)

d. To talk about a series of completed actions and events, for example, demonstration, commentaries, instructions and present-tense stories. (e.g., First I **take** a bowl and **break** two eggs into it.)

In this study, particular emphasis was placed on the prototypical usage (Example 1-a) of the present simple, which refers to "to talk about permanent situations, or things that happen regularly, repeatedly or all the time" (Swan, 2016, p. 30) (see Section 3). Previous studies suggest that L2 learners often encounter challenges in acquiring the third person singular (Dulay & Burt, 1974; Krashen, 1982; Larsen-Freeman, 1975; Goldschneider & DeKeyser, 2001), despite it being considered one of the easiest grammatical rules to learn. Indeed, Ellis (2008) noted that most learners have no difficulty in grasping the rule for the third person singular, yet they still have difficulty in using it accurately.

There are several factors that may contribute to the difficulty in learning to use this morpheme correctly. One source of challenge is that the third person singular is physically and phonologically non-salient (Goldschneider & DeKeyser, 2001; N. Ellis, 2006; Loewen, Erlam, & Ellis, 2009; VanPatten, 1996). The third person singular morpheme undergoes both typographical and phonological alterations based on the voiced and voiceless sounds of the last syllable of the verb to which it is added (Quirk & Greenbaum, 1990):

Typographical alterations:

(a) -*s* after voiceless consonants: cooks, sleeps

(b) -*es* after sibilant consonants (including verbs ending in -s, -ss, sh, -ch, -x, -z): fixes, washes

Phonological alterations:

(a) [s] after voiceless consonants: cooks, sleeps

(b) [z] after voiced segments (excluding sibilants): drives, runs

(c) [iz] after sibilant consonants: watches, brushes

Despite these typographical and phonological alterations, none of these forms are prominently stressed or emphasized in speech or writing, making them less noticeable and attended to during both oral and written interactions.

Another factor contributing to learners' difficulty in acquiring this construction is the challenge in determining when to use the present simple form versus other verb tense combinations. For example, deciding between using the present simple or other constructions such as the present progressive or future tenses poses a challenge for L2 learners. This difficulty may be further compounded by morphosemantic discrepancies between the learners' first and second language (e.g., Bardovi-Harlig, 2001). In Korean, for instance, the present simple can denote both ongoing and habitual events. This can result in a learnability issue for L1 Korean learners of English (Hahn, 2013), as they may use the two forms interchangeably and erroneously.

4.2.4.2. English Present Progressive

The other linguistic target of the current study was the present progressive form. Swan (2016, p. 32–33) outlines four contexts in which the present progressive is typically used, as in Example 2:

(Example 2)

a. To talk about temporary situation that are going to on now or "around now": before during and after the moment of speaking. (e.g., I**'m writing** emails.)
b. To refer to repeated actions and events if they are just happening around the present. (e.g., I**'m travelling** a lot these days.)
c. To talk about development and changes. (e.g., House prices **are going** up again.)
d. To talk about the future. (Come and see us next week if you**'re passing** through London.)

The experimental treatment and assessment tasks aimed to elicit the prototypical usage of this form (Example 2-a), that is, "to talk about temporary situations that are going on now or 'around now': before during and after the moment of speaking" (Swan, 2016, p. 32) (see Section 3). Bardovi-Harlig (2000) suggests that the present progressive is a developmental feature, and learners progress through three fixed stages in acquiring it:

(a) bare progressive (e.g., dancing)
(b) present progressive (e.g., is dancing)
(c) past progressive (e.g., was dancing)

Drawing on Bardovi-Harlig's developmental stages, participants were considered developmentally ready and eligible for participating in this study if they demonstrated the emergence of the bare progressive form in a present progressive context. This criterion aligns with the approach used in other acquisition studies that consider emergence (e.g., Mackey, 1999; Philp, 2003).

As in English, where the present progressive is overtly realized by "*am/are/is + V-ing*", Korean also has an explicit progressive marker, "*-ko iss-*" in its inventory. This morphosyntactic similarity between the participants' native language (L1) and English (L2) may positively influence their acquisition of the *form* of the present progressive. Despite this similarity, learning the *meaning* of the present progressive may still pose issues, as English uses the present progressive only to refer to events simultaneous with speech time. In contrast, Korean is ambiguous since, as mentioned earlier, the simple present tense can denote both ongoing and habitual events. For example, the Korean present tense marker "*-n-*" (or "*-neun-*" when the preceding sound is a consonant) can be interpreted as either habitual or ongoing/progressive, as shown in Example 3-(a) and 3-(b).

(Example 3)

Jiwon-ineun	jagi bang-eseo	gongbuha-*n*-da.
Jiwon-NOM	her room-LOC	study-PRES-DECL

(a) "Jiwon studies in her room (as her habitual behavior)."

(b) "Jiwon is studying in her room (as an ongoing action at the present moment)."

*NOM: nominative, LOC: locative, PRES: present, DECL: declarative.

In addition, the Korean progressive marker "-*ko iss*-" can refer to either an ongoing process or the duration of a state (Kim, 2011; Ko & Ku, 2008; Lee, 2005). In other words, it can be combined with verbs describing both events and state. When combined with verbs describing events, it denotes those that are ongoing, as in Example 4:

(Example 4)

Jiwon-ineun	jaki bang-eseo	kongbuha-*ko iss*-da.
Jiwon-NOM	her room-LOC	study-PRES-DECL

"Jiwon is studying in her room."

When combined with stative verbs, "-*ko iss*-" expresses a durative state. Many state verbs in Korean are compatible with both progressive and simple present markers, as shown in Example 5:

(Example 5)

Jiwon-ineun	dap-eul	a-*n*-da/al-*ko iss*-da.
Jiwon-NOM	answer-ACC	know-PRES-DECL/know-PROGR-DECL

"Jiwon knows the answer."

*PROG: progressive

In summary, Korean learners of English may encounter difficulties in learning the two target features, in that a Korean verb in the simple present is compatible with both habitual and ongoing interpretations. Additionally, unlike English state verbs, the majority of Korean state verbs can be realized in the progressive form. Thus, the present study aims to investigate whether the provision and modality of recasts affect L2 learners' awareness of the obligatory

distinction between the present progressive and habitual aspect in English. These distinctions are grammaticalized by the progressive marker "*be -ing*" and the third person singular morpheme "*-s/-es*" in the present simple tense.

4.2.5. Experimental Treatment Tasks

As in Study 1, all four groups performed three multimodal video-conferencing tasks designed as communicative focused tasks, following Ellis's (2003) guidelines (Appendix A15). These tasks also fulfilled the criteria for tasks as defined by Ellis (2009) and Ellis and Shintani (2014), which was detailed in Study 1 (see Section 3.1.5.2).

To elicit the target constructions (i.e., the third person singular and the present progressive), the tasks were operationalized in an imaginary scenario in which participants acted as FBI agents assigned to observe and report their prime suspects' daily and current activity, so the information could be uploaded to the FBI Suspect Database for real-time viewing and sharing among agents (see Figure 12). Specifically, participants were first given prompts and photos showing the suspect's activities at specific times of the day and asked to orally report the activities based on the given information (e.g., "Every Friday at 4 p.m., he plays basketball."). Then, they were also presented prompts and photos depicting the suspect's current activities and required to verbalize them (e.g., "Now, he is riding a bicycle."). As such, the treatment tasks were designed to create obligatory contexts in which participants had to use the two target constructions to complete the tasks. Considering the participants' proficiency level, the task instructions were provided in their L1, Korean. The treatment materials only included high-frequency words selected from the New General Service List (Brezina & Gablasova, 2015). The tasks were piloted with a population similar

to the participating learners and the data confirmed that the tasks succeeded in eliciting the target structures.

During the treatment tasks, the experimental groups received recasts on their non-target-like utterances of the target features according to their group conditions. For the oral recast group, recasts were only presented orally. The written recast group were provided with written recasts exclusively. The multimodal recast group received recasts in both oral and written modes simultaneously. The no recast group did not receive any feedback. Recasts were given in the form of full declarative recasts (Sheen, 2006) provided by the researcher (see Section 4.2.6 for the detailed characteristics of recasts used in Study 2). Written recasts for the multimodal and written recast groups were displayed as real-time on-screen text using Open Broadcaster software (see Figure 13). The written recasts appeared in a black, monospaced font, specifically using the Consolas font type with a font size of 35. The selection of the monospaced font was made based on its inherent attribute of fixed-width characters, wherein each letter occupies a consistent horizontal space. This typographic choice offered several notable benefits, particularly with regard to the precise control over the physical size of words in the text. By virtue of the fixed-width nature of the font, the length of each word is inherently regulated, resulting in uniformity in the visual appearance of the text.

As in Study 1, the written recasts appeared on the computer screen for ten seconds. Prior to the experiment, the researcher had prepared a corresponding written recast for each question, which was presented on the screen via OBS immediately after participants' errors in the use of each target structure. These prepared written recasts allowed the researcher to provide both oral and written recasts concurrently without the need for real-time typing, ensuring a prompt and contingent presentation of feedback immediately after learners' errors. All treatment sessions were administered and recorded by Zoom.

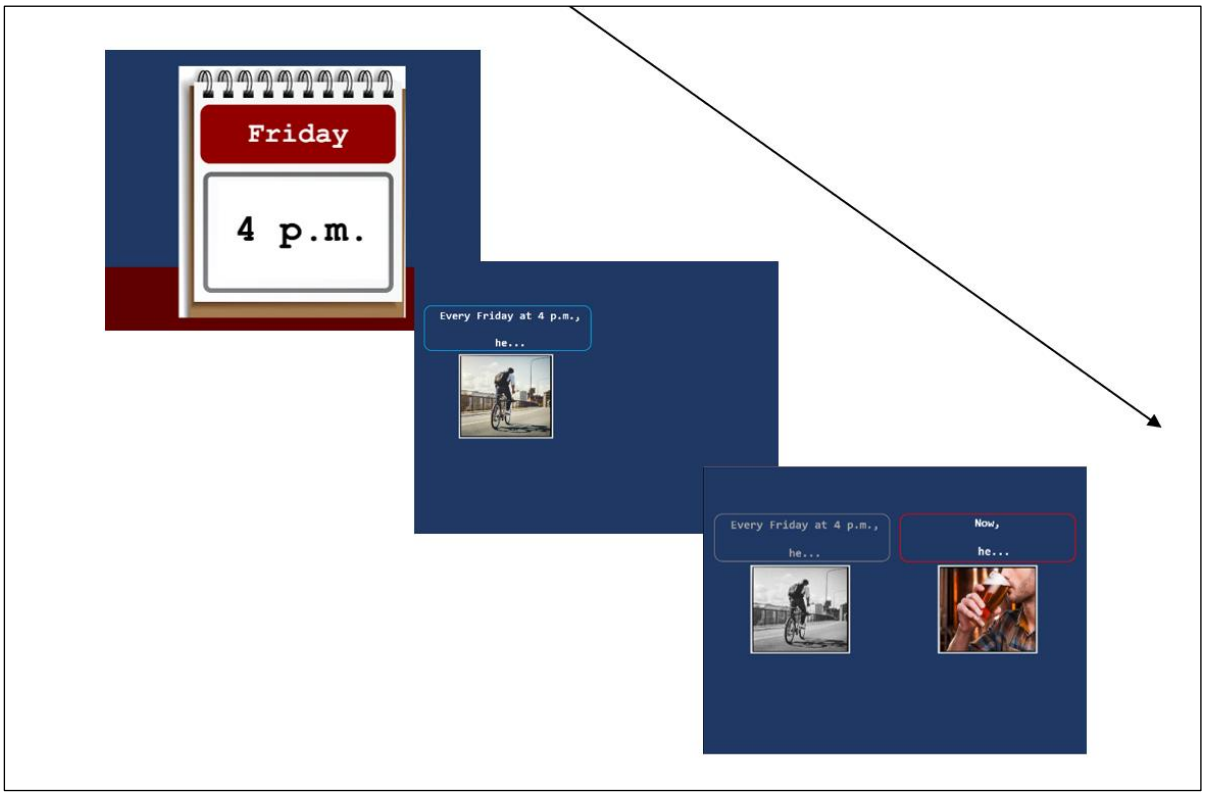


Figure 12. An Example of Experimental Treatment Tasks



Figure 13. Examples of Written Recasts for Each Target Construction

4.2.6. Type of Recasts

While performing the treatment tasks, participants in the experimental groups received recasts in response to all errors they had committed while using the third person singular and the present progressive forms, as illustrated below:

(a) A recast for the non-target-like present third person singular form:

Learner: Every Sunday at 4 p.m., he **play** basketball.

Researcher: He **plays** basketball.

(b) A recast for the non-target-like present progressive form:

Learner: Now, she **brush** her teeth.

Researcher: She **is brushing** her teeth.

Drawing on Sheen's characteristics of recasts (2006, p. 372–375) and previous research (Dulay et al.,1982; Lyster, 1998; Philp, 2003; Roberts,1995), the recasts in this study exhibited the following features: First, in terms of mode and scope, the recasts were declarative and isolated, reformulating only the incorrect part of learners' production without adding new information (Lyster, 1998). Also, the recasts involved no reduction, as they repeated the learners' entire utterances. In regard to length, the recasts were classified as clauses, comprising at least two phrasal constituents, including a finite verb (Philp, 2003). Next, the recasts were one-change recasts, targeting a single error (i.e., the present third person singular or the present progressive). The type of change involved adding the allomorphs (i.e., [s], [z], [iz]) or the present progressive (i.e., *be* and/or *-ing*) or substitution, where the researcher replaced an erroneous utterance with the correct present third person singular or the present progressive forms (Delay et al., 1982). Lastly, the linguistic focus of the recasts was grammar, particularly addressing a single type of grammatical error. Thus, the recasts were delivered in an intensive and focused manner, yet they were elicited in a natural way as participants performed a task contextualized within a real-life context. Syntactic, lexical, pronunciation, and other grammatical errors were not corrected.

4.2.7. Materials

4.2.7.1. General Proficiency Test

The participants' overall English proficiency was assessed using the Oxford Placement Test. This test consisted of 200 questions, with 100 questions in the listening section and 100 questions in the grammar section. In accordance with the test scoring guideline, one point was awarded for each correct answer, resulting in a maximum score of 100 for each section.

4.2.7.2. Assessment Tasks

In Study 2, changes in participants' knowledge of the target constructions were assessed using a SOPT, an OPT, and a WPT. The rationale for including both oral and written assessments was to tap into different aspects of knowledge. Although both measures were likely to elicit procedural knowledge (Anderson, 1993; Dekeyser, 2007), the OPT imposed greater time pressure than the WPT. Thus, participants' performance on the oral production test was expected to be a better indicator of procedural knowledge in the process of automatization. The WPT, on the other hand, allowed more time and opportunities for participants to review and revise their responses and hence deploy their explicit, declarative as well as procedural knowledge during the test. In addition, given that recasts were presented in three different modes during the treatment, utilizing different outcome measures in different modalities was expected to help to investigate whether the assessment modality influenced the relationship between feedback modality and development in the use of the target constructions.

The inclusion of the SOPT allowed for measuring learners' use of the target constructions in a more authentic and less scripted scenario with only key-word prompts provided. The SOPT, being less controlled and more open-ended, allowed participants to

demonstrate their ability to spontaneously produce the target items, contributing to a more ecologically valid assessment. This multimodal approach enhances the validity of the assessment and aligns with previous research that has demonstrated the benefits of employing multiple testing modalities to capture diverse aspects of language proficiency (McDonough, 2005; Philp, 2003). Each test had three versions (pretest, posttest and delayed posttest), which were carefully developed and counterbalanced across participants, modalities, and testing sessions. The OPT and WPT had the same format, differing only in modality, whereas the SOPT was designed more closely to the treatment tasks. Given the learners' proficiency, the instructions were delivered in the participants' L1, Korean. The tests were administered and recorded using Zoom.

4.2.7.2.1. OPT

The aim of the OPT was to measure the participants' ability to produce the target constructions in the oral mode (Appendix A16). In this task, the participants were asked to act as an official government spokesperson representing world leaders. For example, they were presented a scenario where they had to conduct an oral interview with CNN on behalf of Queen Elizabeth II during her official visit to the G7 Summit. During the interview, CNN viewers sent questions about the queen through text messages with brief keywords due to time constraints. The questions were carefully designed to elicit the two target constructions as well as some other constructions which served as distractors. The participants had to answer the viewers' questions orally based on the information which appeared at the bottom of the screen (see Figure 14). They were told that the information was provided by the queen's aide and the information was only visible to them during the interview.

The test consisted of 18 target items and 20 distractors. To ensure the validity of the test and prevent repetition, none of the verbs used in the treatment tasks were included in the test. This approach aimed to avoid learners merely reproducing memorized responses and encouraged them to apply their language skills authentically in the oral context. The target items were divided equally between the third person singular and the present progressive constructions. To create obligatory contexts for the third person singular, nine questions were related to their leaders' regular activities (e.g., "Every morning, he goes to the gym."). On the other hand, to prompt the use of the present progressive form, participants were required to respond to nine questions about their leaders' current activities (e.g., "Now, he is giving a speech."). To increase situational authenticity (Ellis, 2013), relevant video footage was presented before the present progressive questions were given to the participants (Figure 15). To prompt learners to produce each answer in a full sentence, participants were instructed to provide their answers in full sentences to maintain a formal tone as the interview was deemed an official occasion. They were asked to avoid one-word answers (e.g., at 6 a.m.), and instead, use the provided prompts to construct complete sentences (e.g., "Every morning, she gets up at 6 a.m.").

As noted earlier, to avoid potential test effects, three different versions of the test were developed and the sequence of the tests was counterbalanced across the testing sessions. Although different contexts were utilized in each testing session, they elicited the same verbs from the participants. In the case of the third person singular, each allomorph was used an equal number of times in each test version (i.e., three questions for each allomorph in each test version). Additionally, the distractors consisted of regular past tense (10 items) and irregular past tense forms (10 items).



Figure 14. An Example of the OPT



Figure 15. An example of video footage and a prompt for the present progressive question

4.2.7.2.2. WPT

The main objective of the WPT was to gauge the learners' proficiency in producing the target structures in the written mode (Appendix A17). As the OPT, a communicative context was created; for example, participants were asked to engage in a live chat with the BBC on behalf of the German chancellor on her official visit to the UK. During this live chat, BBC viewers sent in questions about the chancellor using brief keywords in text messages. The questions were designed to elicit the target constructions alongside a set of distractors. Similar to the OPT, relevant video footage was presented before presenting the present progressive questions to the participants. In contrast to the OPT, participants had to provide their responses in writing. As with the OPT, three versions of the test were developed. Each version included an equal number of obligatory contexts for each target item: nine contexts for the third person singular and nine contexts for the present progressive. Each allomorph of the third person singular was equally represented in the test items. The three versions of the test were evenly distributed among the participants and assessment sessions. To maintain consistency, the WPT included the same type and number of distractors as the OPT, and none of the verbs used in the treatment tasks were used in the test (Figure 16).



Figure 16. An Example of the WPT

4.2.7.2.3. SOPT

The SOPT was designed to assess the participants' ability to use the target forms in a more spontaneous setting, allowing for an evaluation of their application of the knowledge gained during the treatment in an environment that was less controlled (Appendix A18). While participants were asked to answer short text questions from viewers, a format akin to the OPT and WPT, the SOPT differed in its resemblance to the treatment tasks. In the SOPT, participants, acting as official government spokespersons for world leaders, engaged in oral interview with media outlets such as CNN or BBC. For instance, they took the role of a

spokesperson for President Macron of France during his official visit to the UK. Unlike the OPT and WPT, the SOPT posed a challenge to participants by requiring responses to viewers' questions regarding their leaders' past, everyday, and current activities, solely based on prompts and images, without the provision of specific verbs—a departure from the approach used in the OPT and WPT. Figure 17 illustrates this context. These questions were designed to elicit the two target structures as well as distractors, and none of the verbs used in the treatment tasks were included in the test. Consistent with the OPT and WPT, relevant video footage was presented prior to presenting the present progressive questions to the participants. Since no specific verbs were supplied for use, when participants requested a verb to describe the depicted activities of their leaders, only the base form of the verbs was provided.

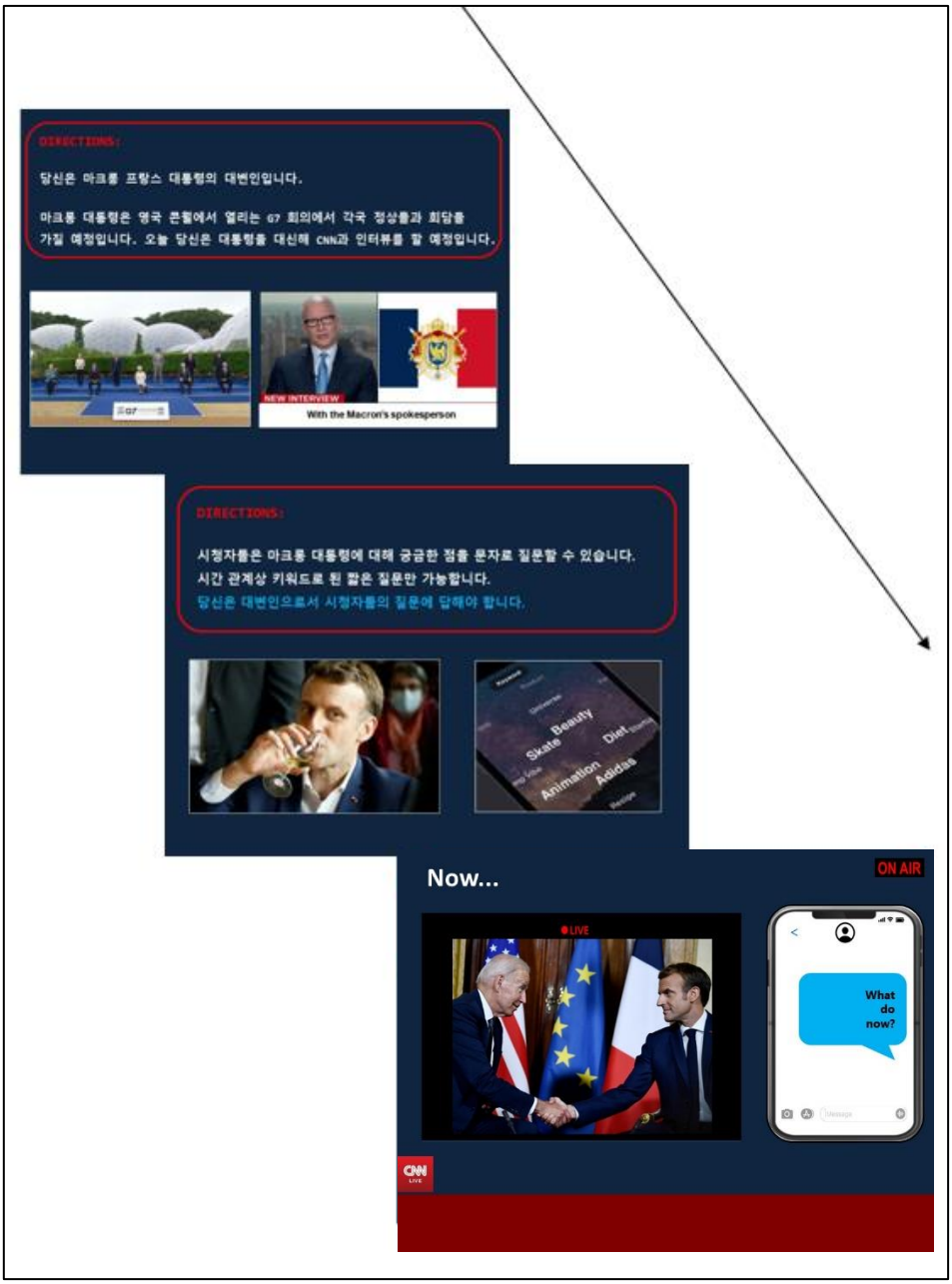


Figure 17. An Example of the SOPT

4.2.7.3. Working Memory Measures

Five distinct WM measures were employed to test different aspects of WM capacity. PSTM was assessed with a NWS task, and VSTM was measured using a forward Corsi block test. The executive functions of updating, task-switching, and inhibitory control were assessed using an AOSPAN task, a color shape task, and a stop signal task, respectively. The NWS

task was presented using PowerPoint and the participants' performance was recorded using a voice recorder. The other four WM measures (i.e., color shape task, stop signal task, forward Corsi block task and AOSPAN) were conducted using the Inquisit 6 Lab software (Version 6.6.1, Millisecond, 2022).

4.2.7.3.1. NWS Task

To assess participants' PSTM, a Korean version of the NWS task (Jung, 2017) was employed. In this task, participants were presented with a set of 32 nonwords, which conformed to the phonotactic rules of Korean, each with a length of 4 to 11 syllables (four sets for each syllable length). Each nonword was aurally presented to the participants in a random order, and their task was to listen and recall them as correctly as possible. Approximately ten seconds were allowed for the participants to give their responses. Each nonword recall was scored as either correct or incorrect. The score for this test was determined by the maximum number of syllables that participants correctly recalled at least twice for each syllable length. The test was conducted individually and took approximately 9–10 minutes to complete. It was administered using an automated PowerPoint presentation, and participants' responses were audio recorded.

4.2.7.3.2. Forward Corsi Block Task

The forward Corsi block task was used to measure participants' VSTM capacity. First, participants were presented with 2–9 identical blocks which were highlighted in different patterns on a computer screen. Each highlighted block appeared as a yellow square. The participants were required to click the blocks in the same order as they were highlighted. The

participants had two trials for each block length (Figure 18). The total score was calculated by counting the number of correctly repeated sequences until the task ended (Kessels et al., 2000).

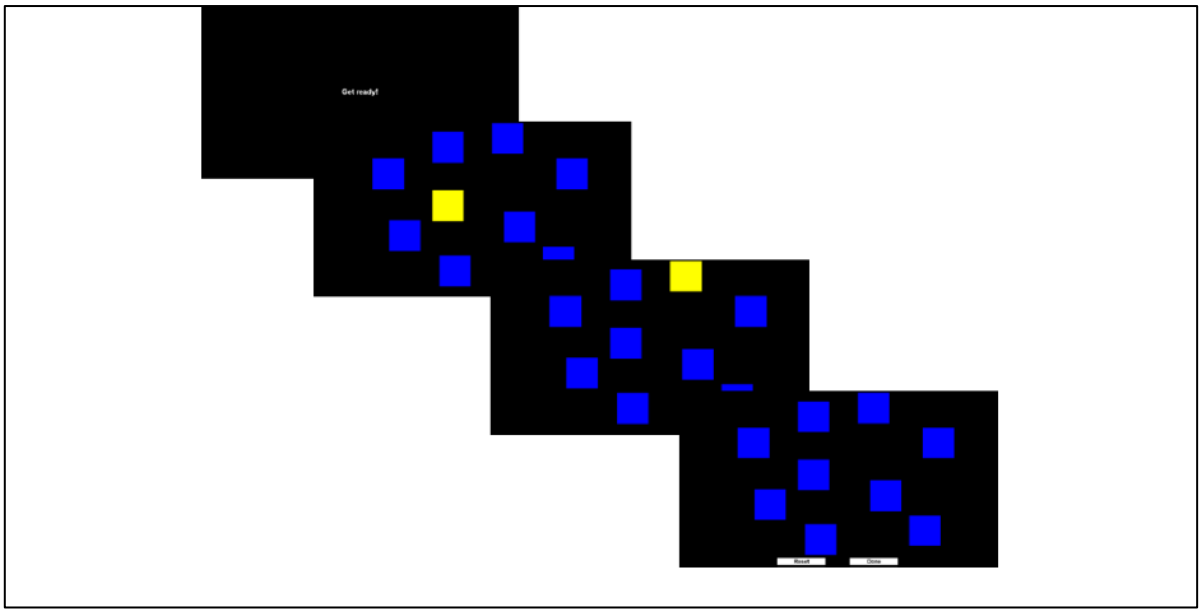


Figure 18. Forward Corsi Block Task

4.2.7.3.3. Color Shape Task

Task-switching ability, an executive function, was measured by a color shape task (Miyake, Emerson, Padilla, & Ahn, 2004). The task involved two different blocks, namely, non-switching and switching blocks. The non-switching block consisted of two separate sub-blocks, either a color or a shape block. In the color block, the participants' response depended on the color of the shape they had seen on the computer screen; they had to press "A" for green and "L" for red. Similarly, in the shape blocks test, the participants were instructed to press "A" for a triangle and "L" for a circle (Figure 19). In switching blocks, the participants had to decide based on either a color or a shape according to a cue letter that was presented

on the screen randomly (‘C’ for color and ‘S’ for shape). That is, if a stimulus with the cue letter ‘C’ appeared on the screen, the participants had to identify the color of the stimulus (i.e., green or red), regardless of its shape. On the other hand, if a stimulus appeared with the cue letter ‘S’, they had to decide the shape (i.e., circle or triangle), irrespective of its color (Figure 20).

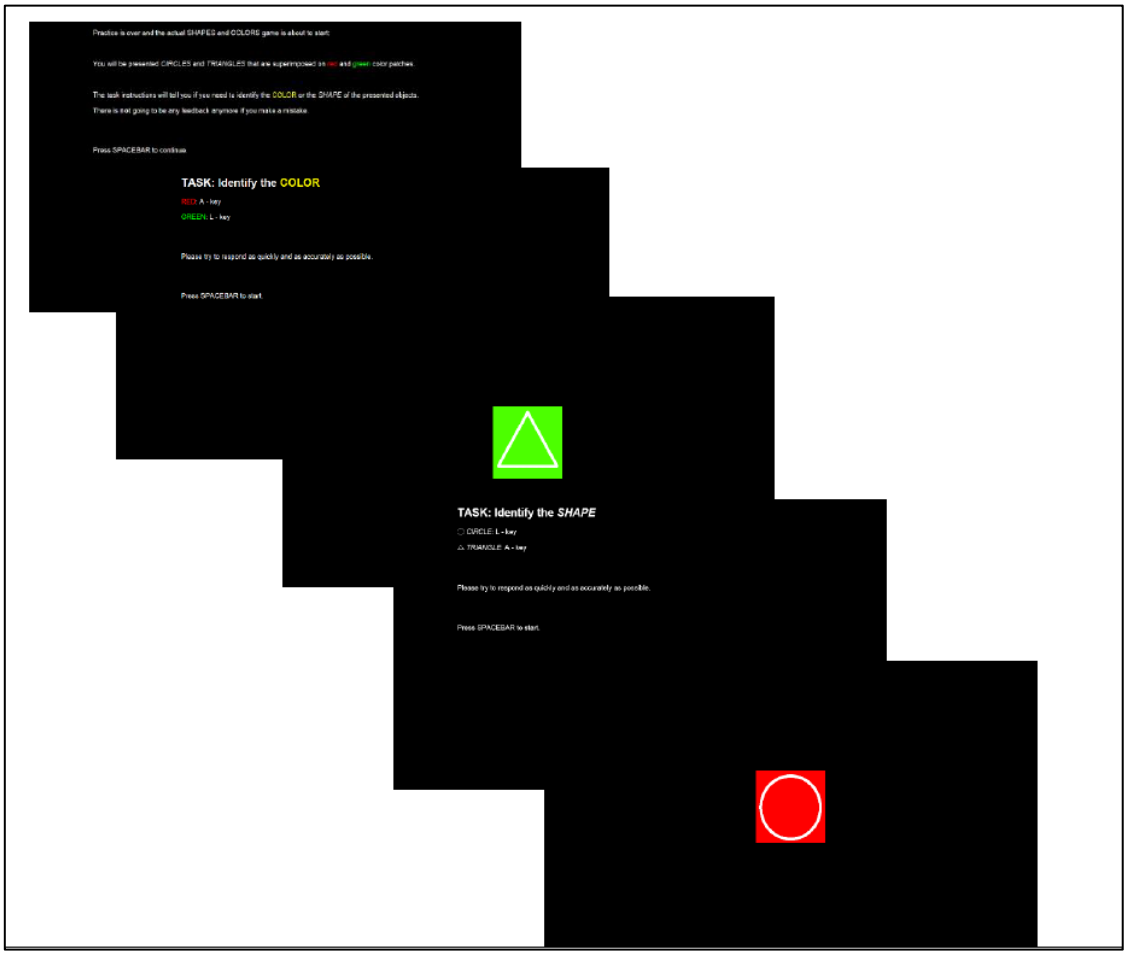


Figure 19. Color Shape Task – Non-switching Block

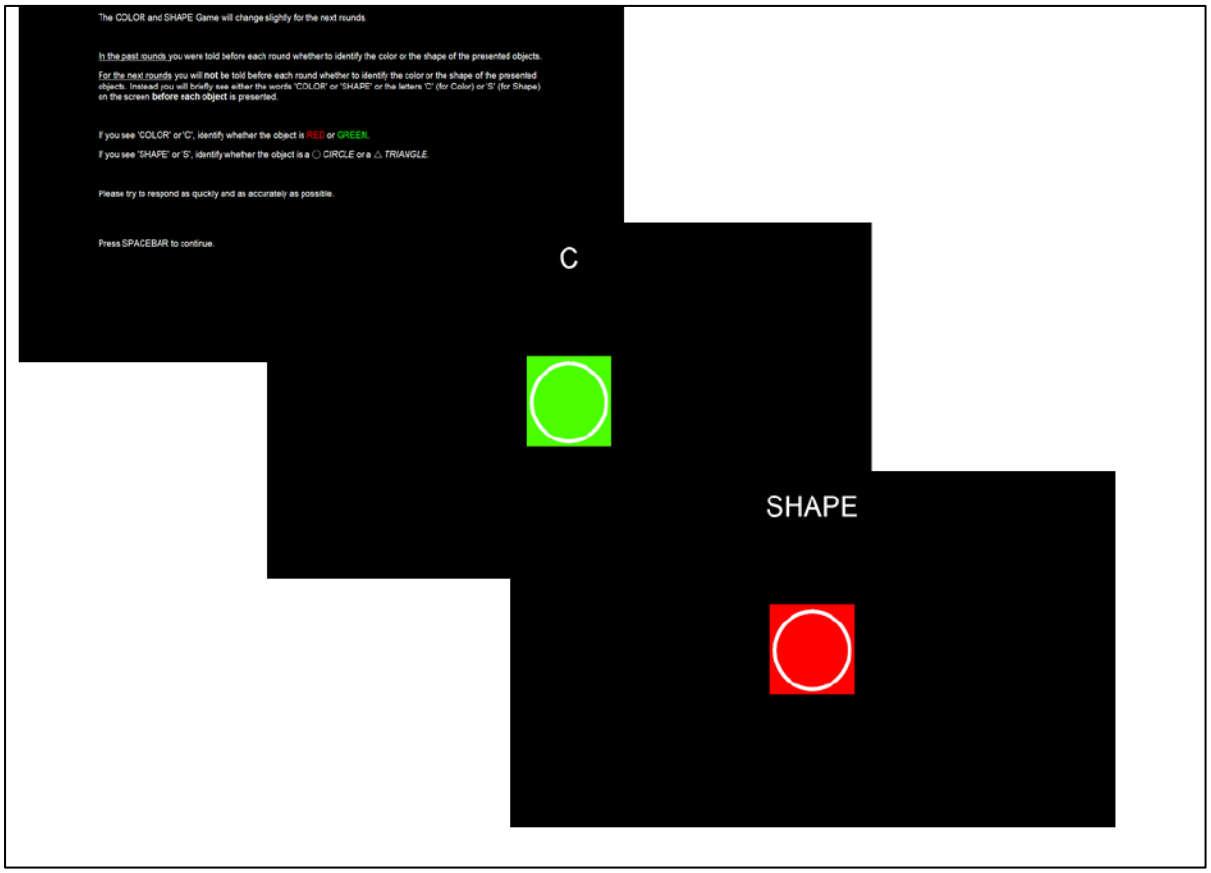


Figure 20. Color Shape Task – Switching Block

4.2.7.3.4. Stop Signal Task

A stop signal task was used as a measure of inhibitory control, another executive function. The participants were asked to respond to an arrow stimulus as quickly and correctly as they could. The arrow stimulus pointing either left or right appeared on the screen, and the participants had to press corresponding buttons on the keyboard; ‘D’ for the arrow pointing to the left and ‘K’ for the arrow pointing to the right (Figure 21). In the case where an arrow stimulus was accompanied by an auditory signal (a beep), the participants were instructed to withhold their response.

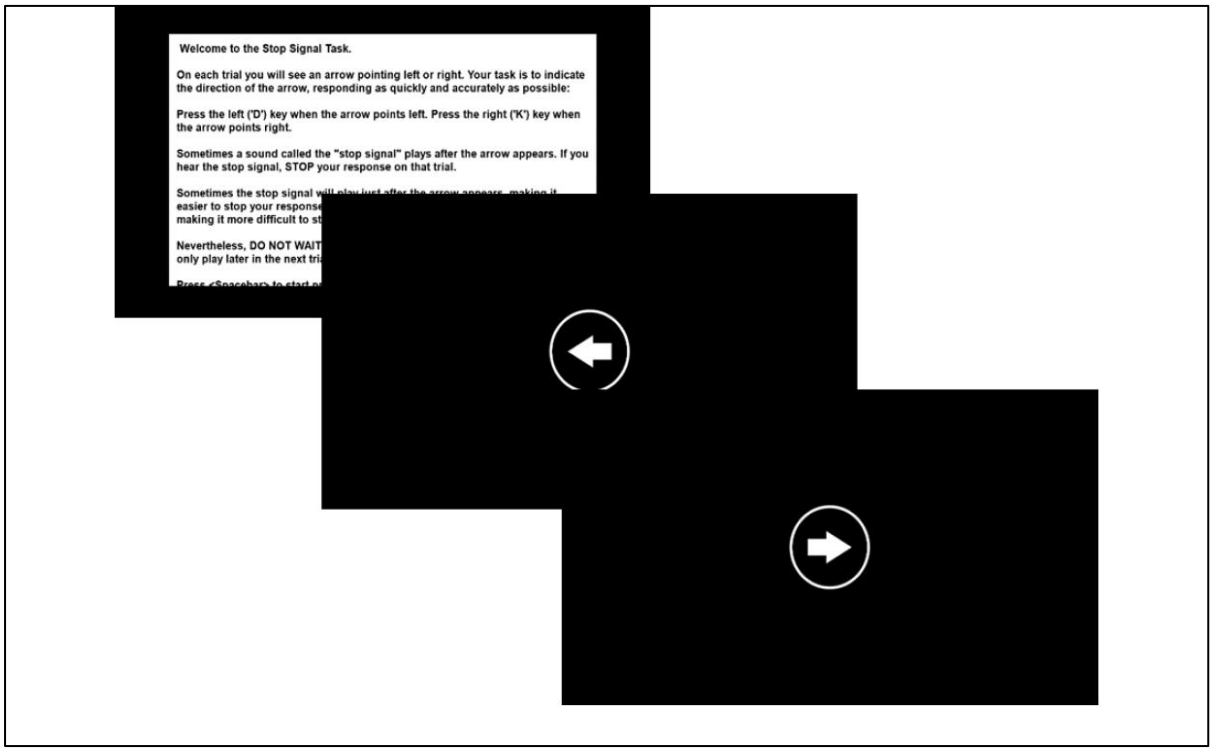


Figure 21. Stop Signal Task

4.2.7.3.5. AOSPAN Task

An automated version of the operation span task (Turner & Engle, 1989) was used to measure the executive function of updating. This task required participants to solve a series of arithmetic equations while attempting to remember a list of unrelated alphabet letters. First, the participants were presented with an arithmetic equation on the screen, followed by one letter (e.g., S, T, K, L, Q, R) at a time. Then, the participants were asked to remember the letter and make a judgement about whether an answer given to the equation was correct or incorrect. At recall, 12 possible alphabet letters in a 4×3 matrix were provided for the participants to select letters in the order they had appeared. The task consisted of three sets, with set sizes ranging from 3 to 7 (75 letters and 75 math problems in total) and the order of set size was random. The total number of letters that were correctly recalled was used as the AOSPAN index (Unsworth et al., 2005).

After completing each equation, the participants were provided with feedback on the accuracy of their responses automatically (Figure 22). In order to guarantee that participants did not trade off between processing (solving arithmetic equations) and storage (remembering letter strings), they were instructed to maintain an 85% accuracy criterion for the math operation (e.g., Unsworth et al., 2005).

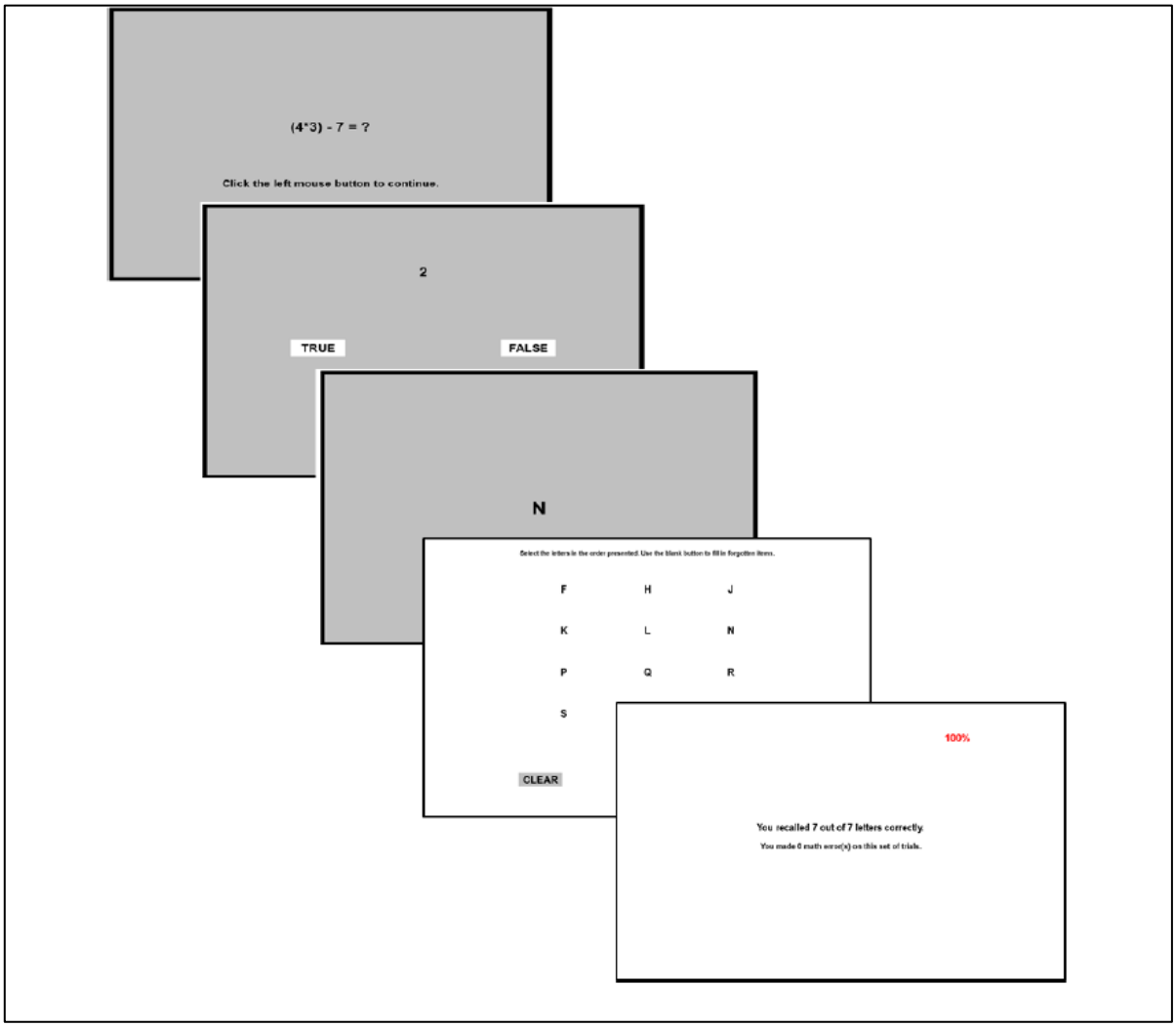


Figure 22. Automated Operation Span Task (AOSPAN)

4.2.7.4. Background Questionnaire

To obtain general information about the participants, a background questionnaire was administered. While the fundamental questions aligned with the background questionnaire used in Study 1 (see Section 3.1.5.4 for a comprehensive overview), Study 2 incorporated supplementary inquiries related to participants' reading behaviors due to the inclusion of the eye-tracking component in Study 2. These additional questions covered aspects such as the participants' visual health and their preferences in reading (e.g., physical books, online content) (Appendix A19). In order to encourage clearer responses, the questionnaire was administered in participants' native language, Korean.

4.2.7.5. Exit Questionnaire

At the end of the experiment, participants were administered an exit questionnaire. The questionnaire inquired about: (a) their awareness of the purpose of the study, (b) whether they noticed and/or focused on any particular linguistic features during the study, (c) whether they consulted any other sources outside the study, and (d) their perceived level of task difficulty. Additionally, open-ended questions were included to obtain insights into their perspectives regarding the usefulness of recasts and the effects of modality (for experimental groups), as well as their overall experience during the treatment and assessment tasks conducted via video-conferencing. Additionally, participants were asked to provide feedback on their overall experiences with eye-tracking research and the procedures of WM measurements. The exit questionnaire was also conducted in Korean (Appendix A20).

4.2.8. Eye-tracking Procedure

A Tobii Pro Nano mobile eye-tracker with a temporal resolution of 60 Hz was employed to record participants' eye movements during the three treatment tasks. Recording was conducted in binocular mode, capturing both the left and right eyes. This screen-based remote eye-tracker was mounted onto a laptop computer equipped with a 15-inch screen with a 1920 × 1080 screen resolution, and directly connected to the computer via a USB 2.0 port. Participants were seated approximately 60 cm away from the computer screen, facing the eye-tracker. Prior to each treatment task, a 9-point calibration grid was used to calibrate the eye-tracker. During calibration, participants were instructed to follow a black dot moving across the computer screen with their eyes. They were asked to maintain the same position as much as possible during the treatment task to ensure accurate detection of eye movements. The Tobii Pro Lab Software (Version 1.194, Tobii Pro AB, 2014) was employed for the recording and analyses of eye-tracking data.

4.2.9. Data Collection Procedure

The total duration of data collection was approximately seven months, with each session held individually for each participant (Figure 23). The five sessions were organized as follows: In the first session, participants signed the consent form after reading the information sheet for the study. Following this, they completed a background questionnaire and took the Oxford Placement Test, which lasted around 40–50 minutes. Subsequently, the pretest was administered, starting with the SOPT, followed by the OPT and the WPT. The duration of the three assessment tests were approximately 17, 15, and 18 minutes respectively.

The second session involved participants completing two treatment tasks, which were multimodal video-conferencing tasks. In the third session, participants performed the third

treatment task, followed by the immediate posttest. The duration of each treatment task was between 13 and 15 minutes. During the treatment tasks, an eye-tracker was used to capture participants' eye movements. In session 4, participants took WM tests. The duration of each WM tests varied as follows: (a) the NWS task lasted 9 to 10 minutes, (b) the color shape task took 40 to 45 minutes, (c) the stop signal task finished in 9 to 10 minutes, (d) the forward Corsi block task was completed in 3 to 5 minutes, and (e) AOSPAN task lasted 35 to 40 minutes. Session 5 took place two weeks later, during which participants completed the delayed posttest and the exit questionnaire. The immediate and delayed posttests had the same format as the pretest. The treatment and assessment sessions were conducted and recorded via Zoom. Participants confirmed that, in the course of data collection, they did not receive any instruction on the target constructions outside of the study. Each session's duration ranged from 2 and 3 hours.

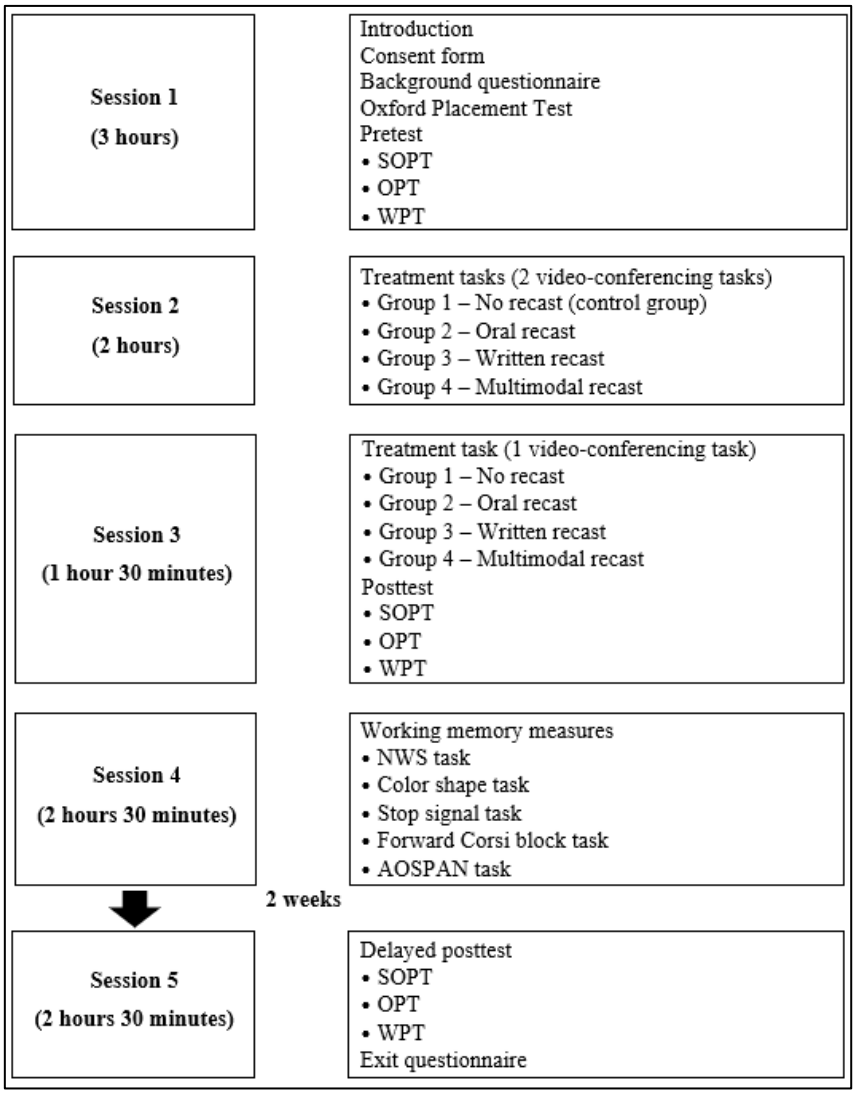


Figure 23. Data Collection Procedure for Study 2

4.2.10. Scoring and Data Analysis

4.2.10.1. Scoring

4.2.10.1.1. Oxford Placement Test

The listening and grammar sections of the Oxford Placement Test each included 100 questions. One point was awarded for each correct response, resulting in a total score of 200.

4.2.10.1.2. OPT, WPT, and SOPT

The scoring process for the present progressive and the third person singular is illustrated in Table 21 and Table 22 respectively. For the present progressive items, three categories were used, inspired by Bardovi-Harlig (2000). In each obligatory context, participants received a maximum score of 2 points for correctly using the present progressive and 1 point for using the bare progressive. Any non-progressive forms were awarded 0 points. On the other hand, the scoring for the third person singular form was binary, with a score of 1 awarded for using the correct form in each required context (e.g., "He cooks", "He watches TV"), with 0 points for any non-third person singular form.

Agreement errors were disregarded for both target constructions. Also, vocabulary errors were ignored. For instance, if a different verb was used while using the correct present progressive form (e.g., "He is using a phone" instead of "He is charging a phone") or the correct third person singular form was accurate (e.g., "He uses a phone" instead of "He charges a phone"), participants received maximum points, indicating their acquisition of the morpheme production rule but its application to an exceptional case (cf., Dulay & Burt, 1974). Pronunciation errors were ignored as the focus of the study was on the development of morphology, not pronunciation. For example, if participants mispronounced "plays" as [pleis] instead of [pleiz], this error was ignored. Spelling errors were also overlooked. For example, when participants misspelled the morphemes (e.g., "He watchs TV"), it was still coded as the production of the correct form. It should also be noted that participants did not receive points for using the target forms in non-obligatory contexts, ensuring that no points were awarded for overusing the target forms (cf., Pica, 1983).

Each production test consisted of 18 target items (9 present progressive items and 9 third person singular) along with 20 distractors. Consequently, the maximum total score for

the target items (excluding distractors) was 18 for the present progressive (2 points × 9 items = 18) and 9 points for the third person singular (1 point × 9 = 9 points).

Table 21. Scoring System for Present Progressive

Present Progressive	
2 points	Use of correct form of the present progressive tense (e.g., He **is cooking**.)
1 point	Use of bare form of the present progressive tense (e.g., *He **cooking**.)
0 point	Use of any non-progressive form, such as 1) incorrect form of verb (e.g., *He **cook**.) 2) different tenses (e.g., He **cooks**.)

Table 22. Scoring System for Third Person Singular

Third Person Singular	
1 point	Use of correct form of the third person singular (e.g., He **cooks**; He **fishes**)
0 point	Use of any non-progressive form, such as 1) incorrect form of verb (e.g., *He **cook**.) 2) different tenses (e.g., He **is cooking**.)

4.2.10.2. Eye-movement Data Analysis

As explained, Tobii Pro Lab Software (Tobii Pro AB, 2014, Version 1.194) was used to analyze participants' eye-movement data. To ensure data quality, fixation durations shorter than 50ms were removed from the dataset based on previous studies (Conklin et al., 2018; Godfroid, 2019, 2020a). Drawing on previous eye-tracking research (e.g., Conklin et al., 2018; Conklin & Pellicer-Sánchez, 2016; Pellicer-Sánchez, 2016; Roberts & Siyanova-Chanturia, 2013; Winke, 2013; Wang & Pellicer-Sánchez, 2022), attention paid to the target

linguistic constructions was examined using five measures: total duration of fixations, first-pass reading time, re-reading time, visit count, and skip rate (Table 23).

Table 23. The Definitions of Five Eye-Tracking Measures for Study 2

Measurement	Definition
Total fixation duration	The sum of all fixation durations on an AOI within the defined TOI
First-pass reading time	The sum of the duration of all fixations on an AOI before exiting the AOI within the defined TOI
Re-reading time	The sum of the duration across all fixations made within an AOI when the eyes revisit the area after the first pass within the defined TOI
Visit count	Total number of visits made to an AOI within the defined TOI
Skip rate	An AOI is considered skipped (i.e., Skip=1) if no fixation occurred in the AOI within defined TOI

To analyze the participants' eye-movement data, two types of AOI were created in the written recasts: one for the third person singular and the other for the present progressive construction (Figure 24). The time duration of each AOIs was fixed at 10 seconds, the same time duration as that of the written recasts. The AOIs were manually drawn, with a fixed height of 100 pixels. The width of each AOI varied depending on the number of letters of each AOI, while ensuring consistent spacing around the text. Table 24 shows the average AOI sizes for each target construction across three treatment sessions. The results of a one-way ANOVA indicated that there were no significant differences in the AOI sizes for the target features across the three treatment sessions, third person singular: $F(2, 18) = .049$, $p = .952$, $\eta^2 = .005$; present progressive: $F(2, 18) = .248$, $p = .783$, $\eta^2 = .027$. TOIs were manually generated based on the presentation time of written recasts. Consistent with Montero Perez

(2019), only the eye-movement data that occurred within the AOIs and during the TOIs were included in the analysis.

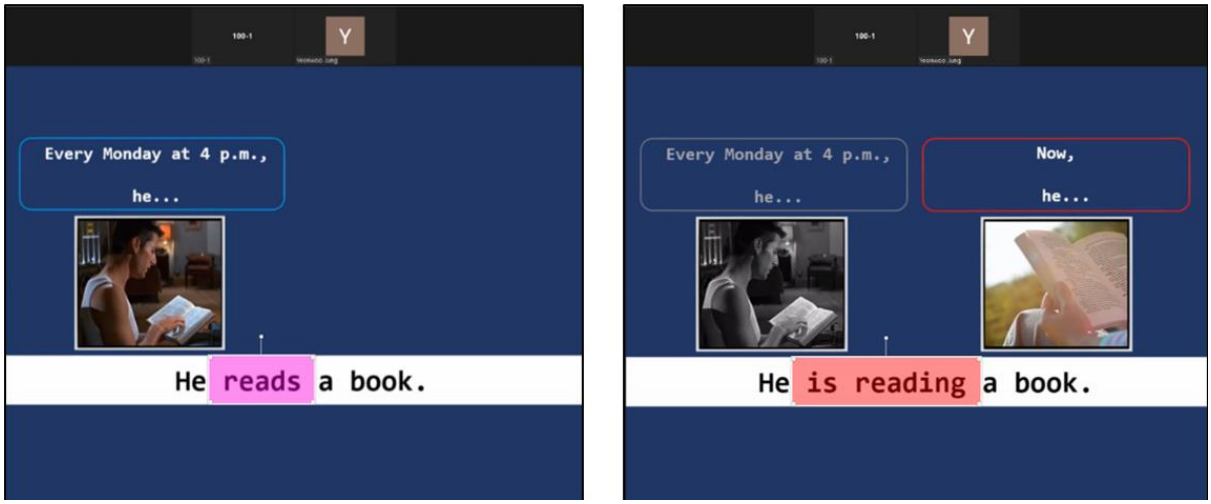


Figure 24. Examples of AOIs of Each Target Construction

Table 24. Average AOI Sizes of Each Target Construction in Each Treatment Session

		Third Person Singular		**Present Progressive**	
		M (SD)	95% CI	*M (SD)*	95% CI
AOI sizes (in pixels)	Time1	248.14 (24.11)	[225.85, 270.44]	402.29 (30.27)	[374.29, 430.29]
	Time2	243.57 (31.21)	[214.70, 272.44]	393.71 (21.83)	[373.52, 413.90]
	Time3	247.86 (35.31)	[215.20, 280.52]	393.14 (28.70)	[366.60, 419.69]
	Total	246.52 (29.08)	[233.29, 259.76]	396.38 (26.14)	[384.48, 408.28]

Note. Time1 = treatment session 1; Time2 = treatment session 2; Time3 = treatment session 3.

The eye-movement data analysis comprised three steps. Initially, total fixation duration, first-pass reading time, re-reading time, number of visits (or visit count), and skip rate were computed for each AOI and averaged for each participant, utilizing the raw mean values of each measure. Subsequently, to accommodate the varying sizes of AOIs across the target constructions and account for the likelihood of larger AOIs attracting more fixations

and visits, the eye-movement data was normalized. Normalization approach, inspired by prior research on subtitles (e.g., Bisson et al, 2014), involved dividing the attention measures by the number of letters in each AOI. Moreover, considering that participants received a different number of recasts, an additional normalization was performed by dividing the attention measurements by the respective number of recasts that each participant had received. The subsequent analyses of the five attention measurements were conducted using these normalized values.

4.2.10.3. Working Memory Measures

4.2.10.3.1. NWS Task

Each nonword that had been recalled was scored as either correct or incorrect. The span length was determined based on the maximum number syllables that participants correctly recalled at least twice for each syllable length, which ranged from 4 to 11.

4.2.10.3.2. Forward Corsi Block Task

To evaluate participants' performance in the forward Corsi block task, a total score was calculated by counting the number of correctly repeated sequences until the test ended. The total score system was selected, as it deemed more reliable than considering the block span alone because it accounts for performance across trials of equal length (Kessels et al., 2000).

4.2.10.3.3. Color Shape Task

The color shape task performance was assessed by measuring the switching cost, which was calculated based on the difference in mean reaction times between non-switching and switching blocks (e.g., Altgassen et al., 2014; Friedman et al., 2006; Gold et al., 2013;

Miyake et al., 2004). To ensure data quality, the reaction times of each participant were trimmed to exclude values that fell outside two standard deviations above and below the mean.

4.2.10.3.4. Stop Signal Task

Participants' performance on this task was assessed in terms of the stop signal reaction time (SSRT). SSRT refers to the time taken by individuals to inhibit their responses and thus serves as an indicator of their inhibitory ability. It is assumed that participants with slower SSRTs are more likely to exhibit difficulties in inhibiting their responses compared to those with faster SSRTs. SSRT values were also trimmed to values within two standard deviations above or below the mean stop signal reaction time for each participant (Congdon et al., 2012; Enticott, Ogloff, & Bradshaw, 2006).

4.2.10.3.5. AOSPAN Task

Previous research has used two scoring systems to calculate scores in the AOSPAN task; absolute AOSPAN score and total score. The absolute AOSPAN score only includes sets with all letters recalled correctly. For instance, if an individual correctly recalls 3 letters in a set size of 3, 4 letters in a set size of 4, and 3 in a set size of 5, their score would be 7 (3+4+0). The total score, on the other hand, reflects the total number of letters recalled in their respective positions within a given string. In this study, the total score was chosen as an index of AOSPAN performance, with a maximum score of 75. This choice is based on the suggestion that the total performance score better reflects WM span, which includes continuous variables (e.g., Friedman & Miyake, 2005; Miyake, 2001; Miyake et al., 1999).

4.2.11. Statistical Analyses

For Study 2, the initial statistical analyses were conducted using IBM SPSS 27.0 for Windows. These analyses aimed to examine the reliability of the tests and compute descriptive and correlational statistics for the data. Cronbach's alpha was employed to determine the internal consistency reliability of the different tests, while Cohen's kappa was used to assess inter-coder reliability. The comparability of the four groups in terms of the number of obligatory contexts created by the treatment tasks and the number of recasts provided during the treatment tasks were also inspected at the outset of the study. Spearman's coefficient was utilized to examine interrelationships between the various test scores. The significance level for Study 2 was set at an alpha level of $p < .05$. Following Plonsky and Oswald (2014), η^2 values of.01,.06, and.14 were considered as small, medium, and large respectively.

The main statistical analyses were carried out using R Version 4.2.1 (R Core Team, 2022), a statistical software. First, the comparability of the four groups in terms of the pretest scores of the three outcome measures was inspected at the outset of the study. Subsequently, a series of mixed-effects models were constructed to explore: (a) the impact of recasts on the development of L2 grammatical knowledge compared to no recasts, (b) the influence of recast modality on participants' development in the use of the target constructions, (c) the effects of written recasts versus multimodal recasts on learners' attentional allocation to the target constructions, and, (d) the moderating effect of WM on the relationship between the type of feedback modality and L2 development in grammatical knowledge.

Mixed-effects regression analyses were employed for the main analyses as they accommodate nested data and allowed for the inclusion of various fixed and random effects in the analyses (Baayen, 2008; Baayenet al., 2008; Cunnings, 2012). For example, unlike

means-based parametric statistical techniques such as t-test and ANOVA, mixed-effects models address differences across subjects and items, enabling generalization of research findings to a larger learner population and different linguistic materials (Baayen et al., 2008; Linck & Cunnings, 2015). In the current study, participants' responses to each test question and their eye-tracking data were collected, nested within each subject. Hence, adding participant and item as random effects could help to better generalize the study findings. Also, mixed-effects models are robust in regard to violations of homoscedasticity and missing data (Cunnings, 2012).

Linear, logistic, and generalized mixed-effects models (GLMMs) were constructed for the statistical analyses. The *lmer* or *glmer* function from the lme4 package (Version 1.1-21; Bates, Maechler, Bolker, & Walker, 2015) was used to construct these models based on the nature of the dependent variables (continuous, binary or count variable). All continuous variables were log-transformed or rescaled before being added to models to address the skewness problem (Godfroid, 2020a). Assumption tests, including the assessment of collinearity, normal distribution of residuals, and homoscedasticity, were performed for linear mixed-effects models using the sjPlot package (v 2.8.4; Lüdecke, 2020). For generalized linear mixed models, the glmmTMB package (v 1.0.1; Brooks et al., 2017) was utilized. The LmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) was employed to obtain *p*-values for linear mixed-effects models. Marginal and conditional R^2 values, measuring variance explained by fixed and random effects, were computed using the r.squared GLMM function in the MuMin package (Barton, 2020), with R^2 values of.06,.16, and.36 interpreted as small, medium, and large, respectively (Plonsky & Oswald, 2014).

Assumption tests for logistic mixed-effects models and GLMMs involved various diagnostic procedures. Collinearity was assessed by examining the correlation matrix or

variance inflation factors for predictor variables. The normal distribution assumption was evaluated using diagnostic plots specific to the type of residuals, such as Q-Q plots or histograms, and statistical tests like the Shapiro-Wilk test or Kolmogorov-Smirnov test. Homoscedasticity, indicating constant variance of residuals, was examined through residual plots and statistical tests such as Levene's test or the Breusch-Pagan test. The sjPlot package (v 2.8.4; Lüdecke, 2020) was employed for conducting these assumption tests in logistic mixed-effects models and GLMMs. These tests provided insights into the validity of the model assumptions and ensured the robustness of the statistical analyses.

Outlier detection procedures were applied to identify potential outliers in the data. Residual analysis, including the plotting of Pearson residuals against predicted values in linear mixed-effects models, was performed to assess the presence of outliers. For logistic mixed-effects models, deviance residuals or standardized residuals were examined. Data points deviating substantially from the expected pattern or displaying unusual patterns were flagged as potential outliers. Cook's distance, a measure of influence, was calculated to identify potentially influential outliers. Additionally, a model criticism approach using the romr.fnc() function from the LMERConvenienceFunctions package (Tremblay & Ransijn, 2020) was utilized to identify outliers. Data points with absolute standardized residuals exceeding 2.5 standard deviations ($|z| > 2.5$) were considered outliers. Sensitivity analyses were conducted by re-running the models with and without the identified outliers to assess their potential impact on the statistical significance of fixed effects. Outliers that significantly affected the fixed effects were excluded from the analyses. These outlier checking procedures ensured the validity and reliability of the statistical models and provided confidence in the interpretation of the results.

For logistic regression models, odds ratios (*ORs*) were used as effect-size measures (Field, Miles, & Field, 2012). An *OR* larger than one indicated a positive relationship, while an *OR* less than one represented a negative association. *ORs* greater than 3 or less than 0.33 were considered strong (Haddock, Rindskopf, & Shadish, 1998).

4.2.11.1. Mixed-Effects Models in Study 2

To address the research questions, mixed-effects models were constructed incorporating Group, Time and their interactions as fixed effects, while random intercepts for participant and item were included as random effects. The modeling began with constructing a null model that only contained the random intercepts of participant and item (e.g., WPT <- lmer(posttest ~ (1|Participant) + (1|Item), data = WPTdata). The best models were constructed using forward selection method and reported based on likelihood ratio tests using the *anova()* function and on Akaike information criterion (AIC) scores. Maximal random-effects structure was adopted as recommended for confirmatory hypothesis testing research to strengthen the generalization of findings (Linck & Cunnings, 2015). Each fixed effect was entered into the null model step-wise (e.g., WPT <- lmer(posttest ~ Group + (1|Participant) + (1Item), data = WPTdata)), and they were kept in the model only if their inclusion significantly improved the model fit according to AIC scores. Additionally, maximal models (Barr et al., 2013) with random slopes for the fixed effects were considered. If convergence issues arose, the models were refit using the optimizing function control = glmerControl(optimizer = "bobyqa") suggested by Linck and Cunnings (2015). If convergence was still not achieved, random effects accounting for the least variance were removed until convergence was attained (Blom, Paradis, & Sorenson Duncan, 2012; Cunnings & Sturt, 2014). The significance of fixed effects was determined using $\chi 2$ statistics.

To investigate the impact of recasts on L2 grammatical development compared to the absence of recasts (RQ1), and the extent to which the modality of recasts affected participants' development in the use of the target constructions (RQ2), a null model was first constructed for each outcome measure. The null model only included test scores as dependent variables and random effects (i.e., participant and item). To the null model, fixed effects (i.e., Group, Time, and their interaction) were added. The models including the fixed effects were then compared to the null model using $\chi 2$ statistics. If a significant difference was revealed, the fixed effects were interpreted as having a significant relationship with the dependent variable. After identifying the fixed effects that improved the null models significantly, maximal random effects structures, including fixed effects and random effects as well as random slope for each fixed effect, were constructed (Barr et al., 2013). However, as mentioned earlier, if the maximal random structure models failed to converge, the random effects that accounted for the least variance were removed from the model (Cunnings & Sturt, 2014) until convergence was achieved. For the continuous dependent variables (e.g., scores of the present progressive in the SOPT, OPT, and WPT), linear mixed-effects models that reached convergence were used for subsequent analyses. For the binary dependent variables, (e.g., scores of the third person singular in the same three outcome measures), logistic mixed-effects models that reached convergence were used for subsequent analyses to determine whether there was an overall Group-by-Time interaction effect. If a significant overall interaction effect was revealed, post hoc analyses were performed for pairwise comparisons, with the post hoc models having the same structure but including data from only two testing times for two of the groups.

To address RQ3, which focused on examining the impact of multimodal recasts and written recasts in directing learners' attention to the target structures, a series of mixed-effects

models were formulated based on the nature of the dependent variables. For continuous dependent variables (total fixation duration, first-pass reading time, re-reading time), linear mixed-effects models were employed. In the case of normalized count and binary variables (visit count and skip rate, respectively), GLMMs with an inverse Gaussian family and log link function were constructed. Each of these models considered attention measurement as a dependent variable. These models aimed to explore the associations between the dependent variables (eye-tracking measures) and the independent variables (i.e., Group, Time and their interaction), all while accounting for random intercepts within participants. Since each participant was presented with the same set of target items but received recasts exclusively for the items that they had used incorrectly, the distribution of recasts varied among participants, leading to different items being addressed. Consequently, due to the nature of this individualized recast process, the variable "item" was not included as a random effect in the models.

Several reasons underpin the selection of GLMMs with an inverse Gaussian family and log function for analyzing both the visit count and the skip rate. As previously mentioned, the normalization process aimed to address the variation stemming from disparities in letter counts of AOIs and the frequency of recasts. This process resulted in the conversion of a count variable (visit count) into a continuous proportional count variable, thus shifting it away from its discrete integer count nature. Given this transformation and the intrinsic continuous proportional characteristics of the variable, it was crucial to identify an appropriate statistical distribution. The choice of the inversion Gaussian distribution was motivated by its suitability for modeling continuous data, offering a superior fit compared to alternative discrete count distributions such as Poisson or negative binomial.

It is important to note that the discussion so far has focused solely on the normalized count variable (visit count). In the case of the normalized binary variable (skip rate), a comparative rationale for the chosen GLMM framework was essential. Despite its binary nature, the skip rate underwent normalization to account for differences in AOI letter counts and recast frequencies. This normalization process transformed the binary variable into a continuous proportional variable. This new representation retained its underlying relationship with other variables, thus allowing for the application of statistical models suitable for continuous data. Similar to the visit count, the skip rate exhibited continuous proportional characteristics after normalization. Hence, the choice of the inverse Gaussian distribution and log link function remained appropriate, as they effectively captured the proportional relationship between predictors and outcome variables, aligning well with the characteristics of the transformed skip rate variable.

Moreover, both visit count and skip rate displayed correlations within clusters, such as multiple visits and skips by the same participant. This clustering effect rendered GLMMs an apt choice, as they could accommodate correlated data while encompassing both within-subject and between-subject factors in the analyses. Taken together, the distinctive attributes of these two variables provide a robust rationale for adopting GLMMs with an inverse Gaussian distribution for the analysis of both the visit count and the skip rate.

For RQ4, which examined the three-way interaction effects of Group, Time, and five attention measurements on the gains of the three outcome measures, a series of linear mixed-effects models (for the present progressive) and logistic mixed-effects models (for the third person singular) were conducted. The models that were originally constructed for the first and second research questions, which addressed the impact of the provision and modality of recasts on L2 grammatical knowledge development, were used as the null models. These null

models included Group and Time as fixed effects, test scores as dependent variables, and participant and item as random effects. To these models, each attention measure was added as an additional fixed effect. A significant three-way interaction would indicate that the extent of gains achieved by the multimodal recast and written recast groups was differentially influenced by the attention measure in each model. Each of these models was first constructed with a maximal random effects structure; that is the models consisted of random effects (group and item) and random slopes for each fixed effect. Those models that failed to converge were restructured by removing the random effects that accounted for the least variance and used for subsequent analyses.

For RQ5, which addressed the impact of individual differences in WM capacity on the effects of oral, written, or multimodal recasts on development in L2 grammatical knowledge, the same procedure conducted for RQ4 was performed. The predictors of interest were the three-way interactions between Group, Time, and WM measures. As with RQ4, a significant three-way interaction would indicate that the extent of gains achieved by the four groups on their respective assessment tasks was differently influenced by the component of WM in the model.

For RQ6, which explored the relationship between participants' attention to the target constructions and their WM capacity, Spearman correlation analysis was used to examine the relationship between the five attention measures (i.e., total fixation duration, first-pass reading time, re-reading time, visit count, and skip rate) and indices of the five WM measures (i.e., NWS task, forward Corsi block task, color switching task, stop signal task, and AOSPAN task). The magnitude of the Spearman correlation coefficient (*rho*) was interpreted as follows: values of 0.10 to 0.29 were considered small, 0.30 to 0.49 were considered medium, and coefficients larger than 0.50 were considered large (Cohen, 1988).

4.3. Analyses and Results

4.3.1. Preliminary Analyses

4.3.1.1. Internal Consistency Reliability

Table 25 provides an overview of the Cronbach's alpha values, which were calculated to assess the internal consistency reliability of the three versions of each assessment test. Concerning the present progressive form, the internal consistency reliability for all three versions of each outcome measure fell within the moderate and good range. Regarding the third person singular form, all three versions of each outcome measure were also found to be reliable.

Table 25. Internal Consistency Reliability

Assessment Test	Present Progressive	Third Person Singular
OPT – Version A	.791	.787
OPT – Version B	.824	.806
OPT – Version C	.800	.821
WPT – Version A	.848	.802
WPT – Version B	.851	.809
WPT – Version C	.832	.817
SOPT – Version A	.808	.800
SOPT – Version B	.806	.799
SOPT – Version C	.792	.803

4.3.1.2. Inter-coder Agreement

The three assessment tests – the OPT, WPT, and SOPT – were coded and scored by the researcher and a second coder. The second coder, who holds a master's degree in applied linguistics, assessed 30% of the three assessment tests. Inter-coder agreement was determined by calculating Cohen's kappa. The results of the inter-coder agreement analyses are presented in Table 26. The inter-coder agreement for the OPT was .98, with a kappa value of .98; for the WPT, it was 1.00, with a kappa value of 1.00; and for the SOPT, it was.99, with a kappa

value of.99. These values indicate a high level of agreement, beyond chance, between the two raters.

Table 26. Inter-coder Agreement on Three Assessment Tasks

Assessment Tasks	Inter-coder Agreement	Cohen's kappa
OPT	0.98	0.98
WPT	1.00	1.00
SOPT	0.99	0.99

4.3.1.3. Number of Obligatory Contexts and Recasts

Table 27 presents descriptive statistics for the number of obligatory contexts created by the treatment tasks for the present progressive and the third person singular during the treatment sessions. Overall, the average number of obligatory contexts remained consistent across all four groups. Table 28 shows descriptive statistics for the number of recasts provided in each treatment session for the two target constructions. Across all experimental groups, there was a decrease in the mean number of recasts from the first to the third treatment sessions, indicating that participants committed fewer errors over time. This tendency was most pronounced in the multimodal recast group, followed by the written recast group and the oral recast group in that order. The results of one-way ANOVAs for the mean number of obligatory contexts and recasts, as presented in Table 29, confirmed that there were no significant differences in the number of obligatory contexts and recasts among the groups during any session for either of the two target constructions.

Table 27. Descriptive Statistics for the Number of Obligatory Contexts Across Treatment Sessions

					95% CI	
	Group	*N*	*M*	*SD*	*Lower*	*Upper*
Present Progressive						
	Control	30	6.70	0.70	6.44	6.96
	Oral	30	6.73	0.52	6.54	6.93
Session 1	Written	30	6.77	0.57	6.55	6.98
	Multi	30	6.67	0.66	6.42	6.91
	Control	30	6.77	0.50	6.58	6.95
	Oral	30	6.80	0.48	6.62	6.98
Session 2	Written	30	6.83	0.46	6.66	7.01
	Multi	30	6.87	0.43	6.70	7.03
	Control	30	6.83	0.38	6.69	6.97
	Oral	30	6.93	0.25	6.84	7.03
Session 3	Written	30	6.87	0.35	6.74	7.00
	Multi	30	6.90	0.31	6.79	7.01
	Control	30	20.30	0.92	19.96	20.64
	Oral	30	20.47	0.73	20.19	20.74
Total	Written	30	20.47	0.82	20.16	20.77
	Multi	30	20.43	0.73	20.16	20.71
Third Person Singular						
	Control	30	6.77	0.57	6.55	6.98
	Oral	30	6.67	0.61	6.44	6.89
Session 1	Written	30	6.53	0.63	6.30	6.77
	Multi	30	6.70	0.53	6.50	6.90
	Control	30	6.77	0.50	6.58	6.95
	Oral	30	6.93	0.25	6.84	7.03
Session 2	Written	30	6.80	0.41	6.65	6.95
	Multi	30	6.90	0.31	6.79	7.01
	Control	30	6.90	0.31	6.79	7.01
	Oral	30	6.80	0.41	6.65	6.95
Session 3	Written	30	6.83	0.38	6.69	6.97
	Multi	30	6.87	0.35	6.74	7.00
	Control	30	20.43	0.73	20.16	20.71
	Oral	30	20.37	1.03	19.98	20.75
Total	Written	30	20.17	1.17	19.75	20.58
	Multi	30	20.47	0.97	20.10	20.53

Note. Control = Control group; Oral = Oral recast group; Written = Written recast group; Multi = Multimodal recast group; the maximum number of obligatory contexts for each target form per each treatment task was 7, resulting in a total of 21 obligatory contexts for each target form over the three treatment sessions.

Table 28. Descriptive Statistics for the Number of Recasts Across Treatment Sessions

					95% CI	
	Group	*N*	*M*	*SD*	*Lower*	*Upper*
Present Progressive						
	Oral	30	5.03	0.49	4.85	5.22
Session 1	Written	30	4.97	0.61	4.74	5.20
	Multi	30	4.87	0.51	4.68	5.06
	Oral	30	3.97	0.56	3.76	4.17
Session 2	Written	30	3.90	0.66	3.65	4.15
	Multi	30	3.63	0.56	3.43	3.84
	Oral	30	3.03	0.49	2.85	3.22
Session 3	Written	30	2.97	0.41	2.81	3.12
	Multi	30	2.77	0.50	2.82	3.02
	Oral	30	12.03	1.40	11.51	12.57
Total	Written	30	11.83	1.51	11.27	12.40
	Multi	30	11.27	1.31	10.78	11.76
Third Person Singular						
	Oral	30	5.37	0.49	5.18	5.55
Session 1	Written	30	5.33	0.48	5.15	5.51
	Multi	30	5.23	0.43	5.07	5.39
	Oral	30	4.07	0.58	3.84	4.28
Session 2	Written	30	4.03	0.61	3.80	4.26
	Multi	30	3.80	0.55	3.59	4.01
	Oral	30	3.03	0.49	2.85	3.22
Session 3	Written	30	3.00	0.37	2.86	3.14
	Multi	30	2.83	0.38	2.69	2.97
	Oral	30	12.47	1.20	12.02	12.91
Total	Written	30	12.37	1.03	11.98	12.75
	Multi	30	11.87	1.07	11.47	12.27

Note. The maximum number of recasts for each target form per each treatment task was 7, resulting in a total of 21 recasts for each target form across the three treatment sessions.

Table 29. Results of One-way ANOVAs on the Number of Obligatory Contexts and Recasts Across Treatment Sessions

		df	*F*	*p*	η_p^2
Number of Obligatory Contexts					
Present Progressive	Session 1	3, 116	0.15	.932	0.00
	Session 2	3, 116	0.25	.861	0.01
	Session 3	3, 116	0.53	.664	0.01
	Total	**3, 116**	**0.29**	**.830**	**0.01**
Third Person Singular	Session 1	3, 116	0.86	.464	0.02
	Session 2	3, 116	1.31	.275	0.03
	Session 3	3, 116	0.43	.735	0.01
	Total	**3, 116**	**0.57**	**.635**	**0.02**
Number of Recasts					
Present Progressive	Session 1	2, 87	0.72	.488	0.02
	Session 2	2, 87	2.65	.076	0.06
	Session 3	2, 87	2.60	.080	0.06
	Total	**2, 87**	**2.39**	**.098**	**0.05**
Third Person Singular	Session 1	2, 87	0.72	.488	0.02
	Session 2	2, 87	2.65	.076	0.06
	Session 3	2, 87	2.60	.080	0.06
	Total	**2, 87**	**2.39**	**.098**	**0.05**

4.3.1.4. Comparability of the Groups

Table 30 and Table 31 present a summary of descriptive statistics depicting each group's performance on the OPT, WPT, and the SOPT. These statistics pertain to the present progressive and third person singular forms, respectively.

Table 30. Descriptive Statistics for the Pretest – Present Progressive

					95% CI	
Test	Group	*N*	*M*	*SD*	*Lower*	*Upper*
OPT	Control	30	3.10	0.61	2.87	3.33
	Oral	30	3.03	0.76	2.74	3.32
	Written	30	3.07	0.45	2.90	3.23
	Multi	30	3.00	0.59	2.78	3.21
WPT	Control	30	3.13	0.63	2.90	3.37
	Oral	30	3.07	0.74	2.79	3.34
	Written	30	3.10	0.48	2.92	3.28
	Multimodal	30	3.03	0.56	2.83	3.24
SOPT	Control	30	2.97	0.67	2.72	3.22
	Oral	30	2.93	0.78	2.64	3.23
	Written	30	2.86	0.63	2.63	3.10
	Multi	30	2.83	0.65	2.59	3.08

Note. OPT max. score = 18; WPT max. score = 18; SOPT max. score = 18.

Table 31. Descriptive Statistics for the Pretest – Third Person Singular

					95% CI	
Test	Group	*N*	*M*	*SD*	*Lower*	*Upper*
OPT	Control	30	1.37	0.67	1.12	1.62
	Oral	30	1.27	0.52	1.07	1.46
	Written	30	1.33	0.55	1.13	1.54
	Multi	30	1.30	0.54	1.10	1.50
WPT	Control	30	1.40	0.62	1.17	1.63
	Oral	30	1.33	0.55	1.13	1.54
	Written	30	1.37	0.56	1.16	1.57
	Multi	30	1.30	0.60	1.08	1.52
SOPT	Control	30	1.33	0.66	1.09	1.58
	Oral	30	1.20	0.48	1.02	1.38
	Written	30	1.30	0.47	1.13	1.47
	Multi	30	1.27	0.52	1.07	1.46

Note. OPT max. score = 9; WPT max. score = 9; SOPT max. score = 9.

To examine whether there were differences among the four groups in terms of their performance on the OPT, WPT, and SOPT at the outset of the study, a series of linear mixed-effects model for the present progressive items and logistic mixed-effects models for the third person singular respectively were constructed. All models incorporated group as a fixed

effect, while participant and item served as random effects. The participants' scores in the relevant pretest were used as the dependent variable in each model. As presented in Tables 32–37, the results of the analyses revealed no significant difference in the four groups' performance, indicating that their pretest scores were comparable.

Table 32. Results for the Linear Mixed-effects Model Examining Performance on the OPT Pretest – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control group as reference level										
(Intercept)	−1.56	0.07	−22.50	<.001***	<.01	<.01	<.01	<.01	<.01	<.01
Oral	0.02	0.12	−0.15	.880						
Written	0.01	0.11	−0.08	.935						
Multi	0.01	0.10	0.12	.906						
Oral group as reference level										
(Intercept)	−1.58	0.09	−16.89	<.001***	<.01	.01	<.01	<.01	.04	.19
Control	0.02	0.12	0.14	.880						
Written	0.01	0.10	0.09	.932						
Multi	0.03	0.10	0.29	.774						
Written group as reference level										
(Intercept)	−1.57	0.08	−18.91	<.001***	<.01	.01	<.01	<.01	.02	.14
Control	0.01	0.11	0.08	.935						
Oral	−0.01	0.10	−0.09	.932						
Multi	0.02	0.10	0.21	.831						
Formula: *lmer(log_OPT_Pretest_PP ~ 1 + Group + (1|Participant) + (Group|Item))*										

Table 33. Results for the Linear Mixed-effects Model Examining Performance on the WPT Pretest – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control group as reference level										
(Intercept)	−1.55	0.07	−22.30	<.001***	<.01	<.01	<.01	<.01	<.01	<.01
Oral	0.02	0.12	−0.15	.889						
Written	0.01	0.11	−0.08	.937						
Multi	0.01	0.11	0.12	.909						
Oral group as reference level										
(Intercept)	−1.57	0.09	−17.01	<.001***	<.01	.01	<.01	<.01	.03	.18
Control	0.02	0.12	0.15	.889						
Written	0.01	0.11	0.08	.936						
Multi	0.03	0.10	0.29	.775						
Written group as reference level										
(Intercept)	−1.56	0.08	−19.04	<.001***	<.01	.01	<.01	<.01	.02	.13
Control	0.01	0.11	0.08	.937						
Oral	−0.01	0.11	−0.08	.936						
Multi	0.02	0.10	0.21	.831						
Formula: *lmer(log_WPT_Pretest_PP ~ 1 + Group + (1|Participant) + (Group|Item))*										

Table 34. Results for the Linear Mixed-effects Model Examining Performance on the SOPT Pretest – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control group as reference level										
(Intercept)	−1.58	0.07	−21.36	<.001***	<.01	<.01	<.01	<.01	<.01	<.01
Oral	−0.03	0.10	−0.32	.794						
Written	−0.01	0.10	−0.14	.900						
Multi	−0.00	0.10	−0.03	.976						
Oral group as reference level										
(Intercept)	−1.61	0.09	−17.21	<.001***	<.01	.01	<.01	<.01	.04	.19
Control	0.03	0.12	0.25	.809						
Written	0.02	0.10	0.16	.874						
Multi	0.03	0.11	0.26	.798						
Written group as reference level										
(Intercept)	−1.59	0.08	−19.05	<.001***	<.01	.01	<.01	<.01	.02	.15
Control	0.01	0.11	0.12	.904						
Oral	−0.02	0.11	−0.16	.874						
Multi	0.01	0.10	0.11	.915						
Formula: *lmer(log_SOPT_Pretest_PP ~ 1 + Group + (1|Participant) + (Group|Item))*										

Table 35. Results for the Linear Mixed-effects Model Examining Performance on the OPT Pretest – Third Person Singular

	Fixed effects					Random effects			
						by participant		*by item*	
	Estimate	*SE*	*z*	*p*	*OR*	*variance*	*SD*	*variance*	*SD*
Control group as reference level									
(Intercept)	–1.81	0.24	–7.42	<.001***	0.16	<.01	<.01	.26	.51
Oral	–0.09	0.25	−0.37	.709	0.91				
Written	–0.03	0.24	–0.12	.902	1.03				
Multi	–0.06	0.25	–0.25	.804	0.94				
Oral group as reference level									
(Intercept)	–1.90	0.25	–7.68	<.001***	0.15	<.01	<.01	.26	.51
Control	–0.09	0.25	−0.37	.709	0.91				
Written	–0.06	0.25	–0.25	.802	0.94				
Multi	–0.03	0.25	–0.13	.900	1.03				
Written group as reference level									
(Intercept)	–1.84	0.25	–7.51	<.001***	0.16	<.01	<.01	.26	.51
Control	–0.03	0.24	−0.12	.902	1.03				
Oral	–0.06	0.25	–0.25	.802	0.94				
Multi	–0.03	0.25	–0.12	.901	0.97				
Formula: glmer(OPT_Pretest_PS ~ 1 + Group + (1|Participant) + (Group|Item))									

Table 36. Results for the Linear Mixed-effects Model Examining Performance on the WPT Pretest – Third Person Singular

	Fixed effects					Random effects			
						by participant		*by item*	
	Estimate	*SE*	*z*	*p*	*OR*	*variance*	*SD*	*variance*	*SD*
Control group as reference level									
(Intercept)	–1.76	0.23	–7.67	<.001***	0.17	<.01	<.01	.21	.46
Oral	–0.06	0.24	−0.24	.807	0.94				
Written	–0.03	0.24	–0.12	.903	0.97				
Multi	–0.09	0.24	–0.37	.713	0.91				
Oral group as reference level									
(Intercept)	–1.82	0.23	–7.84	<.001***	0.16	<.01	<.01	.21	.46
Control	–0.06	0.24	−0.24	.807	1.06				
Written	–0.03	0.24	–0.12	.902	1.03				
Multi	–0.03	0.25	–0.12	.901	0.97				
Written group as reference level									
(Intercept)	–1.79	0.23	–7.76	<.001***	0.17	<.01	<.01	.21	.46
Control	–0.03	0.24	−0.12	.903	1.03				
Oral	–0.03	0.24	–0.12	.902	0.97				
Multi	–0.06	0.24	–0.25	.805	0.94				
Formula: glmer(WPT_Pretest_PS ~ 1 + Group + (1|Participant) + (Group|Item))									

Table 37. Results for the Linear Mixed-effects Model Examining Performance on the SOPT Pretest – Third Person Singular

	Fixed effects					Random effects			
						by participant		*by item*	
	Estimate	*SE*	*z*	*p*	*OR*	*variance*	*SD*	*variance*	*SD*
Control group as reference level									
(Intercept)	–1.87	0.24	–7.68	<.001***	0.15	<.01	<.01	.25	.50
Oral	–0.10	0.25	−0.38	.703	0.91				
Written	–0.06	0.25	–0.25	.802	0.94				
Multi	–0.03	0.25	–0.13	.900	0.97				
Oral group as reference level									
(Intercept)	–1.96	0.25	–7.94	<.001***	0.14	<.01	<.01	.25	.50
Control	–0.10	0.25	−0.38	.703	0.91				
Written	–0.10	0.25	–0.38	.703	1.10				
Multi	–0.06	0.25	–0.26	.798	1.07				
Written group as reference level									
(Intercept)	–1.84	0.25	–7.51	<.001***	0.16	<.01	<.01	.26	.51
Control	–0.03	0.24	−0.12	.902	1.03				
Oral	–0.06	0.25	–0.25	.802	0.94				
Multi	–0.03	0.25	–0.12	.901	0.97				
Formula: glmer(SOPT_Pretest_PS ~ 1 + Group + (1|Participant) + (Group|Item))									

4.3.2. Effects of Recasts and Feedback Modality on Development in L2 Grammatical Knowledge (RQs 1–2)

4.3.2.1. OPT

4.3.2.1.1. Present Progressive

Table 38 summarizes the descriptive statistics for the participants' pretest, posttest, and delayed posttest scores on the OPT. All four groups exhibited pretest–posttest gains with the multimodal group achieving the highest gain scores. On the delayed posttest, all groups showed some declines in their scores.

Table 38. Descriptive Statistics for the OPT Scores – Present Progressive

					95% CI	
		M	*Mean Gain*	*SD*	*Lower*	*Upper*
Control (*N* = 30)	Pretest	3.10	–	0.61	2.87	3.33
	Posttest	3.43	0.33	0.50	3.25	3.62
	Delayed posttest	3.37	0.27	0.72	3.10	3.63
Oral (*N* = 30)	Pretest	3.03	–	0.76	2.74	3.32
	Posttest	9.20	6.17	0.76	8.92	9.48
	Delayed posttest	8.37	5.34	0.67	8.12	8.62
Written (*N* = 30)	Pretest	3.07	–	0.45	2.90	3.23
	Posttest	9.80	6.73	0.76	9.52	10.08
	Delayed posttest	9.00	5.93	0.87	8.67	9.33
Multimodal (*N* = 30)	Pretest	3.00	–	0.59	2.78	3.21
	Posttest	15.17	12.17	1.18	14.73	15.61
	Delayed posttest	14.68	11.68	0.84	14.35	14.98

Note. The maximum score was 18.

The linear mixed-effects model analyses conducted to examine the effect of Group and Time on participants' performance on the OPT identified 10 significant Time-by-Group interactions (see Appendix A21 for full model results). To further investigate group differences, Table 39 presents a summary of post hoc pairwise comparisons concerning these significant interaction effects. The results revealed that all experimental groups displayed greater pretest–posttest and pretest–delayed posttest gains than the control group. As reflected in the effect sizes, the multimodal recast group attained the largest gains compared to the control group, followed by the written recast and oral recast groups in that order. The advantages of all recast groups over the control group were retained during the delayed posttest.

Moving on to modality differences, the multimodal recast group demonstrated more substantial gains from pretest to posttest in comparison to both the oral recast and written recast groups. This distinction was substantiated by large effect sizes. However, no significant difference emerged between the oral recast and written recast groups. Similarly,

during the delayed posttest, the multimodal recast group exhibited superior performance compared to both the unimodal recast groups. In general, the effect sizes for all these significant interactions were larger for pretest–posttest than pretest–delayed posttest gains. Overall, these findings underscore the beneficial impact of recasts provided during video-conferencing tasks on participants' progress in the OPT, with multimodal recasts yielding the greatest improvement.

Table 39. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed−Effects Models Examining Participants' Scores on the OPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control:Oral:Time2	0.95	0.14	6.71	<.001	.12	.14	<.01	<.01	<.01	<.01
Control:Written:Time2	1.05	0.14	7.42	<.001	.14	.17	<.01	<.01	<.01	.02
Control:Multi:Time2	1.86	0.14	13.18	<.001	.39	.39	<.01	<.01	<.01	<.01
Control:Oral:Time3	0.85	0.14	6.04	<.001	.09	.11	<.01	<.01	<.01	<.01
Control:Written:Time3	0.98	0.14	7.01	<.001	.12	.14	<.01	<.01	<.01	.10
Control:Multi:Time3	1.82	0.14	12.93	<.001	.37	.38	<.01	<.01	.03	.19
Oral:Multi:Time2	0.91	0.14	6.48	<.001	.36	.38	<.01	<.01	.07	.26
Written:Multi:Time2	0.81	0.14	5.77	<.001	.36	.39	<.01	<.01	.09	.30
Oral:Multi:Time3	0.97	0.14	6.88	<.001	.36	.36	<.01	<.01	<.01	.09
Written:Multi:Time3	0.83	0.14	5.91	<.001	.33	.34	<.01	<.01	.04	.20
*Formula: lmer(log_OPT_PP ~ 1 + Group*Time + (1|Participant) + (Group|Item))*										

Note. Time1 = pretest (reference level); Time2 = immediate posttest; Time3 = delayed posttest.

4.3.2.1.2. Third Person Singular

Table 40 presents the descriptive statistics for the participants' pretest, posttest, and delayed posttest scores of the third person singular on the OPT. Again, all four groups demonstrated gains from pretest to posttest, with the multimodal recast group attaining the greatest gain scores. On the delayed posttest, all groups showed some reduction.

Table 40. Descriptive Statistics for the OPT Scores – Third Person Singular

					95% CI	
		M	*Mean Gain*	*SD*	*Lower*	*Upper*
Control (*N* = 30)	Pretest	1.37	–	0.67	1.12	1.62
	Posttest	1.40	0.03	0.68	1.15	1.65
	Delayed posttest	1.40	0.03	0.62	1.17	1.63
Oral (*N* = 30)	Pretest	1.27	–	0.52	1.07	1.46
	Posttest	3.30	2.03	0.60	3.08	3.52
	Delayed posttest	2.23	0.96	0.86	1.91	2.55
Written (*N* = 30)	Pretest	1.33	–	0.55	1.13	1.54
	Posttest	3.53	2.20	0.73	3.26	3.81
	Delayed posttest	2.37	1.04	0.67	2.12	2.62
Multimodal (*N* = 30)	Pretest	1.30	–	0.54	1.10	1.50
	Posttest	7.03	5.73	0.72	6.77	7.30
	Delayed posttest	5.93	4.63	0.79	5.64	6.23

Note. The maximum score was 9.

The logistic mixed-effects models employing a binomial family were conducted to examine the impact of Group and Time on participants' learning about the third person singular. These models revealed 10 significant interactions between Group and Time (see Appendix A22 for full model results). The results of the post hoc contrasts for these significant interactions are presented in Table 41. Similar to the findings observed for the present progressive form, all recast groups demonstrated significantly larger gains both in pretest–posttest and pretest–delayed posttest than the control group. These differences were further substantiated in odds ratios, indicating that, in comparison to the control group, the multimodal recast group achieved the most substantial pretest–posttest gains (*OR* = 22.03, *p* <.001), followed by the written recast group (*OR* = 3.77, *p* <.001), and the oral recast group (*OR* = 3.52, *p* <.001). The multimodal recast group also demonstrated the greatest pretest–delayed posttest gains as compared to the control group (*OR* = 11.76, *p* <.001), followed by the written recast group (*OR* = 2.02, *p* =.032), and the oral recast group (*OR* = 1.98, *p* =.038). There was no significant difference between the oral recast and written recast groups. Turning to the effects

of different modalities, the multimodal recast group exhibited superior gains in both pretest–posttest and pretest–delayed posttest compared to the oral recast and written recast groups. The odds of experiencing improvements in OPT gains were 6.26 times higher in the multimodal recast group compared to the oral recast group, and 5.85 times higher than for the written recast group in the posttest. Similarly, in the delayed posttest, the odds of improvements were 5.94 times higher for the multimodal recast group compared to the oral recast group, and 5.83 times higher compared to the written recast group. However, no significant differences in gains were observed between the two unimodal recast groups in the posttest or delayed posttest. In summary, the presence of recasts, regardless of their modality, played a pivotal role in advancing participants' proficiency in using the third person singular form, as demonstrated by their performance on the OPT. Notably, multimodal recasts were more effective in facilitating these gains than their oral and written counterparts.

Table 41. Results of Post Hoc Contrasts for Significant Interactions Identified by Logistic Mixed−Effects Models Examining Participants' Scores on the OPT – Third Person Singular

	Fixed effects						Random effects			
							by participant (intercept)		by item (intercept)	
	Estimate	*SE*	*z*	*p*	*OR*	95% CI	*Variance*	*SD*	*Variance*	*SD*
ControlOral:Time2	1.26	0.32	3.88	<.001	3.52	[1.87, 6.62]	<.01	<.01	.28	.52
ControlWritten:Time2	1.33	0.32	4.12	<.001	3.77	[2.01, 7.05]	<.01	<.01	.17	.41
ControlMulti:Time2	3.09	0.33	9.30	<.001	22.03	[11.53, 42.12]	<.01	<.01	.07	.27
ControlOral:Time3	0.68	0.33	2.07	.038	1.98	[1.04, 3.77]	<.01	<.01	.32	.57
ControlWritten:Time3	0.70	0.33	2.15	.032	2.02	[1.07, 3.82]	<.01	<.01	.12	.35
ControlMulti:Time3	2.46	0.33	7.60	<.001	11.76	[6.25, 22.13]	<.01	<.01	.09	.30
OralMulti:Time2	1.83	0.32	5.81	<.001	6.26	[3.38, 11.60]	<.01	<.01	.24	.49
WrittenMulti:Time2	1.77	0.31	5.64	<.001	5.85	[3.18, 10.78]	<.01	<.01	.06	.25
OralMultimodal:Time3	1.78	0.31	5.68	<.001	5.94	[3.22, 10.96]	<.01	<.01	.22	.47
WrittenMulti:Time3	1.76	0.31	5.68	<.001	5.83	[3.18, 10.69]	<.01	<.01	.09	.31
*Formula: glmer(OPT_PS ~ 1+ Group*Time + (1|Participant) + (Group|Item))*										

4.3.2.2. WPT

4.3.2.2.1. Present Progressive

The extent to which recasts and feedback modality influenced the development in the use of the present progressive was also examined by the WPT. As presented in Table 42, the mean gain scores of the experimental groups were considerably higher than that of the control group. Among the experimental groups, the mean gain score of the multimodal recast group was greater than that of the written recast group, which, in turn, was higher than that of the oral recast group from pretest to posttest as well as from pretest to delayed posttest. On the delayed posttest, however, all four groups displayed slight decreases compared to the posttest.

Table 42. Descriptive Statistics for the WPT Scores – Present Progressive

					95% CI	
		M	*Mean Gain*	*SD*	*Lower*	*Upper*
Control (*N* = 30)	Pretest	3.13	–	0.63	2.90	3.37
	Posttest	3.53	0.40	0.51	3.34	3.72
	Delayed posttest	3.40	0.27	0.50	3.21	3.59
Oral (*N* = 30)	Pretest	3.07	–	0.74	2.79	3.34
	Posttest	9.57	6.50	0.58	9.35	9.78
	Delayed posttest	8.50	5.43	0.68	8.25	8.75
Written (*N* = 30)	Pretest	3.10	–	0.48	2.92	3.28
	Posttest	10.10	7.00	0.84	9.78	10.41
	Delayed posttest	9.10	6.00	0.88	8.77	9.43
Multimodal (*N* = 30)	Pretest	3.03	–	0.56	2.83	3.24
	Posttest	15.80	12.77	1.40	15.28	16.32
	Delayed posttest	15.03	12.00	0.93	14.69	15.38

Note. The maximum score was 18.

The linear mixed-effects carried out to examine the effect of group on participants' performance in the WPT identified 10 significant Time-by-Group interactions (see Appendix A23 for full model results). Table 43 summarizes the results from post hoc pairwise contrasts

for these significant interactions. The results revealed that all recast groups displayed greater pretest–posttest gains compared to the control group. Effect sizes further support these findings, with the multimodal recast group achieving the largest gains, with a large effect size, followed by the written recast and oral recast groups with small- to medium-range effect sizes. This same pattern of advantages of the experimental groups over the control group was maintained in the delayed posttest.

Moving to differences in modalities, the multimodal recast group demonstrated greater pretest–posttest gains than both the oral recast and written recast groups, accompanied by large effect sizes. However, no difference was observed between the two unimodal recast groups. On the delayed posttest, the multimodal recast group once again outperformed both the oral recast and written recast groups, also with large effect sizes. Overall, these results indicate that recasts provided during video-conferencing tasks had a positive impact on participants' improvements in the WPT, with multimodal recasts yielding the greatest benefit.

Table 43. Results of Post Hoc Contrast for Significant Interactions Identified by Linear Mixed−Effects Models Examining Participants' Scores on the WPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control:Oral:Time2	0.98	0.14	6.98	<.001	.13	.15	<.01	<.01	<.01	<.01
Control:Written:Time2	1.05	0.14	7.52	<.001	.14	.17	<.01	<.01	<.01	.03
Control:Multi:Time2	1.91	0.14	13.63	<.001	.41	.41	<.01	<.01	<.01	<.01
Control:Oral:Time3	0.85	0.14	6.13	<.001	.09	.11	<.01	<.01	<.01	<.01
Control:Written:Time3	0.99	0.14	7.12	<.001	.12	.14	<.01	<.01	.01	.10
Control:Multi:Time3	1.87	0.14	13.37	<.001	.38	.38	<.01	<.01	.01	.09
Oral:Multi:Time2	0.93	0.14	6.66	<.001	.39	.40	<.01	<.01	.07	.27
Written:Multi:Time2	0.86	0.14	6.12	<.001	.39	.41	<.01	<.01	.08	.29
Oral:Multi:Time3	1.01	0.14	7.23	<.001	.38	.38	<.01	<.01	<.01	.09
Written:Multi:Time3	0.87	0.14	6.24	<.001	.35	.36	<.01	<.01	.03	.19

*Formula: lmer(log_WPT_PP ~ 1 + Group*Time + (1|Participant) + (Group|Item))*

4.3.2.2.2. Third Person Singular

The study also investigated the extent to which recasts and feedback modality influenced the development of the third person singular usage in the WPT. As illustrated in Table 44, the multimodal recast group outperformed the other three groups. Among the experimental groups, the average gain score of the multimodal recast group was higher than that of the written recast group, which, in turn, exceeded the improvements observed in the gains of the oral recast group from the pretest to the posttest. This trend was maintained on the delayed posttest, with all four groups exhibiting a decline in comparison to the posttest results.

Table 44. Descriptive Statistics for the WPT Scores – Third Person Singular

					95% CI	
		M	*Mean Gain*	*SD*	*Lower*	*Upper*
Control (*N* = 30)	Pretest	1.40	–	0.62	1.17	1.63
	Posttest	1.43	0.03	0.63	1.20	1.67
	Delayed posttest	1.43	0.03	0.62	1.20	1.67
Oral (*N* = 30)	Pretest	1.33	–	0.55	1.13	1.54
	Posttest	3.40	2.07	0.62	3.17	3.63
	Delayed posttest	2.27	0.94	0.83	1.96	2.58
Written (*N* = 30)	Pretest	1.37	–	0.56	1.16	1.57
	Posttest	3.93	2.56	0.80	3.60	4.20
	Delayed posttest	2.43	1.06	0.68	2.18	2.69
Multimodal (*N* = 30)	Pretest	1.30	–	0.60	1.08	1.52
	Posttest	7.70	6.40	0.54	7.50	7.90
	Delayed posttest	6.10	4.80	0.66	5.85	6.35

Note. The maximum score was 9.

The logistic mixed-effects regressions were carried out to explore the impact of Group and Time on the gains that participants generated with respect to the third person singular. These analyses revealed a total of 10 significant interactions between Group and Time (see Appendix A24 for full model results). As summarized in Table 45 once again, all recast groups demonstrated more substantial gains from pretest to posttest compared to the control

group. Examining the odds ratios, it was evident that compared to the control group, the highest pretest–posttest gains were observed in the multimodal recast group (*OR* = 35.74), followed by the written recast group (*OR* = 4.30), and then the oral recast group (*OR* = 3.45). On the delayed posttest, the recast groups also maintained their advantages over the control group. Again, the odds ratios indicated that, in comparison to the control group, the greatest pretest–delayed posttest gains were detected in the multimodal recast group, with the odds of advancing in written production gains being 12.55 times higher compared to the control group. This was followed by the written recast group with 2.03 times higher odds, and the oral recast group with 1.89 times higher odds. Turning to the influence of different modalities, the multimodal recast group exhibited more pronounced gains from pretest to both posttest and delayed posttest compared to the oral recast and written recast groups.

Specifically, the odds of development in the posttest were 10.37 times higher for the multimodal recast group than for the oral recast group, and 8.32 times higher than for the written recast group respectively. This pattern persisted in the delayed posttest, with the odds of written production development being 6.62 times higher for the multimodal recast group compared to the oral recast group, and 6.19 times higher compared to the written recast group. There was no significant gain difference between the oral recast and written recast groups either in the posttest or delayed posttest. In conclusion, these outcomes underscore the favorable impact of recasts on written production gains, irrespective of their modality. Furthermore, the results suggest that multimodal recasts proved more effective in fostering development in written production compared to their oral and written counterparts.

Table 45. Results of Post Hoc Contrast for Significant Interactions Identified by Logistic Mixed-effects Models Examining Participants' Scores on the WPT – Third Person Singular

	Fixed effects						Random effects			
							by participant (intercept)		*by item* (intercept)	
	Estimate	*SE*	*z*	*p*	*OR*	95% CI	*Variance*	*SD*	*Variance*	*SD*
ControlOral:Time2	1.24	0.32	3.88	<.001	3.45	[1.85, 6.42]	<.01	<.01	.21	.46
ControlWritten:Time2	1.46	0.32	4.60	.049	4.30	[2.32, 7.97]	<.01	<.01	.11	.32
ControlMulti:Time2	3.58	0.34	10.46	<.001	35.74	[18.35, 69.62]	<.01	<.01	.05	.23
ControlOral:Time3	0.64	0.33	1.97	.028	1.89	[1.01, 3.57]	<.01	<.01	.28	.52
ControlWritten:Time3	0.71	0.32	2.19	<.001	2.03	[1.08, 3.81]	<.01	<.01	.10	.31
ControlMulti:Time3	2.53	0.32	7.85	<.001	12.55	[6.69, 23.52]	<.01	<.01	.08	.29
OralMulti:Time2	2.34	0.33	7.18	<.001	10.37	[5.50, 19.57]	<.01	<.01	.21	.46
WrittenMulti:Time2	2.12	0.32	6.55	<.001	8.32	[4.43, 15.63]	<.01	<.01	.01	.07
OralMulti:Time3	1.89	0.31	6.08	<.001	6.62	[3.61, 12.15]	<.01	<.01	.18	.42
WrittenMulti:Time3	1.82	0.31	5.90	<.001	6.19	[3.39, 11.30]	<.01	<.01	.06	.25

*Formula: glmer(WPT_PS ~ 1+ Group*Time + (1|Participant) + (1|Item))*

4.3.2.3. SOPT

4.3.2.3.1. Present Progressive

Table 46 summarizes the descriptive statistics for the participants' pretest, posttest, and delayed posttest scores on the SOPT. All four groups exhibited pretest–posttest gains, with the multimodal group attaining the greatest gain scores. On the delayed posttest, all groups demonstrated some decline.

Table 46. Descriptive Statistics for the SOPT Scores – Present Progressive

					95% CI	
		M	*Mean Gain*	*SD*	*Lower*	*Upper*
Control (*N* = 30)	Pretest	2.97	–	0.67	2.72	3.22
	Posttest	3.47	0.50	0.57	3.25	3.68
	Delayed posttest	3.40	0.43	0.77	3.11	3.69
Oral (*N* = 30)	Pretest	2.90	–	0.76	2.62	3.18
	Posttest	9.33	6.43	1.30	8.85	9.82
	Delayed posttest	8.40	5.50	1.40	7.88	8.92
Written (*N* = 30)	Pretest	2.87	–	0.63	2.63	3.10
	Posttest	10.00	7.13	1.05	9.61	10.39
	Delayed posttest	9.07	6.20	0.91	8.73	9.41
Multimodal (*N* = 30)	Pretest	2.83	–	0.65	2.59	3.08
	Posttest	15.57	12.74	1.25	15.10	16.03
	Delayed posttest	15.07	12.24	0.94	14.71	15.42

Note. The maximum score was 18.

The linear mixed-effects model analyses conducted to examine the effect of Group and Time on participants' performance on the SOPT identified 10 significant Time-by-Group interactions (see Appendix A25 for full model results). The results indicated that all recast groups showed greater pretest–posttest and pretest–delayed posttest gains than the control group. Table 47 summarizes the results of the post hoc pairwise comparisons for these significant interaction effects. As reflected in the effect sizes, the multimodal recast group achieved the largest gains when compared to the control group, followed by the written recast

and oral recast groups in that order. All recast groups retained their advantages over the control group on the delayed posttest, with the largest effect size observed in the multimodal recast group, followed by the written recast and oral recast groups. With regard to modality differences, the multimodal recast group showed greater pretest–posttest gains than the oral recast and written recast groups; no difference emerged between the oral recast and written recast groups. On the delayed posttest, the multimodal recast group also outperformed both the oral and written recast groups. Overall, these results indicate that recasts provided during video-conferencing tasks had a positive impact on participants' gains on the SOPT, with multimodal recasts yielding the greatest advantage.

Table 47. Results of Post Hoc Contrast for Significant Interactions Identified by Linear Mixed−Effects Models Examining Participants' Scores on the SOPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control:Oral:Time2	0.98	0.14	7.07	<.001	.13	.15	<.01	<.01	<.01	<.01
Control:Written:Time2	1.07	0.14	7.77	<.001	.15	.19	<.01	<.01	<.01	.04
Control:Multi:Time2	1.92	0.14	13.91	<.001	.42	.42	<.01	<.01	<.01	<.01
Control:Oral:Time3	0.87	0.14	6.29	<.001	.10	.12	<.01	<.01	<.01	<.01
Control:Written:Time3	0.99	0.14	7.17	<.001	.13	.14	<.01	<.01	<.01	.09
Control:Multi:Time3	1.90	0.14	13.72	<.001	.39	.40	<.01	<.01	<.01	.08
Oral:Multi:Time2	0.95	0.14	6.85	<.001	.40	.41	<.01	<.01	.07	.27
Written:Multi:Time2	0.85	0.14	6.15	<.001	.40	.43	<.01	<.01	<.09	.31
Oral:Multi:Time3	1.03	0.14	7.43	<.001	.39	.40	<.01	<.01	<.01	.08
Written:Multi:Time3	0.91	0.14	6.55	<.001	.37	.38	<.01	<.01	.03	.19
*Formula: lmer(log_SOPT_PP ~ 1 + Group*Time + (1|Participant) + (Group|Item))*										

4.3.2.3.2. Third Person Singular

The study also explored the extent to which recasts and feedback modality influenced the development of the use of the third person singular on the SOPT. As presented in Table 48, the mean gain score of the multimodal recast group surpassed that of the other three groups. Among the experimental groups, the average gain score of the multimodal recast group was

notably higher than that of the written recast group which, in turn, exceeded the improvements observed in the oral recast group from the pretest to the posttest. This trend was maintained on the delayed posttest, with all four groups showing a decline compared to the posttest results.

Table 48. Descriptive Statistics for the SOPT Scores – Third Person Singular

					95% CI	
		M	*Mean Gain*	*SD*	*Lower*	*Upper*
Control (*N* = 30)	Pretest	1.30	–	0.65	1.06	1.54
	Posttest	1.37	0.07	0.62	1.14	1.60
	Delayed posttest	1.33	0.03	0.55	1.13	1.54
Oral (*N* = 30)	Pretest	1.20	–	0.48	1.02	1.38
	Posttest	3.37	2.17	0.62	3.14	3.60
	Delayed posttest	2.40	1.20	0.97	2.04	2.76
Written (*N* = 30)	Pretest	1.30	–	0.47	1.13	1.47
	Posttest	3.67	2.37	0.71	3.40	3.93
	Delayed posttest	2.50	1.20	0.68	2.25	2.75
Multimodal (*N* = 30)	Pretest	1.27	–	0.52	1.07	1.46
	Posttest	7.13	5.86	0.82	6.83	7.44
	Delayed posttest	6.07	4.80	0.91	5.73	6.41

Note. The maximum score was 9.

Logistic mixed-effects regressions were performed to examine the influence of Group and Time on participants' advancement in utilizing the third person singular form. These analyses revealed the emergence of 10 significant interactions between Group and Time (see Appendix A26 for full model results). As summarized in Table 49, all recast groups exhibited more pronounced pretest–posttest gains than the control group. The odds ratios underscored this trend, highlighting that, compared to the control group, the highest pretest–posttest gains were achieved by the multimodal recast group (*OR* = 24.06), followed by the written recast group (*OR* = 4.04) and then the oral recast group (*OR* = 3.78). The recast groups also retained their advantages over the control group on the delayed posttest, with the multimodal recast

group achieving the most significant gains (*OR* = 13.20), followed by the oral recast group (*OR* = 2.34), and then the written recast group (*OR* = 2.25). Regarding modality effects, the multimodal recast group generated greater pretest–posttest and pretest–delayed posttest gains compared to the oral recast and written recast groups. The odds of achieving progress in pretest–posttest gains were significantly heightened for the multimodal recast group when compared to both the oral recast and the written recast groups. Specifically, the likelihood was 6.36 times greater compared to the oral recast group and 5.95 times greater when compared to the written recast group. Similarly, during the delayed posttest, the odds of development were 5.65 times higher for the multimodal recast group compared to the oral recast group, and 5.86 times higher compared to the written recast group. No substantial differences in gains were observed between the two unimodal recast groups, either during the posttest or delayed posttest. In summary, the provision of recasts, irrespective of their modality, significantly fostered participants' proficiency in employing the third person singular form, as evidenced in their performance on the SOPT. Again, multimodal recasts were found to be more effective in promoting participants' ability to use the target form in spontaneous oral production in comparison to oral and written recasts.

Table 49. Significant Interactions Identified by Logistic Mixed-effects Models Examining Participants' Scores on the SOPT – Third Person Singular

	Fixed effects						Random effects			
							by participant (intercept)		by item (intercept)	
	Estimate	*SE*	*z*	*p*	*OR*	95% CI	*Variance*	*SD*	*Variance*	*SD*
ControlOral:Time2	1.33	0.33	4.04	<.001	3.78	[1.99, 7.19]	<.01	<.01	.27	.52
ControlWritten:Time2	1.40	0.32	4.29	<.001	4.04	[2.14, 7.63]	<.01	<.01	.17	.42
ControlMulti:Time2	3.18	0.38	9.41	<.001	24.06	[12.45, 46.52]	<.01	<.01	.09	.30
ControlOral:Time3	0.85	0.33	2.53	.011	2.34	[1.22, 4.49]	<.01	<.01	.38	.62
ControlWritten:Time3	0.81	0.33	2.45	.014	2.25	[1.18, 4.29]	<.01	<.01	.12	.35
ControlMulti:Time3	2.58	0.33	7.81	<.001	13.20	[6.93, 25.83]	<.01	<.01	.11	.33
OralMulti:Time2	1.85	0.32	5.76	<.001	6.36	[3.40, 11.89]	<.01	<.01	.27	.52
WrittenMulti:Time2	1.78	0.32	5.63	<.001	5.95	[3.04, 10.48]	<.01	<.01	.09	.30
OralMulti:Time3	1.73	0.32	5.46	<.001	5.65	[3.04, 10.48]	<.01	<.01	.29	.54
WrittenMulti:Time3	1.77	0.31	5.65	<.001	5.86	[3.18, 10.79]	<.01	<.01	.14	.37
*Formula: glmer(SOPT_PS ~ 1+ Group*Time + (1|Participant) + (1|Item))*										

4.3.3. Effects of Multimodal Recasts and Written Recasts on Learners' Attention Allocation (RQ3)

To address Research Question 3, which aimed to investigate the effects of multimodal recasts versus written recasts on drawing learners' visual attention to the target constructions, comparative analyses were conducted on the eye-gaze behavior of these two groups. As was previously explained, the study employed a series of linear mixed-effects model analyses to examine continuous attention metrics (total fixation duration, first-pass reading time, re-reading time), along with a set of generalized linear mixed models for proportional count-based attention metric (visit count) and proportional binary attention measure (skip rate). In these analyses, five different measurements were used to quantify the extent of attention that participants devoted to the target structures: (a) total fixation duration, (b) first-pass reading, (c) re-reading, (d) visit count, and (e) skip rate. A total of 10 sets of models were examined, each encompassing fixed effects for Group, Time, and their interactions, with each attention measure serving as the dependent variable. To account for variability, Participant was incorporated as a random effect. Since participants received recasts on different items, Item could not be included as a random effect.

Prior to proceeding with the primary analyses, I performed preliminary Spearman correlations analyses among the five attention measures, following the suggestion of Godfroid and Hui (2020). The correlation results are summarized in Table 50 for the present progressive form, and in Table 51 for the third person singular form for all groups combined and for each group individually.

For the present progressive form, the analysis conducted for the combined groups demonstrated strong positive correlations across all measurements, with the exception of skip rate, which exhibited strong negative correlations with the other measures.

Similar patterns emerged concerning the third person singular form. The analysis of the combined group once again revealed strong positive correlations among all measurements, except for skip rate, which exhibited strong negative correlations with the other metrics. Contrary to the multimodal recast group, the written recast group displayed varied patterns across the different measures. It is significance that there was no correlation between first-pass reading time, visit count, and skip rate. In contrast to the present progressive form, there was no strong negative correlation between first-pass reading time and re-reading time. On the contrary, a significant positive correlation was observed.

Table 50. Correlation Analysis of Five Attention Measurements – Present Progressive

	Total fixation duration	First-pass reading	Re-reading	Visit count	Skip rate
Combined					
Total fixation duration	1	.693*** (.000)	.775*** (.000)	.687*** (.000)	–.690*** (.000)
First-pass reading		1	.225*** (.000)	.580** (.000)	–.474*** (.000)
Re-reading			1	.525** (.000)	–.554*** (.000)
Visit count				1	–.514*** (.000)
Skip rate					1
Multimodal Recast Group					
Total fixation duration	1	.674*** (.000)	.773*** (.000)	.760*** (.000)	–.483*** (.000)
First-pass reading		1	.139*** (.995)	.469*** (.000)	–.202** (.340)
Re-reading			1	.617*** (.000)	–.410*** (.000)
Visit count				1	–.365*** (.000)
Skip rate					1
Written Recast Group					
Total fixation duration	1	–.049 (.166)	.862*** (.000)	.231*** (.000)	–.626*** (.000)
First-pass reading		1	–.429** (.001)	.001 (.967)	.040 (.254)
Re-reading			1	.255*** (.000)	–.545*** (.000)
Visit count				1	–.250** (.000)
Skip rate					1

Table 51. Correlation Analysis of Five Attention Measurements – Third Person Singular

	Total fixation duration	First-pass reading	Re-reading	Visit count	Skip rate
Combined					
Total fixation duration	1	.678*** (.000)	.915*** (.000)	.625*** (.000)	–.428** (.000)
First-pass reading		1	.371*** (.000)	.455** (.003)	–.270*** (.029)
Re-reading			1	.531*** (.000)	–.389*** (.000)
Visit count				1	–.592*** (.000)
Skip rate					1
Multimodal Recast Group					
Total fixation duration	1	.767*** (.000)	.903*** (.000)	.672*** (.000)	–.447*** (.000)
First-pass reading		1	.440*** (.000)	.612*** (.000)	–.223* (.035)
Re-reading			1	.528*** (.000)	–.505*** (.000)
Visit count				1	–.405*** (.000)
Skip rate					1
Written Recast Group					
Total fixation duration	1	.471*** (.000)	.943*** (.000)	.579*** (.000)	–.307** (.003)
First-pass reading		1	.238* (.024)	.188 (.076)	–.049 (.648)
Re-reading			1	.592*** (.000)	–.300** (.004)
Visit count				1	–.508*** (.000)
Skip rate					1

4.3.3.1. Present Progressive

With regard to how learners directed their visual attention to the present progressive form, Table 52 provides an overview of the descriptive statistics encompassing five eye-tracking measures for both the multimodal recast group and the written recast group across the three treatment sessions. It is important to note that participants in both groups were exposed to the written recast presented on the screen for the same amount of time (10 seconds). Overall, during the three treatment sessions, the multimodal recast group consistently displayed a higher frequency of visits to the target construction. Furthermore, this group demonstrated longer durations of total fixation, first-pass reading time, and re-reading time, all while maintaining a lower skip rate.

Table 52. Descriptive Statistics for Attention Measurements – Present Progressive

		Written Group (N = 30)		Multimodal Group (N = 30)	
		M (SD)	95% CI	*M (SD)*	95% CI
Total fixation duration (ms)	Time1	424.79 (94.43)	[389.52, 460.05]	599.96 (171.78)	[535.81, 664.11]
	Time2	388.75 (84.43)	[357.22 420.28]	537.10 (197.07)	[463.51, 610.68]
	Time3	280.32 (85.10)	[248.54, 312.09]	448.42 (210.89)	[369.67, 527.17]
	Total	364.62 (106.77)	[342.25, 386.98]	537.83 (197.99)	[495.91, 578.85]
First-pass reading (ms)	Time1	121.90 (66.27)	[97.16, 146.65]	227.70 (129.25)	[179.44, 275.96]
	Time2	93.99 (47.75)	[76.16, 111.82]	225.68 (166.89)	[163.36, 287.99]
	Time3	77.17 (21.54)	[69.12, 85.21]	156.90 (93.22)	[118.76, 188.38]
	Total	97.69 (51.66)	[86.87, 108.51]	203.43 (136.21)	[173.79, 230.84]
Re-reading (ms)	Time1	302.88 (68.89)	[277.16, 328.61]	408.83 (154.57)	[351.11, 466.55]
	Time2	294.76 (93.64)	[259.79, 329.72]	348.08 (123.44)	[301.99, 394.18]
	Time3	203.15 (79.30)	[173.54, 232.76]	244.85. (97.57)	[208.42, 281.29]
	Total	266.93 (92.31)	[247.60, 286.26]	333.92 (143.12)	[303.95, 363.90]
Visit count	Time1	0.85 (0.32)	[0.73, 0.97]	1.33 (0.48)	[1.15, 1.51]
	Time2	0.70 (0.22)	[0.62, 0.79]	1.34 (0.43)	[1.18, 1.50]
	Time3	0.66 (0.14)	[0.60, 0.71]	1.27 (0.33)	[1.14, 1.39]
	Total	0.74 (0.25)	[0.68, 0.79]	1.31 (0.41)	[1.23, 1.40]
Skip rate	Time1	0.10 (0.05)	[0.08, 0.12]	0.04 (0.04)	[0.02, 0.05]
	Time2	0.14 (0.08)	[0.11, 0.17]	0.08 (0.21)	[0.01, 0.16]
	Time3	0.19 (0.14)	[0.13, 0.24]	0.11 (0.28)	[0.00, 0.21]
	Total	0.14 (0.10)	[0.12, 0.16]	0.07 (0.20)	[0.04, 0.12]

Note. Time1=treatment session 1; Time2=treatment session 2; Time3=treatment session 3.

To examine the interaction effect of Group and Time on learners' visual attention allocation, linear mixed-effects models were conducted for total fixation duration, first-pass reading, and re-reading time, as well as two GLMMs for visit count and skip rate respectively. The models

included Group and Time as fixed effects and Participant as a crossed random effect to assess differences in visual attention allocation between the multimodal recast and the written recast groups. The results revealed significant differences between the two groups across different time points for total fixation duration, first-pass reading, and re-reading time. Effect sizes for these differences ranged from medium to large.

As shown in Table 53, which presents the results of the linear mixed-effect model analyses, no significant interaction effects were observed between Group and Time in all three attention measures. This indicates that the temporal effect on these three measures remained consistent for both groups. However, Group had a significant positive impact on all three measures with medium- to large effect sizes. This suggests that, on average, the multimodal recast group exhibited longer total fixation duration, first-pass reading, and re-reading durations compared to the written recast group. Figures 25–27 provide visual representation of the changes in the logged means of these three attention measures over time in each group. Taken together, these findings imply that the multimodal recasts group consistently led to enhanced visual attention allocation toward the target feature when compared to the written recast group, suggesting that multimodal recasts were more effective in sustaining learners' overall focus on the target construction than written recasts.

Table 53. Results for the Linear Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Present Progressive

	Fixed effects						Random effects	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*by participant variance*	*SD*
Total fixation duration								
(Intercept)	6.03	0.06	102.91	<.001***	0.38	0.43	.01	.09
Group	0.33	0.08	3.96	<.001***				
Time2	–0.09	0.08	–1.09	.278				
Time3	–0.44	0.08	–5.50	<.001***				
Group:Time2	0.01	0.11	0.12	.904				
Group:Time3	0.08	0.11	0.73	.468				
First-pass reading								
(Intercept)	4.70	0.10	47.64	<.001***	0.34	0.42	.03	.18
Group	0.61	0.14	4.34	<.001***				
Time2	–0.33	0.13	–2.52	.013*				
Time3	–0.39	0.13	–2.97	.004**				
Group:Time2	0.29	0.19	1.55	.125				
Group:Time3	–0.01	0.19	–0.05	.962				
Re-reading								
(Intercept)	5.68	0.07	80.04	<.001***	0.26	0.52	.10	.31
Group	0.25	0.10	2.51	.013*				
Time2	–0.04	0.08	–0.53	.594				
Time3	–0.45	0.08	–5.48	<.001***				
Group:Time2	–0.11	0.11	–0.98	.329				
Group:Time3	–0.09	0.11	–0.78	.440				
*Formula: lmer(log_eye-tracking measure ~ 1 + Group*Time + (1|Participant))*								

Note. The written recast group was used as a reference level; Time1= Treatment session 1 (reference level); Time2 = Treatment session 2; Time3 = Treatment session 3.

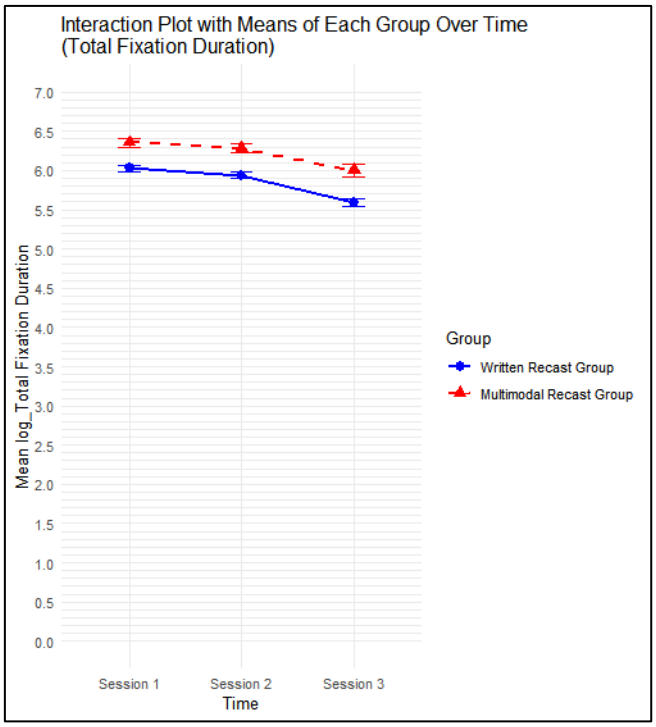


Figure 25. Interaction Plot with Logged Means of Total Fixation Duration for Each Group (Present Progressive)

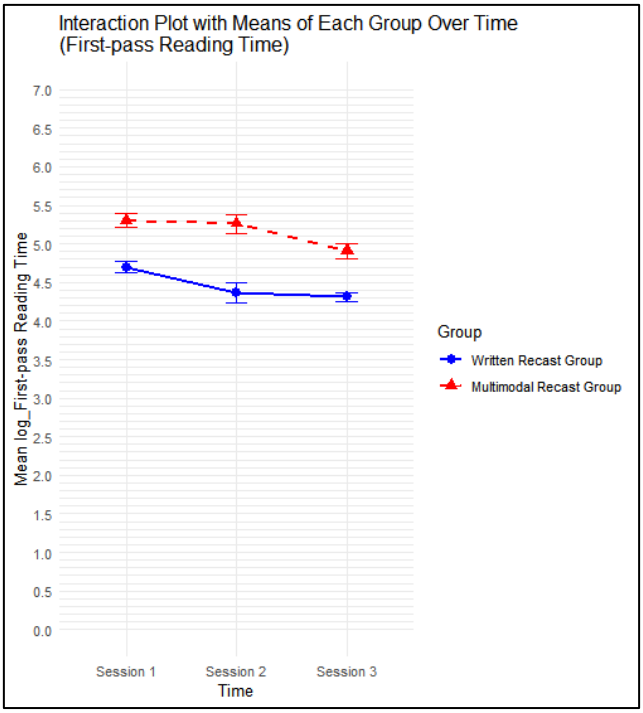


Figure 26. Interaction Plot with Logged Means of First-pass Reading Time for Each Group (Present Progressive)

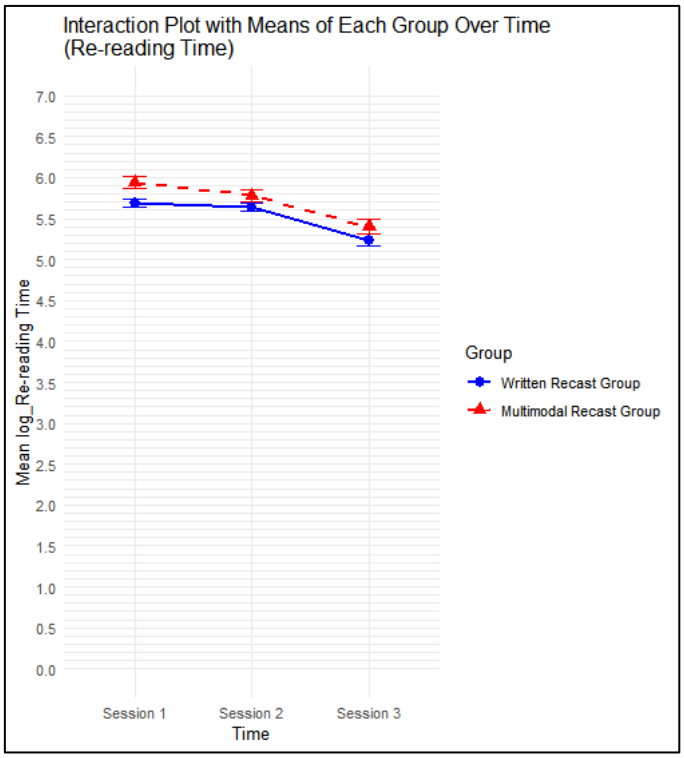


Figure 27. Interaction Plot with Logged Means of Re-reading Time for Each Group (Present Progressive)

Visit count and skip rate were analyzed using GLMMs, employing the inverse Gaussian family distribution with a log link function. Regarding visit count, the results, as summarized in Table 54, revealed significant positive interaction effects between Group and Time 1–Time 2, as well as between Group and Time 1–Time 3. This indicates that participants in the multimodal recast group experienced a smaller decline in the odds of having more visits both from Time 1 to Time 2 and from Time 1 to Time 3 compared to those in the written recast group. The odds associated with the interaction terms were 1.17 and 1.17, respectively. These values suggest that while both groups experienced a decrease in the odds of having visits over time, the multimodal group exhibited a more favorable change in the odds of having more visits as time progressed. Also, there was a significant positive effect of Group on visit count, indicating that the multimodal recast group, on average, had a significantly higher number of

visits, with 1.37 times higher odds of visiting the target feature compared to the written recast group. As depicted in Figure 28, overall, these findings emphasize the effectiveness of the multimodal recast intervention in encouraging more frequent visits to the target construction. Moreover, this intervention demonstrated a more sustained impact on visit count over time, compared to the written recast intervention.

Table 54. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Present Progressive

Total duration	*Estimate*	*SE*	*t*	*p*	*OR*	95% CI for OR
Visit count						
(Intercept)	–0.53	0.05	–10.80	<.001***	0.59	[0.53, 0.65]
Group	0.32	0.07	4.64	<.001***	1.37	[1.20, 1.57]
Time2	–0.13	0.03	–4.11	<.001***	0.88	[0.83, 0.93]
Time3	–0.16	0.03	–5.07	<.001***	0.85	[0.80, 0.91]
Group:Time2	0.15	0.05	3.08	.002**	1.17	[1.06, 1.28]
Group:Time3	0.16	0.05	3.24	.001**	1.17	[1.07, 1.29]
Random effects	*variance*	*SD*				
Participant (Int)	.02	.13				
*Formula: glmer(log_Visit count ~ 1 + Group*Time + (1|Participant))*						

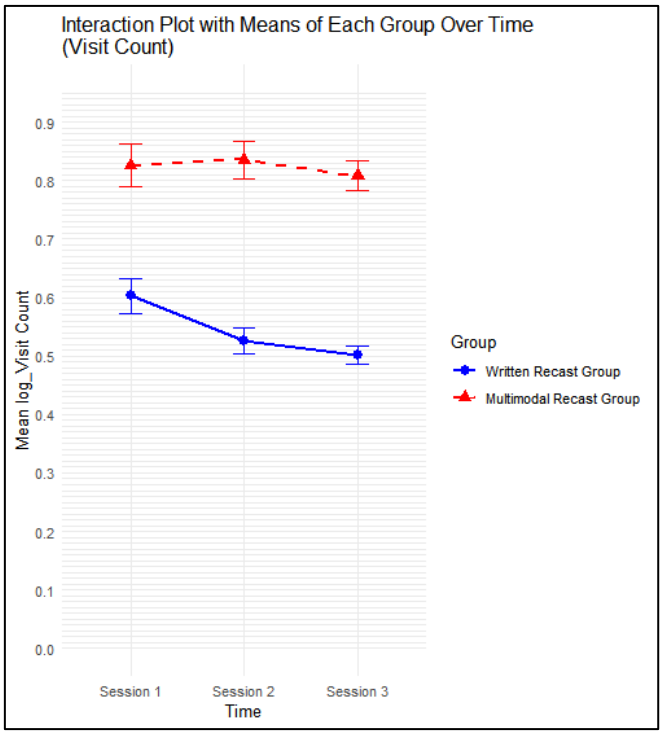


Figure 28. Interaction Plot with Logged Means of Visit Count for Each Group (Present Progressive)

The GLMM results for skip rate are displayed in Table 55, while Figure 29 visually presented the mean skip rate across three treatment sessions for each group. The analysis revealed that the interaction effects between Group and Time did not reach statistical significance, suggesting that the temporal effect on skip rate was consistent for both groups. However, the impact of Group on skip rate was found to be significantly negative. This suggests that participants in the multimodal recast group had a significantly lower skip rate compared to those in the written recast group, in other words, lower odds of skipping (*OR* = 0.91). Additionally, Time 2 and Time 3 showed significant positive effects on skip rate, indicating that both groups experienced a significant increase in the skip rate as time progressed. Overall, these findings lend support to the efficacy of multimodal recasts in reducing skip rates for the target construction, in comparison to written recasts.

Table 55. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Present Progressive

Total duration	*Estimate*	*SE*	*t*	*p*	*OR*	95% CI for OR
Skip rate						
(Intercept)	–2.52	0.14	–17.68	<.001***	0.08	[0.06, 0.11]
Group	–1.13	0.19	–5.98	<.001***	0.91	[0.67, 1.23]
Time2	0.35	0.14	2.57	.010*	0.99	[0.70, 1.41]
Time3	0.56	0.16	3.46	<.001***	1.75	[1.27, 2.40]
Time2:Group2	–0.10	0.16	–0.63	.529	0.91	[0.67, 1.23]
Time3:Group2	–0.01	0.18	–0.04	.969	0.99	[0.70, 1.41]
Random effects	variance	SD				
Participant (Int.)	.81	.90				
*Formula: glmer(log_eye-tracking measure ~ 1 + Group*Time + (1|Participant))*						

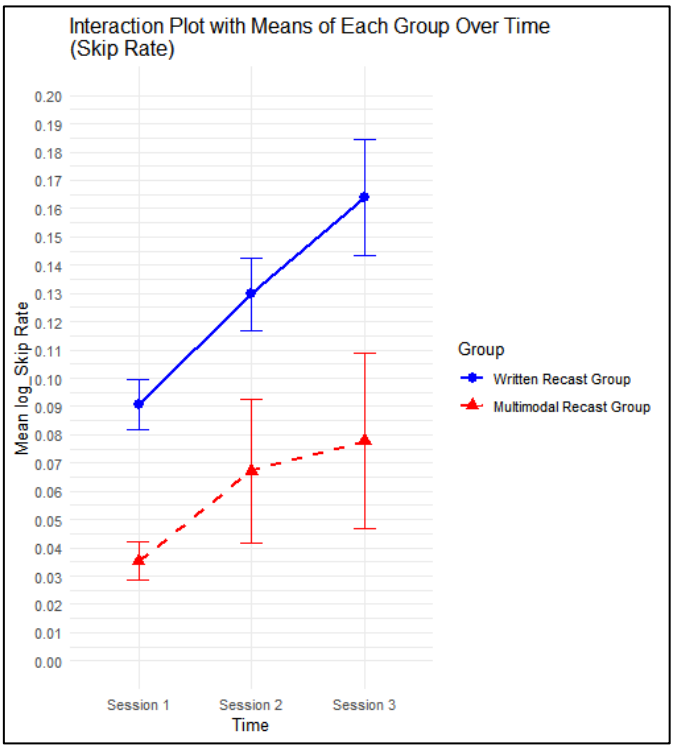


Figure 29. Interaction Plot with Logged Means of Skip Rate for Each Group (Present Progressive)

4.3.3.2. Third Person Singular

The descriptive statistics for five eye-tracking measures by the multimodal recast group and the written recast group concerning the third person singular form are presented in Table 56.

Similar to the findings for the present progressive, the multimodal recasts group, overall, exhibited higher frequencies of visits to the target construction and longer total fixation duration, first-pass reading time, and re-reading time while maintaining a lower skip rate.

Table 56. Descriptive Statistics for Attention Measurements – Third Person Singular

		Written Group (N = 30)		Multimodal Group (N = 30)	
		M (SD)	95% CI	*M (SD)*	95% CI
Total fixation duration (ms)	Time1	605.69 (138.10)	[554.55, 656.82]	796.22 (88.68)	[763.11, 829.34]
	Time2	559.25 (142.73)	[505.96, 612.55]	655.76 (82.84)	[624.83 686.70]
	Time3	461.41 (152.20)	[404.58, 518.24]	538.95 (117.78)	[494.97, 582.93]
	Total	544.34 (153.19)	[512.25, 595.31]	639.20 (162.73)	[605.12, 673.28]
First-pass reading (ms)	Time1	128.74 (35.50)	[115.48, 141.99]	249.88 (37.57)	[235.85, 263.91]
	Time2	158.22 (72.04)	[139.94, 176.50]	196.10 (42.97)	[180.05, 212.15]
	Time3	120.11 (34.90)	[107.08, 133.15]	151.72 (57.12)	[130.39, 173.04]
	Total	136.80 (44.12)	[127.56, 146.04]	179.23 (71.90)	[164.17, 194.29]
Re-reading (ms)	Time1	476.95 (138.36)	[426.58, 527.32]	546.34 (78.71)	[516.95, 575.73]
	Time2	401.04 (124.09)	[354.70, 447.37]	456.33 (89.65)	[422.85, 489.81]
	Time3	341.30 (134.70)	[291.00, 391.59]	387.24 (105.31)	[347.92, 426.56]
	Total	407.54 (140.15)	[378.19, 436.89]	459.97 (118.55)	[435.14, 484.80]
Visit count	Time1	0.86 (0.26)	[0.76, 0.96]	2.01 (0.48)	[1.84, 2.19]
	Time2	0.78 (0.24)	[0.69, 0.87]	1.30 (0.44)	[1.14, 1.47]
	Time3	0.60 (0.20)	[0.52, 0.67]	0.78 (0.31)	[0.67, 0.90]
	Total	0.75 (0.26)	[0.69, 0.80]	1.32 (0.66)	[1.18, 1.46]
Skip rate	Time1	0.20 (0.27)	[0.10, 0.30]	0.02 (0.01)	[0.01, 0.02]
	Time2	0.23 (0.33)	[0.11, 0.36]	0.03 (0.03)	[0.01, 0.04]
	Time3	0.29 (0.45)	[0.13, 0.46]	0.04 (0.05)	[0.03, 0.06]
	Total	0.24 (0.35)	[0.17, 0.32]	0.03 (0.04)	[0.02, 0.04]

The same models were used for the analyses of attention measures for the third person singular form, encompassing three sets of linear mixed-effects models for the continuous dependent variables (total fixation duration, first-pass reading, re-reading) and two sets of GLMMs with an inverse Gaussian family with a log link function for the proportional count

variable (visit count) and the proportional binary variable (skip rate). These models also aimed to explore the associations between the dependent variables (eye-tracking measures) and the independent variables (i.e., Group, Time and their interaction), while accounting for random intercepts within participants.

As shown in Table 57, which presents the results of the linear mixed-effect model analyses, significant negative interaction effects were observed between Group and Time 1 to Time 2, as well as between Group and Time 1 to Time 3, for first-pass reading time. This suggests that the declines in this measure were more pronounced in the multimodal recast group from session 1 to session 2 and to session 3. Similar to the findings for the present progressive, Group exerted a significant positive impact on all three measures with medium to large effect sizes. This implies that, on average, the multimodal recast group exhibited longer periods of fixation, first-pass reading, and re-reading, compared to the written recast group. In essence, these findings collectively suggest that the implementation of multimodal recasts contributed to an overall enhancement in the duration of total fixation, first-pass reading, and re-reading time. Figures 30–32 demonstrate how the three attention measures changed over the course of the three treatment sessions for each group.

Table 57. Results for the Linear Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction – Third Person Singular

	Fixed effects						Random effects	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*by participant variance*	*SD*
Total fixation duration								
(Intercept)	6.37	0.05	140.93	<.001***	0.36	0.66	.03	.17
Group	0.30	0.06	4.66	<.001***				
Time2	–0.08	0.05	–1.78	.078				
Time3	–0.30	0.05	–6.40	<.001***				
Group2:Time2	–0.11	0.07	–1.69	.094				
Group2:Time3	–0.11	0.07	–1.61	.110				
First-pass reading								
(Intercept)	4.82	0.05	95.79	<.001***	0.47	0.56	.01	.11
Group	0.69	0.07	9.66	<.001***				
Time2	0.20	0.06	3.12	.002**				
Time3	–0.07	0.06	–1.11	.268				
Group2:Time2	–0.46	0.09	–4.98	<.001***				
Group2:Time3	–0.48	0.09	–5.26	<.001***				
Re-reading								
(Intercept)	6.11	0.06	104.19	<.001***	0.22	0.56	.04	.21
Group	0.18	0.08	2.21	.029*				
Time2	–0.17	0.06	–2.70	.008**				
Time3	–0.36	0.06	–5.82	<.001***				
Group2:Time2	–0.02	0.09	–0.21	.833				
Group2:Time3	–0.00	0.09	–0.05	.962				
*Formula: lmer(log_eye-tracking measure_PS ~ 1 + Group*Time + (1|Participant))*								

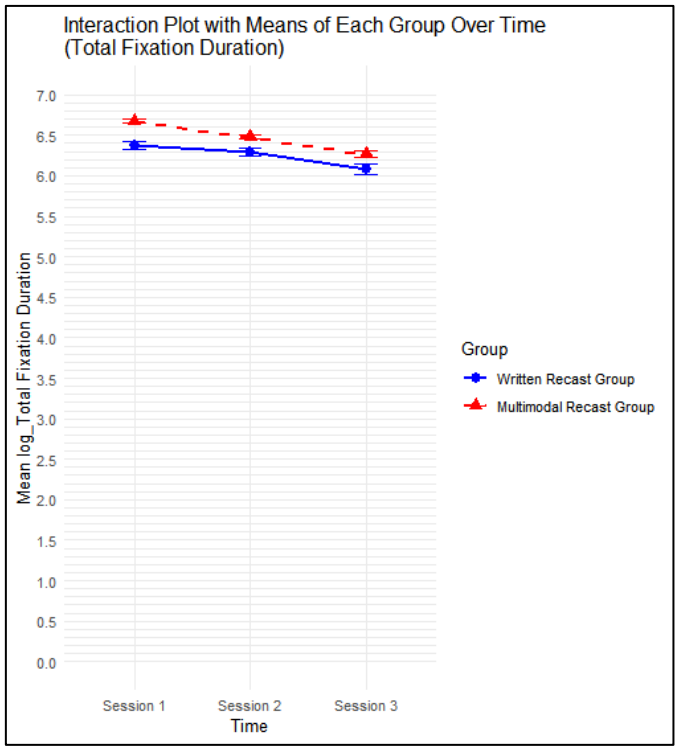


Figure 30. Interaction Plot with Means of Total Fixation Duration for Each Group (Third Person Singular)

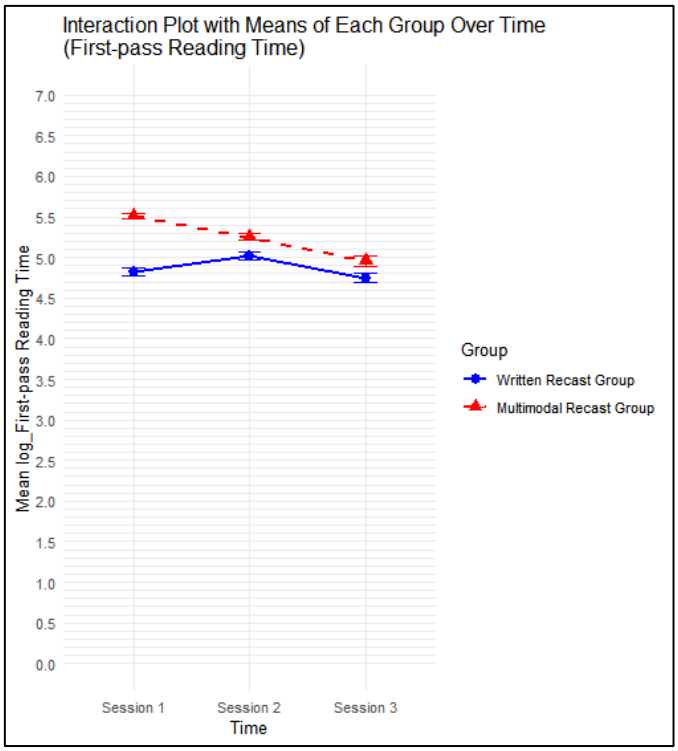


Figure 31. Interaction Plot with Means of First-pass Reading Time for Each Group (Third Person Singular)

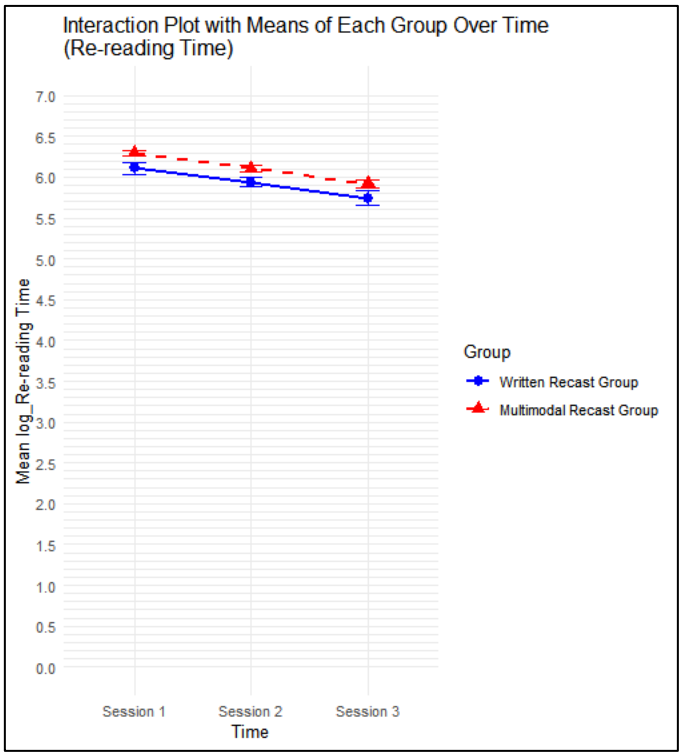


Figure 32. Interaction Plot with Means of Re-reading Time for Each Group (Third Person Singular)

In terms of visit count, the results, as summarized in Table 58, revealed significant negative interaction effects between Group and Time 1 to Time 2, as well as between Group and Time 1 to Time 3. This indicates that the multimodal recast group showed a more prominent reduction in the number of visits from Time 1 and Time 2 as well as from Time 1 and Time 3 in comparison to the written recast group. Group had a significant positive influence on the visit count. This implies that participants in the multimodal recast group exhibited a notably higher visit counts, with the odds of 1.81 compared to those in the written recast group. This finding suggests that the odds of visiting the target construction in the multimodal recast group were 1.81 times higher than those of the written recast group. Figure 33 visualizes the overall patterns of visit count for the three attention measures over the three treatment sessions for each group.

Table 58. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction (Visit Cout) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*t*	*p*	*OR*	95% CI for OR
Visit count						
(Intercept)	–0.48	0.07	–6.94	<.001***	0.62	[0.54, 0.71]
Group	0.59	0.10	6.11	<.001***	1.81	[1.50, 2.19]
Time2	–0.01	0.03	–0.28	.782	0.99	[0.93, 1.06]
Time3	–0.31	0.03	–9.58	<.001***	0.73	[0.69, 0.78]
Group:Time2	–0.44	0.05	–8.29	<.001***	0.65	[0.58, 0.72]
Group:Time3	–0.39	0.05	–7.83	<.001***	0.68	[0.61, 0.74]
Random effects	*variance*	*SD*				
Participant (Int.)	.05	.22				
*Formula: glmer(log_Visit_PS ~ 1 + Group*Time + (1|Participant))*						

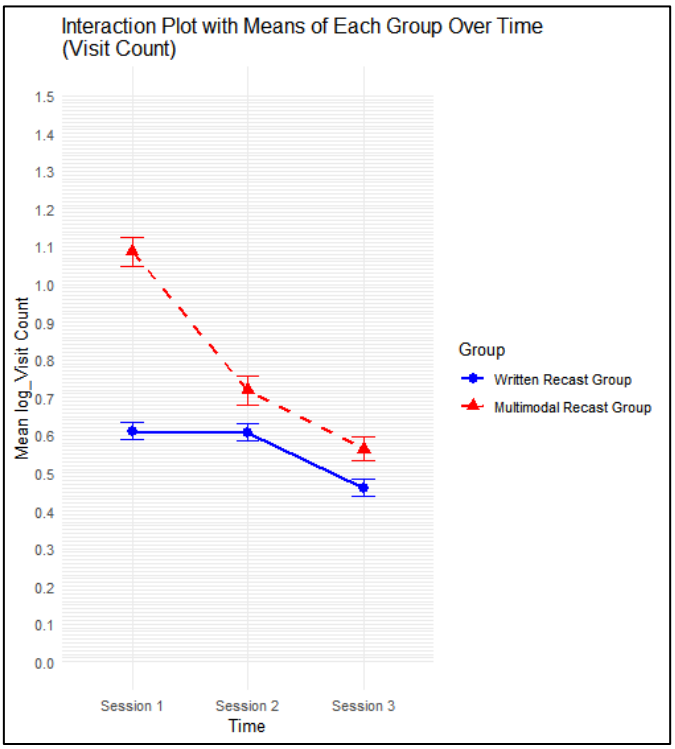


Figure 33. Interaction Plot with Mean of Visit Count for Each Group (Third Person Singular)

The results of GLMM analysis for skip rate are detailed in Table 59, which yielded a significant positive interaction effect between Group and Time 1 to Time 3. This suggests that the increase in skip rate was more prominent in the multimodal recast group session 1 to

session 3, with the odds of 0.54. The impact of Group on skip rate was found to have a significant negative effect. Specifically, the odds of skipping were significantly lower (*OR* = 2.24) for participants in the multimodal recast group in comparison to those in the written recast group. This indicates that participants who received multimodal recasts displayed a substantially reduced tendency to skip the target construction in comparison to their counterparts in the written recast group. Figure 34 demonstrates the overall skip rate patterns for the three attention measures across the three treatment sessions for each group. In summary, these findings also provide support for the effectiveness of multimodal recast in effectively reducing skip rates concerning the target construction when contrasted with the written recast condition.

Table 59. Results for the Generalized Mixed-effects Models Examining Visual Attention Allocated to Target Linguistic Construction (Skip Rate) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*t*	*p*	*OR*	95% CI for OR
Skip rate						
(Intercept)	–2.48	0.17	–14.20	<.001***	0.85	[0.71, 1.01]
Group	–1.76	0.24	–7.24	<.001***	2.24	[1.68, 3.01]
Time2	0.07	0.07	1.00	.316	0.92	[0.77, 1.11]
Time3	0.39	0.07	5.49	<.001***	0.68	[0.57, 0.80]
Group:Time2	0.01	0.07	0.18	.854	0.80	[0.55, 1.15]
Group:Time3	0.36	0.08	4.36	<.001***	0.54	[0.41, 0.72]
Random effects	variance	SD				
Participant (Int.)	.03	.18				
*Formula: glmer(log_Skip_PS ~ 1 + Group*Time + (1|Participant))*						

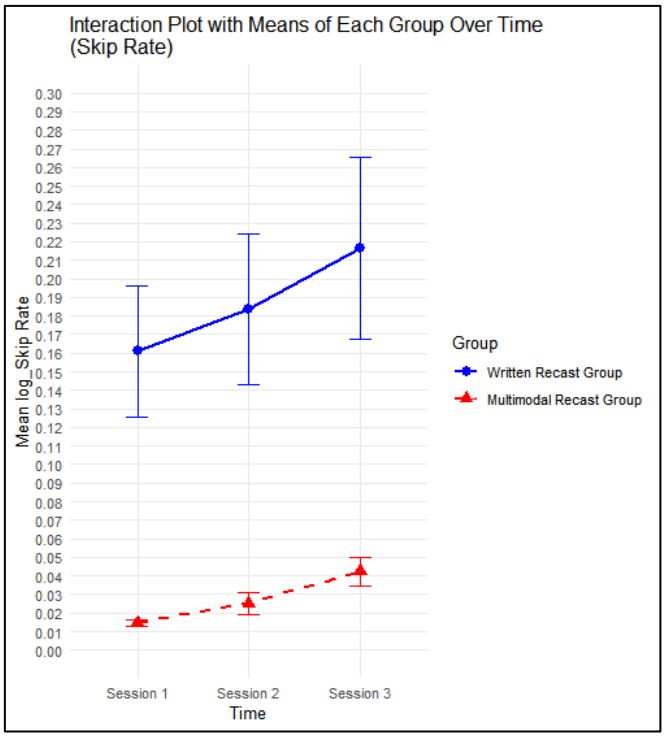


Figure 34. Interaction Plot with Means of Skip Rate for Each Group (Third Person Singular)

These results collectively suggest that participants in the multimodal recast group demonstrated an enhanced inclination to engage more consistently with recasts featuring the target structure, the third person singular. This was evident through both an increased frequency of visits and a decreased tendency to skip content. Ultimately, these findings underscore the advantages associated with employing a multimodal presentation of recasts to enhance engagement and comprehension.

4.3.4. Relationship Between Attention and L2 Grammatical Development (RQ4)

To investigate the interaction effects of Group, Time and five attention measurements on the gains of the OPT, WPT, and SOPT, a series of linear mixed-effects models for the present progressive and logistic mixed-effects models for the third person singular were conducted. As previously explained, the models initially developed for the first and second research

questions, which addressed how the provision and modality of recasts affected L2 grammatical knowledge development over time, served as the null models. These null models included Group and Time as fixed effects, test scores as dependent variables, and participant and item as random effects. To these models, each attention measure was added as an additional fixed effect. The presence of a significant three-way interaction would indicate that the extent of gains achieved by the multimodal recast and written recast groups was differentially influenced by the attention measure in each model. Initially, each of these models was constructed as a maximal random effects structure, comprising random effects (Group and Item) and random slopes for each fixed effect. Models that did not converge were restructured by removing the random effects accounting for the least variance and were used for subsequent analyses.

4.3.4.1. OPT

4.3.4.1.1. Present Progressive

Table 60 provides the results of the linear mixed-effects models, which explored how learners' attention influenced the interaction effects of Group and Time on their knowledge gains in the present progressive, as assessed by the OPT. Significant positive three-way interaction effects were observed between Group, Time, and three attention measures (total fixation duration, re-reading time, and visit count) for learners' posttest and delayed posttest gains, all of which exhibited large effect sizes. Also, there was a positive interaction effect between first-pass reading time and the delayed posttest with a large effect size as well. Conversely, the analysis of skip rate revealed a negative significant three-way interaction at both Time 2 and Time 3. This means that participants in the multimodal recast group who skipped more multimodal recasts experienced lower immediate and delayed oral gains

compared to participants in the written recast group. Overall, these findings suggest that participants who had longer total fixation duration, engaging in more initial reading, devoted more time to re-reading, maintained a higher visit count, and demonstrated a lower skip rate, achieved higher gains in the multimodal recast group, compared to their counterparts in the written recast group.

Table 60. Results for the Linear Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the OPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Total fixation duration										
(Intercept)	–6.21	2.67	–2.33	.020*	.35	.37	<.01	<.01	.07	.26
Group:Time2:Total	1.33	0.61	2.17	.030*						
Group:Time3:Total	2.27	0.66	3.42	<.001***						
First-pass reading										
(Intercept)	–1.27	1.12	–1.14	.255	.35	.36	<.01	<.01	.07	.26
Group:Time2:First	0.01	0.41	0.03	.975						
Group:Time3:First	0.80	0.40	2.02	.044*						
Re-reading										
(Intercept)	–4.76	1.80	–2.65	.008**	.35	.37	<.01	<.01	.07	.26
Group:Time2:RRT	0.90	0.41	2.20	.028*						
Group:Time3:RRT	1.68	0.48	3.46	<.001***						
Visit										
(Intercept)	–1.98	0.31	–6.51	<.001***	.35	.37	<.01	<.01	.07	.26
Group:Time2:Visit	1.08	2.13	2.13	.034*						
Group:Time3:Visit	1.66	3.29	3.29	.001**						
Skip										
(Intercept)	–1.38	0.16	–8.61	<.001***	.35	.36	<.01	<.01	.07	.26
Group:Time2:Skip	–2.09	1.06	–1.97	.049*						
Group:Time3:Skip	–2.84	0.98	–2.89	.004**						
*Formula: lmer(log_OPT ~ Group*Time*log_attention measure + (1|Time) + (Group|Item))*										

Note. The written recast group was used as a reference level; Time1 = pretest (reference level); Time2 = posttest; Time3 = delayed posttest; Total = total fixation duration; First = first-pass reading time; RRT = re-reading time; Visit = visit count; Skip = skip rate.

4.3.4.1.2. Third Person Singular

Table 61 displays the findings obtained from logistic mixed-effects models, which examined the three-way interaction effects between Group, Time and three attention measurements (total fixation duration, first-pass reading time, and re-reading time) concerning participants' knowledge gains related to the third person singular in the OPT. The results revealed the significant interaction effects among Group, Time, and the aforementioned attention measures, particularly in relation to learners' delayed posttest gains. These effects were characterized by odds ratios ranging from small to large magnitudes. The odds ratios indicated that a one-unit increase in total fixation, first-pass reading, and re-reading durations corresponded to 7.35, 1.75, and 9.64 times higher oral production gains on the delayed posttest for the multimodal recast group compared to the written recast group. Overall, these findings suggest that in the multimodal recast group participants who engaged in longer fixation durations, extended initial reading times, and more reanalyzing durations while engaging with the target construction achieved greater delayed gains, compared to participants in the written recast group.

Table 61. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the OPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*z*	*p*	*OR*	95% CI for OR
Total fixation duration						
(Intercept)	–3.91	4.29	–0.91	.362	0.02	[0.00, 8.61]
Group:Time2:Total	0.23	1.68	0.14	.892	1.26	[0.05, 3.33]
Group:Time3:Total	6.60	1.82	3.61	<.001***	7.35	[2.37, 9.43]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.10	.32				
First-pass reading time						
(Intercept)	–0.79	3.32	–0.24	.811	0.45	[0.00, 2.92]
Group:Time2:First	–0.76	1.35	–0.56	.572	0.47	[0.03, 6.52]
Group:Time3:First	2.86	1.38	2.07	.038*	1.75	[1.19, 2.59]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.10	.32				
Re-reading time						
(Intercept)	–4.18	3.55	–1.18	.240	0.02	[0.00, 1.57]
Group:Time2:RRT	0.55	1.32	0.42	.676	1.74	[0.13, 2.27]
Group:Time3:RRT	4.57	1.34	3.40	<.001***	9.64	[7.03, 11.32]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.06	.23				
*Formula: glmer(OPT_PS ~ 1 + Group*Time*log_attention measurement + (1|Participant) + (Group|Item))*						

Positively significant interaction effects were found between Group, Time, and visit count concerning posttest and delayed posttest gains, displaying large effect sizes (Table 62). This implies that participants in the multimodal recast group who visited multimodal recasts more often achieved higher oral production gains in both the posttest and delayed posttest, compared to those in the written recast group. This suggests a substantial disparity in the impact of visit count on oral production gains between these two groups. Furthermore, the odds ratio analysis revealed that oral production gains were significantly predicted by the frequency of visits to the target construction. Specifically, a one-unit increase in visit count

corresponded to 6.95 times higher oral production gains on the posttest and 6.00 times higher gains on the delayed posttest for the multimodal recast group.

In contrast, when examining skip rate, negatively significant three-way interactions emerged at both Time 2 and Time 3. This suggests that participants in the multimodal recast group, who demonstrated a higher tendency to skip the target construction, experienced reduced gains in both the posttest and delayed posttest in comparison to their counterparts in the written recast group. This points to a meaningful distinction in the influence of skip rate on oral production gains between these two groups. Moreover, the odds ratio accentuated a significant negative correlation between skip rate and subsequent oral gains. Specifically, a one-unit increase in skip rate was associated with a 52% decrease in immediate and delayed oral production gains. Overall, these findings underscore that participants who engaged in a higher number of visits and exhibited a lower skip rate achieved heightened gains within the multimodal recast group.

Table 62. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the OPT (Visit count and Skip rate) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*z*	*p*	OR	95% CI for OR
Visit count						
(Intercept)	–1.43	0.26	–5.58	<.001***	0.25	[0.16, 0.39]
Group:Time2:Visit	2.06	1.03	2.00	.046*	6.95	[0.82, 58.63]
Group:Time3:Visit	1.98	0.97	2.03	.042*	6.99	[1.05 46.34]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.10	.30				
Skip rate						
(Intercept)	–2.28	0.38	–6.02	<.001***	0.10	[0.05, 0.21]
Group:Time2:Skip	–0.65	0.30	–2.14	.033*	0.52	[0.29, 0.95]
Group:Time3:Skip	–0.66	0.30	–2.23	.026*	0.52	[0.29, 0.92]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.10	.30				
*Formula: glmer(OPT_PS ~ 1 + Group*Time*log_attention measurement + (1|Participant) + (Group|Item))*						

4.3.4.2. WPT

4.3.4.2.1. Present Progressive

Another series of linear mixed-effect model analyses was performed for the WPT, employing the same models. The results, as summarized in Table 63, also revealed several significant interaction effects with large effect sizes. Similar to the OPT results, the three-way interaction effects were positively significant for total fixation duration, re-reading time, and visit count, both in the posttest and delayed posttest, while first-pass reading time displayed a positive interaction only in the delayed posttest. These findings suggest that the multimodal recast group benefited more from increased fixation duration, active first-pass-reading, and re-reading of the target construction, as well as making repeated visits to it, compared to the written recast group. Increased time devoted to fixation, first-pass reading, re-reading, and

frequent visits to the target structure, particularly in multimodal recasts, were associated with better posttest and delayed posttest performance in both oral and written production tests.

Conversely, the analysis of skip rate revealed a significant negative three-way interaction at both Time 2 and Time 3, suggesting that the effect of skip rate on written gains differed between the multimodal recast and written recast groups. The negative coefficient for the Group, Time, and skip rate interaction indicates that an increase in skip rate for participants in the multimodal recast group was associated with decreased posttest and delayed posttest gains. That is, participants in the multimodal recast group who skipped more multimodal recasts during posttest and delayed posttest tended to have lower written gains compared to participants in the written recast group. Overall, these findings demonstrated that the impacts of attention measurements on learners' written gains were influenced by Group, Time, and the specific attention measurement examined.

Table 63. Results for the Linear Mixed-Effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the Written Production Test – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Total fixation duration										
(Intercept)	–6.34	2.62	–2.42	.015*	.37	.39	<.01	<.01	.07	.26
Group:Time2:Total	1.48	0.60	2.46	.014*						
Group:Time3:Total	2.60	0.65	4.00	<.001***						
First-pass reading										
(Intercept)	–1.27	1.10	–1.16	.246	.37	.39	<.01	<.01	.07	.26
Group:Time2:First	0.10	0.40	0.24	.808						
Group:Time3:First	0.90	0.39	2.30	.021*						
Re-reading										
(Intercept)	–4.86	1.76	–2.76	.006**	.37	.39	<.01	<.01	.07	.26
Group:Time2:RRT	0.93	0.40	2.30	.021*						
Group:Time3:RRT	1.74	0.47	3.66	<.001***						
Visit counts										
(Intercept)	–1.99	0.30	–6.62	<.001***	.37	.39	<.01	<.01	.07	.26
Group:Time2:Visit	1.31	0.50	2.64	.008**						
Group:Time3:Visit	1.84	0.49	3.73	<.001***						
Skip rate										
(Intercept)	–1.97	0.26	–7.57	<.001***	.37	.39	<.01	<.01	.07	.26
Group2:Time2:Skip	–0.32	0.16	–2.00	.045*						
Group2:Time3:Skip	–0.48	0.15	–3.11	.002**						

*Formula: lmer(log_WPT ~ Group*Time*log_attention measure + (1|Participant) + (Group|Item))*

4.3.4.2.2. Third Person Singular

Table 64 depicts the results of logistic mixed-effects models, which explored the three-way interaction effects involving Group, Time and three attention measures (total fixation duration, first-pass reading time, and re-reading time). The results revealed significant three-way positive interactions between Group, Time 2, and all three attention measures. The odds ratios indicated that a one-unit increase in total fixation, first-pass reading time and re-reading durations resulted in 3.45, 1.75, and 2.62 times higher written gains on the posttest for the multimodal recast group compared to the written recast group. These findings suggest that

these three measures had a more favorable impact on learners' immediate gains in the multimodal recast group in comparison to the written recast group.

Table 64. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the WPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*z*	*p*	*OR*	95% CI for OR
Total fixation duration						
(Intercept)	–2.92	4.12	–0.71	.478	0.05	[0.00, 1.67]
Group:Time2:Total	3.54	1.61	2.20	.028*	3.45	[1.50, 7.92]
Group:Time3:Total	0.61	1.57	0.39	.695	1.85	[0.87, 3.92]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	<.01	<.01				
First-pass reading time						
(Intercept)	–0.34	3.26	–0.11	.916	0.71	[0.00, 4.10]
Group:Time2:First	2.88	1.39	2.17	.038*	1.75	[1.20, 2.61]
Group:Time3:First	–0.12	1.33	–0.09	.926	0.88	[0.07, 1.18]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	<.01	.04				
Re-reading time						
(Intercept)	–3.32	3.42	–0.97	.332	0.04	[0.00, 2.85]
Group:Time2:RRT	3.27	1.30	2.50	.012*	2.62	[2.06, 3.33]
Group:Time3:RRT	0.62	1.27	0.49	.627	1.85	[0.16, 2.19]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	<.01	.04				
Formula: glmer(WPT_PS ~ 1 + Group Time*log_attention measurement + (1|Participant) + (Group|Item))*						

Significant positive three-way interaction effects were also detected when incorporating visit count at Time 2 and Time 3 (Table 65). The odds ratio indicated that written production gains correlated significantly with the frequency of visits to the target construction. Specifically, a one-unit increase in visit count corresponded to gains that were 8.83 and 6.99 times higher on the posttest and delayed posttest respectively in the multimodal recast group compared to the written recast group. As for skip rate, significant negative interaction effects were observed at

both Time 2 and Time 3, with odds ratios of 0.51 and 0.52 respectively. This suggests that participants in the multimodal recast group, who demonstrated a higher tendency to skip the target construction, experienced significantly reduced gains in both posttest and delayed posttest in comparison to their counterparts in the written recast group. These findings highlight a substantial difference in the impact of skip rate on written production gains between the two groups. Overall, as with the findings from the OPT, these results emphasize that participants who engaged more frequently and had a lower skip rate achieved greater gains within the multimodal recast group.

Table 65. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the WPT (Visit Count and Skip Rate) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*z*	*p*	*OR*	95% CI for OR
Visit count						
(Intercept)	–1.40	0.23	–6.05	<.001***	0.25	[0.16, 0.39]
Group2:Time2:Visit	2.18	1.09	1.99	.046*	8.83	[1.05, 54.40]
Group2:Time3:Visit	1.95	0.97	2.00	.045*	6.99	[1.05, 46.51]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	<.01	.04				
Skip rate						
(Intercept)	–2.20	0.36	–6.16	<.001***	0.11	[0.06, 0.22]
Group2:Time2:Skip	–0.67	0.32	–2.10	.036*	0.51	[0.27, 0.95]
Group2:Time3:Skip	–0.66	0.30	–2.22	.027*	0.52	[0.29, 0.92]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	<.01	.04				
*Formula: glmer(WPT_PS ~ 1 + Group*Time*log_attention measurement + (1|Participant) + (Group|Item))*						

4.3.4.3. SOPT

4.3.4.3.1. Present Progressive

Table 66 provides the results of linear mixed-effects models investigating the intricate relationship between learners' attention, Group, and Time concerning their knowledge gains in the present progressive on the SOPT. Similar to the findings observed in the OPT and WPT, significant positive three-way interaction effects were identified between Group, Time, and three attention measures (total fixation duration, re-reading time, and visit count) in both posttest and delayed posttest gains, all of which displayed large effect sizes, while first-pass reading time exerted a significant and positive impact on learners' posttest gains. Conversely, the analysis of skip rate revealed negative significant three-way interaction effects on the multimodal recast group's posttest and delayed posttest gains. Overall, these findings indicate that in the multimodal recast group, participants who exhibited longer total fixation duration, invested more time in first reading, re-reading, maintained a higher visit count, and demonstrated a lower skip rate consistently achieved higher gains across various assessment points when compared to their counterparts in the written recast group. These results imply notable contrasts in the relationship between five attention metrics on the SOPT gains between the two groups.

Table 66. Results for the Linear Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the SOPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Total fixation duration										
(Intercept)	–6.30	2.59	–2.44	.015*	.38	.40	<.01	<.01	.07	.27
Group:Time2:Total	1.50	0.59	2.53	.012*						
Group:Time3:Total	1.98	0.64	3.08	.002**						
First-pass reading										
(Intercept)	–1.00	1.08	–0.93	.352	.38	.40	<.01	<.01	.07	.27
Group:Time2:First	0.17	0.40	0.44	.663						
Group:Time3:First	0.79	0.38	2.05	.041*						
Re-reading										
(Intercept)	–5.04	1.74	–2.90	.004**	.38	.40	<.01	<.01	.07	.27
Group:Time2:RRT	0.97	0.40	2.42	.016*						
Group:Time3:RRT	1.20	0.47	2.55	.011*						
Visit										
(Intercept)	–2.03	0.30	–6.85	<.001***	.38	.40	<.01	<.01	.07	.27
Group:Time2:Visit	1.26	0.49	2.57	.010*						
Group:Time3:Visit	1.41	0.49	2.89	.003**						
Skip										
(Intercept)	–1.43	0.16	–9.06	<.001***	.38	.40	<.01	<.01	.07	.27
Group:Time2:Skip	–2.17	1.03	–2.11	.003*						
Group:Time3:Skip	–2.40	0.95	–2.52	.012*						
*Formula: lmer(log_SOPT ~ Group*Time*log_attention measure + (1|Time) + (Group|Item))*										

4.3.4.3.2. Third Person Singular

Table 67 presents the outcomes of logistic mixed-effects models, which examined the three-way interaction effects involving total fixation duration, first-pass reading time, and re-reading time in relation to the knowledge gains for the third person singular in the SOPT. The results revealed significant and positive three-way interactions among Group, Time, and the three attention metrics, particularly influencing posttest gains in the multimodal recast group, displaying medium to large effect sizes. Specifically, in the multimodal recast group, each one-unit increase in total fixation, first-pass reading, and re-reading durations corresponded to 2.75, 1.64, and 1.88 times higher immediate, spontaneous oral production gains on the

posttest when compared to the written recast group. This highlights that learners in the multimodal recast group derived greater benefits from investing more time in fixation, engaging deeply in initial reading, and revisiting the feedback in terms of immediate spontaneous oral production gains, compared to their counterparts in the written recast group.

Table 67. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the SOPT (Total Fixation Duration, First-pass Reading Time, Re-reading Time) – Third Person Singular

Fixed Effects	*Estimate*	*SE*	*z*	*p*	OR	95% CI for OR
Total fixation duration						
(Intercept)	–3.21	4.31	–0.75	.456	0.04	[0.00, 1.80]
Group:Time2:Total	5.62	1.69	3.33	<.001***	2.75	[1.02, 7.39]
Group:Time3:Total	1.23	1.60	0.77	.439	3.46	[1.52, 7.88]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.15	.39				
First-pass reading time						
(Intercept)	–0.54	3.34	–0.16	.872	0.58	[0.00, 3.97]
Group:Time2:First	2.80	1.37	2.04	.041*	1.64	[1.13, 2.37]
Group:Time3:First	0.31	1.34	0.23	.815	1.27	[1.01, 1.85]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.15	.39				
Re-reading time						
(Intercept)	–3.56	3.57	–1.00	.319	0.03	[0.00, 3.01]
Group:Time2:RRT	4.26	1.33	3.21	.001**	1.88	[1.20, 2.69]
Group:Time3:RRT	0.94	1.29	0.73	.468	2.55	[2.06, 3.15]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.15	.39				
*Formula: glmer(SOPT_PS ~ 1 + Group*Time*log_attention measure + (1|Participant) + (Group|Item))*						

Positive significant interaction effects were also found between Group, Time, and visit count concerning posttest and delayed posttest gains, displaying large effect sizes (Table 68). This implies that participants in the multimodal recast group who engaged with multimodal recasts more frequently exhibited higher SOPT gains in both posttest and delayed posttest compared

to those in the written recast group. This suggests a substantial divergence in the impact of visit count on SOPT gains between these two groups. Furthermore, the odds ratio analysis revealed that SOPT gains were significantly predicted by the frequency of visits to the third person singular form. Specifically, a one-unit increase in visit count corresponded to higher immediate gains of 12.11 times and 9.67 times higher gains on the delayed posttest for the multimodal recast group. In contrast, when examining skip rate, significant negative three-way interactions were identified at both Time 2 and Time 3, as indicated in Table 74. This suggests that participants in the multimodal recast group, who demonstrated a greater tendency to skip the target construction, experienced substantially reduced gains in both posttest and delayed posttest in comparison to their counterparts in the written recast group. This indicates a significant distinction in the influence of skip rate on spontaneous oral production gains between these two groups. Furthermore, the odds ratio highlighted a significant negative correlation between skip rate and subsequent SOPT gains. Specifically, a one-unit increase in skip rate was associated with a 46% decrease in posttest gains and a 51% decrease in gains on the delayed posttest. Overall, these findings underscore the finding that participants who engaged in a higher number of visits and exhibited a lower skip rate achieved significant advancements within the multimodal recast group.

Table 68. Results for the Logistic Mixed-effects Models Examining the Moderating Effects of Attention Allocation on Participants' Scores on the SOPT (Visit count and Skip rate) – Third Person Singular

	Estimate	*SE*	*z*	*p*	OR	95% CI for OR
Visit count						
(Intercept)	–1.52	0.27	–5.57	<.001***	0.22	[0.13, 0.37]
Group:Time2:Visit	2.49	1.04	2.40	.017*	12.11	[1.59, 92.25]
Group:Time3:Visit	2.23	0.98	2.27	.023*	9.28	[1.37, 65.86]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.15	.39				
Skip rate						
(Intercept)	–2.24	0.39	–5.82	<.001***	0.11	[0.05, 0.23]
Group:Time2:Skip	–0.79	0.31	–2.57	.010*	0.46	[0.25, 0.83]
Group:Time3:Skip	–0.67	0.30	–2.24	.025*	0.51	[0.28, 0.92]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.15	.39				
*Formula: glmer(SOPT_PS ~ 1 + Group*Time*log_attention measure + (1|Participant) + (Group|Item))*						

4.3.5. Effects of WM in Attention Allocation and L2 Grammatical Knowledge Development (RQ5)

Research question 5 addressed the extent to which individual differences in WM capacity moderated the effects of feedback modality on development of L2 grammatical knowledge. Participants' WM was assessed using five different WM measures: (a) NWS, (b) forward Corsi block task, (c) color shape task, (d) stop signal task, and (e) AOSPAN. The descriptive statistics for each measure are summarized in Table 69. Table 76 shows the results of one-way ANOVAs, which found no significant differences among the four groups on any of the WM measurements.

Table 69. Descriptive Statistics for WM Measures

					95% CI	
		N	*M*	*SD*	*Lower*	*Upper*
NWS	Control	30	8.43	1.01	8.06	8.81
	Oral	30	8.30	1.09	7.89	8.71
	Written	30	8.43	1.14	8.01	8.86
	Multi	30	8.90	1.09	8.49	9.31
Forward	Control	30	60.00	3.96	58.52	61.48
Corsi block	Oral	30	57.33	4.36	55.71	58.96
	Written	30	58.40	5.49	56.35	60.45
	Multi	30	58.37	3.20	57.17	59.56
Color shape	Control	30	384.42	43.65	368.13	400.72
	Oral	30	381.36	44.22	364.85	397.87
	Written	30	401.70	50.35	382.90	420.50
	Multi	30	402.70	52.54	383.08	422.32
Stop signal	Control	30	283.05	20.23	275.49	290.60
	Oral	30	280.58	19.83	273.18	287.98
	Written	30	273.57	16.31	267.48	279.66
	Multi	30	274.05	16.88	267.75	280.35
AOSPAN	Control	30	56.63	3.89	55.18	58.09
	Oral	30	57.17	4.67	55.42	58.91
	Written	30	57.73	4.56	56.03	59.43
	Multi	30	58.13	3.89	56.68	59.59

Note. NWS max. total score = 11 (words); Corsi block max. total score = 88 (blocks); AOSAPN max. total score = 75 (items).

Table 70. Results of One-way ANOVAs on Five WM Measures

	df	***F***	***p***	η_p^2
Nonword	3, 116	1.78	.156	0.04
Forward Corsi block	3, 116	1.94	.127	0.05
Color shape	3, 116	1.65	.181	0.04
Stop signal	3, 116	1.99	.120	0.05
AOSPAN	3, 116	0.71	.549	0.02

Construct validity was assessed by analyzing the correlations among the five WM measures. The results, presented in Table 71, included all four groups as well as individual groups. While no significant correlations emerged among most of the WM measures, there were medium to large positive correlations between the NWS and AOSPAN tasks for the four groups combined (*rho* =.459, p <.001), the control group (*rho* =.852, p <.001), and the oral

recast group (*rho* =.545, *p* =.002). This indicates some degree of overlap in the constructs assessed by these two tasks. Based on this observation, it appears that the forward Corsi block task, as well as stop signal task and color shape task tap into distinctive functions of the central executive.

Table 71. Spearman Correlation Matrix of Five WM Measures

	NWS	Corsi block	Stop signal	Color shape	AOSPAN
Combined (*N* = 120)					
NWS	1	.009 (.925)	–.147 (.109)	.064 (.490)	**.459*** (.000)**
Forward Corsi block		1	–.086 (.350)	.079 (.392)	–.105 (.255)
Stop signal			1	–.002 (.981)	–.126 (.171)
Color shape				1	–.002 (.980)
AOSPAN					1
Control (*N* = 30)					
NWS	1	–.013 (.947)	–.235 (.211)	–.017 (.931)	**.852*** (.000)**
Forward Corsi block		1	–.248 (.187)	–.110 (.563)	–.081 (.671)
Stop signal			1	.164 (.388)	–.122 (.522)
Color shape				1	.119 (.531)
AOSPAN					1
Oral (*N* = 30)					
NWS	1	.232 (.217)	–.167 (.378)	.012 (.948)	**.545*** (.002)**
Forward Corsi block		1	–.301 (.106)	.007 (.969)	–.174 (.357)
Stop signal			1	–.240 (.201)	–086 (.651)
Color shape				1	.192 (.310)
AOSPAN					1
Written (*N* = 30)					
NWS	1	–.241 (.199)	–.140 (.460)	–.114 (.548)	.190 (.315)
Forward Corsi block		1	.029 (.879)	.116 (.541)	–.184 (.331)
Stop signal			1	.217 (.250)	–.153 (.419)
Color shape				1	–.332 (.073)
AOSPAN					1
Multimodal (*N* = 30)					
NWS	1	.260 (.165)	–.085 (.655)	.192 (.310)	.204 (.280)
Forward Corsi block		1	–.022 (.906)	.221 (.240)	.227 (.229)
Stop signal			1	–.005 (.980)	–.335 (.070)
Color shape				1	–.068 (.719)
AOSPAN					1

4.3.5.1. Present Progressive

To examine the potential impact of learners' WM capacity on their use of the present progressive in the OPT, WPT, and SOPT, a set of linear mixed-effects model analyses were performed (see Appendix A32–A42 for a comprehensive presentation of the results for the five WM measures related to the present progressive form). These analyses utilized the same models as were previously employed to explore the relationships between the feedback modality and learning gains in the three outcome measures. The initial models included Group and Time as fixed effects, while participant and item were treated as random effects, creating a null model. Subsequently, each WM measure was introduced as an additional fixed effect to the null model. This allowed for the investigation of significant interactions among Group, Time, and WM measures in relation to the scores attained through the three assessment tests. The dependent variables were test scores. Among the various measures examined, the inclusion of NWS and AOSPAN into the models yielded significant three-way interaction effects. The significant Group–Time–NWS interaction effects were detected when evaluating participants' use of the present progressive in the SOPT, while significant Group–Time–AOSPAN interaction effects were detected when the participants' use of the present progressive was evaluated by the OPT, WPT, and SOPT.

First, the analysis of NWS revealed four significant three-way interactions in the SOPT (see Appendix A27 for full model results). Table 72 presents the results of post hoc contrasts of these significant interactions. The results show that the significant interaction effects were evident in the oral recast group compared to both the control group and the written recast group during Time 2 and Time 3 with medium to large effect sizes. These findings suggest that learners who exhibited larger NWS demonstrated more substantial

pretest–posttest and pretest–delayed posttest improvements in the oral recast group on the SOPT.

Table 72. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of SOPT (NWS) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
ControlOral:Time2	2.37	1.11	2.13	.033	.14	.16	<.01	<.01	<.01	<.01
ControlOral:Time3	2.62	1.11	2.36	.018	.11	.13	<.01	<.01	<.01	<.01
WrittenOral:Time2	2.36	1.07	2.20	.028	.20	.24	<.01	<.01	.09	.30
WrittenOral:Time3	2.85	1.11	2.55	.011	.16	.19	<.01	<.01	.03	.18
*Formula: lmer(log_SOPT ~ Group*Time*log_NWS + (1|Participant) + (Group|Item))*										

In examining the influence of AOSPAN, the results from the linear mixed-effect model analyses concerning the OPT revealed two significant positive three-way interaction effects (see Appendix A28 for full model results). Table 73 illustrates the results of post hoc contrasts for these significant interactions. The results indicated that the significant interaction effects were observed in the oral recast and written recast groups compared to the control group during Time 3 with small to medium effect sizes. The presence of these positive three-way interactions suggests that the influence of Group and the delayed posttest on learners' performance in the OPT was moderated by their AOSPAN scores. Additionally, the medium effect sizes associated with these interactions indicate that the relationships were of substantive magnitude. Overall, these findings show that individuals with higher AOSPAN scores, which involve cognitive functions such as maintaining sequences of tasks and manipulating information, were more capable of effectively utilizing oral and written recasts to enhance oral production skills.

Table 73. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (AOSPAN) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
ControlOral:Time3	4.36	1.88	2.32	.021	.10	.11	<.01	<.01	<.01	<.01
ControlWritten:Time3	–3.80	1.89	–2.01	.044	.13	.14	<.01	<.01	<.01	.10
*Formula: lmer(log_SOPT ~ Group*Time*log_AOSPAN + (1|Participant) + (Group|Item))*										

Likewise, the linear mixed-effect model analyses conducted for AOSPAN on the WPT yielded results analogous to those observed in the OPT (see Appendix A29 for full model results). Expanding upon these findings, Table 74 presents the results of post hoc contrast of these interactions. The outcomes revealed positive significant three-way interaction effects between Group, Time 3, and AOSPAN for the oral recast and written recast groups compared to the control group. These effects were characterized by medium to large effect sizes. These findings indicate that learners' AOSPAN played a significant role in their ability to effectively utilize the oral recasts and written recasts to enhance their oral production skills, particularly concerning their gains from the pretest to the delayed posttest.

Table 74. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of WPT (AOSPAN) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
ControlOral:Time3	–4.08	1.87	–2.19	.029	.10	.11	<.01	<.01	<.01	<.01
ControlWritten:Time3	–3.83	1.88	–2.04	.041	.13	.14	<.01	<.01	.01	.10

The analyses conducted using linear mixed-effect model for the SOPT also revealed four significant positive three-way interaction effects (see Appendix A30 for full model results). Further insights gained from post hoc analyses, presented in Table 75, illustrate the

significant interactions among Group, Time 2, and AOSPAN within the oral recast group relative to the control group with a medium to large effect size. Moreover, interactions involving Group, Time 3, and AOSPAN were observed within both the oral recast and written recast groups. This finding suggests that the influence of AOSPAN on the immediate and delayed gains in the SOPT was significantly stronger for the oral recast group compared to the control group. The delayed gains of the oral recast and written recast groups demonstrated a more substantial impact of AOSPAN in contrast to the control group. Interestingly, a significant negative interaction emerged involving Group, Time 3, and AOSPAN within the multimodal recast group when compared to the oral recast group. This suggests that, while greater AOSPAN capacity appeared advantageous for learners in the oral recast group, its impact on SOPT gains was less pronounced when learners were exposed to multimodal recasts.

Table 75. Results of Post Hoc Contrasts for Significant Interactions Identified by Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of SOPT (AOSPAN) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
ControlOral:Time2	4.16	1.91	2.18	.029	.13	.16	<.01	<.01	<.01	<.01
ControlOral:Time3	6.78	1.84	3.68	<.001	.11	.13	<.01	<.01	<.01	<.01
ControlWritten:Time3	3.82	1.85	2.07	.039	.14	.15	<.01	<.01	<.01	.09
OralMulti:Time3	–3.97	1.90	–2.09	.036	.37	.39	<.01	<.01	.04	.27

4.3.5.2. Third Person Singular

To examine the potential impact of learners' WM capacity on the production of the third person singular in the three assessment tests, a series of logistic mixed-effects model analyses were performed (see Appendix A43–A56 for a comprehensive presentation of the results for the five working memory measures pertaining to the third person singular form). Unlike the

results of the present progressive, these analyses only yielded one significant negative three-way interaction effect. This effect emerged between Group, Time 3, and stop signal task scores concerning the gains in the SOPT for the multimodal recast group compared to the oral recast group (see Appendix A31 for full model results). As shown in Table 76, this interaction exhibited a large effect size. This finding suggests that in the oral recast group, learners performing better on the stop signal task, which assessed the inhibitory control experienced more pronounced pretest–delayed posttest gains in spontaneous oral production in contrast to their counterparts in the multimodal recast group.

Table 76. Significant Interactions Identified by Logistic Mixed-effects Models Examining Participants' Scores on the Spontaneous Oral Production Test – Third Person Singular (Stop Signal)

	Fixed effects						Random effects			
							by participant (intercept)		*by item* (intercept)	
	Estimate	*SE*	*z*	*p*	*OR*	95% CI	*Variance*	*SD*	*Variance*	*SD*
ControlOral:Time2	–8.00	3.08	–2.60	.009	0.00	[0.00, 1.34]	<.01	<.01	.29	.54

4.3.6. Relationship Between Attention and WM (RQ6)

To explore the relationship between learners' visual attention to the target constructions and their WM capacity (RQ6), a set of Spearman correlation analyses was conducted for the multimodal and written recast groups, respectively. The objective of these analyses was to delve into the interrelation between the five eye-tracking metrics and the five measures of WM.

4.3.6.1. Present Progressive

Regarding the present progressive form, the results, presented in Table 77, revealed that learners' visual attention allocation was correlated with different eye-tracking measures depending on their feedback conditions. In the multimodal group, only NWS scores were found to have positive and significant correlations, with medium to large effect sizes, with three attention measures (i.e., first-pass reading time, re-reading time, visit count), while exhibiting a negative correlation with the skip rate.

In the written recast group, on the other hand, several intriguing correlations were identified. First, learners' NWS scores displayed a negative correlation with visit count, suggesting that learners with better PSTM tended to spend less time revisiting the target form while processing the feedback. Forward Corsi block, which measures VSTM, exhibited correlations with various WM measure. Except for first-pass reading time, forward Corsi block task scores showed strong negative correlations with total fixation duration, re-reading time, and visit count. Moreover, this measure displayed a strong positive correlation with skip rate. This implies that individuals with stronger VSTM demonstrated a tendency to allocate less time to fixation, re-reading, and revisiting the target form, opting instead to skip over it.

Overall, the differences observed in the relationship between the five attention metrics and the WM measured across the two groups suggest that the nature of recasts received

(multimodal versus written recasts) can influence the interplay between WM capacity and attentional processes when processing distinct types of feedback. While PSTM appears to have a substantial influence in both groups, other aspects of WM, notably VSTM, may exert more pronounced effects within the written recast group.

Table 77. Correlation Analyses Between Working Memory and Eye-tracking Measures – Present Progressive

		Total duration	First reading	Re-reading	Visit count	Skip rate
Multimodal recast group						
NWS	*rho*	**.454*****	**.355****	**.359****	**.545*****	**–.286*****
	p	**(.000)**	**(.001)**	**(.001)**	**(.000)**	**(.006)**
Forward Corsi block	*rho*	.118	.015	.172	.182	–.003
	p	(.267)	(.888)	(.106)	(.086)	(.980)
Color shape	*rho*	–.067	–.051	–.006	.039	.040
	p	(.531)	(.630)	(.952)	(.717)	(.707)
Stop signal	*rho*	–.092	–.107	–.119	–.136	.128
	p	(.386)	(.314)	(266)	(.200)	(.228)
AOSPAN	*rho*	–.058	–.030	–.065	–.014	–.184
	p	(.589)	(.776)	(.542)	(.893)	(.083)
Written recast group						
NWS	*rho*	–.013	.065	–.001	**–.333****	.014
	p	(.902)	(.544)	(.991)	**(.001)**	(.895)
Forward Corsi block	*rho*	**–.383*****	.118	**–.415*****	**–.505*****	**.557*****
	p	**(.000)**	(.269)	**(.000)**	**(.000)**	**(.000)**
Color shape	*rho*	.074	–.039	.146	.124	**–**.151
	p	(.486)	(.717)	(.170)	(.246)	(.154)
Stop signal	*rho*	.063	.093	**–**.010	.207	.102
	p	(.553)	(.393)	(.927)	(.051)	(.337)
AOSPAN	*rho*	.084	.010	.078	.172	**–**.167
	p	(.433)	(.928)	(.463)	(.106)	(.116)

4.3.6.2. Third Person Singular

Table 78 shows the results of Spearman correlation analyses conducted on the third person singular form. First, in the multimodal group, significant correlations were observed between NWS and color shape task scores with visit count and first-pass reading time, exhibiting a

modest magnitude. This finding suggests that the participants' PSTM and task-switching abilities played a more prominent role in directing their attention allocation, surpassing the significance of other WM indicators in the context of multimodal recasts. This also suggests that learners with enhanced PSTM and task-switching abilities displayed heightened attentiveness to the target constructions when exposed to multimodal recast conditions.

Turning to the written recast group, the findings revealed intriguing associations. Firstly, NWS showed negative correlations of small to moderate magnitude with total fixation duration, re-reading time, and visit count. This pattern indicates that individuals with stronger PSTM tended to allocate less time to fixate on the target form, engage in re-reading, and revisits to the target structure. Secondly, forward Corsi block task scores demonstrated positive correlations with skip rate. This suggests that participants with advanced VSTM abilities were more prone to frequently skip over the target construction in the context of written recasts. Additionally, stop signal task scores displayed both positive correlations with visit count and negative correlations with skip rate. These correlations exhibited varying magnitudes: moderate for visit count and small for skip rate. This indicates that learners with greater inhibitory control within their WM were more inclined to revisit the target structure frequently while refraining from skipping it.

Furthermore, the AOSPAN exhibited small-sized positive correlations with total fixation duration and first-pass reading time, along with a small-sized negative correlation with skip rate. This suggests that learners with adept information updating abilities (as captured by AOSPAN) tended to sustain longer fixation periods during reading, emphasizing the initial reading phase, and also displayed an increased tendency to skip the target construction when provided with written recasts.

Overall, the differences observed in the association between the five attention measures and WM indicators in the two groups suggest that the nature of the recasts provided

(multimodal versus written) influenced how learners' WM capacity intertwined with their attentional processes during the process of understanding and incorporating the feedback.

Table 78. Correlation Analyses Between Working Memory and Eye-tracking Measures – Third Person Singular

		Total duration	First reading	Re-reading	Visit count	Skip rate
Multimodal recast group						
NWS	*rho*	–.014	–.027	–.009	**.229***	.171
	p	(.894)	(.802)	(.935)	**(.030)**	(.108)
Forward Corsi block	*rho*	**–**.002	**–**.184	.117	.088	.125
	p	(.986)	(.083)	(.274)	(.411)	(.241)
Stop signal	*rho*	–.038	.097	–.137	–.109	.081
	p	(.723)	(.365)	(.198)	(.306)	(.448)
Color shape	*rho*	.136	**.242***	.048	.079	**–**.091
	p	(.200)	**(.022)**	(.653)	(.459)	(.396)
AOSPAN	*rho*	.006	–.042	.041	.039	–.023
	p	(.957)	(.692)	(.704)	(.713)	(.826)
Written recast group						
NWS	*rho*	**–.438*****	**–**.147	**–.488*****	**–.294****	.201
	p	**(.000)**	(.168)	**(.000)**	**(.005)**	(.058)
Forward Corsi block	*rho*	–.007	.035	.004	–.097	**.294****
	p	(.949)	(.746)	(.973)	(.364)	**(.005)**
Stop signal	*rho*	.140	**–**.018	.096	**.332*****	**–.228***
	p	(.190)	(.867)	(.369)	**(.001)**	**(.031)**
Color shape	*rho*	.199	.094	.191	**.223***	.072
	p	(.060)	(.378)	(.072)	**(.035)**	(.502)
AOSPAN	*rho*	**.246***	**.228***	.177	.167	**–.243***
	p	**(.019)**	**(.030)**	(.096)	(.115)	**(.021)**

4.3.7. Summary of Results

RQ1. To what extent do recasts affect development in L2 grammatical knowledge during video-conferencing tasks?

The provision of recasts, as opposed to the absence of recasts, positively influenced the development of the present progressive form. Across all three outcome measures (OPT, WPT, and SOPT), all groups exposed to recasts displayed greater improvements in both the

pretest–posttest and pretest–delayed posttest gains when compared to the control group. The effect sizes of these gains ranged from small to medium. The same pattern was evident when analyzing the third person singular. Once again, all experimental groups outperformed the control group on both the posttest and the delayed posttest across all three outcome measures, with effect sizes falling within the small to medium range.

RQ2. To what extent does modality of recasts affect development in L2 grammatical knowledge during video-conferencing tasks?

The results provided strong evidence that multimodal recasts were more effective than either oral or written recasts in enhancing participants' proficiency in utilizing the present progressive and the third person singular forms. This conclusion was supported by the outcomes from all three assessment tasks, including both the immediate posttest and delayed posttest, with effect sizes indicating moderate differences in performance. Notably, there were no significant gain differences between the oral recast and the written recast groups across any of the three outcomes measured. In summary, the data clearly indicate that multimodal recasts played a highly effective role in facilitating development in learners' use of the two target linguistic constructions.

RQ3. To what extent does modality of recasts (in the case of written versus multimodal recasts) affect learners' visual attentional allocation to the target linguistic constructions during video-conferencing tasks, as reflected in participants' eye movements?

Across all five eye-tracking measures (i.e., total fixation duration, first-pass reading time, re-reading time, visit count, and skip rate), significant differences in the visual attention allocation toward the target linguistic structures emerged when comparing the multimodal recast group to the written recast group. These differences were associated with effect sizes

that ranged from small to large. Overall, learners in the multimodal recast group consistently exhibited significantly longer total fixation duration, extended first-pass reading time, and increased re-reading duration for both target structures when compared to the written recast group. Moreover, the multimodal recast group demonstrated a significantly higher number of visits to and a lower skip rate for both target constructions in comparison to the written recast group. Taken together, these findings suggest that, as compared to written recasts, multimodal recasts more effectively captured and retained learners' visual attention to the target constructions.

RQ4. To what extent is learners' visual attention to the target constructions, included in recasts, related to the development in L2 grammatical knowledge? Is this relationship influenced by whether learners are provided with written or multimodal recasts?

Several significant associations were identified between the amount of visual attention directed to the target constructions and the development of L2 grammatical knowledge. The nature of these significant relationships varied across the multimodal recast group and the written recast groups. Regarding the present progressive form, the findings revealed that three attention metrics – total fixation duration, re-reading time, and visit count – significantly predicted immediate and delayed posttest gains for the multimodal recast group across all three outcome measures. Additionally, first-pass reading time emerged as a predictor for the multimodal recast group's delayed posttest gains on the three assessment tasks. In line with this, participants in the multimodal recast group who showed lower skip rate demonstrated significantly higher immediate and delayed posttest gains on all assessments, with large effect sizes. These results suggest that learners who paid more attention to the target form entailed in multimodal recasts demonstrated greater advancements in oral, written, and spontaneous oral production compared to the written recast group.

For the third person singular form, total fixation duration, first-pass reading time, and re-reading time were significant predictors of the delayed posttest gains of the multimodal recast group in the OPT. In contrast, when it comes to gains measured by the WPT and SOPT, these same attention measures positively influenced the immediate gains of the multimodal recast group. This implies that the duration of the impact of these attention measures might vary depending on the type of assessment being conducted. The frequency of visits to the target form had a significant positive impact on the relationship between multimodal recasts and development in the OPT, WPT, and SOPT, both on the posttest and delayed posttest. In parallel, participants in the multimodal group displaying a lower skip rate exhibited significantly greater improvements in both immediate and delayed posttest scores across all assessments. These findings emphasize the pivotal role of attention allocation in determining the effectiveness of multimodal recasts, indicating that learners who paid closer attention to these recasts exhibited more substantial gains in their grammatical knowledge across various assessment measures.

RQ5. To what extent do individual differences in WM capacity moderate the effects of oral, written, or multimodal recasts on development in L2 grammatical knowledge during video-conferencing tasks?

To explore the potential influence of WM on the development of L2 grammar across different feedback modalities, five distinct WM measures were employed: the NWS task, forward Corsi block task, color shape task, stop signal task, and AOSPAN task. These measures were aimed to discern whether various WM functions had differential effects on L2 grammatical development, depending on the feedback modality. Among these measures, significant Group–Time–NWS interaction effects were detected when assessing participants' use of the present progressive in the SOPT, all with medium- to large effect sizes. Specifically, in the

SOPT, participants in the oral recast group with higher NWS scores showed greater pretest–posttest gains than the control group. Moreover, in terms of the pretest–delayed posttest gains, participants with higher NWS scores in the oral recast group displayed more significant progress than either the control group or the written recast group.

In terms of AOSPAN, a measure of the updating function, those with higher AOSPAN scores in both the oral recast and written recast groups demonstrated greater gains from the pretest to the delayed posttest across all three assessments compared to the control group. Furthermore, participants who had higher AOSPAN scores in the oral recast group were more likely to display greater improvements from the pretest to the delayed posttest in the SOPT than their counterparts in the multimodal recast group.

Concerning the third person singular, the results indicated that learners who excelled in the stop signal task, which assessed inhibitory control, tended to experience greater progress in the SOPT compared to their counterparts in the multimodal recast group.

RQ6. To what extent is learners' visual attention to the target constructions related to their WM capacity? Is this relationship influenced by whether learners are provided with written or multimodal recasts?

To investigate potential correlations between the distribution of visual attention to target constructions and individual differences in WM, a series of Spearman correlation analyses were conducted. These analyses aimed to shed light on how WM might interact with visual attention paid to target structures, considering the feedback modality. In the multimodal recast group, focusing on the present progressive form, significant positive correlations of medium- to large effect sizes were observed between NWS scores and four attention metrics (total fixation duration, first-pass reading time, re-reading time, and visit count). Furthermore, participants with higher NWS scores exhibited a reduced tendency to skip over the target

construction. These findings suggest that individuals with superior PSTM tended to allocate more visual attention to the target form.

In contrast, in the written recast group, NWS scores displayed a moderate negative correlation with visit count, indicating that learners with stronger PSTM tended to spend less time revisiting the target form when processing written recasts. The forward Corsi block task, assessing VSTM, demonstrated correlations with various attention measures. Apart from first-pass reading time, forward Corsi block task scores showed strong negative correlations with total fixation duration, re-reading time, and visit count. This measure displayed a strong positive correlation with skip rate, implying that learners with stronger VSTM abilities tended to allocate less time to fixation, re-reading, and revisiting, showing tendency to skip over the target form.

Moving onto the third person singular form, in the multimodal recast group, significant correlations were found between NWS scores and color shape task scores with visit count and first-pass reading time, respectively, with modest magnitudes. This indicates that learners with higher PSTM and efficient inhibitory control tended to spend more time visiting and showing greater initial attention during feedback processing.

In the written recast group, intriguing associations emerged. Firstly, the moderate negative correlations between NWS scores and total fixation duration, re-reading time, and visit count suggest that learners with better PSTM spent less time on total fixation and re-reading, with fewer revisits to the target form. Secondly, there was a small-sized positive correlation between forward Corsi block task scores and skip rate, suggesting that participants with greater VSTM were more likely to skip over the target structure. Additionally, stop signal task scores displayed a positive correlation with visit count and a negative correlation with skip rate, with varying magnitudes (moderate for visit count and small for skip rate). This implies that those with better inhibitory control visited the third

person singular form more frequently and were less likely to skip over the form. The AOSPAN exhibited small-sized positive correlations with total fixation duration and first-pass reading time, along with a small-sized negative correlation with skip rate. This indicates that individuals with superior updating abilities invested more time in total fixation and initial reading, exhibiting a lower rate of skipping. In summary, the differences observed in the relationships between the five attention measures and WM indicators in the two groups suggest that the nature of the recasts provided (multimodal versus written) influenced how learners' WM capacity shaped their attention allocation to the target constructions.

4.3.8. Interim Discussion

In extending Study 1, this study further examined whether the modality of recasts affected learners' visual attention to the target constructions and promoted L2 grammatical development employing eye-tracking methodology. Another goal of the present investigation was to shed light on whether individual differences in WM influenced learners' allocation of visual attentional resources to the recasts and development in the use of L2 grammatical knowledge. Study 2 differed from Study 1 in that it involved: (a) the recruitment of less proficient learners of English, (b) the use of eye-tracking technique, and (c) the employment of five different WM measures.

4.3.8.1. Effects of Recasts on L2 Grammatical Development (RQ1)

The first research question focused on assessing the impact of recasts, regardless of their modality, on the development of L2 grammatical knowledge in the context of video-conferencing. The no recast group, serving as the control group, did not exhibit significant changes in the accuracy of using the target constructions, the present progressive and the third person singular, in any of the assessment tasks over time. This suggests that the

observed improvements in the test scores of the experimental groups arose from the treatment they had received. In contrast, the provision of recasts positively influenced the improvement of learners' use of the target constructions, as assessed by all three assessment tests. Furthermore, this positive impact of recasts extended not only to the posttest but also to the delayed posttest for both target structures. These findings underscore the pedagogical value of incorporating recasts in video-conferencing L2 learning environments. The extension of this beneficial effect to the delayed posttest highlights the enduring advantages of recasts over time, suggesting that recasts contributed to a sustained enhancement in learners' language competence and retention.

These findings are consistent with previous literature in supporting the effectiveness of recasts in promoting L2 grammatical development, particularly concerning the development of the present progressive (e.g., Mackey & Gass, 2005; Lyster & Mori, 2006) and the third person singular (e.g., Zhao, 2015; Zhao & Ellis, 2022). As in Study 1, the results of this study support prior research that has demonstrated the benefits of recast for L2 grammatical development in text-based (e.g., Baralt, 2013, 2015; Kourtali & Borges, 2023; Yilmaz, 2012; Yilmaz & Yuksel, 2011) and in video-based SCMC interactions (Monteiro, 2014; Rassaei, 2017, 2019).

While the findings of the present study are consistent with those of Study 1, where recasts, regardless of their modality, facilitated the development of the knowledge related to *wh*-question formation, it is worth noting that in Study 2, recasts had a positive effect not only on the salient form (i.e., the present progressive) but also on the less salient structure (i.e., the third person singular) in the context of video-conferencing tasks. In particular, previous studies in SCMC literature have suggested that written recasts are less effective when addressing low salient features, such as the past tense (Loewen & Erlam, 2006), the English zero article (Sauro, 2009), Turkish locative case morphemes (Yilmaz, 2012), and the

third person singular *-s* (Kourtali, 2022). However, the present study provided favorable evidence for the potential of on-screen written recasts in promoting the acquisition of less salient grammatical features in SCMC L2 learning contexts, although they were not perfectly identical and comparable to the conventional form of text-chat recasts.

Another significant finding is the transferability of learning gains across modalities, as was also noted in Study 1. Similar to Martin's (2018) research, which involved metalinguistic feedback delivered through three modalities (i.e., text-based, voice-based, and video-based), participants who engaged in tasks in the written modality demonstrated improved oral production of the target form, and vice versa. Adding to these findings, the present study revealed that gains attributed to the provision of recasts in one modality could also be transferred to other modalities. Specifically, the written recast groups outperformed the control group not only in the WPT but also in the OPT and SOPT. Similarly, the oral recast group achieved higher gains than the control group in the OPT, SOPT, and the WPT. Moreover, the multimodal recast group also exhibited greater gains in the OPT, SOPT, and the WPT compared to the control group.

In conjunction with the findings from Study 1, Study 2 also confirmed the facilitative role of recasts in improving both declarative and procedural knowledge, as has also been found in earlier research (Ellis et al., 2006; Révész, 2012). Nevertheless, to establish the robustness of learning gains' transferability across varied modalities and assessments, additional empirical evidence is imperative. Consequently, further research that encompasses a diverse array of SCMC modalities and employs various outcome measurement modes is warranted.

4.3.8.2. Effects of Feedback Modality on L2 Grammatical Development (RQ2)

The second research question aimed to determine whether the modality of recasts – oral, written or multimodal – affects their effectiveness in promoting learners' knowledge development in using the target constructions. Consistent with the findings from Study 1, the results of Study 2 demonstrated that the multimodal recast group achieved the highest pretest–posttest and pretest–delayed posttest gains in both target structures across all outcome measures. However, there were no significant gain differences between the two unimodal recast groups in all three assessment tests. These results indicate that the modality of recasts had a significant impact on learners' development of proficiency in the use of the target constructions, with multimodal recasts generating the greatest improvements.

Following Study 1, the success of multimodal recasts in Study 2 also lends support to the predictions of the Dual Coding Theory (Paivio, 1986) and the Cognitive Theory of Multimedia Learning (Mayer, 2014). As predicted by these theories, learners in the present study likely formed stronger and more elaborate mental representations of the target structures when exposed to multimodal recasts, as the feedback engaged both oral and visual channels. Moreover, the presentation of complementary information through multiple modalities likely reduced the cognitive load imposed by each modality (Sweller, 2005), thereby facilitating deeper processing of the input, and ultimately leading to improved learning and retention of the target structures.

Also, the operationalization of multimodal recasts as captioned oral recasts may have aided learners in segmenting and parsing the oral input (Gass et al., 2019; Vanderplank, 1988), allowing more attentional resources to be directed toward the target forms in the feedback. This increased focus on the target structures may have further contributed to the overall effectiveness of the multimodal recasts. This mirrors the benefits of captions in recognizing linguistic forms by segmenting speech, enhancing the salience of unfamiliar

words, and establishing connections between auditory and written word forms (Bisson et al., 2014; Peters, 2019; Winke et al., 2010). This study represents the first investigation into the effectiveness of the captioned oral recasts for promoting L2 grammatical development in the context of multimodal video-conferencing tasks. To establish and strengthen the empirical basis supporting the potential of this innovative form of recasts for L2 development, further research is warranted.

4.3.8.3. Effects of Modality of Recasts on Allocation of Visual Attention (RQ3)

To investigate the impact of different modalities of recasts on learners' visual attention to the target linguistic constructions during video-conferencing tasks, eye-tracking technology was employed to concurrently gauge the extent of visual attention directed toward the target constructions in the feedback. Given that using multiple measures to gain a holistic understanding of various cognitive processes is recommended (Conklin et al., 2018; Godfroid, 2020), learners' eye movements were analyzed using five eye-tracking indices encompassing both early and late measures: total fixation duration (late), first-pass reading time (early), re-reading time (late), visit count (late), and skip rate (early).

The findings revealed notable distinctions between the multimodal recast and the written recast groups across all eye-movement measures for both target constructions. Specifically, the multimodal recast group exhibited significantly longer fixation duration, first-pass reading time, re-reading time, a higher number of visits, and a lower skip rate, not only for the present progressive but also for the third person singular. These results highlight the effectiveness of the multimodal recast, which combined oral and written input, in drawing learners' visual attention to the target forms compared to their written-only counterparts. The inclusion of both oral and written information likely heightened the salience and noticeability

of the target forms, resulting in increased attention toward the corrected forms in the feedback.

Given that increased attention to a specific area typically implies higher saliency and greater cognitive processing (Conklin et al., 2018), the patterns observed in the attention measures provide compelling evidence of a deeper level of engagement with the target forms within the multimodal recast group. The longer duration of first-pass reading indicates increased initial attention and cognitive involvement, suggesting that learners in the multimodal recast group actively allocated cognitive resources to process and understand the target forms from the start of their reading. It is also plausible that the inclusion of both oral and written input played a role in this augmented initial processing time, as learners simultaneously attended to both forms of information, resulting in more comprehensive and intricate cognitive processing. The higher re-reading time suggests that learners reanalyzed the target features to reinforce their understanding and consolidate their knowledge, signifying a deliberate effort to fully grasp and internalize the corrected linguistic constructions. Moreover, the longer total fixation duration suggests that learners dedicated a greater amount of visual attention to the target linguistic constructions. Given that longer fixation durations are indicative of more processing (Godfroid, 2020), with more time spent fixating on the target forms, learners in the multimodal recast group appeared to have extended opportunities to process and integrate these target-like structures into their mental representations.

Furthermore, the multimodal recast group demonstrated a higher number of visits to the target forms compared to the written recast group, indicating active engagement in comprehending the input provided. This group also exhibited a lower skip rate, suggesting that learners were less likely to overlook or disregard the target forms while reading the feedback under the multimodal intervention. This implies that multimodal recasts effectively

captured learners' attention for the target constructions and minimized the tendency to ignore them, indicating a higher level of attentiveness and a greater focus on the target linguistic constructions.

In terms of the correlation between the five attention measurements, the Spearman correlation analyses revealed that for the present progressive, the multimodal recast group demonstrated robust associations among all attention measures, with skip rate exhibiting negative correlations with the other metrics. These significant correlations among the measures mean that participants who directed more attention to specific measures, such as longer fixation durations, also tended to invest more attentional resources to other measures, such as higher visit counts while maintaining a lower skip rate. This substantiates the efficacy of multimodal recasts in directing learners' attention toward the target construction, highlighting that learners exposed to multimodal recasts adopted a consistent and engaged approach, resulting in interconnected attention patterns across diverse measures.

The written recast group, on the other hand, displayed distinct patterns in their relationships among attention measures for the present progressive form. Their reading behavior could be attributed to several factors that shaped their input processing strategies. Specifically, the significant negative correlation between first-pass reading time and re-reading time indicates that learners who invested more time in attending and understanding the target structure during their initial reading phase tended to spend less time re-reading the structure, suggesting a trade-off between initial comprehension efforts and the subsequent revisitation for reanalysis. This implies that the written recast group prioritized a thorough understanding of the target construction during the first pass, which reduced the need for re-reading to reanalyze the input. Consequently, this led to an shorter total fixation duration overall for this group. Moreover, there was a lack of relationships between first-pass reading time and visit count as well as skip rate. Again, the written recast group appeared to invest

more time and cognitive effort in understanding the content upon first encountering it. This prioritization of initial comprehension probably results in reduction of the need for later visits. Conversely, by investing more time in grasping the content during their initial reading, learners may have identified segments that were clear and familiar to them. These segments might have been skipped during subsequent readings because they were perceived as being less relevant or in less need of additional attention.

When examining the attention metrics for the third person singular form, the multimodal recast group, again, demonstrated a consistent distribution of attention, reflected in significant correlations among all five attention measurements. These results closely paralleled the strong correlation patterns observed in the analysis of the present progressive form within this group. In contrast, the correlation patterns for the written recast group differed from those observed in the present progressive. While the lack of correlation between first-pass reading time and visit count as well as skip rate remained consistent, significant positive correlations emerged between first-pass reading time and re-reading time as well as total fixation duration. These changes in correlation might have been influenced by the varying salience of the two target constructions. As has been previously discussed, the present progressive form was deemed more salient than the third person singular form due to its conspicuous form or communicative significance, making it more noticeable.

Consequently, for the present progressive form, the absence of a significant correlation between first-pass reading time and re-reading time might suggest that learners in the written recast group felt less compelled to revisit or re-read this form. This could be because the construction's salience allowed them to grasp it sufficiently during their initial reading, reducing the necessity for subsequent revisits or reanalysis. Conversely, for the less salient third person singular form (i.e., low physical salience and communicative value), the observed significant positive correlation between first-pass reading time and re-reading may

indicate that learners who spent more time initially attending to this form also tended to revisit it more during subsequent readings, possibly perceiving this form as less immediately clear or more complex. This may have prompted them to reanalyze it to ensure better understanding. Consequently, this led to an overall increase in total fixation duration, resulting in a significant correlation between first-pass reading time and total fixation duration for the third person singular form as well.

The observed differences in attention metrics correlations between the two groups could also be attributable to the heightened cognitive load associated with the written format of the feedback. Compared to the multimodal presentation of oral and written input, processing written input alone generally requires higher level of cognitive effort. In the absence of oral cues, the written recast group likely had to invest more cognitive resources in tasks such as visually decoding the text, comprehending the feedback, and subsequently integrating the input. This increased cognitive demand could have influenced their reading behavior, leading to variations in how they distributed their attention across different aspects of the written feedback. Consequently, learners may have strategically allocated their cognitive resources, focusing more on specific elements of the text while potentially minimizing attention directed toward other aspects due to limited cognitive capacity, resulting in discrepancies in the attention metrics.

The distinct correlation patterns that may have arisen from the potentially differential cognitive load between the multimodal recast group and the written recast group can be elucidated by Sweller's (Sweller, 1988, 2005a; Chandler & Sweller, 1991) Cognitive Load Theory. This theory posits that when auditory and visual information complement and correspond to each other, the cognitive load on WM decreases, compared to when a single channel is used (Sweller, 1988, 1994). In the case of the multimodal recast group, the inclusion of both oral and written input enabled learners to engage in dual coding of the input

(Paivio, 1986, 2007). Consequently, the processing load required in each channel (i.e., aural and visual) was less likely to exceed the learners' cognitive capacity (Mayer & Fiorella, 2014). This, in turn, may have facilitated the processing of linguistic input and allowed learners to leverage their full capacity for processing information (Mayer, 2014), ultimately enhancing their ability to decode the input. As a result, the multimodal recast group might have been more effective in consistently allocating their cognitive resources across various attention measures. This optimized cognitive processing could account for the observed consistent and interconnected patterns in the attention measures within this group.

Given that the multimodal recast group achieved significantly greater improvements in all three assessment tasks at both posttest and delayed posttest compared to the written recast group, these findings support the idea that multimodal recast facilitated a more comprehensive and effective processing of the input, leading to improved comprehension and potential acquisition of the target forms. The relationship between the learners' visual attention to the target constructions and their subsequent learning gains related to the target features will be discussed in the next section.

4.3.8.4. Relationship Between Attention and L2 Development (RQ4)

Research Question 4 delved into the potential relationship between how participants distributed their visual attention to target grammatical constructions and their learning gains of these structures across three different tests (i.e., OPT, WPT, SOPT). The predictive role of eye-movement measures in this relationship varied, depending on the type of target features and the feedback modality. Specifically, there were more pronounced associations between attention measurements and the performance of the multimodal recast group compared to the written recast group.

The results revealed that all five attention measures had stronger relationships with the multimodal recast group's gains across all three outcome measures for both target features. These findings suggest that learners who devoted more attention to the target forms presented in multimodal recasts demonstrated greater development in oral, written, and spontaneous oral production compared to the written recast group. Given the superior performance of the multimodal recast group on all outcome measures in both posttest and delayed posttest assessments for both target constructions, these findings imply that learners in the multimodal recast group may have engaged in more extensive processing of the input, leading to a better understanding and retention of the rules associated with the target structures. Consequently, this underscores the significant influence of recast modality on the potential relationship between learners' visual attention and their subsequent development in L2 grammar learning.

The findings of the present study align with prior research that has demonstrated a positive connection between increased attention to specific linguistic features and L2 development in SCMC learning environments (e.g., Smith, 2010, 2012; Smith & Renaud, 2013). For instance, Smith (2010) observed that participants tended to fixate more on lexical recasts than on grammatical ones, resulting in more accurate production of lexical items. This showed the importance of total time participants spent fixating on items, as it correlated with their ability to remember and use the target items effectively. It should also be noted that grammatical recasts seemed to attract less attention from learners compared to those targeting lexical items. Similarly, Smith and Renaud (2013) found that participants achieved higher posttest scores for items they had fixated on, albeit without reaching statistical significance. Nonetheless, there was a noticeable trend indicating a positive relationship between the number of fixations and posttest success, implying that fixating on a target item predicted better posttest performance than not fixating on it. The study revealed that learners' fixation

on a target item for at least 200 milliseconds emerged as the strongest predictor of posttest success, regardless of the target's number, type, or difficulty. The total fixation duration, however, did not show a correlation with posttest gains.

The positive link between the degree of attention given to multimodal recasts, operationalized as captioned oral recasts, and subsequent L2 grammatical development, aligns with previous studies investigating the impact of captions on L2 grammatical development. For example, Lee and Révész (2020) found significant correlations between attention measurements and improvements in the use of the target construction for their enhanced captions group, which displayed longer fixation durations and more visits to the target feature (the present perfect) compared to the unenhanced captions group. These findings point to a positive correlation between the time spent on target grammatical structures in captions and subsequent L2 grammatical development.

It is worth noting that Lee and Révész's study also hinted that the relationship between attention allocation and L2 grammatical development might vary across specific target structures. For example, in the case of the past simple, none of the attention measures exhibited a significant relationship with learning gains. The researchers speculated that the degree of physical salience of the linguistic constructions played a role in the attention-development relationship. In contrast, the present study showed the potential benefits of using multimodal recasts, operationalized as captioned oral recasts, in capturing learners' attention, not only for the more salient form (the present progressive) but also for the less salient form (the third person singular).

Both Lee and Révész's research and the present study have shown the significance of attention allocation in the development of L2 grammatical knowledge. However, these studies differed in several critical aspects, potentially influencing the relationships observed between attention and L2 development. One notable difference lies in the assessment tasks

used to evaluate learners' progress. Lee and Révész's study included an oral productive test, a written productive test, and a receptive test, while the present study focused solely on meaning-oriented production tests. Consequently, the choice of different outcome measures between the two studies could have an impact on the associations observed between attention and development. In the present study, eye-tracking indices demonstrated significant relationships with gains in all three outcome measures, all of which were meaning-oriented production tests. This implies that these relationships could have been different if a receptive test had been integrated into the assessment framework. Therefore, further research is warranted to incorporate both productive and receptive tests to yield a more comprehensive understanding of the intricate relationship between attention and L2 grammatical development, particularly in the context of video-conferencing utilizing multimodal recasts. Such investigations would provide a more holistic perspective on the impact of attention allocation on L2 acquisition.

In conclusion, beyond reaffirming prior research findings regarding the advantages of increased attention to text-based recasts and captions for L2 grammatical development, the present study provides additional corroborative evidence that multimodal recasts effectively heightened learner attention to L2 grammatical features and contributed to their acquisition. As was discussed earlier, previous SCMC research predominantly focused on enhancing the salience of recasts through input enhancement, either orally or typographically (e.g., Sachs & Suh, 2007; Saggara & Abbuhl, 2013), without integrating various SCMC modalities to amplify recast salience. As the initial effort into this under-researched area, the current study has demonstrated that multimodal recasts successfully amplified both the salience and corrective force of recasts, thus capturing a greater degree of learner attention toward the target forms present in multimodal recasts when compared to unimodal recasts. This highlights the potential efficacy of utilizing the modality of recasts to effectively engage

learners' attention with grammatical recasts in the context of multimodal video-based SCMC interactions.

4.3.8.5. Moderating Effects of WM in L2 Grammatical Knowledge Development (RQ5)

This study sought to investigate the role of WM in shaping the effectiveness of different feedback conditions (i.e., oral recasts, written recasts, multimodal recasts, no recasts) in facilitating learners' acquisition of the present progressive and the third person singular. To accomplish this, five different WM measures were employed, including the NWS task for PSTM assessment, the forward Corsi block task for VSTM evaluation, and executive functions evaluation (updating, task-switching, inhibitory control) through the AOSPAN task, the color shape task, and the stop signal task, respectively. The results revealed that the impact of these various WM measures on learners' acquisition of the target constructions differed across different groups and time points.

Regarding the present progressive, the study revealed a significant influence of NWS and AOSPAN on learners' gains in the SOPT and the three assessments. The findings pertaining to learners' development in the SOPT revealed that the impact of NWS was particularly prominent in the oral recast group compared to both the control group and the written recast groups. This observation suggests that oral recasts were more favorable to learners with higher PSTM, especially when they were engaged in the rapid and spontaneous language use during the SOPT. Recent theoretical accounts in SLA propose that tasks conducted in the oral modality place a greater burden on learners' WM (e.g., Gilbert et al., 2016; Kormos, 2014; Williams, 2012), due to the heightened temporal demands inherent in producing spoken language and processing auditory input in this medium. The SOPT, in particular, demanded rapid process and reproduction of spoken language, requiring not only retrieval of the correct linguistic forms but also active maintenance and manipulation of these

forms within WM. This is where PSTM becomes a crucial cognitive resource, as learners with superior PSTM abilities were more adept at efficiently retaining and manipulating phonological information. Consequently, this made it easier for them to grasp the corrected linguistic forms and successfully produce them during the SOPT.

The more prominent impact of PSTM on the oral group was also found in Zalbidea and Sanz's (2020) study, which examined the effects of computerized tasks and feedback in oral versus written modalities on L2 grammar development. Their research revealed a positive correlation between learners' PSTM and their progress in the speaking group (oral modality) across various productive and receptive outcome measures, when compared to the writing group (written modality). When considering these findings alongside the results of the present study, it seems evident that the more significant influence of PSTM in the oral recast group compared to the multimodal recasts and written recast groups can also be explained by the varying memory demands associated with different modalities. The nonvisual and ephemeral nature of oral linguistic input is claimed to impose greater processing demands, since linguistic forms in the oral input are available for noticing only for a very brief period (Gilbert et al., 2016), which imposes a greater processing burden compared to written feedback (Sagarra & Abbuhl, 2013). This heightened processing demand may explain why individuals with greater PSTM capacity excelled in retaining and manipulating target forms during oral tasks.

On the other hand, the absence of PSTM effects in the writing group in Zalbidea and Sanz's study, as well as in both the multimodal recast and written recast groups in the present study, implies that the features of the written modality, such as the presence of visual linguistic input and longer processing time, offer "a temporal extension of WM" (Schoonen et al., 2009, p. 81). As a result, participants in the multimodal recast and written recast groups appeared to be better able to notice, maintain and rehearse the target form, continuously

strengthening memory traces throughout the experiment (e.g., Baddeley, 2003a). Consequently, the SOPT gains achieved in these groups seemed to be less susceptible to individual differences in PSTM, even when dealing with the substantial attentional demands associated with L2 input and output processing requirements during spontaneous oral production tasks. This may suggest that when written input is available, it acts as a leveler for L2 learners with varying PSTM capacities. In other words, written input can make it easier for learners with lower PSTM to perform as well as those with higher PSTM in terms of acquiring L2 grammar, ultimately diminishing the predictive power of phonological WM capacity, especially in the context of spontaneous oral production tasks.

The significant impact of PSTM on learning gains in the SOPT aligns with previous research findings that demonstrated the positive influence of NWS on learners' development in oral production (e.g., Payne & Ross, 2005; Révész, 2012; Trofimovich et al., 2017; White, 2020; Zhao, 2015). As was previously explained, learners with higher PSTM capacity likely had the ability to maintain the target-like form presented in oral recasts in their short-term memory for a longer time. This, in turn, increased the likelihood that long-term memory traces could be effectively encoded. These resulting long-term memory traces could then serve as the foundation for implicit learning processes, facilitating the proceduralization of their L2 knowledge (N. Ellis, 2005). Increased procedural knowledge, in turn, might have enabled learners with high PSTM scores to perform better on the SOPT, since spontaneous oral production tasks typically draw more extensively on procedural knowledge than explicit knowledge.

In the present study, the updating function of the central executive, operationalized as AOSPAN, played a moderating role in the performance of the oral and written recast groups across the three assessments concerning the learning of the present progressive. AOSPAN scores significantly influenced the gains of the oral and written recast groups on the OPT and

WPT compared to the control group. Also, the impact of AOSPAN on gains in the SOPT was particularly pronounced for both the oral and written recast groups in comparison to the control group. Moreover, AOSPAN had a more prominent effect when comparing the SOPT gains of the oral recast group to those of the multimodal recast group. Overall, these findings confirm the substantial role played by the ability to update and manage information in WM in shaping how learners in the oral and written recast groups harnessed the advantages offered by their respective recast interventions.

The heightened impact of AOSPAN on the gains displayed by the oral and written recast groups suggests that learners in these groups relied more extensively on their updating abilities. Updating capacity, which is integral for processing and integrating new information, thereby adapting to changing task demands, might have been employed by learners in these two groups to effectively replace old and irrelevant information with new and more pertinent input, thereby incorporating the feedback provided through their respective modalities. Consequently, this probably led them to obtain higher gains of the present progressive across the three outcome measures.

The more significant impact of AOSPAN on the learning gains of the two unimodal recast groups, as compared to the multimodal group, may also have resulted from differences in cognitive load. Updating abilities involve continuous monitoring, revision, and integration of new information into existing mental representations (Baddeley, 1996; Miyake et al., 2000). In the present study, this process included adding, deleting, or modifying linguistic input in WM to keep it current and germane to the ongoing task, which was crucial for incorporating the target-like form in the feedback into learners' language production skills and successful task completion. In the case of multimodal recasts, where cognitive resources were divided between two modalities (oral and written input), learners may not have needed to rely as heavily on their updating abilities as the learners who received unimodal recasts.

This is because those in the multimodal recast group were dealing with two sources of input simultaneously, and the integration of feedback from these two sources might have necessitated less reliance on step-by-step updating processes typically associated with processing a single source of input. Consequently, learners in the multimodal recast group may have adopted a more holistic or integrative approach, enabling them to effectively process the feedback without needing to engage in meticulous incremental updating of their mental representations. Taken together, these findings suggest that engagement with oral recasts as well as written recasts may require more substantial cognitive resources, particularly in terms of AOSPAN, which can significantly affect these groups' overall gains across the three outcome measures.

The observed moderating role of updating abilities in the efficacy of recasts on L2 learning aligns with previous research that yielded similar findings regarding the significant impact of updating capacity on the effects of recasts in face-to-face contexts (e.g., Ahmadian, 2020; Goo, 2012). For instance, Goo (2012) found that learners' updating capacity significantly predicted the acquisition of English *that*-trace filter in the recast group compared to the metalinguistic feedback group. Similarly, Ahmadian (2020) also reported that learners' updating abilities were significantly correlated with the development of productive and receptive knowledge in the recast group compared to the explicit correction group. Contrary to these two studies, Yilmaz (2013a) found that WM-including updating ability had a positive and significant correlation with the performance of the explicit correction group compared to the recast group. Additionally, Liao and Zhang's (2022) study, which compared the effects of explicit corrections, metalinguistic feedback, and analogy-based CF on the acquisition of the third person singular *-s*, indicated that learners' updating abilities only predicted the effects of analogy-based CF but not explicit corrections and metalinguistic feedback. These findings

collectively suggest that the predictive power of updating may depend on the explicitness or type of feedback.

The connection between updating abilities and L2 learning outcomes, particularly observed in the case of the more salient form in the present study, also merits discussion. This finding appears to resonate with Yilmaz's (2013a) study, where effects of updating ability were attested for the more salient form but not for the less salient one. Yilmaz speculated that the absence of any significant influence of updating capacity for the less salient form in his study might be attributed to the fact that no substantial improvements were detected in that specific target form. This may imply that this component of WM plays a crucial role in recognizing patterns within linguistic input and establishing successful representations of these patterns. However, given that all experimental groups in the present study exhibited significant gains in using the third person singular (less salient form), further research is needed to gain a better understanding of the relationship between target structure salience and its connection to learners' updating capacities.

In the context of SCMC or computerized L2 learning environments, several studies have reported results contradictory to the present study (e.g., Baralt, 2010, 2015; Martin, 2018; Zalbidea & Sanz, 2020) in that they found no effects for updating ability. For example, Baralt (2010, 2015) found that while participants learned the Spanish past subjunctive in the text-based SCMC, the construct of executive WM, including updating ability, did not affect the extent to which learners benefited from recasts. Similarly, Martin's (2018) investigation, which explored the role of updating capacity in the efficacy of metalinguistic feedback for L2 grammatical learning across text-based, voice-based, and video-based SCMC settings, showed that all groups successfully acquired Spanish past subjunctive irrespective of their WM capacity. Zalbidea and Sanz (2020) also found that learners' updating abilities did not contribute to their posttest gains either in the speaking or the writing group across oral

production and written production tests as well as acceptability judgement tasks targeting linguistic forms of varying perceptual salience. These contradicting results indicate the complexity of the relationship between updating capacity and L2 grammatical development in various SCMC or computerized contexts and feedback types.

Clearly, while the present study revealed a significant moderating effect of learners' updating capacity on L2 grammatical development during multimodal video-conferencing tasks, the results do not provide a definitive conclusion when extrapolated to other SCMC contexts or different types of feedback. Therefore, further research is needed to attain a fuller understanding of the role of updating ability in L2 grammatical development within the context of SCMC, taking the influence of varying SCMC modalities and feedback types into consideration.

The present study also yielded an intriguing finding regarding the third person singular, which differed from the observations for the present progressive. The analyses demonstrated a more pronounced influence of stop signal task scores on the improvements in spontaneous oral production skills in the oral recast group, as compared to the multimodal recast group. This effect exhibited a large magnitude, underscoring the potential significance of inhibitory control abilities in facilitating the development of spontaneous oral production, particularly in the context of oral recasts. Inhibitory control, as assessed by the stop signal task, involves the capacity to suppress irrelevant or distracting information or responses while maintaining focus on the current task. This ability appeared crucial in the SOPT, which required learners to process and produce spoken language promptly without being hindered by incorrect linguistic forms or other distractions.

Inhibitory control might have also played a role in error correction, helping learners filter out inappropriate words or syntactic structures and facilitating the selection of correct forms. Consequently, learners with stronger inhibitory control likely excelled in maintaining

precision and accuracy during rapid oral production tasks, resulting in more significant improvements in their spontaneous oral production skills. Furthermore, inhibitory control may have been instrumental in resolving conflicts arising from competing linguistic forms, especially when multiple grammatical options were available for completing a particular task. In the present study, learners had to select the most appropriate allomorphs within the third person singular form. Inhibitory control appeared to assist learners in making these choices and suppressing alternative forms.

In addition, the presence of written records of feedback made the comprehension and correction of errors straightforward in the case of multimodal recasts. That is, learners in the multimodal recast group were more likely to easily notice negative feedback and efficiently employ their cognitive resources to identify errors, their locations, and how to rectify them. Consequently, the noticing of multimodal recasts, if it occurred, may not have consumed cognitive resources to the same extent. In contrast, the third person singular form in oral recasts appeared less noticeable and prominent due to the transient nature of the oral feedback. This group did not benefit from written feedback that explicitly pointed out errors and presented the correct form. In the absence of written cues, learners in the oral recast group had to rely more on their internal cognitive resources, including inhibitory control, to process and internalize the feedback. They had to infer the errors and identify the correct form themselves. This process involved comparing the orally presented target-like form with their own utterance to pinpoint similarities and differences, extracting and applying the underlying rules presented in the feedback to their original utterance, and subsequently producing a modified response to the prompted question.

Given that the significant impact of inhibitory control on the effectiveness of oral recast was particularly evident when oral recasts targeted the less salient third person singular form, as opposed to the more noticeable form (the present progressive), it seems that this task

appeared particularly challenging for the oral recast group in comparison to the multimodal recast group. As a result of these factors, learners in the oral recast group seemed to have a higher demand for inhibitory control to filter out distracting or incorrect linguistic forms, maintain focus on the target form, and make the necessary adjustments in their oral production. Thus, as Kane and Engle (2000) argued, learners with high inhibitory control were more able to resist proactive interference and have more cognitive resources available for processing an attention-demanding target structure.

However, Yilmaz and Sagdic (2019) reached a conclusion contradictory to the findings of the present study. They delved into the relationship between inhibitory control and the effects of feedback timing (immediate versus delayed) in text-based SCMC. Their findings indicated that inhibitory control did not significantly impact learning gains, as measured by an oral production test and a GJT, on either of the outcome measures. Yilmaz and Sagdic attributed their results to the characteristics of text-based SCMC. Text-based interactions often unfold at a slower pace compared to oral interactions, affording learners additional time to process input and formulate their responses. Also, the permanence of written feedback in text-based SCMC may have influenced the results. Learners could revisit and review written feedback, potentially alleviating the cognitive load on WM. The combination of a slower interaction pace and the availability of persistent written feedback may have mitigated the impact of inhibitory control in the text-based context, rendering its role less prominent compared to that of the oral recast context observed in the present study.

The discrepancy in findings between the present study and Yilmaz and Sagdic's (2019) research raises the issue that the significance of inhibitory control in L2 grammar learning might be context-dependent. This implies that variables such as the specific interaction context, the modality of feedback, and the cognitive demands imposed by these factors play a crucial role in shaping the influence of inhibitory control on learning outcomes.

However, it is important to acknowledge that the existing body of research in this area remains limited, underscoring the need for further investigation to gain deeper insights into how learners manage their cognitive resources when confronted with these factors within SCMC settings.

4.3.8.6. Relationship Between WM and Attention (RQ6)

Research Question 6 set out to investigate the association between learners' visual attention to target constructions involved in recasts and their WM capacity, considering whether this relationship was influenced by the modality of recasts. To achieve this goal, Spearman correlation analyses were employed to examine the interconnections between the eye-tracking and WM measures.

For the present progressive, the multimodal recast group's results displayed moderate to large positive correlations between NWS scores and four attention measures: total duration of fixation, first-pass reading time, re-reading time, and visit count. Also, a small negative correlation emerged between NWS scores and skip rate. These findings indicate learners with stronger PSTM abilities tended to exhibit more attentive reading behaviors when encountering this grammatical structure, enhancing their focus and attentiveness during the processing of multimodal recasts. Given that these correlations manifested in both early measures (first-pass reading, skip rate) and late measures (total fixation duration, re-reading time, visit count), learners with higher NWS scores not only demonstrated initial attentiveness but also sustained their attention throughout their reading process. The negative correlation between NWS scores and skip rate suggests that learners with superior phonological WM were less inclined to skip over the present progressive form, indicating that they engaged in more meticulous reading behaviors and were less likely to overlook important details.

Conversely, in the written recast group, there was a moderate negative correlation between NWS scores and visit count. This suggests that as PSTM increased, the frequency of revisiting the target form decreased. This implies that stronger PSTM might have enabled learners to process and retain linguistic information more effectively, reducing the need for repetitive visits to comprehend and retain the target form. As a result, learners with higher PSTM scores in this group seemed to manage their attention more efficiently, potentially enhancing their overall comprehension and engagement with the target construction.

The contrasting relationships between NWS scores and visit count in the multimodal recast and the written recast group could be attributed to the presence or absence of auditory cues in the recasts. The multimodal recast group received feedback enriched with both auditory and visual cues, making the target feature more salient and memorable (Mayer, 2014; Paivio, 2007). Consequently, learners perceived the content as highly engaging and important, possibly prompting those with higher PSTM abilities to revisit the target construction more frequently, recognizing the added value in engaging more deeply with the multimodal information. In contrast, the written recast group was only provided linguistic input that was primarily text-based, lacking the auditory cues available in the multimodal recasts. This may have resulted in a lower perceptual salience, affecting learners' motivation to revisit the feedback, as they might have not perceived the same level of value in engaging with the input when sensory cues were absent.

With regard to visuospatial WM as measured by the forward Corsi block task, the written recast group demonstrated moderate to large negative correlations between the forward Corsi block task scores and total fixation duration, re-reading time, and visit count. Moreover, a strong positive correlation emerged between forward Corsi block task scores and skip rate. These findings suggest that in the written recast group, attention allocation to the target form was significantly influenced by the learners' VSTM. The negative correlations

between VSTM and attention metrics suggest that learners with stronger VSTM might have been better equipped to handle the cognitive demands posed by the single sensory modality in the written recast group, leading to reduced visual attention allocation. In contrast, the absence of significant correlation between VSTM and attention measures in the multimodal recast group might be due to the enhanced salience of the target form achieved through the multimodal presentation of the feedback, which attracted learners' attention without relying heavily on visuospatial WM. In other words, given that this WM component, as expected, plays a role in pedagogic L2 tasks that draw on visual memory (Kellogg, 1996, 2007), the sensory-rich environment may have successfully attracted attention to the target form, diminishing the importance of VSTM in managing attention.

Turning to the third person singular form, the multimodal recast group demonstrated significant positive correlations between NWS scores and visit count, as well as color shape task scores and first-pass reading time, displaying a modest degree. Similar to the findings for the present progressive form, these correlations reaffirmed the crucial role of learners' PSTM capacities in shaping their revisiting patterns. This indicates that learners with stronger PSTM were still inclined to revisit the third person singular form and devote more time to engaging with it. In addition, task-switching abilities, as measured by the color shape task, emerged as an influential factor during the initial reading phase for the third person singular form. This finding indicates the learners' need to shift their cognitive focus efficiently between different linguistic aspects or rules when encountering this structure. The introduction of task-switching abilities as a prominent factor suggests that the cognitive demands during the initial reading phase varied between the two target constructions.

Several factors are likely to have contributed to the prominence of task switching for the third person singular form, which was not observed for the present progressive form in this group. Firstly, its relatively low physical and phonological salience may have imposed a

higher cognitive load on learners (Goldschneider & DeKeyser, 2001; N. Ellis, 2006; Loewen et al., 2009; VanPatten, 1996). This required learners to allocate more cognitive resources to process and comprehend the form effectively, necessitating frequent task switching between different forms or rules. It is also possible that the variability of the third person singular form, with its different allomorphs, may have demanded learners adapt and switch between these variations depending on the context. This linguistic variability may have required learners to shift their cognitive focus promptly to handle these variations effectively during initial reading.

In the written recast group, a more extensive set of correlations emerged between WM metrics and attention measurements. Negative correlations between NWS scores and total fixation duration, re-reading time, and visit count suggested that participants with stronger PSTM tended to allocate less time to fixating on, re-reading, and revisiting the target construction. Moreover, their superior PSTM may have facilitated more extensive rehearsal of the target form in short-term memory, ultimately reducing the perceived need for prolonged fixation, re-reading, or repeated visits. This heightened efficiency in processing could have been especially advantageous when dealing with a less salient form like the third person singular.

Secondly, the positive correlation between forward Corsi block task scores and skip rate indicates that learners' visuospatial WM continued to have a significant influence on their skipping behavior. This suggests that individuals with higher VSTM capacities were more adapt at processing and retaining the visual information of the target form. They tended to skip over the target structure more frequently, relying on their superior VSTM. This observation echoes the findings related to the present progressive form, highlighting that the substantial role of VSTM in facilitating attention to both grammatical structures in the written recast group.

Thirdly, the correlations between stop signal task scores and the measures of visit count and skip rate revealed intriguing dynamics. The positive correlation with visit count and negative correlation with skip rate suggest that learners with greater inhibitory control, as measured by the stop signal task, were more inclined to revisit the target structure frequently and were less prone to skipping it. This behavior suggests that learners with superior inhibitory control had an enhanced ability to maintain their focus and engage in more thorough processing. Interestingly, unlike the absence of significant correlation between inhibitory control and attention measurements for the present progressive form in the written recast group, the emergence of this significant correlation for the third person singular implies that inhibitory control may have played a more significant role in managing attention allocation for this specific grammatical feature. This finding underscores the complex interaction between cognitive capacities and linguistic input, suggesting that the degree of salience and complexity of the form may influence the extent to which different aspects of WM are brought into play when processing the linguistic input. Therefore, further research is needed to gain a more fine-tuned understanding of these potential relationships.

The positive correlation between color shape task scores and visit count in the written recast group implies that participants who performed better in the color shape task, which assessed task-switching abilities, tended to revisit the target construction more frequently. As with the multimodal recast group, which demonstrated a significant positive correlation between the color shape task scores and first-pass reading time, one plausible interpretation is that the inherent linguistic variability of the third person singular form, characterized by different allomorphs depending on the context, likely demanded learners to adapt and switch between these variations when they encountered them. Thus, learners with stronger task-switching abilities might have been more adept at shifting their cognitive focus flexibly between the various forms and rules associated with the third person singular, particularly

when faced with linguistic variability. This adaptability in task switching might have led them to revisit the form more frequently to ensure they processed it correctly in its different variations, resulting in a positive correlation between task-switching abilities and visit count.

Finally, the AOSPAN scores exhibited positive correlations with total fixation duration and first-pass reading time, along with a negative correlation with skip rate in the written recast group. This suggests that learners with better information updating abilities (as captured by AOSPAN) tended to sustain longer fixation periods, especially the initial reading phase, and a reduced tendency to skip the target construction when provided with written recasts. This indicates that high updating ability, by allowing learners to better monitor and code the relevance of incoming information (Miyake et al., 2000), might have facilitated learners' attention to the written input and helped them perceive it as relevant.

The differential impact of AOSPAN on attention allocation between the two groups may be indicative of how sensory cues and feedback modality influenced learners' cognitive processes. This distinction becomes evident when comparing the multimodal recast group, who received feedback combining visual and auditory cues to enhance the salience and memorability of the target construction, with the written recast group that relied solely on written feedback. The absence of additional auditory cues in the written recast group appeared to place a heavier demand on learners to actively update and monitor written information to grasp and retain the target feature.

Significant correlations between attention metrics and WM measures have also been reported in previous eye-tracking research in learning L2 grammar using captions (Indrarathne & Kormos, 2018; Lee & Révész, 2021). For instance, Lee and Révész (2021) found that participants with higher NWS and AOSPAN scores tended to pay more attention to the target form under the enhanced captions condition. This suggests that when the target feature was emphasized through textual enhancement, learners with greater WM capacity

were more inclined to focus their attention on it. Similarly, Indrarathne and Kormos's (2018) study revealed that in the most explicit learning condition (explicit metalinguistic explanation + instruction to pay attention to the textually enhanced part), a composite score of WM measures that combined PSTM, updating, task-switching, and inhibitory control strongly predicted the extent of attention participants devoted to the target feature. In other words, when learners' attention was deliberately directed by interventions, those with higher WM capacity exhibited increased visual attention toward the target construction.

Together with the results from the present study, these findings highlight the strong association between the level of attention devoted to the target constructions and WM capacities. However, this relationship is evidently influenced by several factors, such as the type of instruction and the manipulation of input salience, which means that the precise nature of this relationship is unclear, especially in the SCMC L2 learning context. The aforementioned studies were not conducted in the SCMC framework, and their primary focus was not on the impact of different modalities but rather on textual enhancement (enhanced versus unenhanced) and varying instructional conditions (implicit versus explicit). Moreover, the use of composite WM scores by Indrarathne and Kormos (2018) poses challenges when making direct comparisons with the present research. To the best of my knowledge, this is the first study that has established a connection between learners' WM capacity, feedback modality, and visual attention allocation to L2 grammatical structures in the multimodal video-based SCMC. In summary, these collective findings emphasize the need for further exploration in the domain of SCMC, with a focus on the examination of various SCMC modalities and the assessment of how different facets of WM and feedback modes influence the allocation of visual attention by L2 learners.

CHAPTER 5. GENERAL DISCUSSION AND CONCLUSION

The main purpose of this study was to investigate the effectiveness of multimodal recasts in promoting L2 learners' grammatical development in the context of video-conferencing tasks in comparison to oral recasts, written recasts, and no recasts. This research was motivated by the widespread use of SCMC platforms in the realm of language learning and teaching. It sought to address a critical and timely need, considering that these online communication technologies are increasingly becoming a principal means of instruction. In particular, in light of the growing prominence of video-conferencing as a central mode of interaction in a wide range of educational settings (Peterson & Sachs, 2016), this research not only adds to the evidence supporting the effectiveness of video-conferencing in L2 grammar learning, but also prompts us to expand this line of investigation.

As regards recasts, prior SCMC eye-tracking studies exploring the impact of recasts on L2 development have predominantly been conducted in text-based SCMC, employing only a single modality and a limited range of attention measures. Therefore, this research also endeavored to examine how multimodal recasts affect learners' visual attention allocation to target linguistic constructions and their subsequent L2 development in comparison to written recasts, employing various early and late attention metrics.

Last but not least, while previous cognitively oriented L2 research in this domain has primarily focused on individual differences in executive WM, less empirical focus has been directed toward phonological and visuospatial short-term memory. Hence, an additional objective of this study was to delve into the influence of executive, phonological, and visuospatial WM in relation to the effects of feedback modality on L2 grammatical development and its impact on the allocation of visual attention by learners.

This chapter commences with the provision of an overview of the key findings from the thesis. As the results and discussions pertaining to each research question (RQ) have been detailed in Chapter 3 (Study 1), as well as in Chapter 4 (Study 2), the general discussion is organized thematically rather than by individual RQs. Following this, the chapter proceeds to explore the theoretical, methodological, and pedagogical implications arising from these findings. Finally, the chapter concludes with a discussion of the limitations of the present study and suggestions for future research directions.

5.1. Effects of Recasts (RQ1) and Recast Modalities (RQ2) on L2 Grammatical Development – Study 1 and Study 2

The aims of RQ1–2 centered on conducting a comparative analysis between the absence of recasts and their provision, as well as assessing the impact of different recast modalities on learners' knowledge development in the formation of *wh*-questions (Study 1) and the correct use of the present progressive and the third person singular (Study 2). An experimental design comprising a pretest–posttest–delayed posttest was employed to investigate participants' learning gains after exposure to various feedback types, including oral recasts, written recasts, multimodal recasts, and no recasts. The overall results from both studies unequivocally demonstrated the beneficial role of recasts in promoting development in the use of the target constructions compared to no recasts.

With a specific focus on the differential effects of SCMC modalities, the findings from Study 1 and Study 2 may suggest that the effectiveness of multimodal recasts tends to remain consistent across different proficiency levels and target constructions. This hints at a more robust and versatile impact of multimodal recasts on the development of L2 grammatical knowledge when compared to unimodal recasts. In Study 1, participants were classified as English users comparable to bands A2–B1 (intermediate) according to the

CEFR, with the target construction involving the *wh*-question formation. In Study 2, less proficient users of English, falling into A2 band (elementary), were included and the target constructions were the use of the present progressive and the third person singular. Regardless of the proficiency levels and the type of target constructions, the superiority of multimodal recasts in fostering grammatical knowledge gains over unimodal recasts was consistent. These findings underscore the potential advantages of incorporating multimodal recasts in L2 instruction.

5.2. Effects of Feedback Modality on Learners' Visual Attention Allocation (RQ3) and Relationship Between Attention and L2 Grammatical Development (RQ4) – Study 2

Grounded in the premise that individual differences in WM and attentional capacity play a role in various cognitive processes, both in early stages (e.g., noticing, pattern recognition), and later stages (e.g., retrieval, error avoidance), thereby directly influencing SLA (Robinson, 1995; Skehan, 2016), the present study explored the impact of feedback modality on learners' distribution of visual attention to target constructions and its correlation with the acquisition of these forms. The eye-tracking data demonstrated that across all five eye-tracking measures – total fixation duration, first-pass reading time, re-reading time, visit count, and skip rate – the multimodal recast group, when compared to the written recast group, consistently exhibited longer total fixation duration, extended first-pass reading time, increased re-reading duration, a higher number of visits, and a lower skip rate for both target structures. These findings highlight the superior ability of multimodal recasts to capture and sustain learners' visual attention, regardless of the inherent salience of the target forms.

The results revealed the significant influence of recast modality on the relationship between learners' visual attention and their subsequent progress in L2 grammar learning. All five attention measures had a stronger relationship to the gains of the multimodal recast group

across all three outcome measures for both target features, in comparison to the written recast group. In other words, beyond confirming previous research findings regarding the advantages of increased attention to text-based recasts and captions for L2 development (e.g., Lee & Révész, 2020; Smith, 2010, 2012; Smith & Renaud, 2013), the present study provides further empirical evidence that multimodal recasts have the potential to heighten learners' attention to L2 grammatical structures effectively and to contribute significantly to the acquisition of such structures in the context of video-based SCMC. To the best of my knowledge, this is the first study to explore the relationship between various SCMC modalities of recasts, the amount of attention allocated to target features, and subsequent L2 grammatical development in video-based SCMC. The findings of this study underscore the need for further exploration in this domain, with the aim of obtaining a more global understanding of the potential impact of various feedback modalities, including the multimodal provision of CF, in drawing learners' attention to target constructions and how their attention allocation associates with their subsequent L2 acquisition in SCMC settings.

It is worth highlighting that while the attention metrics displayed significant correlations with the gains of the multimodal recast group on each of three assessments, this might have been because all of these tests were meaning-oriented production tests. As noted earlier, these outcome measures and the experimental treatment tasks elicited similar contexts for the productive use of the target forms. Thus, the results might be a reflection of the principle of transfer-appropriate processing (Lightbown, 2008), which suggests that recall is more likely to occur when the learning task and its retrieval task closely resemble each other. Thus, it is crucial to acknowledge that these relationships might have differed if the assessment had included a form-oriented test or a receptive test. This highlights the importance of incorporating a range of assessments, including tests eliciting meaning-oriented, form-oriented, or receptive uses of the target forms, to provide insights into whether

the nature of the assessment indeed influenced the relationship between visual attention, feedback modality, and L2 grammatical development, especially in multimodal video-conferencing SCMC.

5.3. Moderating Effects of WM on L2 Grammatical Development (RQ5) – Study 2

The potential influence of WM on the outcomes of pedagogical interventions, such as CF in second or foreign language learning, is well documented (Pawlak, 2017). Given the important role of WM capacity in L2 acquisition (Linck et al., 2014), as well as its established relationship with CF (Granena & Yilmaz, 2018; Li, 2017), Study 2 was designed to explore how the contributions of WM abilities to L2 development might exhibit variations across different feedback modalities and assessments in the context of video-based SCMC.

This study also contributes to previous research exploring the effects of WM and feedback modality by conducting a comprehensive examination of various functions of WM and their distinct roles in L2 acquisition. To achieve this, a battery of five WM measures were employed, covering different aspects of WM: the NWS task for PSTM assessment; the forward Corsi block for VSTM evaluation; and executive functions assessed through tasks measuring updating, task-switching, and inhibitory control using the AOSPAN task, color shape task, and stop signal task, respectively. The multifaceted WM assessment revealed that the impact of these measures on learners' acquisition of the target constructions varied across different groups and types of target constructions.

Let us first consider the findings obtained for the present progressive form. Firstly, the results revealed a significant correlation between learners' PSTM and the improvements of the oral recast group in the SOPT, aligning with previous research that highlights the positive influence of PSTM on oral production development (e.g., Payne & Ross, 2005; Payne & Whitney, 2002; Révész, 2012; Trofimovich et al., 2017; White, 2020; Zalbidea & Sanz,

2020; Zhao, 2015). This finding suggests that, due to the greater cognitive demand imposed by processing oral input and stemming from the absence of written input and its transient nature (Gilbert et al., 2016), learners with higher PSTM capacity were likely to retain the target-like forms presented in oral recasts better for more extended periods. This increased the likelihood of effectively encoding these forms into long-term memory, facilitating the proceduralization of L2 knowledge. Such an effect seems particularly beneficial for tasks that heavily rely on procedural knowledge, as seen in the SOPT.

These findings contribute to the existing body of research on the relationship between WM and SCMC modality, providing evidence that the impact of PSTM can vary depending on the feedback modality. So far, however, the association between learners' PSTM capacity and various SCMC modalities and their interaction effects on subsequent L2 development in video-based interactions remains relatively unexplored. Further investigations are warranted to optimize feedback modalities for diverse aspects of L2 learning, accommodating learners with varying cognitive profiles and facilitating their L2 development.

The present study also revealed that the updating function of the central executive, as assessed by the AOSPAN task, emerged as a moderating factor in the performance of the oral recast and written recast groups across all three assessments, particularly concerning the learning of the present progressive form. This suggests that the significant role of updating capacity in L2 grammatical development may vary depending on feedback modality. However, it is important to note that this observation may not be applicable to other SCMC contexts and feedback types, as previous SCMC studies have yielded contradictory results. For instance, in text-based SCMC, Baralt (2010, 2015) revealed that the ability to update information did not significantly influence the extent to which learners benefited from recasts when learning the Spanish past subjunctive. Similarly, Martin's (2018) study, which explored the role of updating capacity in the effectiveness of metalinguistic feedback for L2

grammatical learning across text-based, voice-based, and video-based SCMC settings, found that all groups successfully acquired the Spanish past subjunctive regardless of their WM capacity. These inconsistent findings highlight the complex and context-specific nature of the relationship between updating capacity and the development of L2 grammatical knowledge in various SCMC settings using different forms of feedback.

Regarding the acquisition of the third person singular form, this study revealed an intriguing finding that differed from what had been observed in the learning of the present progressive. The analyses pointed to a more pronounced influence of stop signal task scores on the oral recast group's performance in the SOPT, particularly when compared to the multimodal recast group. This suggests the potential significance of inhibitory control abilities in facilitating the development of spontaneous oral production skills, especially within the oral recast group.

The substantial impact of inhibitory control on the effectiveness of oral recasts, particularly when dealing with the less salient third person singular form as opposed to the more conspicuous present progressive form, highlights the distinct challenge faced by the oral recast group. Inhibitory control is assumed to play a crucial role in L2 acquisition, as successful L2 use requires learners to inhibit information from competing language domains, such as syntax, semantics, and phonology (Gass & Lee, 2011). In this sense, in the case of learning the third person singular form, inhibitory control has particular significance, as learners need to select the most appropriate allomorphs within this form. Thus, inhibitory control appeared to assist learners in making these choices and suppressing alternative forms while processing this attention-demanding target form (Kane & Engle, 2000). During meaning-oriented interactive tasks, learners should rely on their inhibitory control to benefit from immediate feedback after their errors, by suppressing processing for meaning and selecting the feedback as the object of attention. In such contexts, learners must prioritize

where they need to direct their attention; there is competition between focusing on form and focusing on meaning (Skehan, 1998). It can be hypothesized that the production of third person singular in the SOPT would pose a greater challenge to the oral recast group than to the multimodal recast group because of their inability to rely on additional written cues.

In the context of text-based SCMC, the findings from Yilmaz and Sagdic's (2019) study contradict the results of the present study. Their research found that in text-based SCMC, inhibitory control did not significantly affect the acquisition of Spanish noun–adjective gender agreement, as measured by an oral production test and a GJT. The researchers speculated that this result could be attributed to the slower interaction pace and the availability of persistent written feedback, potentially rendering its impact less prominent. The discrepancy between the findings of this study and Yilmaz and Sagdic's (2019) research suggests that the specific interaction context, feedback modality, and cognitive demands associated with these factors might play a pivotal role in shaping the impact of inhibitory control on L2 grammar learning. However, it is important to acknowledge that only a limited number of studies have explored the moderating impact of inhibitory control on the effects of feedback modality on L2 grammar learning. Further exploration is needed to ascertain whether these relationships can be replicated in other SCMC research settings where various SCMC modalities can be utilized and directly compared.

5.4. Moderating Effects of WM on Attention Allocation (RQ6) – Study 2

To investigate the association between WM capacity and the visual attention of the multimodal recast and written recast groups, a series of Spearman correlation analyses was conducted. Overall, the two groups exhibited somewhat different relationships concerning the two target constructions. When examining the present progressive form, the attention allocation in the multimodal recast group was primarily associated with NWS scores,

reflecting the involvement of PSTM. In contrast, the written recast group's attention allocation displayed connections not only with NWS scores but also with scores from the forward Corsi block task, indicative of a role for VSTM. In the case of the third person singular form, a different pattern emerged. The multimodal recast group exhibited correlations between their attention allocation and NWS scores, as well as their performance in the color shape task, which tapped into their task-switching abilities. Conversely, the written recast group demonstrated correlations between their attention and all five WM metrics. These findings highlight the nuanced interplay between various WM functions, the allocation of visual attention, and the feedback modalities during the processing of recasts. In other words, feedback modality is closely intertwined with WM and how learners manage and direct their attention while processing feedback.

Previous eye-tracking research in L2 grammar learning, especially employing various WM measures, has also reported significant correlation between attention metrics and WM measures (e.g., Indrarathne & Kormos, 2018; Lee & Révész, 2021). These studies have shown that the relationship between WM and attention can be influenced by variables such as the type of intervention (e.g., implicit versus explicit instruction or enhanced versus unenhanced captions), and the intrinsic salience of target structures. The present study builds upon these previous findings by providing additional empirical evidence that demonstrates the intricate relationship between WM, attention measures, and feedback modalities employed in the context of multimodal, video-based SCMC. One important conclusion drawn from this line of research is that in order to obtain a more comprehensive understanding of the complex role of WM abilities in distribution of learners' attention and the effects of feedback modalities, it is necessary to incorporate a range of measurements that assess different functions of WM.

5.5. Implications of the Study

5.5.1. Theoretical Implications

This thesis has several noteworthy theoretical implications. First and foremost, it has provided evidence for the effectiveness of recasts in promoting the development of L2 grammatical knowledge in video-based SCMC contexts, reaffirming their beneficial role in L2 grammar development, as was also observed in prior SCMC research (e.g., Baralt, 2013; Kourtail & Révész, 2020; Monteiro, 2014; Rassaei, 2017, 2019; Yilmaz, 2012; Yilmaz & Yuksel, 2011). The novel contribution of this study lies in its confirmation of the efficacy of multimodal recasts, particularly in video-conferencing SCMC tasks—an area that has been relatively underexplored. By establishing the effectiveness of multimodal recasts in L2 development, this research further substantiates the Interaction Hypothesis (Gass & Mackey, 2006; Long, 1996), which suggests that a negotiated interaction wherein learners receive and respond to CF contributes to L2 development, hold also true in the context of multimodal feedback.

With a particular focus on the effectiveness of multimodal recasts, their superiority over unimodal recasts aligns well with the Dual Coding Theory (Paivio, 1986, 2007) and the Cognitive Theory of Multimedia Learning (Mayer, 2005a, 2014). These theoretical frameworks emphasize the advantages of engaging multiple sensory channels, such as auditory and visual systems, to facilitate information processing, ultimately leading to better learning outcomes and retention. In particular, in the context of L2 grammatical development through multimodal feedback, the findings of this study suggest that the sensory-modality approach (Baddeley, 1999; Baddeley, Eysenck, & Anderson, 2009) offers a more suitable framework for understanding information processing from multimodal input compared to the presentation-mode approach (Paivio, 1986, 2006). The sensory-modality approach distinguishes auditory and written input as processed in separate channels, one auditory and

the other visual. In contrast, the presentation-mode approach combines auditory and written input into a single channel (verbal), while treating visual input (e.g., pictures, graphs) as a separate channel (nonverbal).

As was previously noted, Mayer (2014) called for further research to clarify whether the sensory-modality approach or the presentation-mode approach better characterizes these channels. While previous studies in the field of SLA using audiovisual or multimodal linguistic input have predominantly adhered to the presentation-mode approach (e.g., Dang, Lu, & Webb, 2022; Lee & Révész, 2018, 2020, 2021; Wang & Pellicer-Sánchez, 2022), the present findings encourage a shift toward considering the sensory-modality approach. Recognizing the benefits of treating auditory and written input as separate channels, as opposed to conflating them into a single channel, may guide the design and implementation a more effective multimodal language learning materials and strategies, thereby enhancing learners' overall L2 development.

The advantages of multimodal recast observed also lend support to Wickens' (2007) Multiple Resource Model, which posits that different modalities draw upon different cognitive resources, each of which has limited capacity. By presenting information across modalities (e.g., oral + visual) rather than redundantly within a single modality (e.g., oral + oral), the cognitive load on learners can be reduced, facilitating the integration and organization of new information. The findings from this study suggest that in the case of multimodal recasts, the engagement of multiple sensory channels and simultaneous activation of multiple resources effectively manage cognitive load. This results in the improved encoding and retention of corrected forms in memory, ultimately enhancing the processing and integration of linguistic information during the L2 learning process.

This study challenges Wickens' (2007) model, which suggests that competition for similar or identical resources may hinder input processing. According to this model, tasks

engaging the same processing stage (e.g., listening and reading) would result in greater interference than carrying out perception and response tasks (e.g., listening and writing), as both tasks in the former combination involve perception. However, the increased learning gains observed in the multimodal recast group, wherein learners listened to and read recasts simultaneously, indicate that such competition may not impede input processing, but rather facilitate the integration and organization of new information, leading to improved learning outcomes and retention.

The benefits of multimodal recasts in this study provide empirical evidence for the Redundancy Facilitation Hypothesis (Mayer et al., 2014), which proposes a "reverse redundancy effect" (p. 654), where redundant information aids in fundamental cognitive processing. This might have been particularly critical for non-native speakers who have not yet automated certain linguistic aspects, as this redundancy might have helped decrease cognitive burden. In the case of multimodal recasts, the redundant written information likely enhanced learners' processing of the oral input, increasing the attention they paid to it. This assistance in recognizing unfamiliar input and compensating for less-advanced listening abilities can explain the advantages of multimodal recasts. The redundant presentation of information may have facilitated information processing, making it easier for learners to acquire and retain the target forms.

Another significant theoretical implication of this study pertains to the role of attention in L2 grammatical development. The results strongly support the idea that the amount of attention allocated by L2 learners to grammatical structures positively correlates with the extent of grammatical development. This reinforces the view that attention, operationalized here as eye fixations on the target constructions presented in the recasts, plays a crucial role in predicting SLA. This conclusion is aligned with the findings of existing SCMC and captioning studies that also recognize the significance of learners' attention in

input processing and language development (e.g., Lee & Révész, 2020; Smith, 2010, 2012; Smith & Renaud, 2013). Indeed, it is a generally accepted theoretical assumption that not all input that is available is processed by learners (Corder, 1976), and learner attention is necessary for input to be processed further to become intake with the potential to promote learning (Shmidt, 1990, 1993b, 2001). In line with this theoretical assumption, one goal of the current research was to shed light on the association between the amount of attention allocated to target linguistic constructions and development in L2 grammatical knowledge. By using three assessment tests and various early and late eye-tracking measures, this study has shown that the amount of attention L2 learners paid to grammatical constructions had a positive correlation with the extent of grammatical development they displayed.

The results showing that the multimodal recast group exhibited more attention than the written recast group appears to support Sweller's (Sweller, 1988, 2005; Chandler & Sweller, 1991) Cognitive Load Theory. This theory posits that, when auditory and visual information complement and correspond to each other, the cognitive load on WM decreases, compared to when a single channel is used (Sweller, 1988, 1994). Van Patten (2007) claimed that learners can attend to unfamiliar linguistic forms only when they have sufficient attentional resources while reading or listening. It is assumed that the provision of target-like forms through multimodal recasts may have imposed less cognitive load on the learners. Consequently, they likely had more attentional resources available to focus on the linguistic features in the input. This optimized cognitive processing may explain the observed consistent and a greater overall degree of attention to the target forms in the multimodal recast group as well as the greater learning gains of this group compared to the written recast group.

This study also presents a significant theoretical contribution by delving into the potentially distinctive roles of various WM abilities in predicting L2 learning. It adopted a

comprehensive approach, assessing different functions of WM, including executive control, PSTM, and VSTM. The multifaceted assessment of WM highlights the intricate and context-dependent influence of these measures on the acquisition of target constructions. Specifically, significant associations were found between PSTM and the SOPT gains in the oral recast group. Updating abilities emerged as a moderating factor influencing the performance of the oral and written recast groups across the three assessments, particularly concerning the learning of the present progressive form. Also, regarding the acquisition of the third person singular, the analyses revealed a more pronounced impact of inhibitory control on the improvement of spontaneous oral production skills in the oral recast group. Adding to the positive links between WM and L2 processing and learning outcomes, which are well documented in several narrative reviews (Jackson, 2020; Jeon & Yamashita, 2014; Juffs & Harrington, 2011; Wen, 2016; Williams, 2012) and meta-analyses (Linck et al., 2014; Shin, 2020), the findings of the present study collectively suggest that these WM functions may have varying roles in L2 acquisition and that their effects seem to be intricately linked to the type of feedback modality, significantly influencing the overall progress and development of L2 competence.

The various processes involved in oral and written interactions, both of which are prevalent in L2 learning contexts, have been claimed to draw on varying degrees of WM involvement (e.g., Gilabert et al., 2016; Grabowski, 2010; Kellogg, 2007). While prior research has tested this proposition in the context of L2 morphosyntactic development through computerized grammar practice with oral and written feedback (e.g., Saggara & Ahbbul, 2013a; Zalbidea & Sanz, 2020), as well as in video-conferencing interactions with audio-based, text-based, and video-based CF (e.g., Martin, 2018), it is important to note that only Zalbidea and Sanz's study incorporated a range of WM metrics assessing executive, phonological, and visuospatial WM abilities. Conversely, other research employed a single

WM measure, such as the L2 reading span task in Saggara and Ahbbul's (2013) study or the operation letter span task in Martin's research, both primarily tapping into executive WM. To date, no research has thoroughly examined how individual differences in executive, phonological, and visuospatial WM abilities influence L2 grammar learning outcomes resulting from various SCMC feedback modalities, including multimodal feedback, in video-based SCMC. As the initial study in this domain, this research revealed that the engagement of these WM functions varied across different CF modalities, as the findings of the study illustrated above. Nonetheless, further empirical investigations, encompassing a range of WM measures addressing executive, phonological, and visuospatial components, are essential to obtain a more fine-grained understanding of each WM component' contributions in conjunction with feedback modalities.

Finally, the present study concurs with the theoretical proposition that learners' WM may influence their attention allocation (Robinson et al., 2012) by identifying significant correlations between attention metrics and WM measures. The different patterns of association observed in the multimodal recast group and the written recast group substantiate the idea that learners' WM is linked to the efficient distribution of attentional resources. This finding, coupled with an exploration of how cognitive resources impact actual L2 performance, helps provide a more fine-grained understanding of the extent to which different WM components contribute to shaping learners' allocation of visual attention and their opportunities for L2 development across various modalities.

5.5.2. Methodological Implications

The present study yielded several methodological implications. First, the design adopted in this study allows for the concurrent use and comparison of various SCMC modalities during video-based interactions. In previous comparative research using different SCMC modalities

for L2 learning, each modality was used separately in diverse interaction contexts (i.e., oral CF during audio/video-based SCMC or written CF during text chat), hindering a direct comparison of these modalities within the same interaction contexts (e.g., Martin, 2018; see also Bryfonski & Ma, 2020). This study adds an additional dimension to our understanding of the effectiveness of various SCMC modalities.

The research findings also highlight the benefits of incorporating a variety of eye-tracking measures when investigating the relationship between attention metrics and feedback modalities, as well as the link between visual attention to target constructions and subsequent learning gains of these structures. In this study, both early and late measures displayed more significant associations with the learning gains of the multimodal recast group across all three assessments for both target constructions, when compared to the written recast group. This implies that both early and late attention measures may be correlated with L2 grammatical development, and these correlations could vary depending on the feedback modalities utilized in video-based SCMC. In addition, it is also worth noting that early and late measures, such as first-pass reading time and re-reading time, exhibited somewhat different associations. For instance, in the case of the present progressive, first-pass reading time only predicted the multimodal recast group's pretest–delayed posttest gains, whereas re-reading time predicted both pretest–posttest and pretest–delayed posttest gains across all three assessment tasks. So far, previous SCMC research on recasts using eye-tracking has exclusively focused on text-based SCMC, often using a limited set of attention measures, such as the number of fixations and total fixation durations. Therefore, the combination of both early and late measures is expected to provide a multifaceted view of participants' eye-gaze behavior, offering converging evidence for research (Conklin et al., 2018; Godfroid, 2020).

Another noteworthy methodological implication arising from this study is its focus on the impact of WM on initial morphosyntactic achievement among elementary-level learners

with minimal prior knowledge of the target structures. The emphasis on elementary-level learners can benefit future research, as it provides a clear and focused investigation, devoid of potential confounding variables associated with advanced language skills or extensive prior knowledge. Considering that WM is viewed as central to many important cognitive functions that play a role in regulating L2 development, particularly during its early stages (e.g., Skehan, 2016; Wen et al., 2017), this study, by commencing with this specific group, laid the groundwork for exploring the connection between various WM components and initial L2 acquisition within the realm of multimodal video-based SCMC. Given existing evidence suggesting a diminishing impact of cognitive capacities on L2 performance as proficiency increases (e.g., Serafini & Sanz, 2016), future research may be inclined to explore how the interaction between different feedback modalities and individual WM abilities evolves as L2 learners advance to higher proficiency levels.

5.5.3. Pedagogical Implications

Some useful pedagogical implications are deducible from the findings of this study. One primary pedagogical implication is the encouragement of the use of multimodal CF to facilitate L2 grammar learning in video-conferencing environments. As noted earlier, a key novelty of this research lies in the utilization of multimodal recasts in the form of captioned oral recasts, which seamlessly combine oral recasts with on-screen written recasts delivered simultaneously. The noteworthy aspect here is that the multimodal recast group achieved significantly greater improvements in all three assessment tests at both posttest and delayed posttest, as compared to the other groups. These findings reinforce the assumption that multimodal recast likely alleviates learners' cognitive burden, ultimately leading to better understanding and acquisition of the target forms. Moreover, the positive impact of multimodal recasts on L2 development underscores their potential to raise learners'

awareness of linguistic input, and aid in decoding and segmenting speech, reducing cognitive load (Frumuselu, 2018; Montero Perez, 2020b).

The on-screen written recasts within multimodal recasts are provided in a contingent manner, immediately following learners' erroneous productions. This contingency, which is often challenging to achieve in conventional text-chat CF due to the absence of strict turn adjacency and a split negotiation routine (Smith, 2003), is considered a crucial factor contributing to the effectiveness of recasts. The juxtaposition of ill-formed learner output and target-like models within contingent recasts is known to promote learners' noticing (Lai et al., 2008; Long, 2007; Ziegler, 2017). The immediate on-screen written feedback has shown its advantages in promoting L2 development, yielding the second most significant gains in the written recast group following the gains from the multimodal recasts. This insight further enriches our existing knowledge regarding the importance of feedback contingency in L2 learning. In summary, the highest gains achieved through multimodal recasts both in oral and written production underscore the potential of utilizing multimodal pedagogical interventions to foster overall L2 grammatical development. This implication is especially pertinent in the current educational trend, where a wide array of SCMC modalities, including multimodal video-conferencing, are being widely utilized for language instruction.

It is evident that multimodal recasts offer the most significant benefits and appear to have the potential to provide a versatile approach capable of adapting to various circumstances, thus fostering a dynamic and effective learning environment. However, it is crucial to acknowledge that oral and written recasts also proved to be advantageous when compared to the no recast condition across all assessments. This finding corroborates earlier findings indicating that L2 development can transfer between different modes (e.g., Baralt, 2013; Martine, 2019; Payne & Whitney, 2002). Recognizing this transferability across different modalities underscores the importance of embracing a holistic approach to language

learning, wherein both oral and written forms of feedback can be deployed and can contribute to an overall L2 learning process. Research endeavor, such as the current study, can inform curriculum developers and educators about optimal design when structuring courses for L2 learners (Blake 2015; Blake, Wilson, Cetto, & Pardo-Ballester, 2008; Cerezo, Baralt, Suh, & Leow, 2014).

This study also sheds light on the intricate interplay between learners' executive, phonological, and visuospatial WM capacities, and the learning gains derived from different SCMC feedback modalities, targeting L2 grammatical structures with varying degrees of salience. These findings advance our understanding of the distinct contributions of various WM components to L2 grammar proficiency. Moreover, they provide insights into how the relationship between WM and linguistic outcomes can differ depending on the types of modalities and the specific grammatical structures being targeted. In other words, this research highlights the way in which individual differences interact with linguistic and contextual variables (Dekeyser, 2012). As such, the findings of this study may pave the way for profile-matched pedagogical options, enabling tailored instruction that accommodates the varying WM capacities of L2 learners, thereby enhancing the overall effectiveness of language education.

5.6. Limitations and Future Research Directions

The present study has several limitations, each of which offers opportunities for refinement and expansion in future research. These limitations can be broadly categorized into four main areas: 1) material selection, 2) generalization of findings, 3) data collection methods, and 4) target constructions.

One limitation arises from the exclusive reliance on meaning-oriented production tests to gauge L2 grammatical knowledge development. Although this choice aligns with the

nature of treatment, exploring whether the benefits of multimodal recasts observed extend to diverse assessment tasks, such as form-oriented assessments (e.g., GJTs) or receptive tests (e.g., fill-in-the-blank tests) would allow a more comprehensive understanding of how different modalities affect L2 grammatical knowledge development.

Secondly, this study focused on the influence of WM on initial morphosyntactic achievement. The participant sample was confined to elementary-level learners with minimal prior knowledge of the target structures. This intentional selection aimed to ensure a focused investigation, minimizing potential confounding variables linked to advanced language skills or extensive prior knowledge. While this approach facilitated a clear and focused exploration, future research is encouraged to broaden the scope of these findings by extending these results to more advanced proficiency levels and diverse learner populations. Exploring the persistence of WM effects over time and its role in learning more intricate morphosyntactic structures could offer valuable insights into the ongoing developmental aspects of language acquisition.

A third limitation pertains to the eye-tracking system employed for eye-movement data collection. While the use of a remote eye-tracker enhances ecological validity compared to head-and-chin-rest eye-tracker by allowing learners to move naturally, it may introduce challenges related to less precise gaze capture due to head and positional shifts. An additional issue concerns the resolution of the eye-tracker. In this study, an eye-tracker with a temporal resolution of 60 Hz was used. Future research could benefit from a higher-resolution eye-tracker for more accurate fixations and saccade data.

The fourth limitation is its exclusive focus on morphological development. Future research endeavors may seek to ascertain whether similar findings can be replicated in other video-conferencing environments targeting the acquisition of different linguistic features. For example, some previous video-based SCMC research has examined the effects of various CF

types on L2 phonological development (e.g., Bryfonski & Ma, 2020; Parlak & Ziegler, 2017; Saito & Akiyama, 2017). Among them, Bryfonski and Ma (2020) employed video-conferencing to compare the effects of explicit (metalinguistic input) and implicit CF (recasts) on L2 Mandarin tone perception and production. Their findings indicated that learners who received implicit feedback exhibited more significant improvements in tone perception compared to those in the explicit feedback group. Therefore, it would be worthwhile to investigate whether incorporating multimodal implicit or explicit feedback, in conjunction with various SCMC modalities, could yield similar results regarding phonological development during video-based interactions. Such an exploration has the potential to offer educators valuable insights into selecting the most suitable modality for their educational objectives and adapting it to learners' proficiency levels in the context of video-conferencing language learning settings.

REFERENCES

Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research*, *7*(3), 347–376. https://doi.org/10.1191/1362168803LR127OA

Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 29–51). Oxford, UK: Oxford University Press.

Aeen, J. B., Rassaei, E., Riasati, M. J., & Zamanian, M. (2022). EFL learners' modified output and their perceptions of recast during face-to-face, audio-based, and video-based synchronous mobile-mediated communication. *TESL-EJ, 25*(4).

Ahmadian, M. J. (2020). Explicit and implicit instruction of refusal strategies: Does working memory capacity play a role? *Language Teaching Research*, *24*(2), 163–188.

Al-Surmi, M. (2012). Learners' noticing of recasts of morpho-syntactic errors: Recast types and delayed recognition. *System, 40*(2), 226–236.

Altgassen, M., Vetter, N. C., Phillips, L. H., Akgün, C., & Kliegel, M. (2014). Theory of mind and switching predict prospective memory performance in adolescents. *Journal of Experimental Child Psychology, 127*, 163–175.

Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research, 12*, 183–210.

Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition, 28*, 543–574.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Arroyo, D. C., & Yilmaz, Y. (2018). An open for replication study: The role of feedback timing in synchronous computer-mediated communication. *Language Learning, 68*(4), 942–972.

Atkins, P. W., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*(4), 537–552.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *Psychology of learning and motivation* (Vol. 2, pp. 89–195). : Academic Press.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Clarendon.

Baddeley, A. D. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences, 93*(24), 13468–13472.

Baddeley, A. D. (1999). *Essentials of human memory*. Hove, UK: Psychology Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/s1364-6613(00)01538-2

Baddeley, A. D. (2003a). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208.

Baddeley, A. D. (2003b). Working memory: Looking back and looking forward. *Nature, 4*, 829–839.

Baddeley, A. D. (2007). *Working memory, thought, and action*. New York: Oxford University Press.

Baddeley, A. D. (2010). Working memory. *Current Biology*, *20*(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A. D. (2015). Working memory in second language learning. In Z. Wen, M. B. Mota, & A. McNeill, A. (Eds.), *Working memory in second language acquisition and processing*, 17–28. Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093595-012

Baddeley, A. (2017). *Exploring working memory: selected works of Alan Baddeley.* Routledge.

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

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 8* (pp. 47–89). New York: Academic Press.

Baddeley, A., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake, & P. Shah (Eds.), *Models of working memory* (pp. 28–61). Cambridge University Press.

Baralt, M. (2010). *Task complexity, the cognition hypothesis, and interaction in CMC and FTF environments* [Unpublished doctoral dissertation]. Georgetown University, Washington, D.C

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

Baralt, M. (2015). Working memory capacity, cognitive complexity and L2 recasts in online language teaching. In Z. Wen, M. B. Mota, & A. McNeill, A. (Eds.), *Working memory in second language acquisition and processing* (pp. 248–269). Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093595-012

Bardovi-Harlig, K. (2000). Tense and aspect in second language acquisition: Form, meaning, and use. *Language Learning: A Journal of Research in Language Studies*, *50*, 1.

Bardovi-Harlig, K. (2001). Another piece of the puzzle: The emergence of the present perfect. *Language learning*, *51*, 215–264.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Barton, K. (2020). *MuMIn: Multi-Model Inference*. R package version 1.43.17. Retrieved from 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

Berglund, T. Ö. (2012, August 22–25). Corrective feedback and noticing in text-based second language interaction. Paper presented at the EUROCALL Conference, Gothenburg, Sweden.

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics, 23*(4), 509–533.

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*(02), 399–418. https://doi.org/10.1017/s0142716412000434

Blake, R. J. (2008). *Brave new digital classrooms: Technology and foreign language learning*. Washington D.C: Georgetown University Press.

Blom, E., Paradis, J., & Sorenson Duncan, T. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular *–s* in child L2 English. *Language Learning, 62*, 965–994.

Bower, J., & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTandem. *Language Learning and Technology*, *15*(1), 41–71.

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

Brooks, M. E., Kristensen, K., van Bentham, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed. *The R*

Journal, 9(2), 378–400. https://doi.org/10.3929/ethz-b-000240890
Retrieved from https://journal.r-project.org/archive/2017/RJ-2017-066/index.html

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

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: Cambridge University Press.

Carpenter, H., Jeon, K. S., MacGregor, D., & Mackey, A. (2006). Learners' interpretations of recasts. *Studies in Second Language Acquisition, 28*(2), 209–236. https://doi.org/10.1017/S0272263106060104

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293 -332.

Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139524681

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language learning*, *27*(1), 29–46.

Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition, 7*, 1–14.

Choi, M.Y. (2000). *Effects of recasts on irregular past tense verb morphology in web-chat*. Unpublished master's thesis, University of Hawaii, Honolulu.

Cintrón-Valentín, M. C., & Ellis, N. C. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition, 37*, 197–235. doi:10.1017/S0272263115000029

Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in second language acquisition: Physical form, learner attention, and instructional focus. *Frontiers in Psychology, 7*, 1284. https://doi.org/10.3389/fpsyg.2016.01284

Cintrón-Valentín, M., & García-Amaya, L. (2021). Investigating textual enhancement and captions in L2 grammar and vocabulary: An experimental study. *Studies in Second Language Acquisition, 43*(5), 1068-1093.

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.

Clark, R. C., & Mayer, R. E. (2011). E-learning and the science of instruction: *Proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: Pfeiffer.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van Gompel (Ed.), *Eye movements: A window on mind and brain* (pp. 341–371). Amsterdam, the Netherlands: Elsevier.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New York: Academic Press.

Congdon, E., Mumford, J. A., Cohen, J. R., Galvan, A., Canli, T., & Poldrack, R. A. (2012). Measurement and reliability of response inhibition. *Frontiers in Psychology, 3,* 1–10.

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language acquisition research. *Second Language Research, 32*(3), 453–467. doi:10.1177/02676583166 37401

Conklin, K., Pellicer-Sánchez, A., & Carroll, G. (2018). *Eye-tracking: A guide for applied linguistics research.* Cambridge, UK: Cambridge University Press.

Conway, A., Jarrold, C., Kane, M., Miyake, A., & Towse, J. (Eds.). (2007). *Variation in working memory.* Oxford, UK: Oxford University Press.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161–170.

Corder, S. P. (1982). *Error analysis and interlanguage* (2nd ed.). Oxford, UK: Oxford University Press.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104,* 163–191.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Cuevas, H. M., Fiore, S. M., & Oser, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science, 30*(6), 433–464.

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*, 369–382.

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language, 75*, 117–139.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Dang, T. N. Y., Lu, C., & Webb, S. (2022). Incidental learning of collocations in an academic lecture through different input modes. *Language Learning*, *72*(3), 728–764.

Dai, B. (2013). *Individual differences in learners' working memory, noticing of L2 forms in recasts and their L2 development in task-based interactions* [Unpublished doctoral dissertation]. The Chinese University of Hong Kong.

DeKeyser, R. (2001). Automaticity and automatization. In Peter Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151) Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524780.007

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

DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, *62*, 189–200.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge, UK: Cambridge University Press.

Doughty, C., & Long, M. H. (Eds.). (2003). *Handbook of second language acquisition*. New York: Basil Blackwell.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). New York: Cambridge University Press.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice.* Springer Science & Business Media.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*, 257–303.

Dulay, H., & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning, 24*, 37–53.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford, UK: Oxford University Press.

d'Ydewalle, G., & Van de Poel, M. (1999). Incidental foreign language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research, 28*(3), 227–244.

Egi, T. (2007). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, *29*(4), 511–537.

Ellis, N. (2002). Frequency effects in language processing: A Review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*(2), 143–188. https://doi.org/10.1017/S0272263102002024

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*, 305–352.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164–194. doi: 10.1093/applin/aml015.

Ellis, R. (1994). *The study of second language acquisition*. Oxford, UK: Oxford University Press.

Ellis, R. (2003). Task-based language learning and teaching. Oxford: Oxford University Press.

Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 339–360). Oxford, UK: Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford, UK: Oxford University Press.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*, 221–246.

Ellis, R. (2012). *Language teaching research and language pedagogy*. Oxford, UK: John Wiley & Sons.

Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research, 20*(3), 405–428. https://doi.org/10.1177/1362168816628627

Ellis, R. (2017). Position paper: Moving task-based language teaching forward. *Language Teaching*, *50*(4), 507–526.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, *28*, 339–368. https://doi.org/10.1017/S0272263106060141

Ellis, R., & Sheen, Y. (2006). Re-examining the role of recasts in SLA. *Studies in Second Language Acquisition*, *28*, 575–600.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. New York: Routledge.

Enticott, P. G., Ogloff, J. R., & Bradshaw, J. L. (2006). Associations between laboratory measures of executive inhibitory control and self-reported impulsivity. *Personality and Individual Differences*, *41*, 285–294.

Erlam, R., & Loewen, S. (2010). Implicit and explicit recasts in L2 oral French interaction. *The Canadian Modern Language Review*, *66*(6), 887–916.

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

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, *37*(4), 581–590.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179.

Frumuselu, A. D. (2018). The implications of Cognitive Load Theory and exposure to subtitles in English Foreign Language (EFL). *Translation and Translanguaging in Multilingual Contexts*, *4*(1), 55–76. http://doi.org/10.1075/ttmc.00004.fru

Gass, S. M. (1988). Integrating research areas: A framework for second language studies1. *Applied Linguistics, 9*(2), 198–217. https://doi.org/10.1093/applin/9.2.198

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gass, S. M. (2003). Input and interaction. In C. Doughty, & M. Long (Eds.), *The handbook of second language acquisition* (pp. 224–255). Oxford, UK: Blackwell. https://doi:10.1002/9780470756492.ch9

Gass, S. M., & Lee, J. (2011). Working memory capacity, inhibitory control, and proficiency in a second language In M. S. Schmid & W. Lowie (Eds.), *Modeling bilingualism: From structure to chaos* (pp. 59–84). Amsterdam, Netherlands: John Benjamins.

Gass, S. M., & Lewis, K. (2007). Perceptions about interactional feedback: 381 differences between heritage language learners and non-heritage language learners. In Mackey A. (Ed.), *Conversational interaction in second language acquisition* (pp. 79–99). Oxford, UK: Oxford University Press.

Gass, S. M., & Mackey, A. (2006). Input, interaction, and output: An overview. *AILA Review, 19*, 3–17.

Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. Vanpatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175–200). Mahwah, NJ: Lawrence Erlbaum and Associates.

Gass, S. M., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 180–206). Routledge.

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.). (2018). *Salience in second language acquisition*. New York, Routledge.

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/10.10125/44684

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410–419.

Gilabert, R., Manchón, R., & Vasylets, L. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics, 36*, 117–135.

Godfroid, A. (2019). Investigating instructed second language acquisition using L2 learners' eye-tracking data. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 44–57). New York: Routledge.

Godfroid, A. (2020a). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. New York: Routledge.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye-tracking. *Studies in Second Language Acquisition, 35*(03), 483–517. http://doi.org/10.1017/s0272263113000119

Godfroid, A., Housen, A., & Boers, F. (2010). A procedure for testing the noticing hypothesis in the context of vocabulary acquisition. In M. Putz & L. Sicola (Eds.), *Cognitive*

processing in second language acquisition (pp. 169–197). Amsterdam: John Benjamins.

Godfroid, A., & Hui, B. (2020). Five common pitfalls in eye-tracking research. *Second Language Research, 36*(3), 277–305. http://doi.org/10.1177/0267658320921218

Godfroid, A., & Uggen, M. S. (2013). Attention to irregular verbs by beginning learners of German: An eye-movement study. *Studies in Second Language Acquisition, 35*(2), 291–322.

Godfroid, A., Winke, P., & Conklin, K. (2020). Exploring the depths of second language processing with eye tracking: An introduction. *Second Language Research, 36*(3), 243–255.

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *The Journal of Neuroscience, 33*, 387–396.

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 51*(1), 1–50. https://doi.org/10.1111/1467-9922.00147

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition, 34*, 445–474. https://doi.org/10.1017/S0272263112000149

Goo, J. (2014a). Corrective feedback and phonological short-term memory in the acquisition of the English *that*-trace filter: A conceptual replication of Goo (2012). *English 21, 27*, 231–257.

Goo, J. (2014b). Type of feedback, phonological short-term memory, and L2 development. *English Language & Literature Teaching, 20*(4), 45–66.

Goo, J. (2016). Corrective feedback and working memory capacity: A replication. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 279–302). Amsterdam, the Netherlands: John Benjamins.

Goo, J. (2020). Research on the role of recasts in L2 learning. *Language Teaching*, *53*, 289–315.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. *Implicit and Explicit Learning of Languages*, 443–482.

Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition, 35*, 127–165.

Grabowski, J. (2010). Speaking, writing, and memory span in children: Output modality affects cognitive performance. *International Journal of Psychology, 45*(1), 28–39. https://doi.org/10.1080/00207590902914051

Granena, G., & Yilmaz, Y. (2018). Aptitude-treatment interaction in L2 learning: A research synthesis. *Studies in English Education, 4*, 803–830. https://doi.org/10.22275/SEE.23.4.02.

González-Lloret, M. (2014). The need for needs analysis in technology-mediated TBLT. *Technology-mediated TBLT: Researching technology and tasks*, *6*, 23–50.

Gurzynski-Weiss, L., Al Khalil, M., Baralt, M., & Leow, R.P. (2016). Levels of awareness in relation to type of recast and type of linguistic item in synchronous computer mediated communication: A concurrent investigation. In R. P. Leow, L. Cerezo, & M.

Baralt (Eds.), *A psycholinguistic approach to technology and language learning* (pp. 151–170). Berlin: De Gruyter Mouton.

Gurzynski-Weiss, L., & Baralt, M. (2015). Does type of modified output correspond to learner noticing of feedback? A closer look in face-to-face and computer-mediated task-based interaction. *Applied psycholinguistics, 36*(6), 1393–1420.

Gurzynski-Weiss, L., Henderson, C. & Jung, D. (2018). Examining timing and type of learner modified output in relation to perception in face-to-face and synchronous task-based chat. In M. J. Ahmadian & M.P. García Mayo (Eds.), *Recent perspectives on task-based language learning and teaching* (pp. 53–76). Boston, MA/Berlin: De Gruyter.

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

Hahn, H. R. (2013). The acquisition of progressive morphology by Korean learners of English: L1 transfer and learnability. *English Teaching*, *68*(2), 105–130.

Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition, 32*, 465–491.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 input. *TESOL Quarterly, 36*, 543–572.

Hampel, R., & Stickler, U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. *ReCALL, 24*(2), 116–137.

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied, 12*(2), 67–86.

Hummel, K. M., & French, L. M. (2010). Phonological memory and implications for the second language classroom. *Canadian Modern Language Review, 66*(3), 371–391.

Indrarathne, B., & Kormos, J. (2016). Relationship between attentional processing of input and working memory: An eye-tracking study. In *12th BAAL Language Learning and Teaching SIG: Crossing Boundaries: Language Learning and Teaching Inside and Outside the Classroom*.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions: An eye-tracking study. *Studies in Second Language Acquisition, 3*9(3), 401–430.

Indrarathne, B., & Kormos, J. (2018). The role of working memory in processing L2 input: Insights from eye-tracking. *Bilingualism: Language and Cognition, 21*(2), 355–374.

Inquisit 6 [Computer software]. (2022). Retrieved from https://www.millisecond.com.

Issa, B. I. (2019). Examining the relationships among attentional allocation, working memory, and second language development: An eye-tracking study. In *The Routledge handbook of second language research in classroom learning* (pp. 464–479). Routledge.

Isreal, J., Chesney, G., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for a multiple capacity view of attention. *Psychophysiology, 17,* 259–273.

Issa, B., Morgan-Short, K., Villegas, B., & Raney, G. (2015). An eye-tracking study on the role of attention and its relationship with motivation. In L. Roberts, K. McManus, N. Vanek, & D. Trenkic (Eds.), *EUROSLA Yearbook, 15*, pp. 114–142. Amsterdam: John Benjamins.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*, 541–577.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly, 34*, 239–278.

Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching, 10*, 89–101.

Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. *Neuroscience, 139*(1), 39–50. https://doi.org/10.1016/j.neuroscience.2005.07.002.

Jelani, N. A. M., & Boers, F. (2018). Examining incidental vocabulary acquisition from captioned video. *ITL-International Journal of Applied Linguistics*, *169*(1), 169–190.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language learning*, *64*(1), 160–212.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language teaching*, *44*(2), 137–166.

Jung, J. (2017). *Effects of task complexity, glossing, and working memory capacity on L2 reading and L2 learning* [Unpublished dissertation]. University College London Institute of Education.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–448.

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed.) (pp. 247–262). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.013

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336–358. https://doi.org/10.1037/0278-7393.26.2.336

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.). *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Mahwah, NJ: Erlbaum.

Kellogg, R. T. (2007). *Fundamentals of cognitive psychology.* Sage Publications, Inc.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*, 382–397.

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., & De Haan, E. H. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology, 7*, 252–258.

Kim, J. H. (2019). The relative effect of recasts on L2 Korean learners' accuracy development of two different forms and its relationship with language analytic ability. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/1362168819859910

Kim, J. H. (2021). The relative effect of recasts on L2 Korean learners' accuracy development of two different forms and its relationship with language analytic ability. *Language Teaching Research*, *25*(3), 451–475.

Kim, J. H., & Han, Z. (2007). Recasts in communicative EFL classes: Do teacher intent and learner interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 269–297). Oxford, UK: Oxford University Press.

Kim, S., & Cho, Y. W. (2017). The effects of recasts and working memory on Korean EFL learners' past tense accuracy. *English Teaching, 72*(4), 105–132.

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

Kormos, J. (2014). Differences across modalities of performance. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning–Insights from and for L2 writing* (pp. 193–216). Amsterdam: John Benjamins.

Kourtali, 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

Kourtali, N. E., & Borges, L. (2023). The effects of feedback timing on L2 development in written SCMC. *Computer Assisted Language Learning*, 1–29.

Kourtali, N. E., & Révész, A. (2020). The roles of recasts, task complexity, and aptitude in child second language development. *Language Learning, 70*, 179–218.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, UK: Pergamon.

Krashen, S. (1983). Newmark's "Ignorance Hypothesis" and current second language acquisition theory. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 135–153). Rowley, MA: Newbury House.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. London/New York: Longman.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Lado, B. (2017). Aptitude and pedagogical conditions in the early development of a nonprimary language. *Applied Psycholinguistics*, *38*(3), 679–701.

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/

Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology, 10*(3), 102–120. http://dx.doi.org/10.10125/44077

Larsen-Freeman, D. E. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*(4), 409-419.

Larsen-Freeman, D. E., & Long, M. H. (1991). *An introduction to second language acquisition research.* London/New York: Longman.

Lee, E. (2013). Corrective feedback preferences and learner repair among advanced ESL students. *System, 41*(2), 217–230

Lee, H., & Mayer, R. E. (2018). Fostering learning from instructional video in a second language. *Applied Cognitive Psychology, 32*(5), 648–654. https://doi.org/10.1002/acp.3436

Lee, M., & Révész, A. (2018). Promoting grammatical development through textually enhanced captions: An eye-tracking study. *The Modern Language Journal, 102*(3), 557–577.

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

Lee, M., & Révész, A. (2021). The role of working memory in attentional allocation and grammatical development under textually-enhanced, unenhanced and no captioning conditions. *Journal for the Psychology of Language Learning, 3*(1), 6–25.

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

Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from linguistics and psychology* (Cambridge Applied Linguistics Series, 111–137). Cambridge, UK: Cambridge University Press.

Leibowitz, H., & Post, R. (1982). The two modes of processing concept and some implications. In J. Beck (Ed.), *Organization and representation in perception* (pp. 343–363). Hillsdale, NJ: Erlbaum.

Leow. R. P. (2001). Attention, awareness, and foreign language behavior. *Language Learning, 51*, 113–155.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach* (1st ed.). New York: Routledge.

Leow, R., & Driver, M. (2021). Cognitive Theoretical perspectives of corrective feedback. In H. Nassaji & E. Kartchava (Eds.), *The Cambridge handbook of corrective feedback in second language learning and teaching* (Cambridge Handbooks in Language and Linguistics, pp. 65–84). Cambridge, UK: Cambridge University Press. https://doi:10.1017/9781108589789.004

Li, H., & Iwashita, N. (2019). The role of recasts and negotiated prompts in an FL learning context in China with non-English major university students. *Language Teaching Research*, 1–15. https://doi.org/10.1177/1362168819839727

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, 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, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal, 97*, 634–654.

Li, S. (2014). The interface between feedback type, L2 proficiency, and the nature of the linguistic target. *Language Teaching Research, 18*(3), 373–396.

Li, S. (2015). Working memory, language analytic ability and L2 recasts. In Z. Wen, M. B. Mota, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp.139–159). Multilingual Matters. https://doi.org/10.21832/9781783093595-012

Li, S. (2017). The effects of cognitive aptitudes on the process and product of L2 interaction: A synthetic review. In L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors* (pp. 41–70). John Benjamins.

Li, S., Ellis, R., & Zhu, Y. (2019). The associations between cognitive ability and L2 development under five different instructional conditions. *Applied Psycholinguistics, 40*(3), 693–722. doi:10.1017/S0142716418000796

Li, S., & Fu, M. (2018). Strategic and unpressured within-task planning and their associations with working memory. *Language Teaching Research*, *22*(2), 230–253.

Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in System. *System, 84*(1), 93–109.

Liao, Y., & Zhang, W. (2022). Corrective feedback, individual differences in working memory, and L2 development. *Frontiers in Psychology*, *13*, 811748. https://doi.org/10.3389/fpsyg.2022.811748

Lightbown, P. M. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 177–196). Cabridge, UK: Cambridge University Press.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language acquisition. *Studies in Second Language Acquisition*, *12*, 429–448.

Lightbown, P. M., & Spada, N. (1997). Learning English as a second language in a special school in Quebec. *Canadian Modern Language Review, 53*, 315–355.

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning, 65*, 185–207.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861–883.

Liu, Y., & Wickens, C. D. (1992). Visual scanning with or without spatial uncertainty and divided and selective attention. *Acta Psychologica, 79*, 131–153.

Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, *19*(1), 1–14. https://doi.org/10.1080/09588220600803311

Loewen, S., & Inceoglu, S. (2016). The effectiveness of visual input enhancement on the noticing and L2 development of the Spanish past tense. Studies in *Second Language Learning and Teaching, 6*(1), 89–110.

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361–377). Oxford, UK: Oxford University Press.

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

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285–329. https://doi:10.1017/S0261444818000125

Loewen, S., & Reinders, H. (2011). *Key concepts in second language acquisition*. Basingstoke, UK: Palgrave Macmillan

Logan, G. D. (1985). On the ability to inhibit simple thoughts and actions: II. Stop-signal studies of repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 675–691.

Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.

Long, M. H. (1981). Input, interaction, and second-language acquisition. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences: Vol. 379. Native language and foreign language acquisition* (pp. 259–278). New York: New York Academy of Sciences.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*, 126–141.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam and M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–100). Clevedon, UK: Multilingual Matters.

Long, M.H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251–285. http://dx.doi.org/10.1017/S0272263100009165

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: John Benjamins.

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). New York: Academic Press.

Long, M. H. (2006). Recasts in SLA: The story so far. In M. Long (Ed), *Problems in SLA* (pp.75–116). Mahwah: Erlbaum.

Long, M. H. (2007). *Problems in SLA.* Mahwah, NJ: Erlbaum.

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Malden, MA: Wiley Blackwell.

Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, *82*(3), 357–371. https://doi.org/10.1111/j.1540-4781.1998.tb01213.x

Lüdecke, D. (2020). *sjPlot: data visualization for statistics in social sciences*. R package version 2.8.4. Retrieved from https://CRAN.R-project.org/package=sjPlot

Lüdecke, D. (2023). *sjPlot: Data Visualization for Statistics in Social Science.* R package version 2.8.15. Retrieved from https://CRAN.R-project.org/package=sjPlot

Lyster, R. (1998a). Form in immersion classroom discourse: In or out of focus? *Canadian Journal of Applied Linguistics*, *1*(1–2), 53–82.

Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *48*(2), 183–218.

Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, *20*, 51–81. https://doi.org/10.1017/S027226319800103X

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, *26*(3), 399–432.

Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, *59*(2), 453–498.

Lyster, R., Lightbown, P. M., & Spada, N. (1999). A response to Truscott's 'What's wrong with oral grammar correction'. *Canadian Modern Language Review*, *55*(4), 457–467. https://doi.org/10.3138/cmlr.55.4.457

Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition, 28*, 321–341. https://doi.org/10.1017/S0272263106060128

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., & Ranta, L. (2013). Counterpoint piece: The case for variety in corrective feedback research. *Studies in Second Language Acquisition, 35*(1), 167–184. https://doi.org/10.1017/S027226311200071X

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., and Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching. 46*(1) 1–40. https://doi: 10.1017/S0261444812000365

Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*(4), 557–587. https://doi.org/10.1017/S0272263199004027

Mackey, A. (2002). Beyond production: Learners' perceptions about interactional processes. *International Journal of Educational Research (Special issue on the role of interaction in instructed language learning), 37,* 379–394.

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. (2007). The role of conversational interaction in second language acquisition. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 1–26). Oxford, UK: Oxford University Press.

Mackey, A. (2012). *Input, interaction, and corrective feedback in L2 learning*. Oxford, UK: Oxford University Press.

Mackey, A., Abbuhl, R., & Gass, S. M. (2012). Interactionist approach. In S. M. Gass & A. Mackey (Eds.) *The Routledge handbook of second language acquisition*. New York: Routledge, 7–24.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Erlbaum.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, *22*, 471–49. https://doi.org/10.1017/S0272263100004010

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford, UK: Oxford University Press.

Mackey, A., & Goo, J. (2012). Interaction approach in second language acquisition. In C.A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Oxford, UK: John Wiley & Sons. https://doi.org/10.1002/9781405198431.wbeal0551

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

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). Amsterdam: John Benjamins.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*(3), 501–533.

Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning*, *62*(3), 704–740.

Markham, P. (1999). Captioned videotapes and second language listening word recognition. *Foreign Language Annals, 32*, 321–328.

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*. Washington, DC: Georgetown University.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*(3), 379–413. https://doi.org/10.1017/S0272263112000125

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85–139.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge, UK: Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.

Mayer, R. E. (2014). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1–24). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.002

Mayer, R., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (Cambridge Handbooks in Psychology, pp. 279–315). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.015

Mayer, R. E., Lee, H., & Peebles, A. (2014). Multimedia learning in a second language: A cognitive load perspective. *Applied Cognitive Psychology, 28*(5), 653–660.

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology, 24*(2), 222.

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. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 323 – 338). Oxford, UK: Oxford University Press.

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

Meskill, C. (1996). Listening skills development through multimedia. *Journal of Educational Multimedia and Hypermedia*, *5*(2), 179–201. Department of Educational Theory and Practice, University at Albany, State University of New York, Albany, USA.

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing, 45*, 31–45.

Michel, M., & O'Rourke, B. (2019). What drives alignment during text chat with a peer vs. a tutor? Insights from cued interviews and eye-tracking. *System, 83*, 50–63.

Michel, M. & Smith, B. (2014). *Alignment and Structural Convergence During SCMC: An Eye Tracking Study*. Paper presented at the American Association for Applied Linguistics conference, Portland, OR.

Michel, M., & Smith, B. (2017). Eye-tracking research in computer-mediated language learning. In S. Thorne, & S. May (Eds.), *Language, education and technology: Encyclopedia of language and education* (pp. 1–12). Cham, Switzerland: Springer International.

Michel, M., & Smith, B. (2018). Measuring lexical alignment during L2 chat interaction: An eye-tracking study. In S. M. Gass, P. Spinner, &J. Behney (Eds.), *Salience in second language acquisition* (pp. 244–268). New York: Routledge.

Michel, M., & Smith, B. (2019). Measuring lexical alignment during L2 chat interaction: An eye-tracking study. In S. Gass, P. Spinner, & J. Behney (2019) (Eds.), *Salience in second language acquisition* (pp. 244–268). Routledge.

Mifka-Profozic, N. (2013). *The effectiveness of corrective feedback and the role of individual differences in language learning*. Peter Lang.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General, 130*, 163–168.

Miyake, A., Emerson, M. J., & Friedman, N. P. (1999). Good interactions are hard to find. *Brain & Behavioral Sciences, 22*, 108–109.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J.-C. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, 115, 123–142.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Miyake, A., & Shah, P. (1999). Models of working memory. In *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 442–481). Cambridge, UK: Cambridge University Press.

Monteiro, K. (2014). An experimental study of corrective feedback during video conferencing. *Language Learning & Technology, 18*(3), 56–79. Retrieved from https://llt.msu.edu/issues/october2014/monteiro.pdf

Montero Perez, M. (2019). Pre-learning vocabulary before viewing captioned video: An eye-tracking study. *Language Learning Journal, 47*(4), 460–478. https://doi.org/10.1080/09571736.2019.1638623

Montero Perez, M. (2020a). Incidental vocabulary learning through viewing video: The role of vocabulary knowledge and working memory. *Studies in Second Language Acquisition, 42*(4), 749–773. http://doi.org/10.1017/S0272263119000706

Montero Perez, M. (2020b). Multimodal input in SLA research. *Studies in Second Language Acquisition*, *42*(3), 653–663. http://doi.org/10.1017/S0272263120000145

Montero Perez, M., Peters, E., & Desmet, P. (2013a). Is less more? Effectiveness and perceived usefulness of keyword and full captioned video for L2 listening comprehension. *ReCALL, 26*, 21–43.

Montero Perez, M., van den Noortgate, W., & Desmet, P. (2013b). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System, 41*, 720–739.

Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning, 31*(1–2), 1–6. https://doi:10.1080/09588221.2017.1375960

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94*, 156–163.

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

Muñoz, C., Pujadas, G., & Pattemore, A. (2023). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Language Research, 39*(1), 13–37.

Nakatsukasa, K. (2016). Efficacy of gestures and recasts on the acquisition of locative prepositions. *Studies in Second Language Acquisition, 38*(4), 771–799. https://www.jstor.org/stable/26330942

Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language learning*, *59*(2), 411–452.

Nassaji, H. (2015). *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. London: Bloomsbury. https://doi.org/10.5040/9781474219068

Nassaji, H. (2016). Anniversary article: Interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research, 20*, 535–562. https://doi.org/10.1177/1362168816644940

Nassaji, H. (2017). The effectiveness of extensive versus intensive recasts for learning L2 grammar. *The Modern Language Journal, 101*(2), 353–368.

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27*(1), 95–106.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learner. *Language Learning, 51,* 719–758

Örnberg Berglund, T. (2012). Corrective feedback and noticing in text-based second language interaction. In L. Bradley & S. Thouësny (Eds.), CALL: Using, Learning, Knowing, EUROCALL Conference, Gothenburg, Sweden, 22–25 August 2012, Proceedings (234–239). doi:10.14705/rpnet.2012.000058

Örnberg Berglund, T. (2013). Text-based chat and language learning. Opportunities and challenges. In A. Sundberg, C. Rosén, & P. Simfors. (Eds.), ASLA 2012 Proceedings (pp. 139–149). Uppsala: ASLA

O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer mediated communication and language learning. *Computer Assisted Language Learning, 21*(3), 227–251.

O'Rourke, B. (2012). Using eye tracking to investigate gaze behavioor in synchronous computer-mediated communication for language learning. In M. Dooly & R. O'Dowd (Eds.), *Researching online interaction and exchange in foreign language education: Methods and issues*, (pp. 305–341). Frankfurt: Peter Lang.

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 27–42). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.004

Paivio, A. (1983). The mind's eye in arts and science. *Poetics*, 12, 1–18.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

Paivio A. (1990). *Mental representations: A dual coding approach* (2nd ed.). Oxford, UK: Oxford University Press.

Paivio A. (2006). *Mind and its evolution: A dual coding theoretical approach.* Mahwah, NJ: Erlbaum.

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). New York: Psychology Press.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36*, 573–595.

Pattemore, A., & Muñoz, C. (2020). Learning L2 constructions from captioned audio-visual exposure: The effect of learner-related factors. *System, 93*, 102303.

Pattemore, A., & Muñoz, C. (2022). Captions and learnability factors in learning grammar from audio-visual input. *JALT CALL Journal*, *18*(1), 83–109.

Pawlak, M. (2017). Overview of learner individual differences and their mediating effects on the process and outcome of interaction. In L. Gurzynski-Weiss L. (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors* (pp. 19–40). Amsterdam: John Benjamins.

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.

Payne, J. S., & Ross, B. M. (2005). Synchronous CMC, working memory, and L2 oral proficiency development. *Language Learning and Technology, 9*(3), 35–54.

Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal, 20*(1), 7–32.

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition, 38*(1), 97–130. http://doi.org/10.1017/S0272263115000224

Pellicer-Sánchez, A., & Siyanova-Chanturia, A. (2018). Eye movements in vocabulary research: Current findings and future directions. ITL – *International Journal of Applied Linguistics, 169*(1), 5–29. https://doi.org/10.1075/itl.00004.pel

Pellicer-Sánchez, A., Tragant, E., Conklin, K., Rodgers, M., Serrano, R., & Llanes, Á. (2020). Young learners' processing of multimodal input and its impact on reading comprehension. *Studies in Second Language Acquisition, 42*(3), 577–598. https://doi.org/10.1017/S0272263120000091

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin, 144*, 48–76.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 67–86). Amsterdam: John Benjamins. https://doi.org/10.1075/swll.11.14per

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly, 53*(4), 1008–1032.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, 1–27. http://doi.org/10.1017/s0272263117000407

Petersen, K., & Sachs, R. (2016). The language classroom in the age of networked learning. In R. P. Leow, L. Cerezo & M. Baralt (Eds.). *A psycholinguistic approach to technology and language learning* (pp. 3–22). Berlin/Boston: De Gruyter Mouton.

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

Philp, J., & Iwashita, N. (2013). Talking, turning in and noticing: Exploring the benefits of output in task-based peer interaction. *Language Awareness, 22*(4), 353–370.

Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning, 33*(4), 465–497.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44*(3), 493–527. https://doi.org/10.1111/j.1467-1770.1994.tb01115.x.

Pica, T., Holliday, L., Lewis, N. E., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition, 11*, 63–90. http://dx.doi.org/10.1017/S027226310000783X

Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45–141). Adelaide: National Curriculum Resource Centre, AMEP.

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

Pienemann, M., & Mackey, A. (1993). An empirical study of children's ESL development and rapid profile. In P. McKay (Ed.), *ESL development: Language and literacy in schools* (pp. 115–259). Canberra: Commonwealth of Australia and National Languages and Literacy Institute of Australia.

Plonsky, L., & Brown, D. (2015). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research, 31*(2), 267–278. https://doi.org/10.1177/0267658314536436

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

Plotnikova, A. (2017). Intentional and incidental learning of English grammar through captioned video exposure [Unpublished master's thesis]. Barcelona, Spain: Universitat de Barcelona.

Prendergast, C. (2011). *A Study of negotiation routines in SCMC NS-NNS tandem sessions using eye-tracker gaze data*. [Unpublished final-year dissertation]. Dubin, Trinity College Dublin.

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin, 124*, 123–164.

Psychological Corporation, The. (1997). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: Author.

Quirk, R., & Greenbaum, S. (1990). *A student's grammar of the English language.* Harlow: Longman, Ltd.

R Core Team. (2022). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Ranta, L., & Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The Awareness-Practice-Feedback sequence. *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*, *12*, 141–160.

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): Recasts during mobile-mediated audio and video interactions: Learners' responses, their interpretation, and the development of English articles. *Computer Assisted Language Learning 35* (2), 114–140. https://doi.org/10.1080/09588221.2019.1671461

Rassaei, E. (2022). *The effects of recasts on L2 grammar: A meta-analysis*. *The Language Learning Journal*. https://doi.org/10.1080/09571736.2022.2097298.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506. http://doi.org/10.1080/17470210902816461

Reichle, E. D., Rayner, K., & Pollatsek, A. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement control during reading. *Cognitive Systems Research, 7,* 4–22.

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

Révész, A., Michel, M., & Lee, M. (2017). Investigating IELTS academic writing task 2: Relationship between cognitive writing processes, text quality, and working memory. *IELTS Research Reports Online Series*. https://www.ielts.org/en-us/for-researchers/research-reports/ielts_online_rr_2017-3

Révész, A., & Han, Z. (2006). Task content familiarity, task type and efficacy of recasts. *Language Awareness, 15*(3), 160–179.

Révész, A., Sachs, R., & Hama, M. (2014). The effects of task complexity and input frequency on the acquisition of the past counterfactual construction through recasts. *Language Learning, 64*(3), 615–650. https://doi.org/10.1111/lang.12061

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. *System, 10*(5), 133.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition, 35,* 213–235.

Roberts, M. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Tech. Rep. No. 9, pp. 162–

182). Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*, 283–331.

Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Malden, MA: Blackwell Publishing Ltd.

Robinson, P., Mackey, A., Gass, S. M., & Schmidt, R. (2011). Attention and awareness in second language acquisition. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247–267). Abingdon, UK: Routledge.

Rodgers, M. P. H., & Webb, S. (2017). The effects of captions on EFL learners' comprehension of English-language television programs. *CALICO Journal*, *34*(1), 20–38. http://doi.org/10.1558/cj.29522

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. *Synthesizing research on language learning and teaching*, *13*, 133–164. https://doi.org/10.1075/lllt.13.09val

Sachs, R. (2010). *Individual differences and the effectiveness of visual feedback on reflexive binding in L2 Japanese* [Unpublished doctoral dissertation]. Georgetown University, Washington, DC.

Sachs, R., & Suh, B. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 199–227). Oxford, UK: Oxford University Press.

Sagarra, N. (2007). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In Mackey, A. (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies*. Oxford, UK: Oxford University Press.

Sagarra, N. (2013). Working memory in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley-Blackwell. doi:10.1002/9781405198431.wbeall286.

Sagarra, N., & Abbuhl, 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

Sagarra, N., & Abbuhl, R. (2013b). Computer-delivered feedback and L2 development: The role of explicitness and working memory. In C. Sanz and B. Lado (Eds.), *Individual differences, l2 development, and language program administration: From theory to application* (pp. 53–70). Cengage Learning.

Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language Learning*, *67*(1), 43–74.

Saito, K. & R. Lyster (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning 62*(2), 595–633.

Salaberry, M. R. (2000). L2 morphosyntactic development in text-based computer-mediated communication. *Computer Assisted Language Learning*, *13*(1), 5–27.

Sanz, C., Lin, H. J., Lado, B., Stafford, C. A., & Bowden, H. W. (2016). One size fits all? Learning conditions and working memory capacity in *ab initio* language development. *Applied Linguistics, 37*(5), 669–692.

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition, 34*(4), 591–626.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, *13*(1), 96–120. Retrieved from https://llt. msu.edu/vol13num1/sauro.pdf

Saxton, M. (1997). The Contrast Theory of negative input. *Journal of Child Language, 24*(1), 139–161. https://doi.org/10.1017/S030500099600298X

Saxton, M. (2000). Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. *First Language*, *20*, 221–252.

Schachter, J. (1986). In search of systematicity in interlanguage production. *Studies in Second Language Acquisition, 8*, 119–134.

Schachter, J. (1986). Three approaches to the study of input. *Language Learning, 36*, 211–225. https://doi.org/10.1111/j.1467-1770.1986.tb00379.x

Schachter, J. (1991). Corrective feedback in historical perspective. *Second Language Research, 7*, 89–102.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11(2)*, 17–46.

Schmidt, R. (1993a). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206–226.

Schmidt, R. (1993b). Consciousness, learning, and interlanguage pragmatics. *Interlanguage Pragmatics, 21*(42), 1–31.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11–26.

Schmidt, R. (1995). Consciousness and foreign language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu, HI: University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.

Schmidt, R. (1998). The centrality of attention in SLA. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139524780

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524780.003

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. *Perspectives on individual characteristics and foreign language education*, *6*, 27.

Schmidt, R. W., & Frota, S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237–326). Rowley, MA: Newbury House.

Schoonen, R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 77–101). Multilingual Matters.

Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition, 38*(4), 607–646. https://doi.org/10.1017/S0272263115000327

Sharwood Smith, M. (1986). Comprehension versus acquisition: Two ways of processing input. *Applied Linguistics, 7,* 239–256.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA. *Studies in Second Language Acquisition, 15,* 165–179.

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8*(3), 263–300. https://doi.org/10.1191/1362168804lr146oa

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

Sheen, Y. (2007a). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301–322). Oxford, UK: Oxford University Press.

Sheen, Y. (2007b). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*(2), 255–283. https://doi.org/10.1002/j.1545-7249.2007.tb00059.x

Sheen, Y. (2008). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41,* 255–83.

Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, *32*(2), 203–234.

Sheen, Y. (2011). *Corrective feedback, individual differences and second language learning*. London: Springer.

Sheen, Y., & Ellis, R. (2011). Corrective feedback in language teaching. In E. Hickel (Ed.), *Handbook of research in second language teaching and learning* (pp. 593–610). Routledge.

Shi, L., & Stickler, U. (2021). Eyetracking a meeting of minds: Teachers' and students' joint attention during synchronous online language tutorials. *Journal of China Computer-Assisted Language Learning, 1*(1), 145–169. https://doi.org/10.1515/jccall-2021-2006.

Shi, L., Stickler, U., & Lloyd, M. E. (2017). The interplay between attention, experience, and skills in online language teaching. *Language Learning in Higher Education, 7*(1), 205–208. https://doi.org/10.1515/cercles-2017-0007.

Shin, J. (2020). A meta-analysis of the relationship between working memory and second language reading comprehension: Does task type matter? *Applied Psycholinguistics, 41*(4), 873–900. https://doi.org/10.1017/S0142716420000272

Shintani, N., & Aubrey, S. (2016). The effectiveness of synchronous and asynchronous written corrective feedback on grammatical accuracy in a computer-mediated environment. *The Modern Language Journal, 100*(1), 296–319. https://doi.org/10.1111/modl.12317

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguistic explanation on learners' explicit and implicit knowledge of the

English indefinite article. *Journal of Second Language Writing, 22*(3), 286–306. https://doi.org/10.1016/j.jslw.2013.03.011

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 17–40). Amsterdam: John Benjamins.

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

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. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In F. M. Hult (Ed.), *Directions and prospects for educational linguistics* (pp. 79–97). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-90-481-9136-9_6

Smith, B. (2012). Eye tracking as a measure of noticing: A study of explicit recasts in SCMC. *Language Learning & Technology, 16*(3), 53–81. http://dx.doi.org/10.10125/443

Smith, B., & Renaud, C. (2013). Using eye tracking as a measure of foreign language learners' noticing of recasts during computer-mediated writing conferences. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 147–166). Amsterdam: John Benjamins.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*, 1657–1661.

Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15*, 205.

Spada, N., & Lightbown, P. M. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *Modern Language Journal, 83*(1), 1–22. https://doi.org/10.1111/0026-7902.00002

Spada, N., & Lightbown, P.M. (2009). Interaction research in second/foreign language classrooms. In A. Mackey & P. Charlene (Eds.), *Multiple perspectives on interaction* (pp. 157–175). New York: Routledge.

Stickler, U., & Shi, L. (2015). Eye movements of online Chinese learners. *CALICO, 32*(1), 52–81. https://doi.org/10.1558/calico.v32i1.25964

Stickler, U., & Shi, L. (2017). Eyetracking methodology in SCMC: A tool for empowering learning and teaching. *ReCALL,* 1–18. https://doi.org/10.1017/S0958344017000040.

Sue, & Jiang. (2020). The effects of oral corrective feedback on Chinese English learners' acquisition of the simple past tense-moderation effects of the working memory. *Foreign Language Teaching*, (1), 76–81.

Swan, M. (2005). *Practical English usage*. Oxford, UK: Oxford University Press.

Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues* (pp. 327–366). Washington, DC: American Psychological Association. https://doi.org/10.1037/13273-012.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input and Second Language Acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*(2), 158–164. https://doi.org/10.3138/cmlr.50.1.158

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson* (pp. 125–144). Oxford, UK: Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom SLA* (pp. 64–81). New York: Cambridge University Press.

Swain, M. (2001). Integrating language and content teaching through collaborative tasks. *Canadian Modern Language Review, 58*(1), 44–63. https://doi.org/10.3138/cmlr.58.1.44

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, N.J: Erlbaum.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*(3), 320–337.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J. (1994). Cognitive Load Theory, learning difficulty and instructional design. *Learning and Instruction, 4*, 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

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, UK: Cambridge University Press.

Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning and Technology*, 14(2), 50–73.

Tabbers, H. K., Martens, R. L., & Van Merriënboer, J. J. G. (2004). Multimedia instructions and Cognitive Load Theory: Effects of modality and cueing. *British Journal of Educational Psychology*, *74*, 71–81. https://doi.org/10.1348/000709904322848824.

Tanenhaus, M. (2007). Eye movements and spoken language processing. In R.P.G. van Gompel, M.H. Fischer, W.S. Murray, & R.L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 444–469). Oxford, UK: Elsevier.

Teng, F. (2018). Incidental vocabulary learning for primary school students: The effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher, 46*,113–136. https://link.springer.com/article/10.1007%2Fs13384-018-0279-6

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

Tobii AB[Computer software]. (2014). Tobii Pro Lab (Version 1.194). Danderyd, Sweden: Tobii AB.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 183–203.

Tremblay, A., & Ransijn, J. (2020). *LMERConvenienceFunctions: Model selection and post-hoc analysis for (g) lmer models (R package version 3.0)* [Computer software]. https://cran.r-project.org/web/packages/LMER ConvenienceFunctions

Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed), *Conversational interaction in second language acquisition* (pp. 171–195). Oxford, UK: Oxford University Press.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Unsworth, N., Heitz, R., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498–505.

Van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2015). Learning new grammatical structures in Task-Based Language Learning: The effects of recasts and prompts. *The Modern Language Journal*, *99*(2), 246–262. https://doi.org/10.1111/modl.12211

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, *40*, 191–210.

Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *ELT Journal, 42*(4), 272–281. https://doi.org/10.1093/elt/42.4.272

Van Lommel, S., Laenen, A., & d'Ydewalle, G. (2006). Foreign-grammar acquisition while watching subtitled TV programmes. *British Journal of Educational Psychology, 76*, 243–258.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

VanPatten, B. (2004). Input and output in establishing form-meaning connections. In B. Van Patten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form-meaning connection in second language acquisition* (pp. 29–47). Mahwah, NJ: Erlbaum.

Varnosfadrani, A. D., & Basturkmen, H. (2009). The effectiveness of implicit and explicit error correction on learners' performance. *System*, *37*, 82–98.

Verbruggen, F., Logan G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods*, 40, 479–483.

Wacha, R. C., & Liu, Y. T. (2017). Testing the efficacy of two new variants of recasts with standard recasts in communicative conversational settings: An exploratory longitudinal study. *Language Teaching Research*, *21*(2), 189–216. https://doi.org/10.1177/136216881562426

Wang, A., & Pellicer-Sánchez, A. (2022). Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study. *Language Learning*, *72*(3), 765–805. https://doi.org/10.1111/lang.12495

Wang, Y. (2006). Negotiation of meaning in desktop videoconferencing-supported distance language learning. *ReCALL, 18*(1), 122–146.

Warschauer, M. (1996a). Comparing face-to-face and electronic communication in the second language classroom. *CALICO Journal, 13*(2), 7–26.

Warschauer, M. (1996b). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3–20). Tokyo: Logos.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The Modern Language Journal*, *81*(4), 470–481.

Webb, S., & Chang, A. C. S. (2020). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition, 44*(1), 35–56. https://doi.org/10.1017/S0272263120000297

Wen, Z. (2015). Working memory in second language acquisition and processing: The phonological/executive model. In Z. Wen, M. B. Mota, & A. McNeill, A. (Eds.), *Working memory in second language acquisition and processing*, 41–62. Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093595-012

Wen, Z. (2016). Working memory and second language learning: Towards an integrated approach. In Z. Wen, M. B. Mota, & A. McNeill, A. (Eds.), *Working memory in second language acquisition and processing.* Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093595-012

Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching, 50*(1), 1–31.

Wen, Z., Mota, M. B., & McNeill, A. (Eds.) (2015). *Working Memory in Second Language Acquisition and Processing.* Bristol: Multilingual Matters. https://doi.org/10.21832/9781783093595-012

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics, 8*(2), 95–110. https://doi.org/10.1093/applin/8.2.95

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: language and cognition, 6*(2), 129–141. https://doi.org/10.1017/s1366728903001081

White, M. J. (2021). Phonological working memory and non-verbal complex working memory as predictors of future English outcomes in young ELLs. *International Journal of Bilingualism, 25*(1), 318–337. https://doi.org/10.1177/13670069209481

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.). *Attention and performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2007). Attention to the second language. *International Review of Applied Linguistics in Language Teaching, 45*, 177–191. https://doi.org/10.1515/iral.2007.008

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159–177.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30*, 599–616.

Williams, J. N. (2012). Working memory and SLA. In S. Gass & A. Mackey (Eds.), *Handbook of Second Language Acquisition* (pp. 427-441). Oxford, UK: Routledge/Taylor Francis.

Winke, P. M. (2013). The Effects of Input Enhancement on Grammar Learning and Comprehension. *Studies in Second Language Acquisition, 35*(02), 323–352. https://doi.org/10.1017/s0272263112000903

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*, 65–86.

Winke, P., Godfroid, A., & Gass, S. (2013). Introduction to the special issue: Eye-movement recordings in second language research. *Studies in Second Language Acquisition, 35*(2), 205–212. https://doi.org/10.1017/S027226311200085X.

Woolfolk, A. E. (1998). *Educational psychology*. Boston: Allyn & Bacon.

Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition, 32*(2), 235–263.

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. (2013a). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics, 34*, 344–368. https://doi.org/10.1093/applin/ams044

Yilmaz, Y. (2013b). The relative effectiveness of mixed, explicit and implicit feedback. *System, 41*, 691–705.

Yilmaz, Y. (2016). The role of exposure condition in the effectiveness of explicit correction. *Studies in Second Language Acquisition, 38*(1), 65–96.

Yilmaz, Y., & Granena, G. (2016). The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback. *Bilingualism: Language and Cognition, 19*, 147–161. https://doi.org/10.1017/S136672891400090X

Yilmaz, Y., & Granena, G. (2019). Cognitive individual differences as predictors of improvement and awareness under implicit and explicit feedback conditions. *Modern Language Journal, 103*(3), 686–702. https://doi.org/10.1111/modl.12587

Yilmaz, Y., & Sagdic, A. (2019). The interaction between inhibitory control and corrective feedback timing. *International Journal of Applied Linguistics, 170*(2), 204–227. https://doi.org/10.1075/itl.19010.yil

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

Zalbidea, J. (2017). 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *The Modern Language Journal, 101*(2), 335-352. https://doi.org/10.1111/modl.12389

Zalbidea, J., & Sanz, C. (2020). Does learner cognition count on modality? Working memory and L2 morphosyntactic achievement across oral and written tasks. *Applied Psycholinguistics, 41*(5), 1171–1196. https://doi.org/10.1017/S0142716420000442

Zhao, Y., (2015). *The effects of explicit and implicit recasts on the acquisition of two grammatical structures and the mediating role of working memory* [Unpublished doctoral dissertation]. University of Auckland, New Zealand.

Zhao, Y., & Ellis, R. (2022). The relative effects of implicit and explicit corrective feedback on the acquisition of 3rd person *-s* by Chinese university students: A classroom-based study. *Language Teaching Research, 26*(3), 361–381. https://doi.org/10.1177/1362168820903343

Zhao, W., & Wang, J. J. (2016). The impact of recasts on the production of English questions. *Foreign Languages in China, 13*(3), 73–82. https://doi.org/10.13564/j.cnki.issn.1672-9382.2016.03.010

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition, 38*(3), 553–586. https://www.jstor.org/stable/26330932

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.

APPENDICES

List of Appendices

Appendix A1: Information Sheet for Study 1

Appendix A2: Consent Form for Study 1

Appendix A3: Developmental Stages in L2 English Question Formation with Data from the Present Study (stages adapted from Pienemann et al., 1988).

Appendix A4: A Sample of Experimental Treatment Tasks in Study 1

Appendix A5: A Sample of Oral Production Test (OPT) in Study 1

Appendix A6: A Sample of Written Production Test (WPT) in Study 1

Appendix A7: A Sample of Spontaneous Oral Production Test (SOPT) in Study 1

Appendix A8: Background Questionnaire for Study 1

Appendix A9: Exit Questionnaire for Study 1

Appendix A10. Results for the Ordinal Logistic Mixed-effects Models Examining Participants' Scores on the OPT

Appendix A11. Results for the Ordinal Logistic Mixed-effects Models Examining Participants' Scores on the WPT

Appendix A12. Results for the Ordinal Logistic Mixed−Effects Models Examining Participants' Scores on the SOPT

Appendix A13: Information Sheet for Study 2

Appendix A14: Consent Form for Study 2

Appendix A15: A Sample of Experimental Tasks in Study 2

Appendix A16: A Sample of OPT in Study 2

Appendix A17: A Sample of WPT in Study 2

Appendix A18: A Sample of SOPT in Study 2

Appendix A19: Background Questionnaire for Study 2

Appendix A20: Exit Questionnaire for Study 2

Appendix A21: Results for the Linear Mixed-effects Model Examining Performance on the OPT – Present Progressive

Appendix A22. Results for the Logistic Mixed-effects Models Examining Performance on the OPT – Third Person Singular

Appendix A23. Results for the Linear Mixed-effects Model Examining Performance on the WPT – Present Progressive

Appendix A24. Results for the Logistic Mixed-effects Models Examining Performance on the WPT – Third Person Singular

Appendix A25. Results for the Linear Mixed-effects Model Examining Performance on the SOPT – Present Progressive

Appendix A26. Results for the Generalized Linear Mixed-effects Models Examining Performance on the SOPT – Third Person Singular

Appendix A27. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (NWS) – Present Progressive

Appendix A28. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (AOSPAN) – Present Progressive

Appendix A29. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (AOSPAN) – Present Progressive

Appendix A30. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (AOSPAN) – Present Progressive

Appendix A31. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Stop Signal) – Third Person Singular

Appendix A32. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (NWS) – Present Progressive

Appendix A33. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (NWS) – Present Progressive

Appendix A34. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Forward Corsi Block) – Present Progressive

Appendix A35. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Forward Corsi Block) – Present Progressive

Appendix A36. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Forward Corsi Block) – Present Progressive

Appendix A37. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Color Shape Task) – Present Progressive

Appendix A38. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Color Shape Task) – Present Progressive

Appendix A39. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Color Shape Task) – Present Progressive

Appendix A40. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Stop Signal Task) – Present Progressive

Appendix A41. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Stop Signal Task) – Present Progressive

Appendix A42. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Stop Signal Task) – Present Progressive

Appendix A43. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (NWS) – Third Person Singular

Appendix A44. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (NWS) – Third Person Singular

Appendix A45. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (NWS) – Third Person Singular

Appendix A46. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Forward Corsi Block) – Third Person Singular

Appendix A47. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Forward Corsi Block) – Third Person Singular

Appendix A48. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Forward Corsi Block) – Third Person Singular

Appendix A49. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Color Shape) – Third Person Singular
Appendix A50. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Color Shape) – Third Person Singular
Appendix A51. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Color Shape) – Third Person Singular
Appendix A52. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Stop Signal) – Third Person Singular
Appendix A53. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Stop Signal) – Third Person Singular
Appendix A54. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (AOSPAN) – Third Person Singular
Appendix A55. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (AOSPAN) – Third Person Singular
Appendix A56. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (AOSPAN) – Third Person Singular

Appendix A1: Information Sheet for Study 1

Institute of Education

UCL

Information Sheet for Korean ESL Learners

- **Title of Study:** An investigation of the communication and interactional patterns of Korean ESL learners in video-conferencing tasks
- **Department:** Communication, Culture and Media
- **Name and contact details of the researcher:** Yeonwoo Jung [redacted]
- **Supervisor:** Professor Andrea Révész

My name is Yeonwoo Jung, and I would like to invite you to participate in a PhD research project. Before you decide, it is important for you to understand the purpose of the research and the involvement required. Please carefully read the following information and feel free to seek clarification or request additional details.

Who is carrying out the research?

I am a doctoral student in the Department of Communication, Culture and Media at the UCL Institute of Education. I am interested in second language learning, teaching, and assessment, and this study is part of my PhD research.

Why are we doing this research?

The aim of the present study is to examine the online communication and interactional patterns of Korean ESL (English as a Second Language) learners' during one-on-one multimodal video-conferencing tasks, conducted via synchronous computer-mediated communication (SCMC) platforms, such as Zoom. The findings of the research will contribute to the understanding of the effectiveness of video-conferencing in improving English learners' online communication competence and provide suggestions to facilitate Korean ESL learners' English communication skills.

Why am I being invited to take part?

You are being invited to take part as this study is looking to involve those who are learners of English as a second language.

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Institute of Education

What will happen if I choose to take part?

If you choose to participate, you will engage in a series of five individual sessions.

At the outset of the experiment, you will be requested to fill out a brief questionnaire about yourself and complete a general English proficiency test, along with a set of pretests (oral production, written production, and spontaneous oral production tests). Approximately one and a half hours will be designated for the completion the questionnaire and the pretests.

During the intervention phase, you will be involved in three individual sessions. In each session, you will take part in a communicative video-conferencing task, with your performance recorded. These sessions will last approximately 15–20 minutes each.

Upon completion of the sessions, you will be required to undertake a posttest, lasting approximately 1 hour, and your test performance will be recorded. After one week, you will be asked to participate in a delayed posttest, and your performance will also be recorded. Following the delayed posttest, you will be administered a brief post-task questionnaire, expected to take around 10 minutes. The entire experiment is anticipated to span approximately 4 hours. Completing all activities is crucial for achieving the study's objectives, as failure to do so would impede the utilization of the collected data.

Will anyone know I have been involved?

The recordings of your performance will exclusively be used for analysis and conference presentations. No other use will be made without obtaining your written permission, and access to the original recordings will be restricted to project personnel only. These recordings will undergo transcription, with stringent measures to ensure your name or any personal details are not included. Any data acquired from you will be securely stored, maintaining confidentiality throughout every phase of the project and beyond. To preserve anonymity, a unique identifier will be employed to anonymize your identity and data. The digital format of the data will be securely maintained and subjected to the anonymization process. Confidentiality will be upheld unless there exist compelling and legitimate reasons necessitating a breach. In such instances, we will promptly inform you of any decisions that might impact your confidentiality.

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Could there be problems for me if I take part?

Engaging in this study is not expected to cause you any disadvantages or discomfort. You have the right to withdraw from the study at any time, without providing a reason and without any impact on you. If you decide to withdraw, all data collected from you will be promptly deleted.

What will happen to the results of the research?

The results might be presented at professional conferences and in published in research journals. We may archive the dataset in a digital depository for open access. Upon request, a summary of the study's overall results will be provided at the completion of the research. Your identity will not be disclosed in any report or publication.

Do I have to take part?

The decision to participate in entirely yours. We hope that if you choose to join, you will find it to be a valuable experience.

Contact for further information

Research designs often necessitate that the complete intent of the study is not disclosed before participation. While we have described the general nature of the tasks you will perform, the full intent of the study will only be revealed to you upon the completion of the study.

If you have any further questions before deciding to take part, you can reach me at [redacted], [redacted] or Professor Andrea Révész at [redacted]

Data Protection Privacy Notice

The data controller for this project will be University College London (UCL). The UCL Data Protection Office provides oversight of UCL activities involving the processing of personal data, and can be contacted at data-protection@ucl.ac.uk. UCL's Data Protection Officer can also be contacted at data-protection@ucl.ac.uk.

Further information on how UCL uses participant information can be found here: www.ucl.ac.uk/legal-services/privacy/participants-health-and-care-research-privacy-notice

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Institute of Education

If you are concerned about how your personal data is being processed, or if you would like to contact us about your rights, please contact UCL in the first instance at data protection@ucl.ac.uk.

Thank you very much for taking the time to read this information sheet and for your consideration to participate!

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Appendix A2. Consent Form for Study 1

Institute of Education

UCL

Consent Form

An investigation of the communication and interactional patterns of Korean ESL learners in video-conferencing tasks

If you are happy to participate in this study, please complete this consent form and return to Yeonwoo Jung in person or at the address below.

	Yes	No
I have read and understood the information leaflet about the research.	☐	☐
I agree to the arrangements for my participation as described in the Information Sheet.	☐	☐
I agree to be filmed during the treatment sessions.	☐	☐
I understand that if any of my words are used in reports or presentations they will not be attributed to me.	☐	☐
I understand that I can withdraw from the project at any time, and that if I choose to do this, any data I have contributed will not be used.	☐	☐
I understand that I can contact Yeonwoo Jung at any time and request for my data to be removed from the project database.	☐	☐

Name_____Signed _____

Date _____

Yeonwoo Jung
UCL Institute of Education
20 Bedford Way London WC1H 0AL
[redacted]

Yeonwoo Jung
Department of Communication, Culture and Media
UCL Institute of Education
20 Bedford Way London WC1H 0AL

:

Appendix A3: Developmental Stages in 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	*Wh*-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	*Yes/no* questions with auxiliary or copula	Is the lab recorded in real time? Have you experienced a similar incident?
	Yes/no questions with modal	Can you recall the situation?
	Wh- questions with copula (not auxiliary)	Where is your head office? *What are your position?
Stage 5: Inversion	Auxiliary (e.g., *is*) in 2nd position	*What are you discover at the scene? Where were you trained?
	Do operator (e.g., *does/do*) in 2nd position	When did you discover the scene? Who do you work for?
	Modal (e.g., may) in 2nd position	Who will help you?
Stage 6: Complex question	Negative questions with *Do* 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 A4: A Sample of Experimental Treatment Tasks in Study 1







The company recently succeeded in creating dinosaurs using genetic engineering. However, some key dinosaur eggs were stolen!



The FBI said it is actively investigating the incident.





Welcome!

You are an FBI agent

investigating this incident.





As you may know,

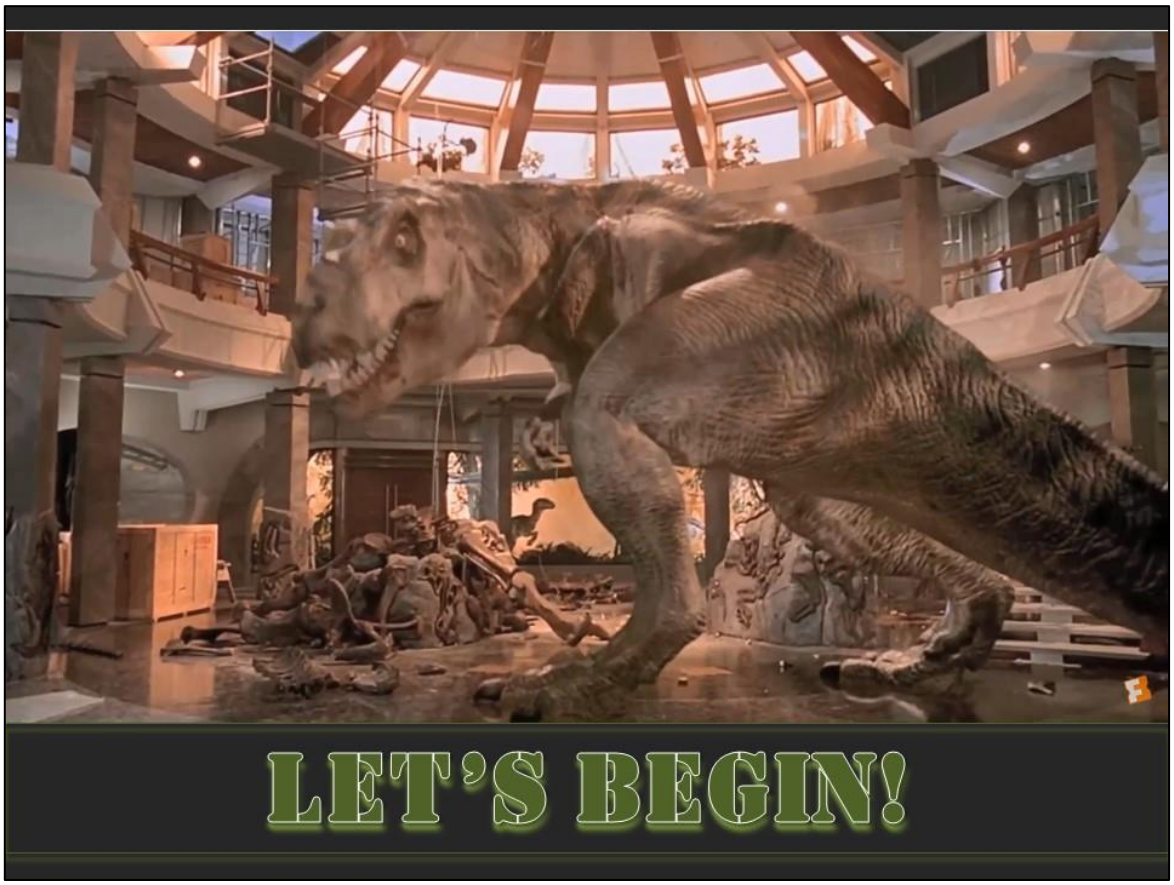
the key dinosaur eggs

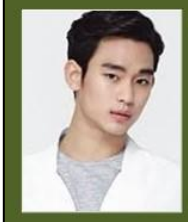
were stolen yesterday!


Here,
we have THREE witnesses.
Please INTERVIEW them
and obtain as much
information as possible!
FEDERAL
FBI
BUREAU of INVESTIGATION
DEPARTMENT of INVESTIGATION
FBI
SPECIAL AGENT


You are an FBI agent!
Your Role
1. First, INTERVIEW three witnesses, and
2. FILL IN a "Crime Report Form" with the information obtained.
3. Then, SHARE your report orally with your colleague, Yeonwoo.
FBI
CRIME REPORT FORM
1. What?
2. How many?
3. When?
4. What weapons?
5. Who?
6. How?
7. Storage Place?
8. Victim?
9. Why?


WHAT? HOW? WHEN?
WHO? WHERE? WHO?
WHEN? WHY? WHAT? HOW?
WHY? WHERE? HOW? WHAT? WHO? WHERE?
WHO? WHERE? WHO? WHY? HOW? HOW? WHERE?
WHAT? WHY? WHAT?
WHO? WHERE? WHEN?
WHERE? WHO?
WHY? HOW? WHAT?
WHO? WHERE?
WHY? WHAT? HOW?
HOW? WHO? WHEN?
WHAT? WHERE?
WHY? WHEN? WHAT?
WHERE? WHO? WHY? HOW?
HOW? WHERE?
WHAT? WHY?
WHEN? WHO? WHAT?
WHERE? WHAT? HOW?
WHO? WHY? WHERE?
WHAT? WHEN?
WHERE? WHO? WHAT?
WHERE? WHY? HOW? WHEN? WHAT? WHO? WHEN? WHY? WHERE? WHEN? HOW? HOW? WHERE? WHO? WHAT? WHY? WHAT? WHEN? WHERE? HOW? WHAT? WHEN? WHO? WHERE?


LET'S BEGIN!



Directions:

No. 1
Mr. Kim

Ask Mr. Kim the questions below.

FBI

1. What / your position?
2. How many thieves?
3. When / incident / occur?



Directions:

No. 2
Mr. Leo

Ask Mr. Leo the questions below.

FBI

4. What weapon / they / use?
5. Who / report / crime?
6. How / hide / their faces?



Directions:

No. 3
Ms. May

Ask Ms. May the questions below.

FBI

7. Where / you / keep / eggs?
8. Who / they / attack?
9. Why / shoot / him?

Well done! Now,

please **SHARE your report orally**

with your colleague, Yeonwoo.

FBI

CRIME REPORT FORM

1. What?	2. How many?	3. When?
4. What weapons?	5. Who?	6. How?
7. Storage Place?	8. Victim?	9. Why?

FBI

CRIME REPORT FORM

1. What?	2. How many?	3. When?
 Three dinosaur eggs	 Two thieves	 11 p.m.
4. What weapons?	**5. Who?**	**6. How?**
 Guns	 Security guard	 Masks
7. Storage Place?	**8. Victim?**	**9. Why?**
 Lab cabinet	 Senior Manager	 He tried to take off their masks.

Great!

Thanks to your hard work, we successfully completed the crime report form!



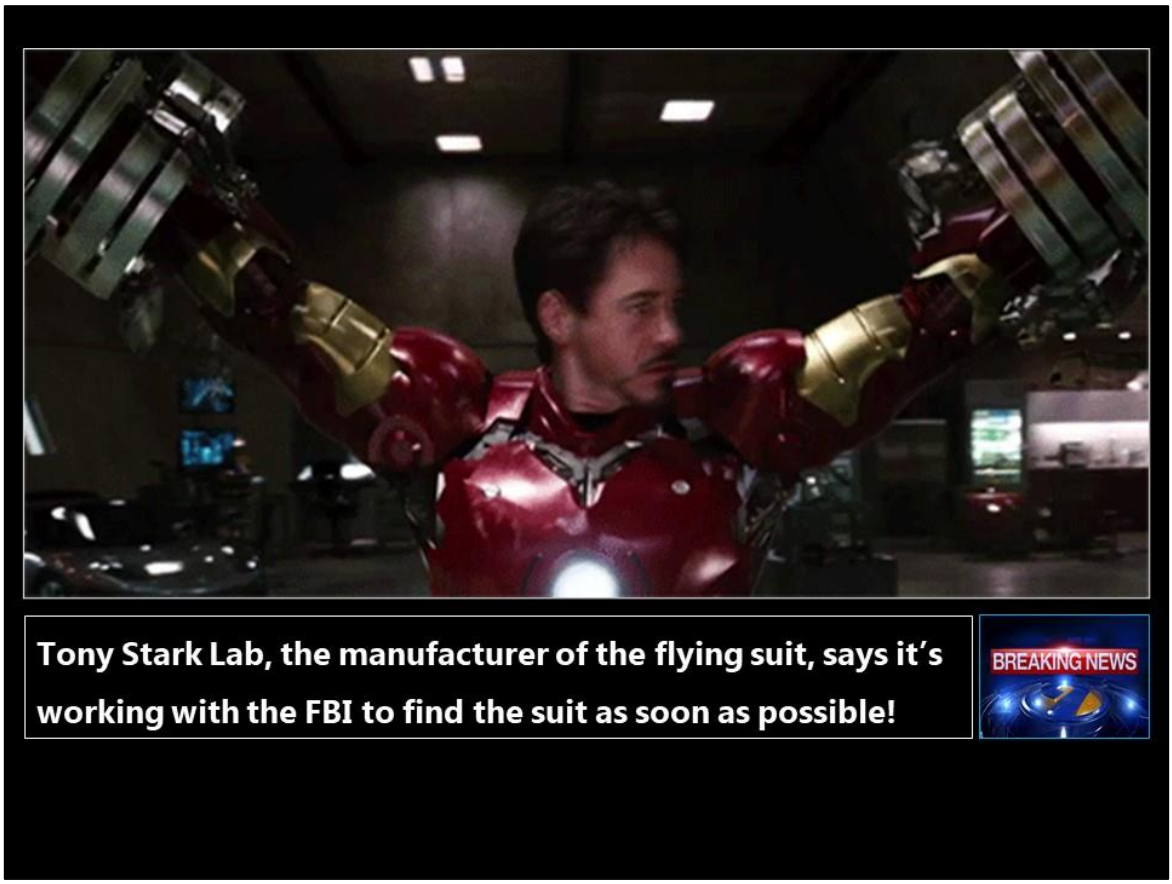
Thank you!
I will see you next time!

Appendix A5: A Sample of Oral Production Test (OPT) in Study 1



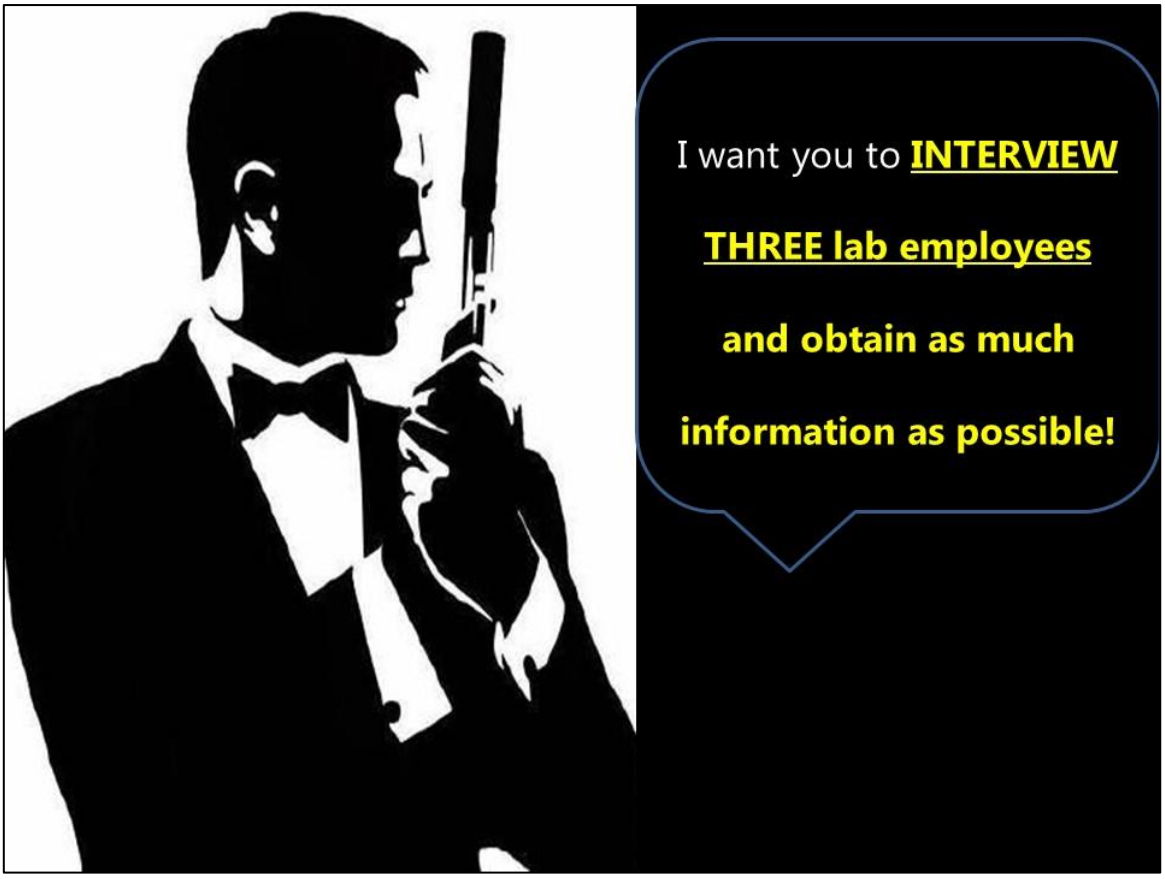



The flying suit was produced as part of the government's space project.
BREAKING NEWS


Tony Stark Lab, the manufacturer of the flying suit, says it's working with the FBI to find the suit as soon as possible!
BREAKING NEWS


Welcome!
You are an FBI agent
investigating this incident.


As you may know,
a flying suit called 'Iron
Man' was stolen recently!



Your Role

You are an FBI agent.

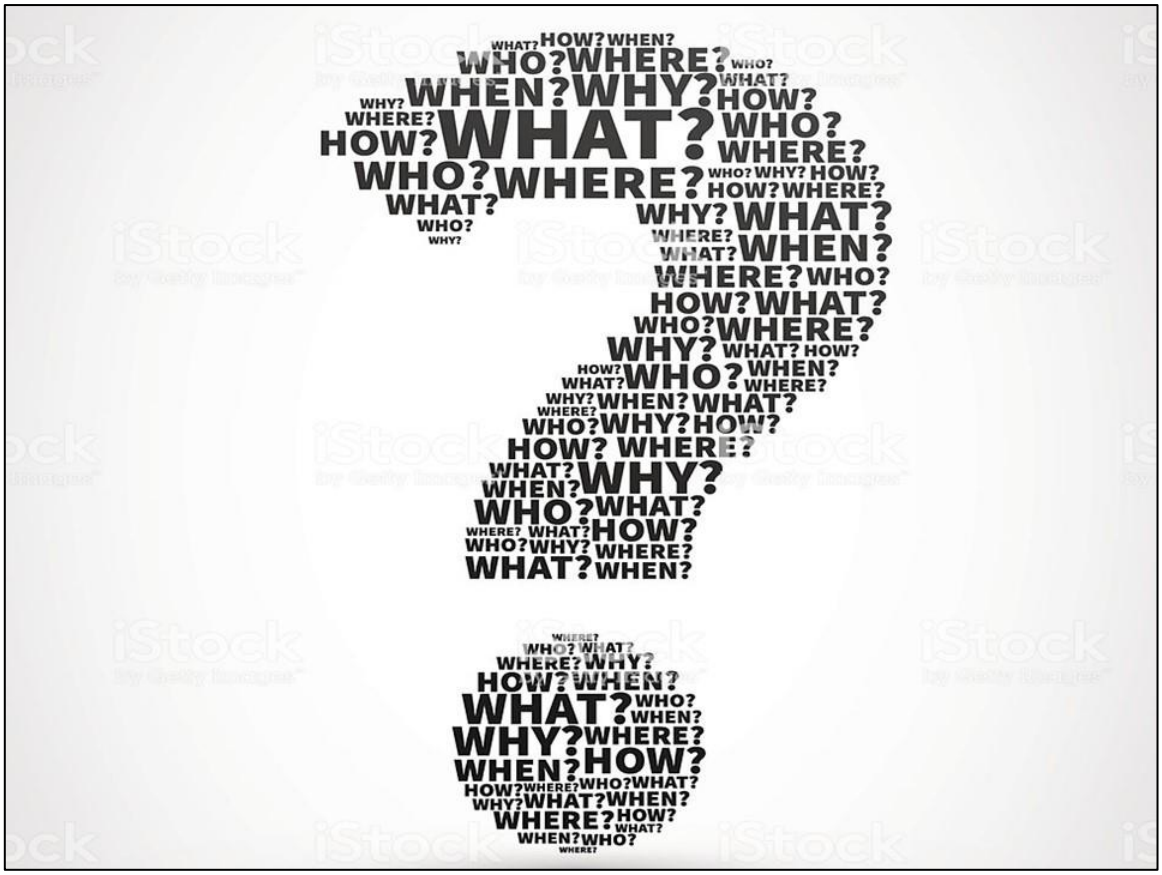
You need to

interview THREE lab

employees and

gather as much information

as possible!


WHAT? HOW? WHEN?
WHO? WHERE? WHO?
WHEN? WHY? WHAT?
WHY? HOW?
WHERE? WHAT? WHO?
HOW? WHERE?
WHO? WHERE? WHO? WHY? HOW?
WHAT? HOW? WHERE?
WHO? WHY? WHAT?
WHERE? WHEN?
WHAT? WHERE? WHO?
HOW? WHAT?
WHO? WHERE?
WHY? WHAT? HOW?
HOW? WHO? WHEN?
WHAT? WHERE?
WHY? WHEN? WHAT?
WHERE?
WHO? WHY? HOW?
HOW? WHERE?
WHAT? WHY?
WHEN?
WHO? WHAT?
WHERE? WHAT? HOW?
WHO? WHY? WHERE?
WHAT? WHEN?
WHERE?
WHO? WHAT?
WHERE? WHY?
HOW? WHEN?
WHAT? WHO?
WHEN?
WHY? WHERE?
WHEN? HOW?
HOW? WHERE? WHO? WHAT?
WHY? WHAT? WHEN?
WHERE? HOW?
WHAT?
WHEN? WHO?
WHERE?


ARE YOU READY?





No. 1

Mr. Thor

Directions: Ask Mr. Thor the questions below.

1. What / your position?

2. How long / thief / stay?

3. When / leave / lab / yesterday?

4. Can / tell / working hours?

5. Everyone / have / access to the lab?



No. 1
Mr. Thor

Directions: Ask Mr. Thor the questions below.

6. What / you / doing / at that time?

7. How / thief / break into / lab?

8. Report / the incident / to the police?

9. Change / password / regularly?

10. Who / call / the FBI / first?



No. 2
Mr. Jackson

Directions: Ask Mr. Jackson the questions below.

1. When / discover / crime scene?

2. Where / you / keep / suit?

3. What / cause / incident?

4. Why / lab / be left open?

5. How / steal / suit?



No. 2
Mr. Jackson

Directions: Ask Mr. Jackson the questions below.

6. Who / in charge of security?

7. All visitors / can / enter / lab?

8. Have experienced / similar incident?

9. Who / report / security issue to?

10. Security guard / on duty / at the time?



No. 3
Ms. Pepper

Directions: Ask Ms. Pepper the questions below.

1. Who / you / make / suit for?

2. Why / your team / be chosen for this project?

3. How many engineers / in your lab?

4. You / be trained / on security rules?

5. Have seen / anyone suspicious?



No. 3

Ms. Pepper

Directions: Ask Ms. Pepper the questions below.

6. Can / monitor / lab / when it is empty?

7. Lab / be recorded / in real time?

8. Where / thief / hide / suit?

9. What / see / at the scene?

10. Anyone / have entered / lab illegally?

Well done!

Your interview was very successful!



GOOD BYE

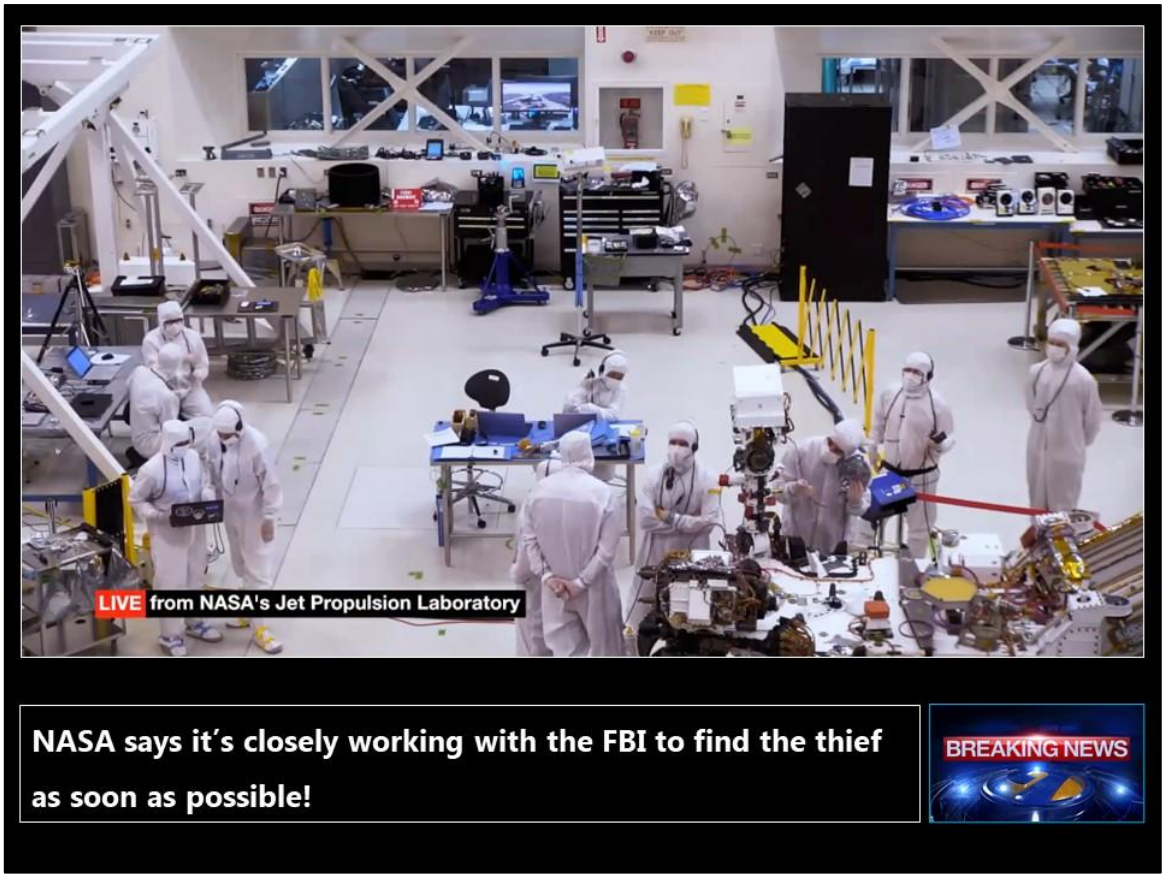
Thank you!

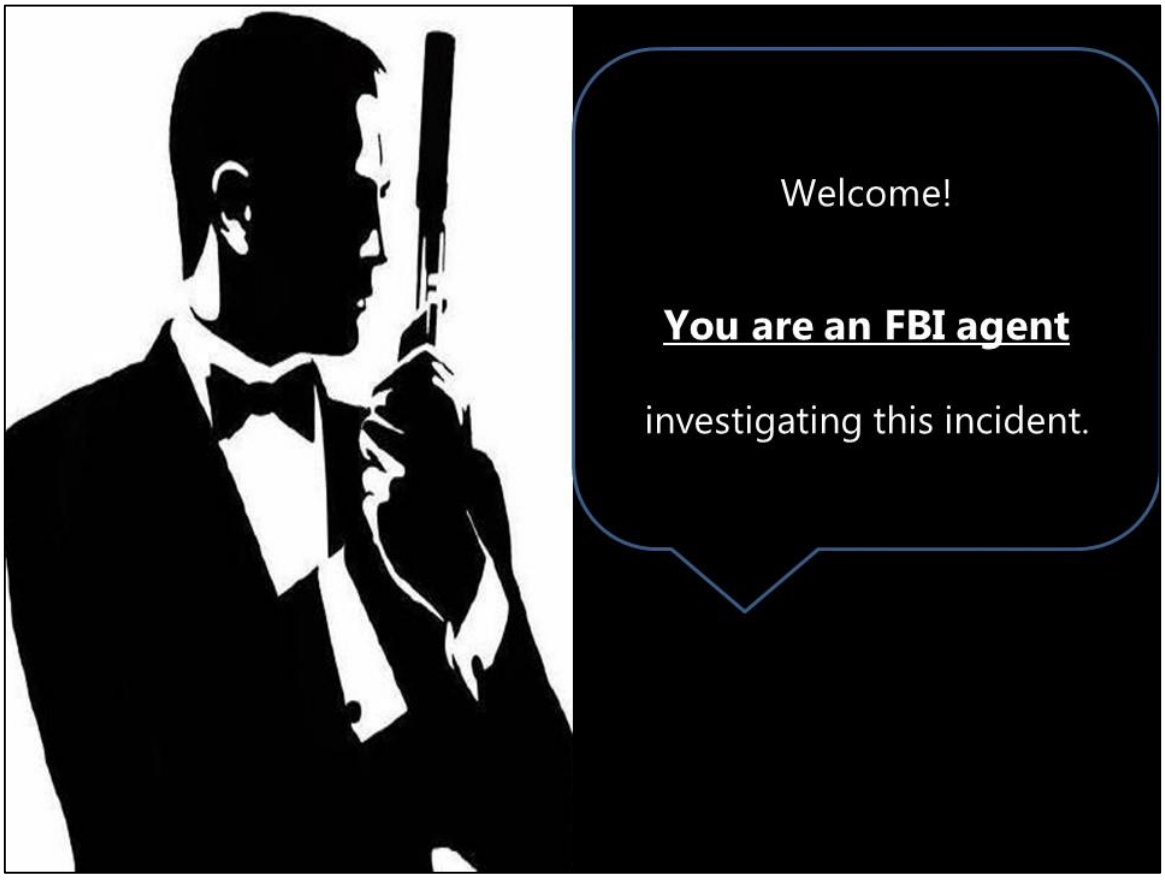
I will see you next time!

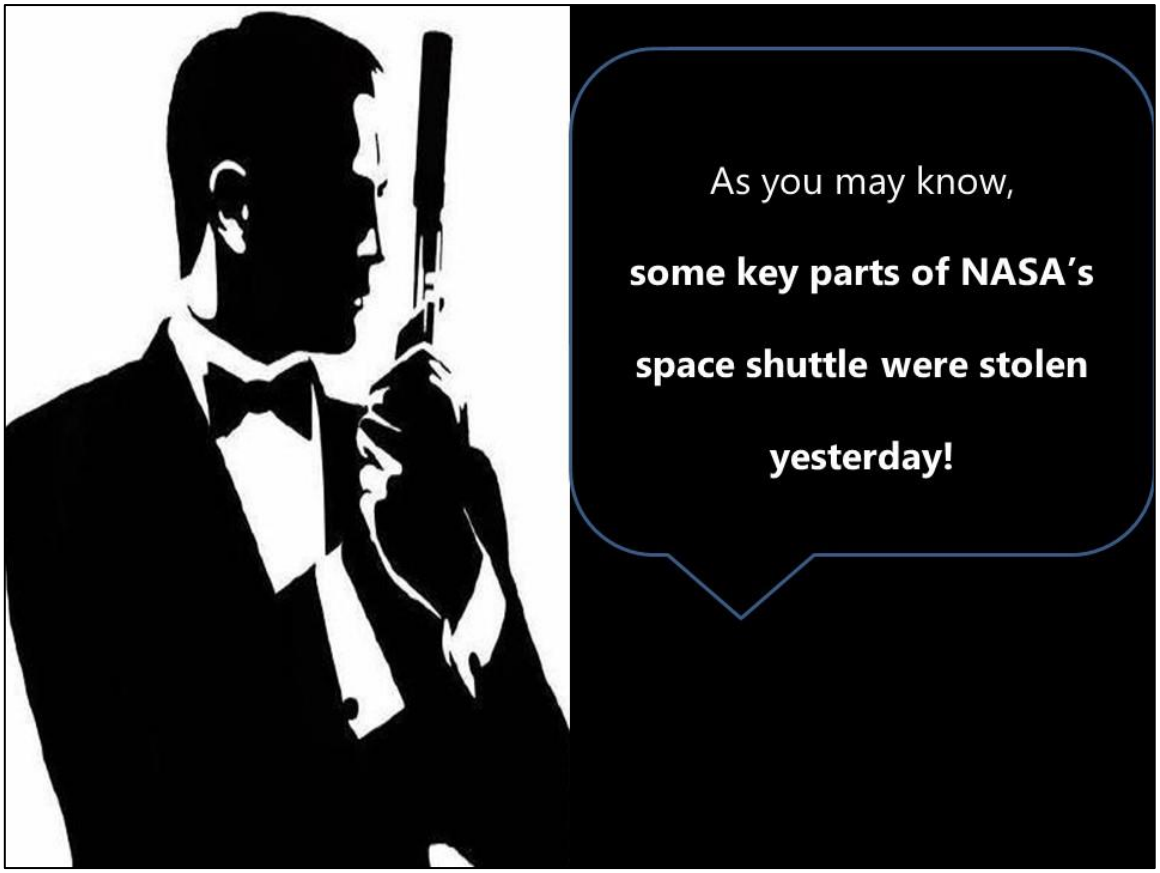
Appendix A6: A Sample of Written Production Test (WPT) in Study 1

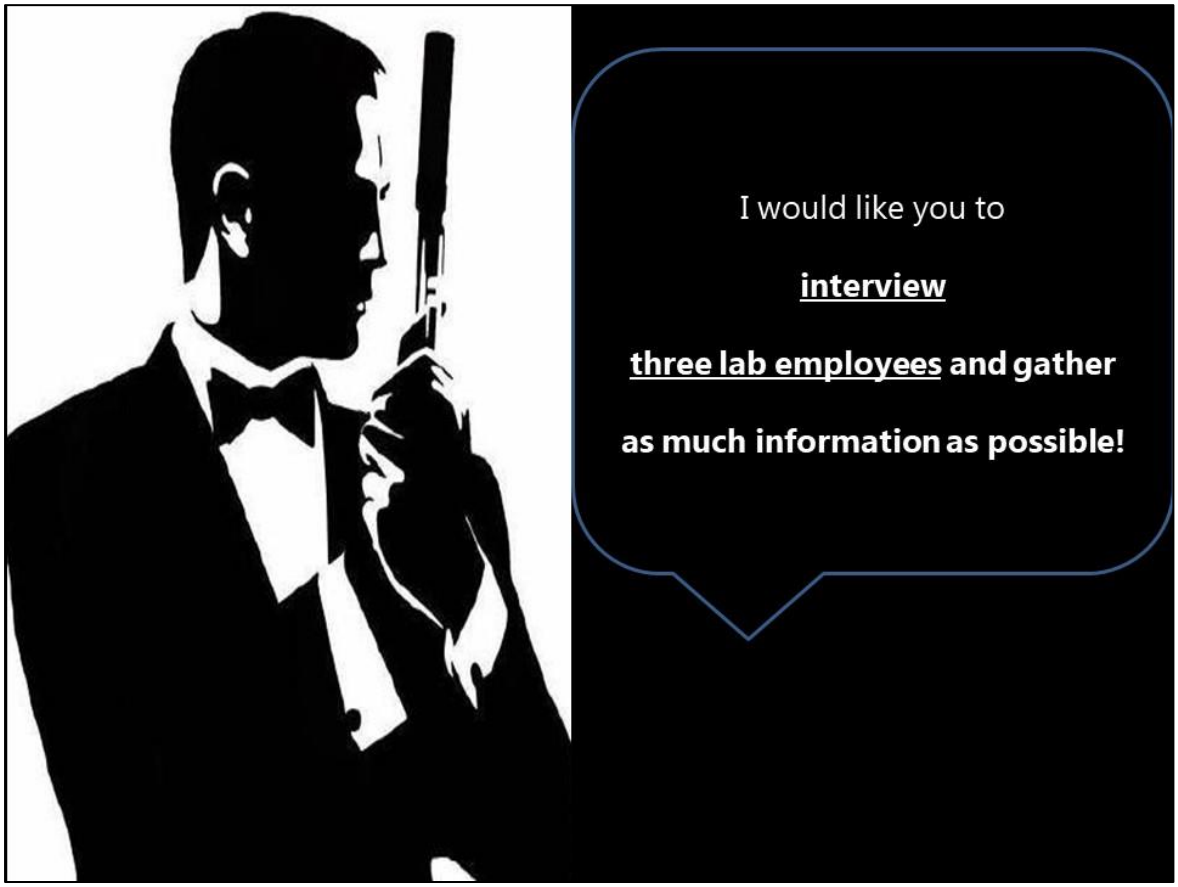





LIVE from NASA's Jet Propulsion Laboratory
NASA says it's closely working with the FBI to find the thief as soon as possible!
BREAKING NEWS


Welcome!
You are an FBI agent
investigating this incident.

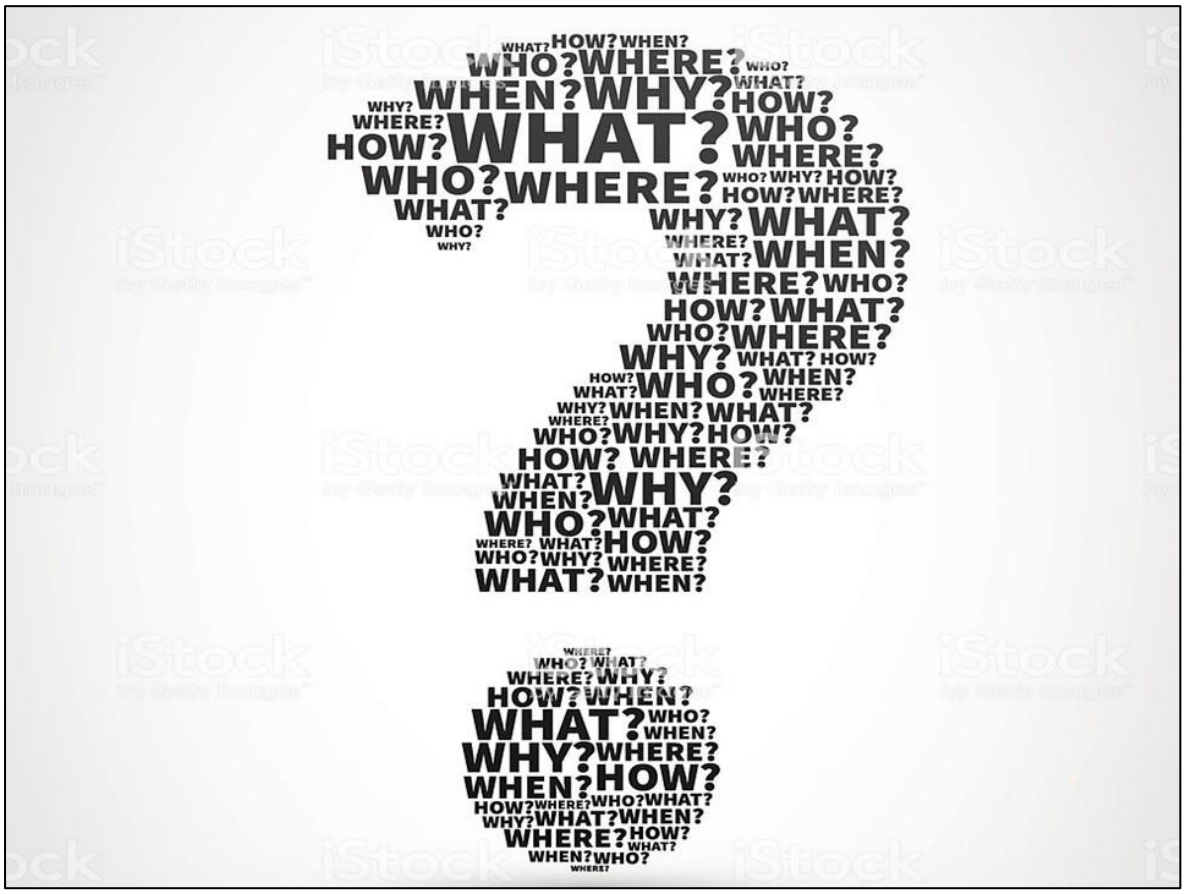

As you may know,
some key parts of NASA's space shuttle were stolen yesterday!


I would like you to
interview
three lab employees and gather as much information as possible!

Your Role

You are an FBI agent.

You need to interview THREE lab employees and gather as much information as possible!







No. 1

Mr. Lim



Directions: TYPE the following questions in the chat box.

1. What / your position?
2. How / incident / occur?
3. When / discover / crime scene?
4. See / anything unusual?
5. Who / report / incident?
6. Experiencing / financial difficulties?
7. Can / bring / anyone / to the lab?



No. 2
Mr. Kirk

Directions: TYPE the following questions in the chat box.

2-1. What / doing / at the time of the incident?
2-2. Where / watch / the movie?
2-3. Can / prove / your alibi?
2-4. Have brought / someone / to the lab?
2-5. How often / work / overtime?
2-6. Report / this / to the government?
2-7. The lab / be open to / all employees?



No. 3
Mr. Gosling

Directions: TYPE the following questions in the chat box.

3-1. Have lost / any key components / before?
3-2. Why / lose / them?
3-3. You / be trained / on security rules?
3-4. Have talked to anyone / about the project?
3-5. Who / you / work with / most closely?
3-6. Have / any questions?
3-7. Anyone / can / access / lab?

Well done!

Your interview was very successful!



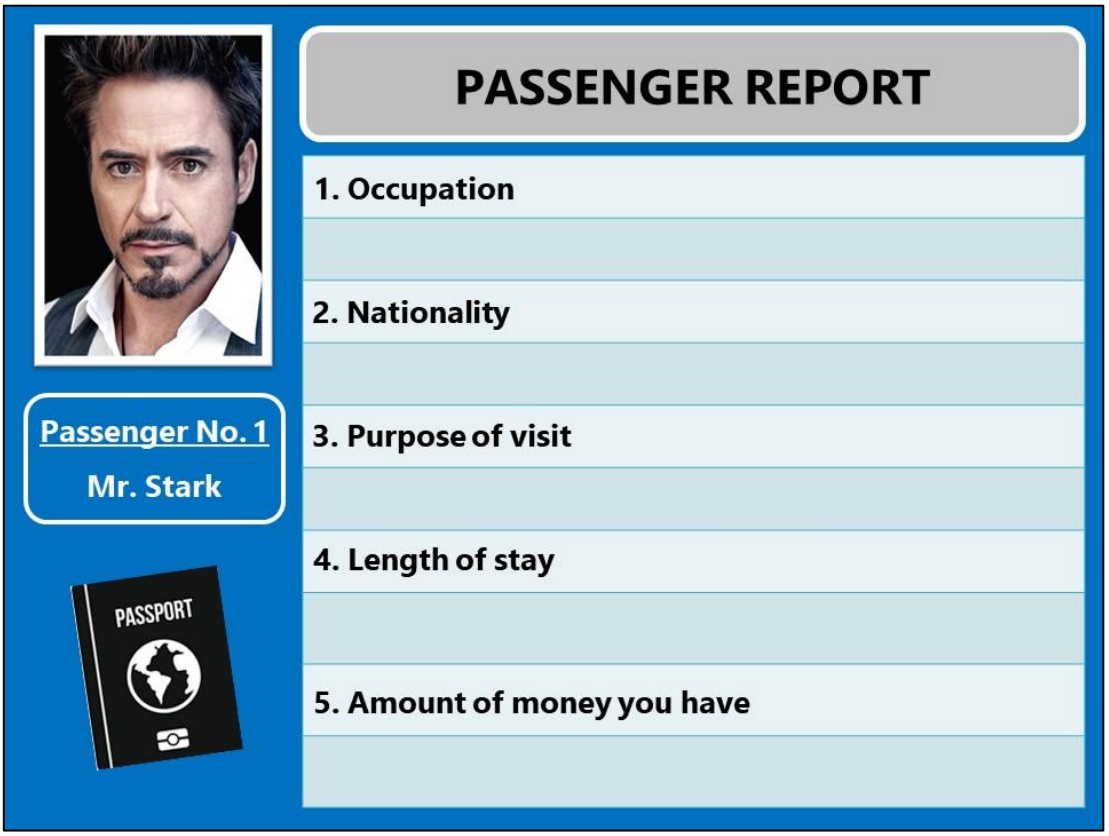
GOOD BYE

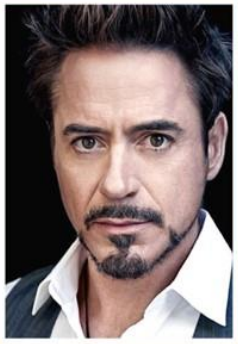
Thank you!

I will see you next time!

Appendix A7: A Sample of Spontaneous Oral Production Test (SOPT) in Study 1







Passenger No. 1



PASSENGER REPORT

6. Purpose of carrying a large amount of cash

7. Need an interpreter?

8. Country of last visit

9. Your intended address in the UK

10. Anything to declare



Passenger No. 2



PASSENGER REPORT

1. Full name

2. Date of birth

3. Any restricted items

4. Explain the purpose of the hammer

5. Place of last departure



Passenger No. 2



PASSENGER REPORT

6. Number of visits to the UK
7. Importing any vegetables or meat into the UK?
8. Return ticket
9. Travelling alone or with other people
10. Any family members in the UK

THANK YOU FOR YOUR COOPERATION!

Appendix A8: Background Questionnaire for Study 1

설 문 지 (Questionnaire)

참가자 번호: _____

귀하께서는 각 항목별로 어느 부분에 해당이 되는지 **각 해당란에 체크**해 주시기 바랍니다.

1. 귀하의 성별은 무엇입니까?	**남성**	**여성**	**기타**		
	☐	☐	☐		
2. 귀하의 연령은 어떻게 되십니까?	**10 대**	**20 대**	**30 대**	**40 대**	**50 대**
	☐	☐	☐	☐	☐
3. 표준 영어 능력 시험 점수가 있습니까??	**토익**	**토플**	**아이엘츠**	**기타**	
	☐	☐	☐	☐	
4. 영어를 언제 배우기 시작하셨습니까?	**10 대**	**20 대**	**30 대**	**40 대**	**50 대**
	☐	☐	☐	☐	☐
5. 총 영어 교육 기간을 적어주세요.	_____년_____ 개월				
6. 영어권 국가에서 생활한 경험이 있나요?	**예**	**아니요**			
	☐	☐			
7. 있다면 국가 이름과 체류 기간 및 영어권 국가에서의 교육 기간을 적어주세요.					
● 국가: ① ② ③	● 체류기간/교육 기간 ① ② ③				

8. 1(초급)부터 5(고급)까지의 척도에서 각각의 영어 능력을 평가해보세요					
	1	**2**	**3**	**4**	**5**
① 말하기	☐	☐	☐	☐	☐
② 듣기	☐	☐	☐	☐	☐
③ 읽기	☐	☐	☐	☐	☐
④ 쓰기	☐	☐	☐	☐	☐
9. 한국어와 영어 외에 사용하는 언어가 있나요?	**예**	**아니요**			
	☐	☐			
10. 있다면 사용하는 언어를 적어 주세요. 그리고 각 언어에 대해 본인의 레벨을 체크해 주세요.					
	초급	**중급**	**고급**		
① 언어:	☐	☐	☐		
② 언어:	☐	☐	☐		
③ 언어:	☐	☐	☐		

11. 하루에 소셜 네트워킹 사이트는 (SNS) 얼마나 사용하시나요?

① 1 시간 미만

② 1–2 시간

③ 2–3 시간

④ 4 시간 이상

12. 어떤 SNS 애플리케이션이나 플랫폼을 선호하시나요? (예: Facebook, Instagram, Twitter)

① Facebook

② Instagram

③ Twitter

④ 카카오톡

⑤ 기타:

13. 어떠한 유형의 SNS 커뮤니케이션을 선호하시나요?

① 텍스트 기반 (예: 문자, 이메일, 실시간 문자 채팅)

② 음성 기반 (예: 전화)

③ 비디오 기반 (예: 화상 채팅)

④ 혼합 형태 (텍스트, 음성, 비디오)

⑤ 기타:

14. **실시간 커뮤니케이션 플랫폼 중에 어떤 것을 선호하시나요?**

① Zoom

② Microsoft Teams

③ Skype

④ 기타_____

		매우 편함	편함	중립	불편함	매우 불편함
15.	SNS 나 인터넷을 사용하는 것에 대해 얼마나 편안한가요?	☐	☐	☐	☐	☐
16.	하루에 인터넷은 얼마나 사용하시나요?	1 시간 미만	2 시간	3 시간	4 시간 이상	기타
		☐	☐	☐	☐	☐
17.	언어 습득을 위해 언어 학습 앱이나 온라인 리소스를 사용한 경험이 있나요?	예	아니요			
		☐	☐			

18. 있다면 학습 경험에 대해 간략하게 적어 주세요.

그 밖에 하고 싶은 말은 무엇인가요?

(자유롭게 서술하셔도 좋습니다)

소중한 의견 감사합니다!

Appendix A9: Exit Questionnaire for Study 1

출구 설문조사 (Exit Questionnaire)

1. 이 연구의 목적이 무엇이라고 생각합니까?

_____

2. 이 연구를 통해 영어 문법이나 단어를 배운 것이 있습니까?

_____

3. 이 연구에 참여하는 동안 특정한 영어 문법이나 단어에 주의를 기울였습니까?

_____

4. 주의를 기울였다면 그 이유는 무엇입니까?

_____

5. 연구 밖에서 영어 문법이나 단어를 찾아보거나 공부한 적이 있습니까?

_____

6. 귀하의 영어 답변에 대해 피드백을 받는다면 어떤 형태의 피드백을 받고 싶습니까?

① 말로 받는 피드백

② 글로 받는 피드백

③ 말과 글로 동시에 받는 피드백

④ 기타 _____

7. 고르신 피드백을 선호하는 이유는 무엇입니까?

_____

8. 화상회의를 통한 영어 학습이 효과가 있다고 생각하십니까?

_____

9. 그렇게 생각하는 이유는 무엇입니까?

_____

10. 피드백을 받으셨나요? (Yes → 11 번 질문 / No → 13 번 질문)

_____

11. 피드백이 특정한 영어 문법이나 단어에 주의를 기울이게 했습니까?

_____

12. 피드백이 귀하의 영어 학습에 도움이 됐다고 생각하십니까?

_____

13. 이 연구에 대해 남기시고 싶은 말씀이 있다면 적어주세요.

_____

_____

_____

_____

Appendix A10. Results for the Ordinal Logistic Mixed-effects Models Examining Participants' Scores on the OPT

Fixed effects	*Estimate*	*SE*	*z*	*p*	*OR*	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 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]
OrallWritten: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 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	.305	1.31	[0.76, 2.24]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		1.64	1.17			

Appendix A11. Results for the Ordinal Logistic Mixed-effects Models Examining Participants' Scores on the WPT

Fixed effects	*Estimate*	*SE*	*z*	*p*	*OR*	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 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]
OrallWritten: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 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		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.70	.84			

Table A12. Results for the Ordinal Logistic Mixed−Effects Models Examining Participants' Scores on the SOPT

	Fixed effects				Random effects *by participant*	
	Estimate	*SE*	*t*	*p*	*variance*	*SD*
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***		
ControlOral:Time3	0.03	0.01	4.29	<.001***		
ControlWritten:Time3	0.03	0.01	4.29	<.001***		
ControlMulti:Time3	0.08	0.01	12.76	<.001***		
Oral 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***		
OralControl: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 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***		
R2m: 0.87 *R2c:*0.91						
*Formula: lmer(SOPT ~ 1 + Group*Time + (1|Participant))*						

Appendix A13: Information Sheet for Study 2

Institute of Education

UCL

Information Sheet for Korean EFL Learners

- **Title of Study:** An investigation of the communication, interactional patterns, and on-screen eye-movements of Korean EFL learners in video-conferencing tasks
- **Department:** Communication, Culture and Media
- **Name and contact details of the researcher:** Yeonwoo Jung [redacted]
- **Supervisor:** Professor Andrea Révész

My name is Yeonwoo Jung, and I would like to invite you to participate in a PhD research project. Before you decide, it is important for you to understand the purpose of the research and the involvement required. Please carefully read the following information and feel free to seek clarification or request additional details.

Who is carrying out the research?

I am a doctoral student in the Department of Communication, Culture and Media at the UCL Institute of Education. I am interested in second language learning, teaching, and assessment, and this study is part of my PhD research.

Why are we doing this research?

The purpose of the present study is to explore the online communication, interactional patterns, and on-screen eye movements of Korean EFL (English as a Foreign Language) learners during video-conferencing tasks. These tasks will be conducted through synchronous computer-mediated communication (SCMC) platforms, such as Zoom. The study will utilize an eye-tracker to examine cognitive processes and viewing behaviors. An additional aim is to assess whether the online communication and interactional patterns, as well as viewing processes of EFL learners, are influenced by their working memory capacity. The outcomes of this research are expected to advance our understanding of the effectiveness of video-conferencing tasks in capturing EFL learners' attention, enhancing their online communication competence, and examining the role of learners' working memory in mediating their processing of the input.

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Institute of Education

Why am I being invited to take part?

You are being invited to take part as this study is looking to involve those who are learners of English as a foreign language.

What will happen if I choose to take part?

If you choose to participate, you will engage in a series of five individual sessions.

At the outset of the experiment, you will be requested to fill out a brief questionnaire about yourself and complete a general English proficiency test, along with a set of pretests (oral production, written production, and spontaneous oral production tests). Approximately one and a half hours will be designated for the completion the questionnaire and the pretests.

During the intervention phase, you will be involved in three individual sessions. In each session, you will take part in a communicative video-conferencing task, with your performance recorded. Also, while performing the tasks, your eye-movements will also be recorded. These sessions will last approximately 20–30 minutes each. Upon completion of the sessions, you will be required to undertake a posttest, consisting of oral production, written production, and spontaneous oral production tests, lasting approximately 1 hour, and your test performance will also be recorded. During the fourth sessions, you will carry out a set of memory activities, which will require you to remember and repeat nonwords, and particular visual patterns. After two weeks, you will be asked to participate in a delayed posttest, which includes oral production, written production, and spontaneous oral production tests, and your performance will also be recorded. Following the delayed posttest, you will be administered a brief post-task questionnaire, expected to take around 10 minutes. Completing all activities is crucial for achieving the study's objectives, as failure to do so would impede the utilization of the collected data.

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Will anyone know I have been involved?

The recordings of your performance will exclusively be used for analysis and conference presentations. No other use will be made without obtaining your written permission, and access to the original recordings will be restricted to project personnel only. These recordings will undergo transcription, with stringent measures to ensure your name or any personal details are not included. Any data acquired from you will be securely stored, maintaining confidentiality throughout every phase of the project and beyond. To preserve anonymity, a unique identifier will be employed to anonymize your identity and data. The digital format of the data will be securely maintained and subjected to the anonymization process. Confidentiality will be upheld unless there exist compelling and legitimate reasons necessitating a breach. In such instances, we will promptly inform you of any decisions that might impact your confidentiality.

Could there be problems for me if I take part?

Engaging in this study is not expected to cause you any disadvantages or discomfort. You have the right to withdraw from the study at any time, without providing a reason and without any impact on you. If you decide to withdraw, all data collected from you will be promptly deleted.

What will happen to the results of the research?

The results might be presented at professional conferences and in published in research journals. We may archive the dataset in a digital depository for open access. Upon request, a summary of the study's overall results will be provided at the completion of the research. Your identity will not be disclosed in any report or publication.

Do I have to take part?

The decision to participate in entirely yours. We hope that if you choose to join, you will find it to be a valuable experience.

Contact for further information

Research designs often necessitate that the complete intent of the study is not disclosed before participation. While we have described the general nature of the tasks you will perform, the full intent of the study will only be revealed to you upon the completion of the study.

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Institute of Education

If you have any further questions before deciding to take part, you can reach me at [redacted], [redacted] or Professor Andrea Révész at [redacted]

Data Protection Privacy Notice

The data controller for this project will be University College London (UCL). The UCL Data Protection Office provides oversight of UCL activities involving the processing of personal data, and can be contacted at data-protection@ucl.ac.uk. UCL's Data Protection Officer can also be contacted at data-protection@ucl.ac.uk.

Further information on how UCL uses participant information can be found here: www.ucl.ac.uk/legal-services/privacy/participants-health-and-care-research-privacy-notice

If you are concerned about how your personal data is being processed, or if you would like to contact us about your rights, please contact UCL in the first instance at data protection@ucl.ac.uk.

Thank you very much for taking the time to read this information sheet and for your consideration to participate!

UCL Institute of Education
20 Bedford Way, London WC1H 0AL

Appendix A14: Consent Form for Study 2

Institute of Education

UCL

Consent Form

An investigation of the communication, interactional patterns, and on-screen eye-movements of Korean EFL learners in video-conferencing tasks

If you are happy to participate in this study, please complete this consent form and return to Yeonwoo Jung in person or at the address below.

	Yes	No
I have read and understood the information leaflet about the research.	☐	☐
I agree to the arrangements for my participation as described in the Information Sheet.	☐	☐
I agree to be filmed during the treatment sessions.	☐	☐
I understand that if any of my words are used in reports or presentations they will not be attributed to me.	☐	☐
I understand that I can withdraw from the project at any time, and that if I choose to do this, any data I have contributed will not be used.	☐	☐
I understand that I can contact Yeonwoo Jung at any time and request for my data to be removed from the project database.	☐	☐

Name_____Signed _____

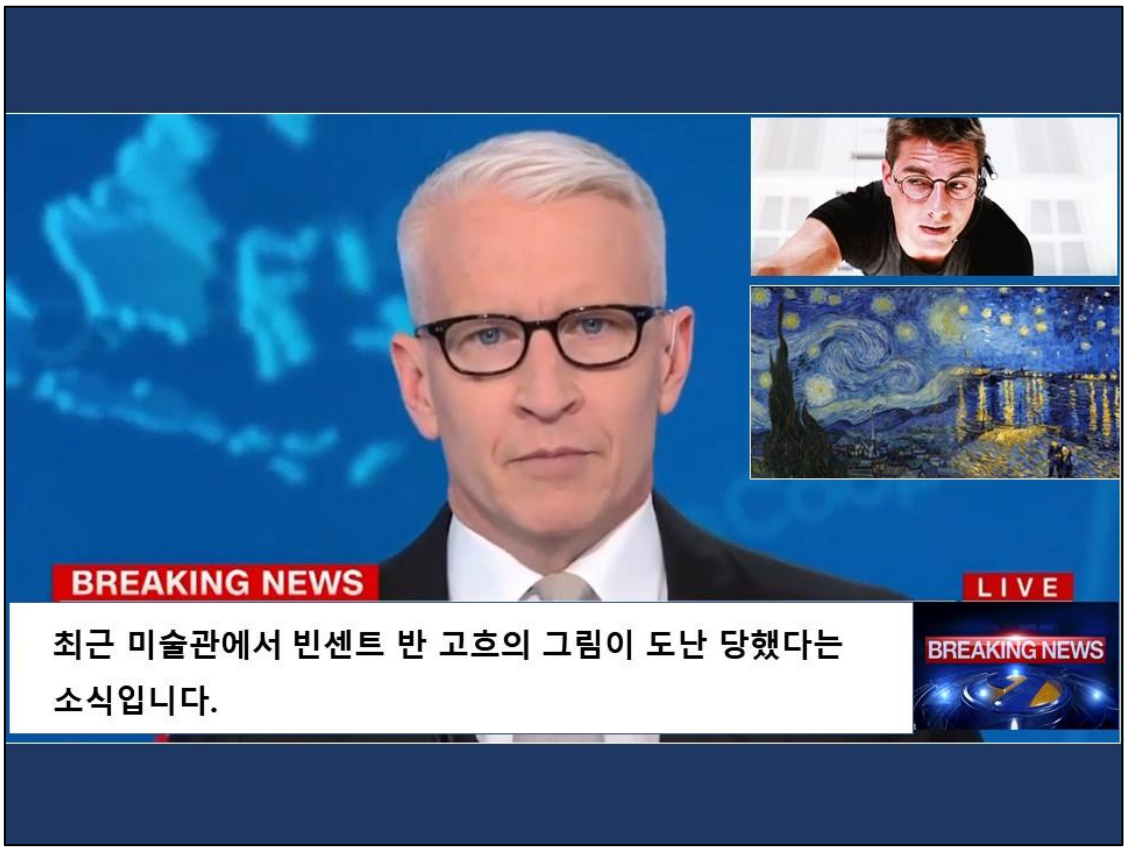
Date _____

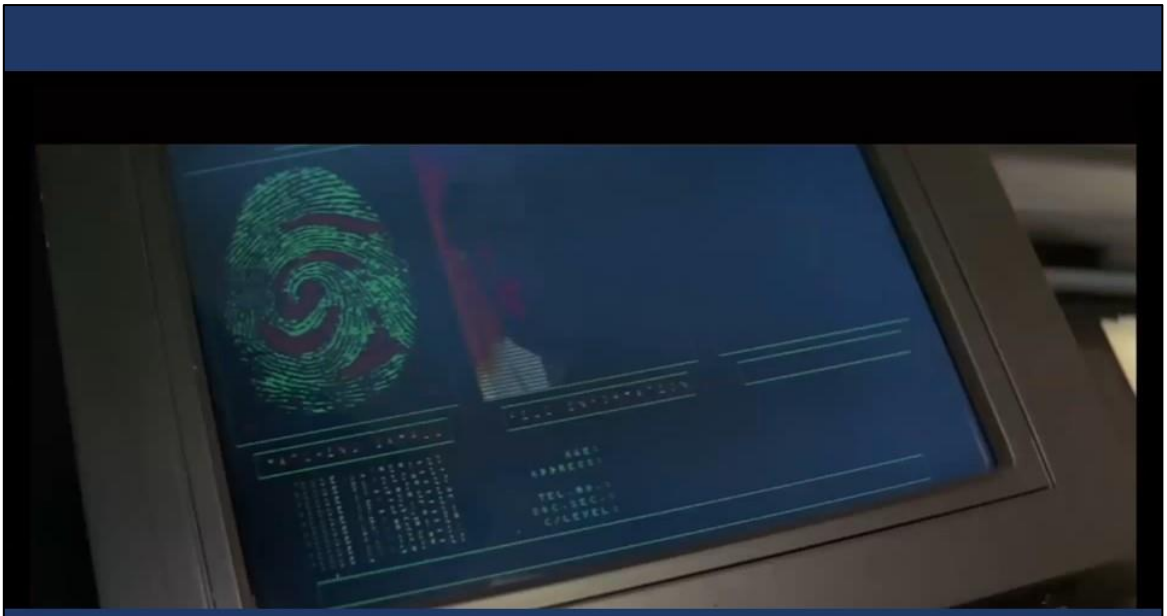
Yeonwoo Jung
UCL Institute of Education
20 Bedford Way London WC1H 0AL
[redacted]

Yeonwoo Jung
Department of Communication, Culture and Media
UCL Institute of Education
20 Bedford Way London WC1H 0AL

Appendix A15: A Sample of Experimental Tasks in Study 2







그림을 되찾기 위해 FBI는 유력한 용의자를 감시 중이라고 밝혔습니다.






환영하네!
자네는 이 사건을 수사 중인
FBI 미술 범죄 전담반 수사관일세.
FBI
DEPARTMENT OF JUSTICE
FEDERAL BUREAU OF INVESTIGATION
ART CRIME TEAM


유력 용의자는
톰 헌트일세.
Name:
톰 헌트
Age:
42
Address:
런던



DIRECTIONS:

당신은 이 사건을 수사 중인 FBI 미술 범죄 전담반 수사관입니다.

당신의 임무는 유력 용의자를 매일 감시하고 그 정보를 모든 요원들이 실시간으로 공유할 수 있도록 FBI 데이터베이스에 업로드하는 것입니다.





DIRECTIONS:

1) 용의자가 매일 오후 4시에 하는 활동을 구두로 보고하세요.
2) 용의자가 현재 하고 있는 활동을 구두로 보고하세요.

Every Monday at 4 p.m.,

he...



Now,

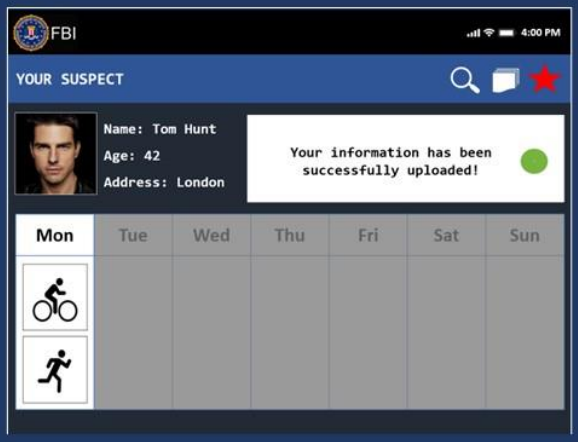
he...



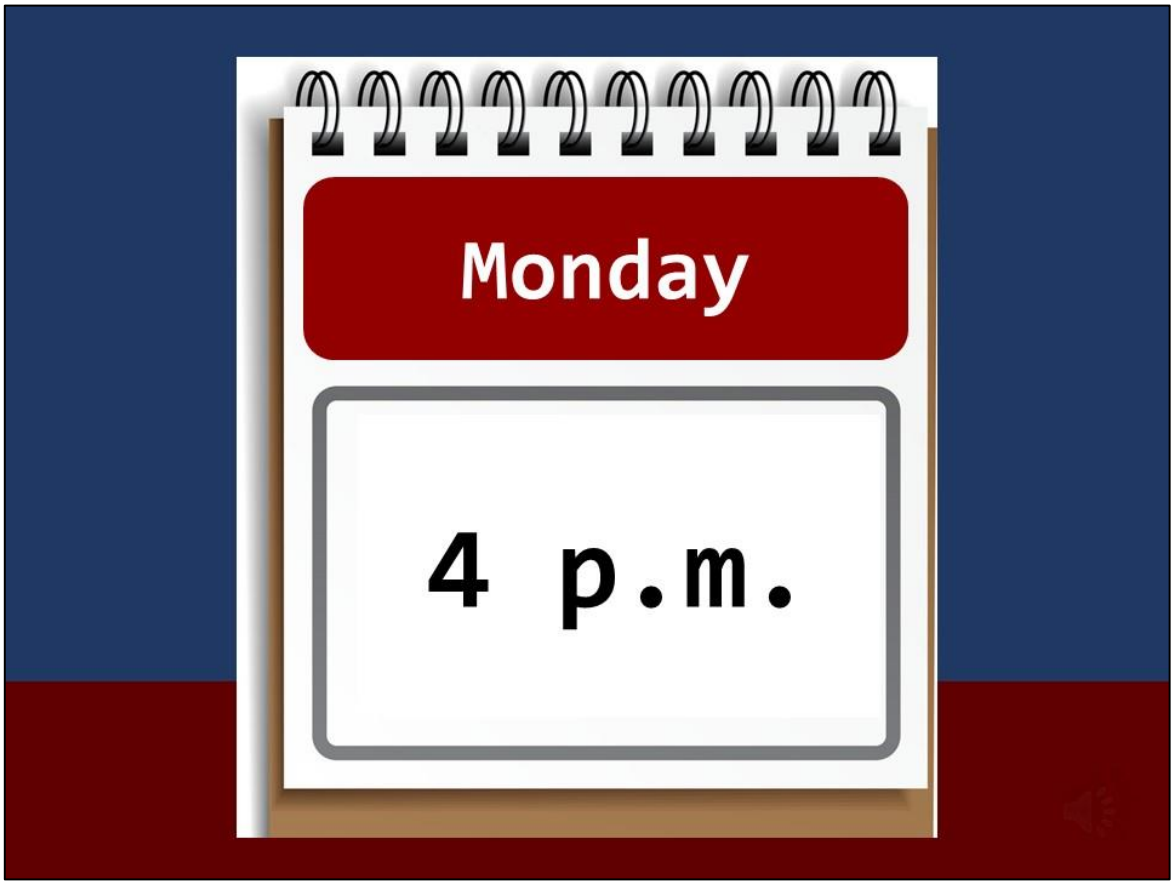
DIRECTIONS:

1) 용의자가 매일 오후 4시에 하는 활동을 구두로 보고하세요.
2) 용의자가 현재 하고 있는 활동을 구두로 보고하세요.
3) 당신이 보고한 정보는 모든 FBI 요원들이 실시간으로 용의자의 정보를 공유할 수 있는 FBI 데이터베이스에 업로드 될 것입니다.






Let's start watching!
FBI
SUSPECT
database
Home
Emergency
Safety Tips
Checklist


Monday
4 p.m.

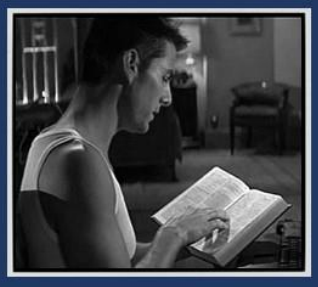
Every Monday at 4 p.m.,

he...



Every Monday at 4 p.m.,

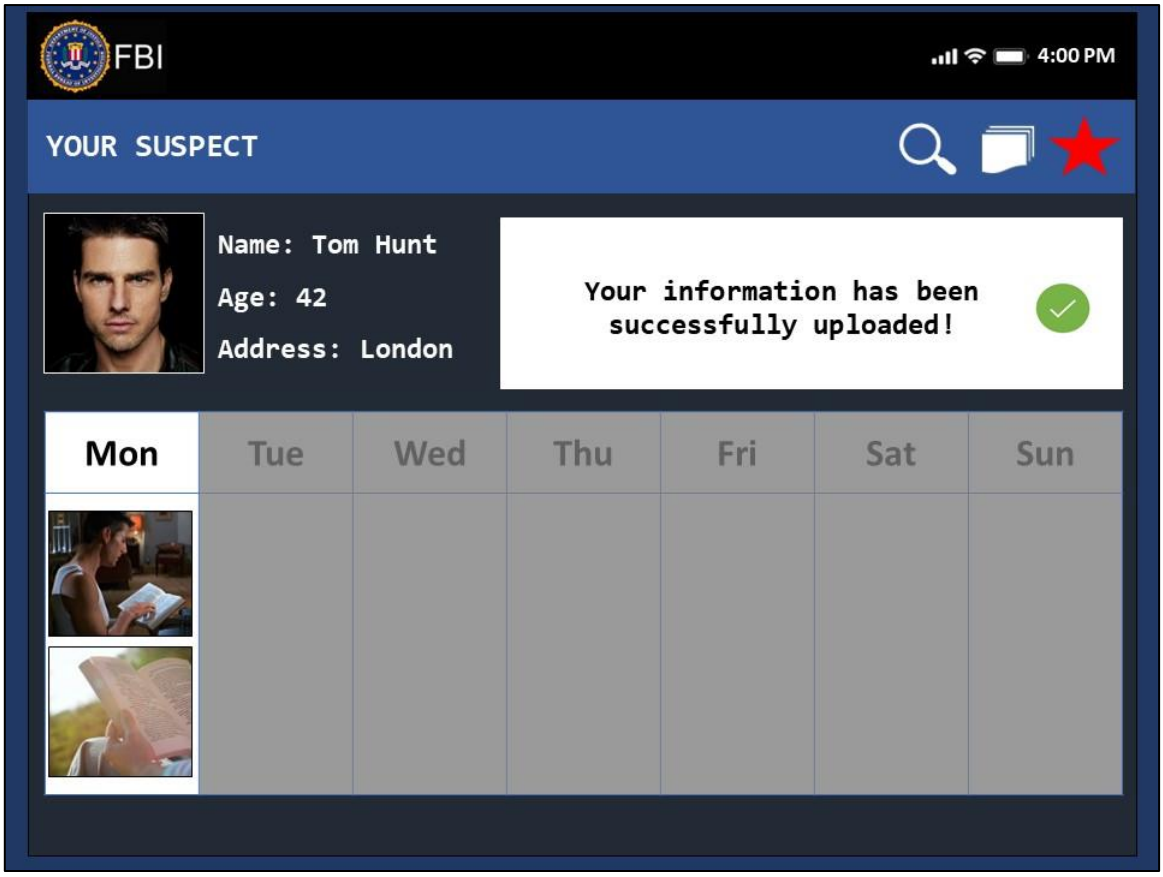
he...



Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Tuesday
4 p.m.

Every Tuesday at 4 p.m.,

he...



Every Tuesday at 4 p.m.,

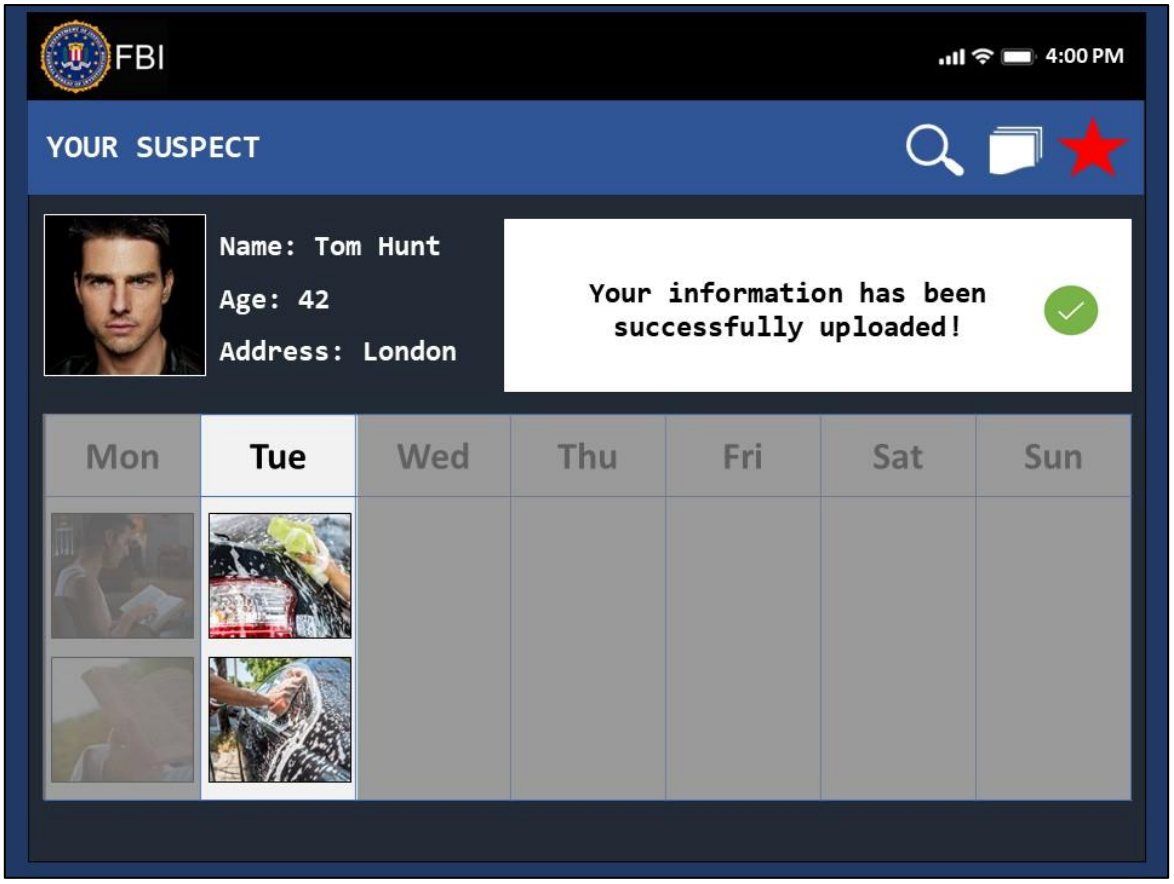
he...

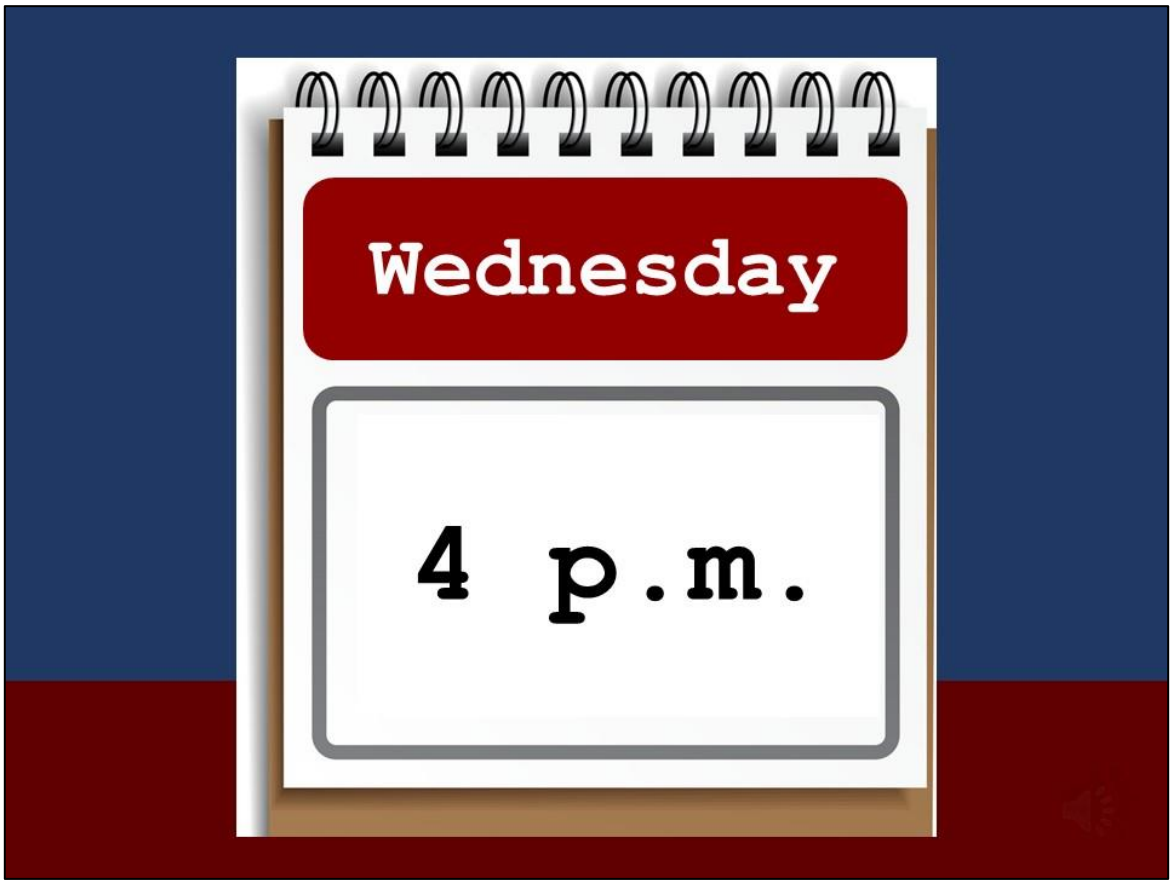


Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Wednesday
4 p.m.

Every Wednesday at 4 p.m.,

he...



Every Wednesday at 4 p.m.,

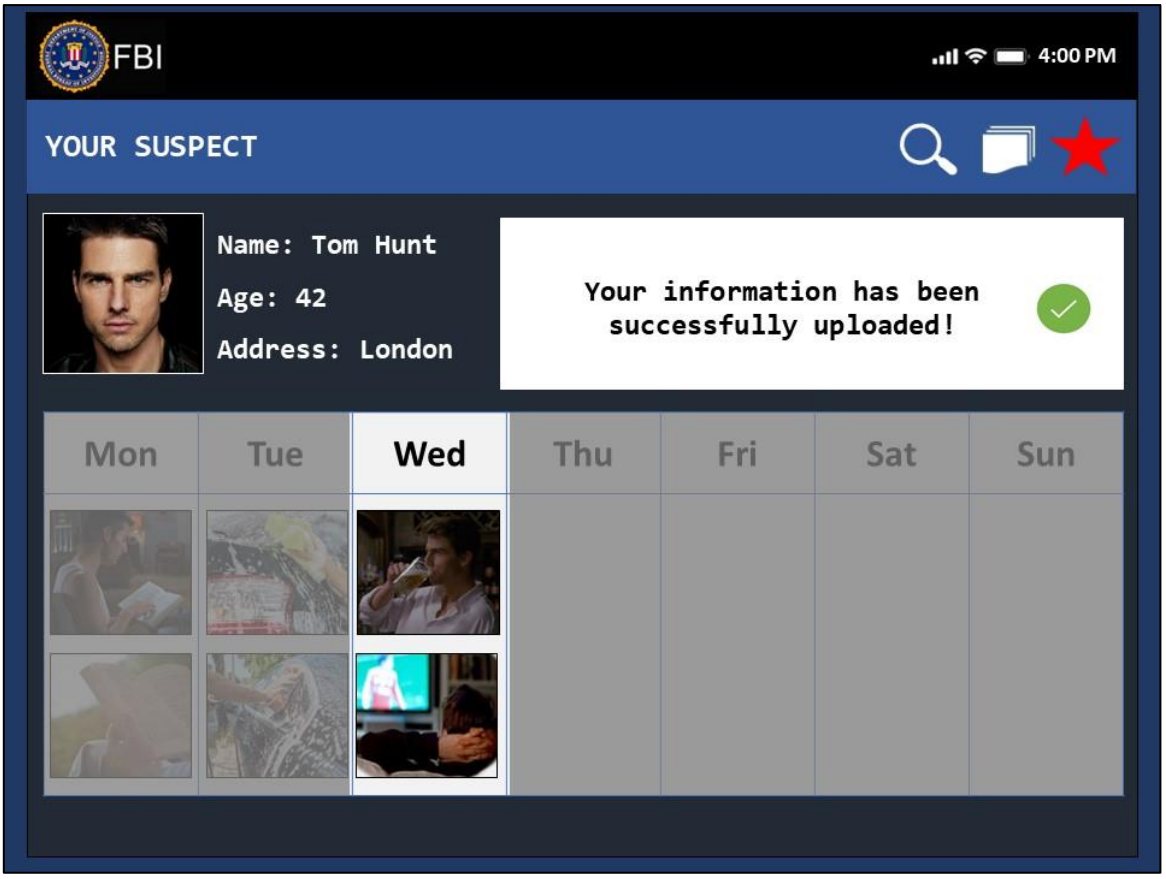
he...

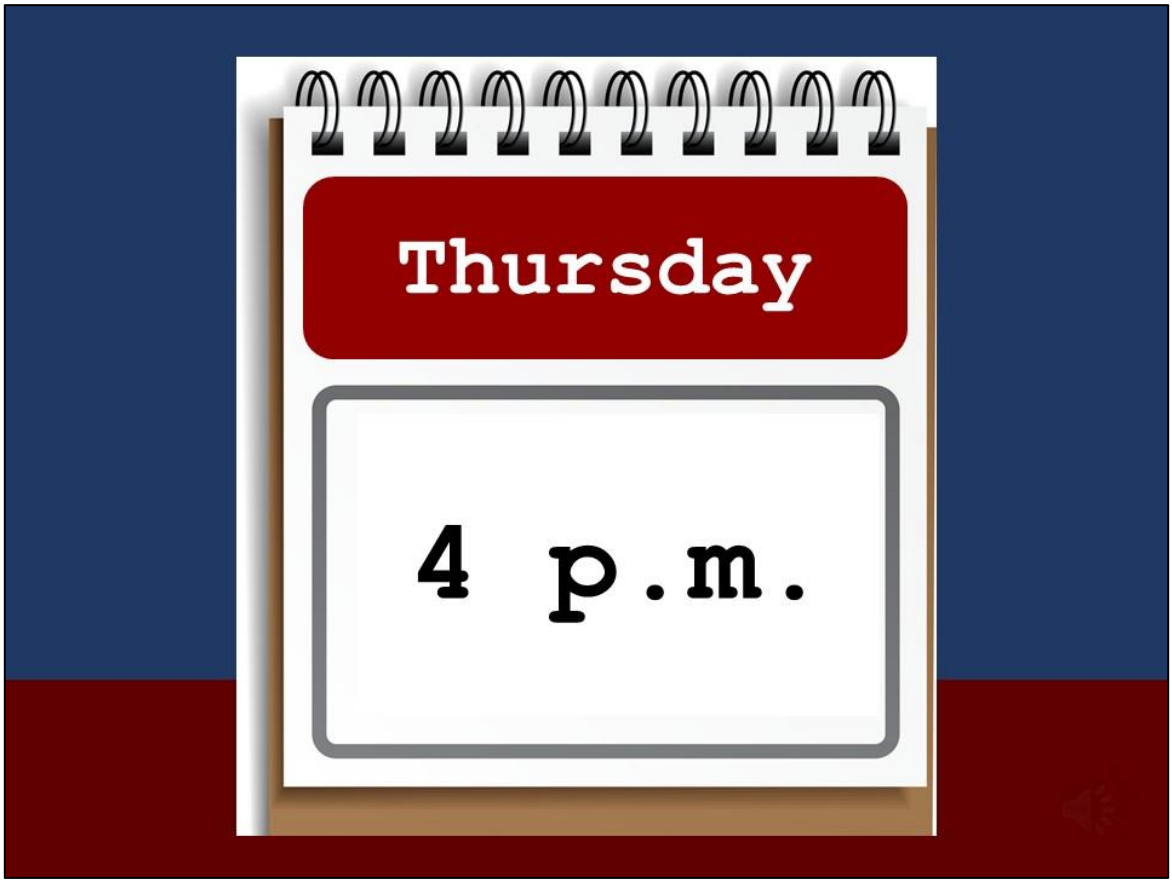


Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Thursday
4 p.m.

Every Thursday at 4 p.m.,

he...



Every Thursday at 4 p.m.,

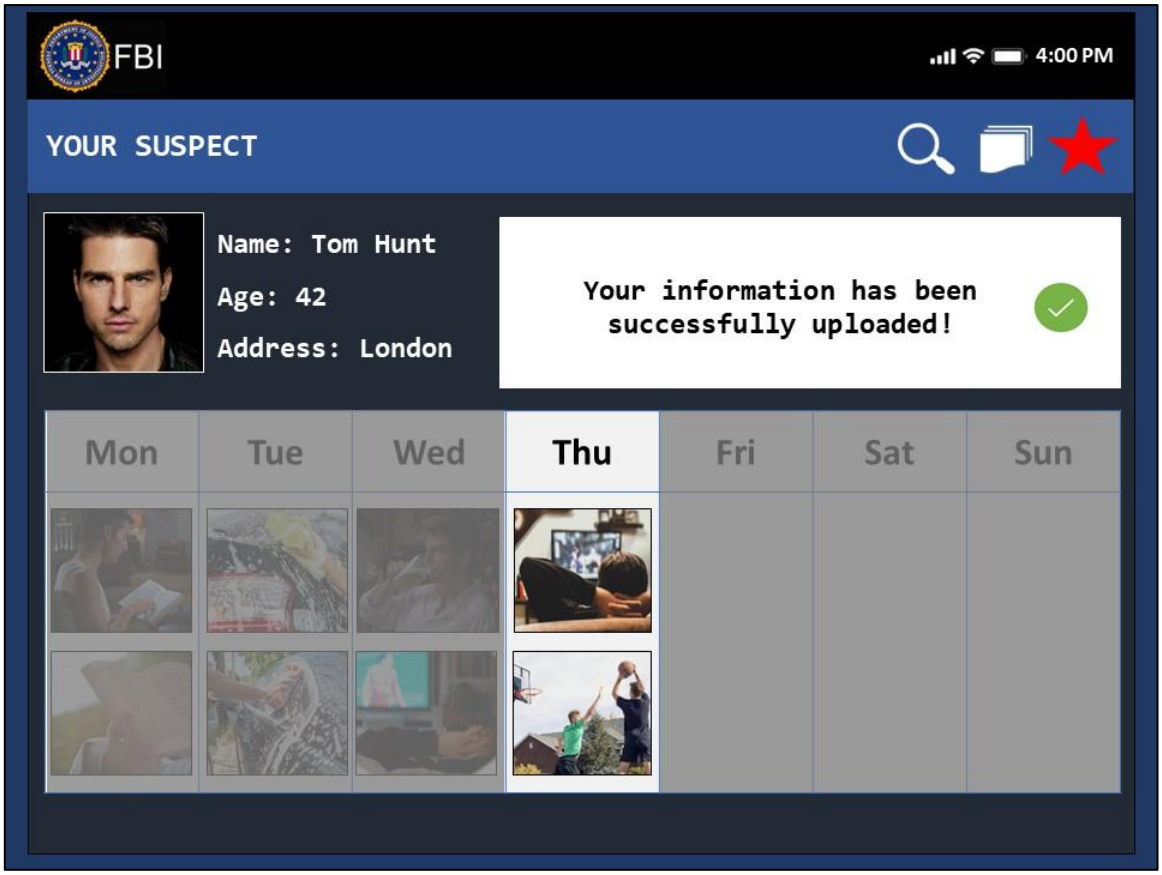
he...

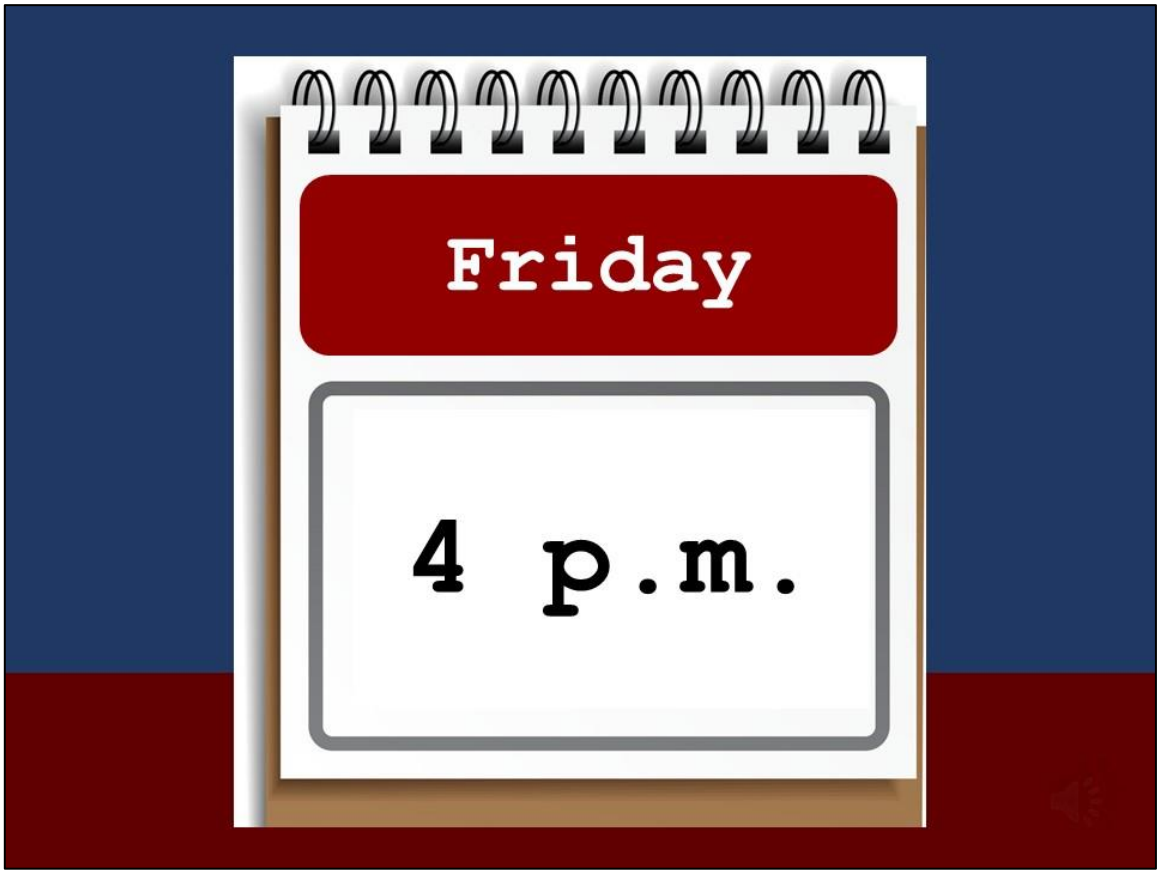


Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Friday
4 p.m.

Every Friday at 4 p.m.,

he...



Every Friday at 4 p.m.,

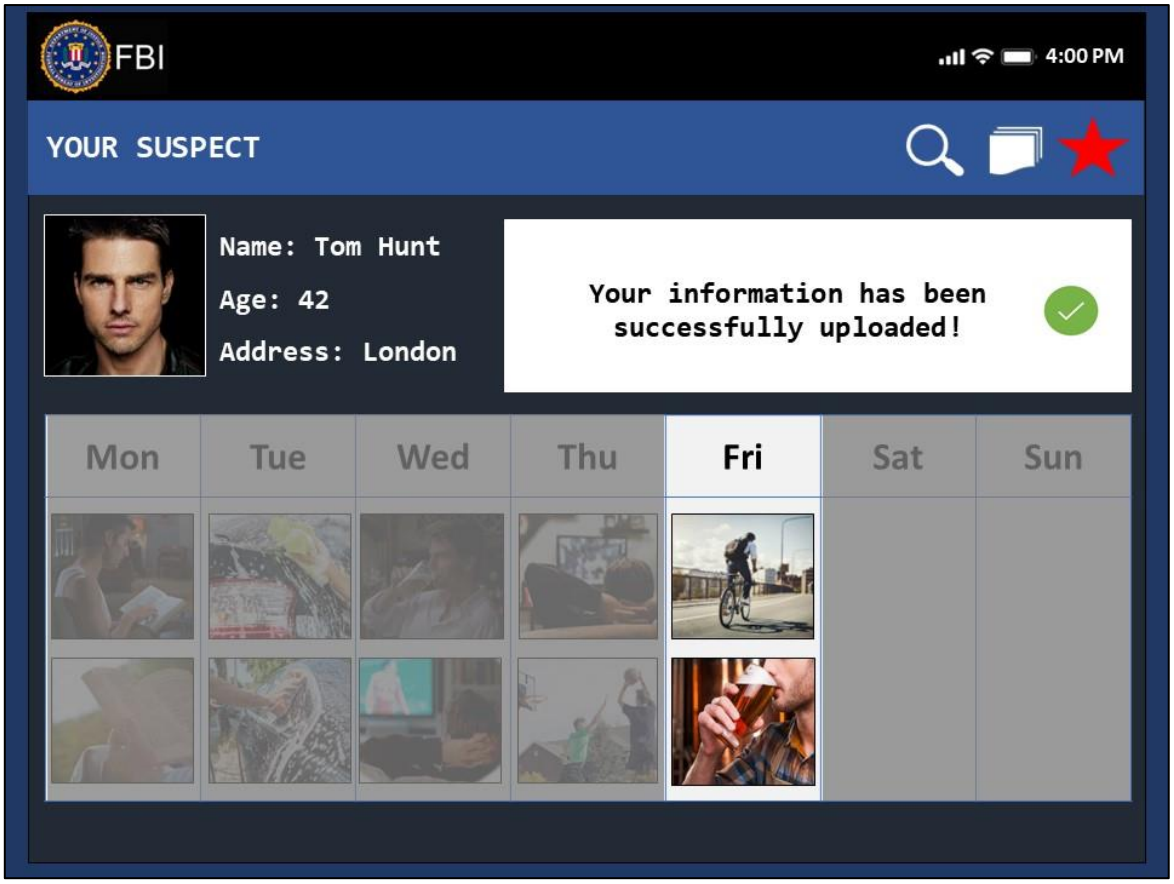
he...



Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Saturday
4 p.m.

Every Saturday at 4 p.m.,

he...



Every Saturday at 4 p.m.,

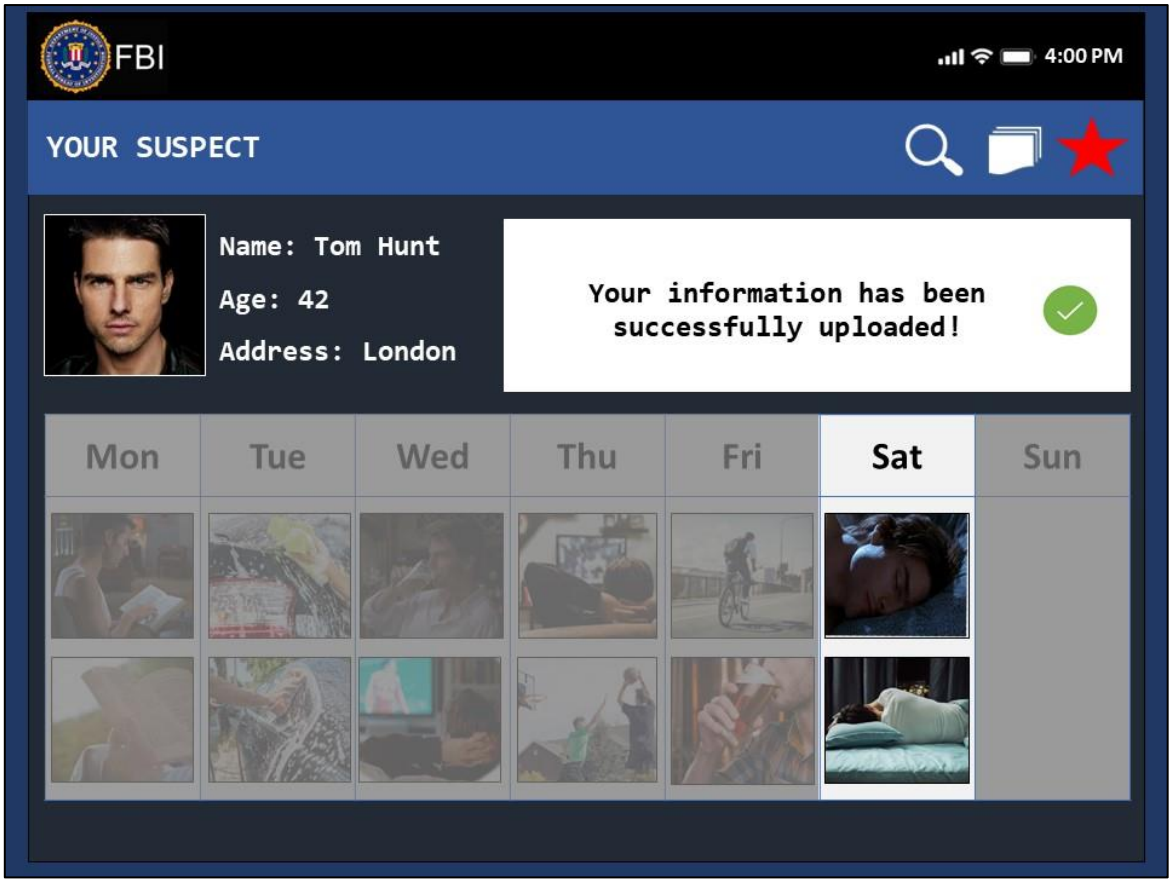
he...

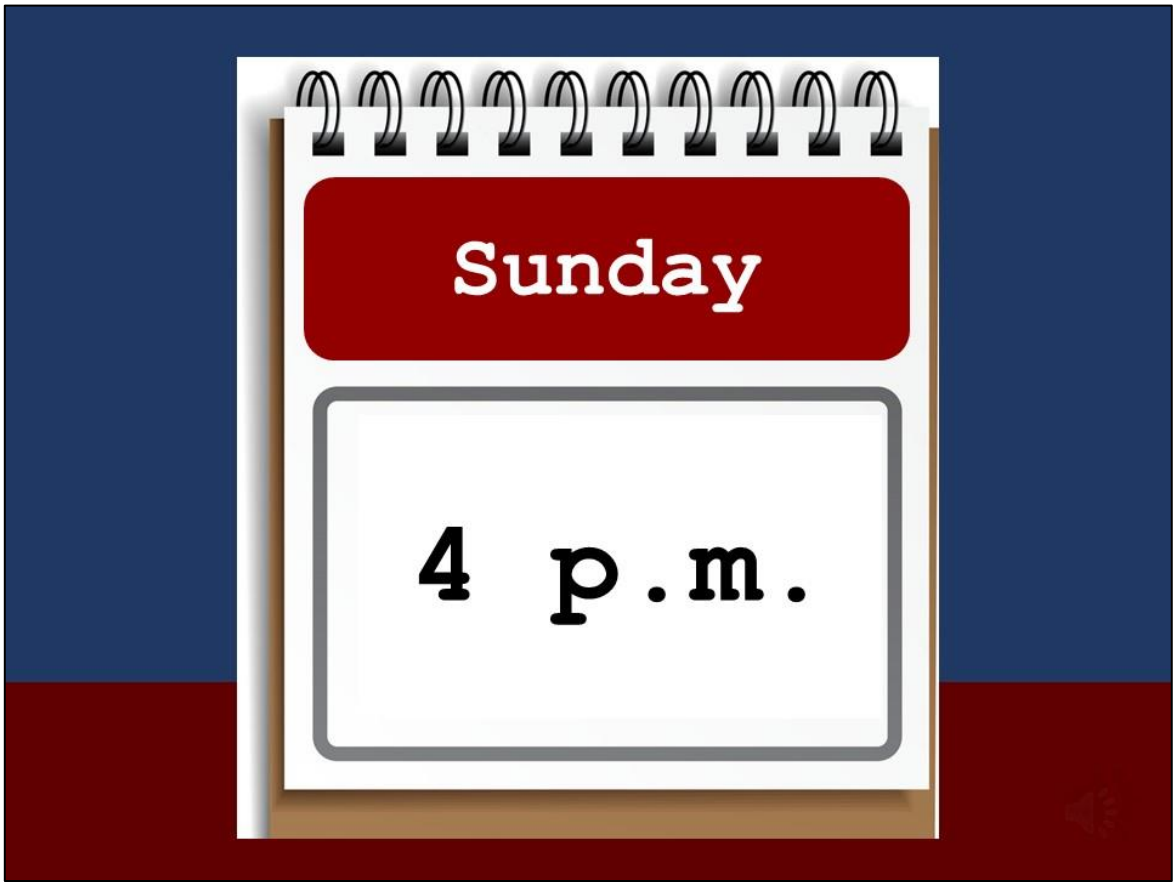


Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


Sunday
4 p.m.

Every Sunday at 4 p.m.,

he...



Every Sunday at 4 p.m.,

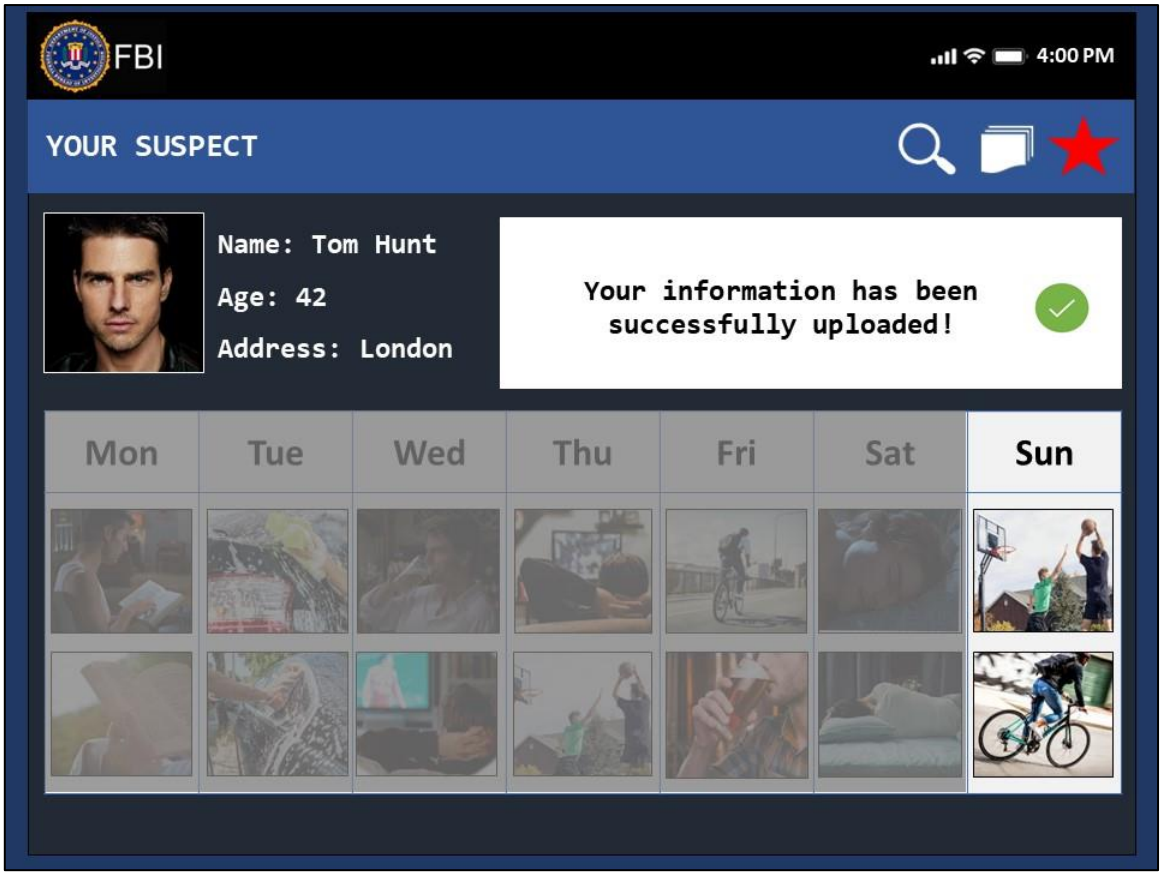
he...

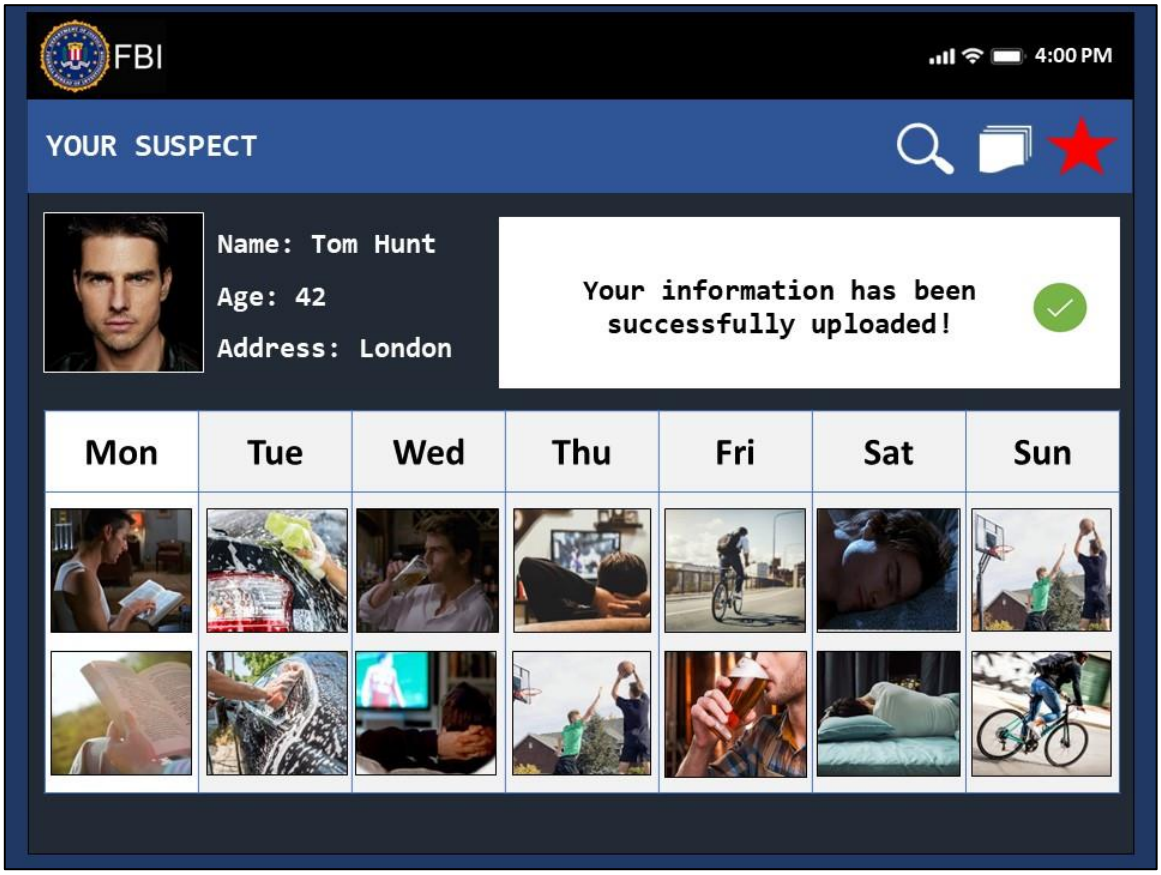


Now,

he...




FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


FBI
4:00 PM
YOUR SUSPECT
Name: Tom Hunt
Age: 42
Address: London
Your information has been successfully uploaded!
Mon
Tue
Wed
Thu
Fri
Sat
Sun


훌륭하네!
성공적인 임무 수행을
축하하네!

Appendix A16: A Sample of OPT in Study 2

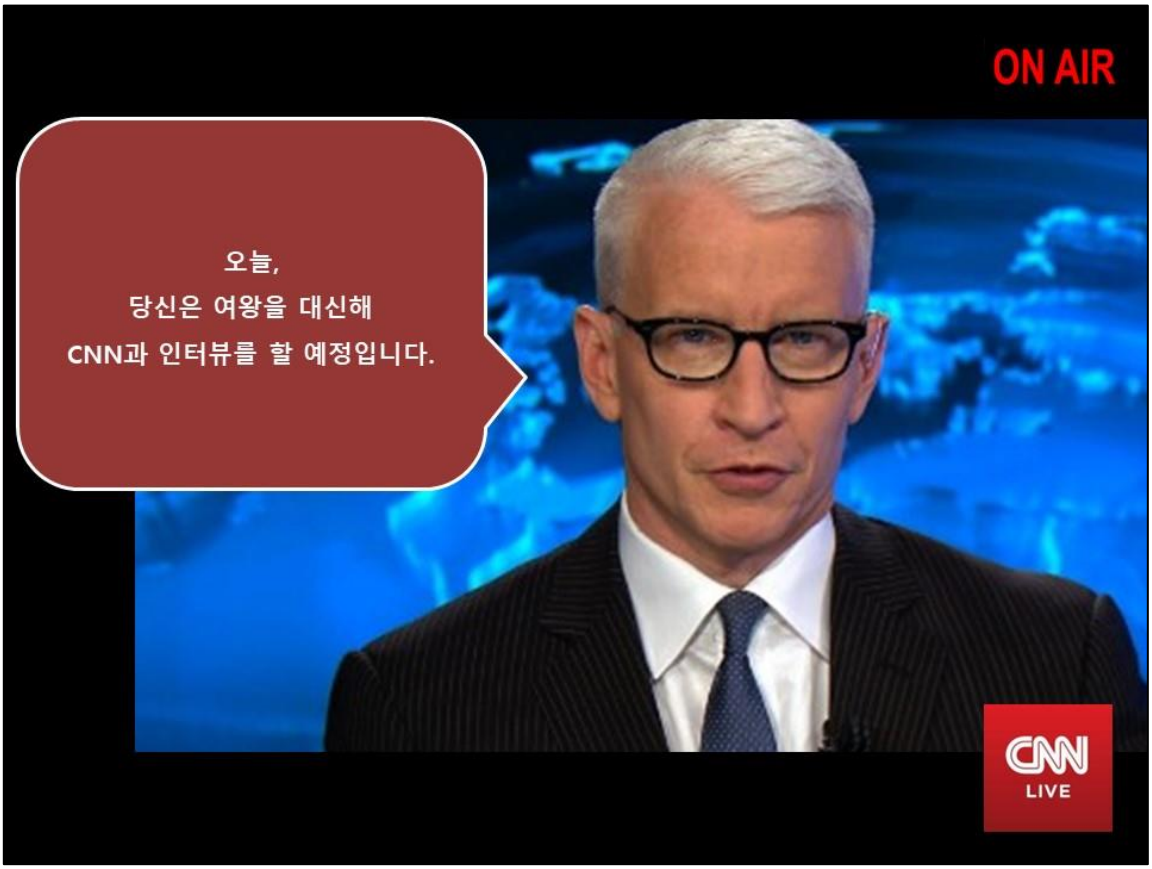


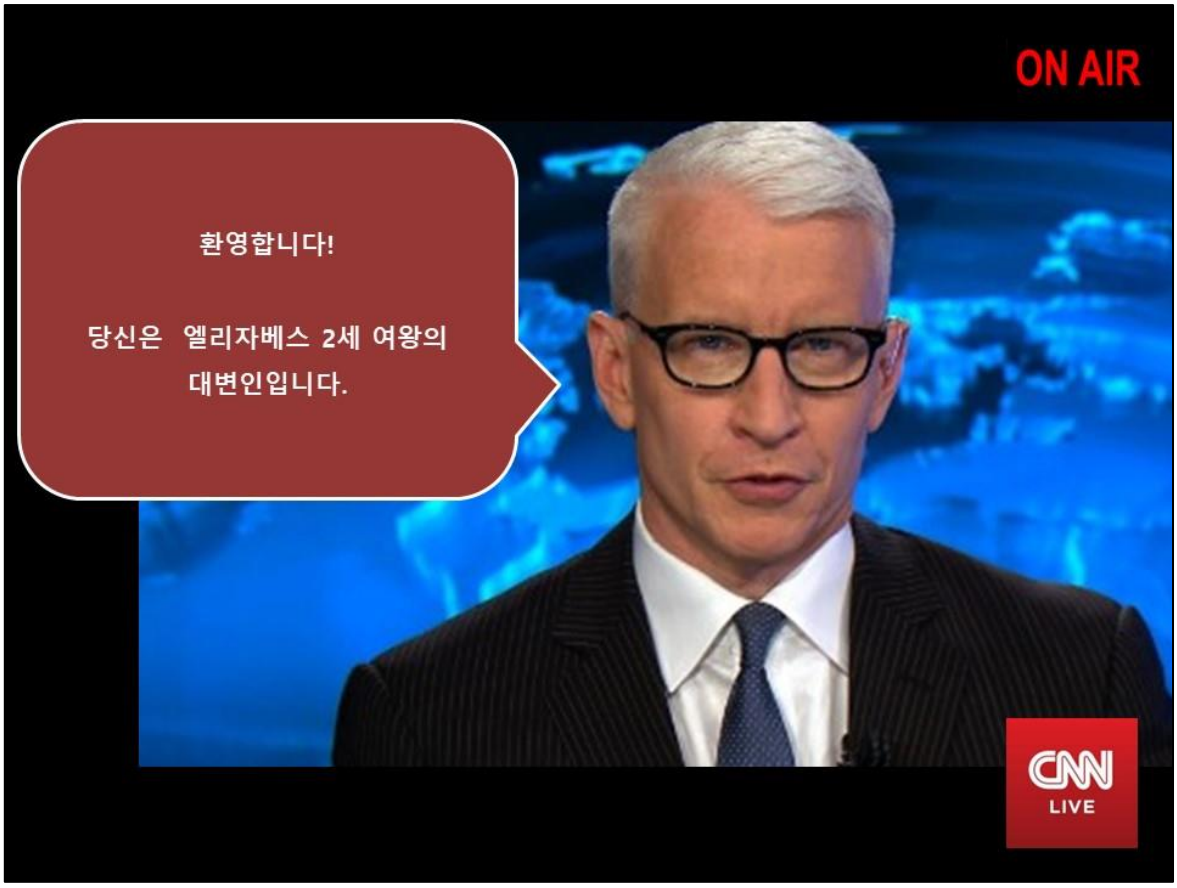



INTERVIEW WITH CNN
CNN
BREAKING NEWS


BREAKING NEWS
오늘 엘리자베스 2세 여왕은 G7 회의에서
각국 정상들과 회담을 가질 예정입니다.
CNN
LIVE


SMS
sending...
BREAKING NEWS
엘리자베스 여왕에 대해 궁금한 점이 있나요?
당신의 질문을 CNN 뉴스에 보내세요!
CNN
LIVE


ON AIR
오늘,
당신은 여왕을 대신해
CNN과 인터뷰를 할 예정입니다.
CNN
LIVE



DIRECTIONS:

당신은 영국 엘리자베스 2세 여왕의 대변인입니다.

여왕은 영국 콘월에서 열리는 G7 회의에서 각국 정상들과 회담을 가질 예정입니다. 오늘 당신은 여왕을 대신해 CNN과 인터뷰를 할 예정입니다.





With the Queen's spokesperson

DIRECTIONS:

1) 시청자의 질문은 화면에 제시될 것입니다.
2) 제공된 정보를 바탕으로 질문에 대답하세요.
3) 공식 인터뷰이므로, 반드시 완성형 문장으로 답하세요.



No.	✗
No, she can't.	✗
No, she can't speak German.	○

DIRECTIONS:

1) 시청자의 질문은 화면에 제시될 것입니다.



DIRECTIONS:

1) 시청자의 질문은 화면에 제시될 것입니다.
2) 주어진 정보를 바탕으로 질문에 대답하세요.

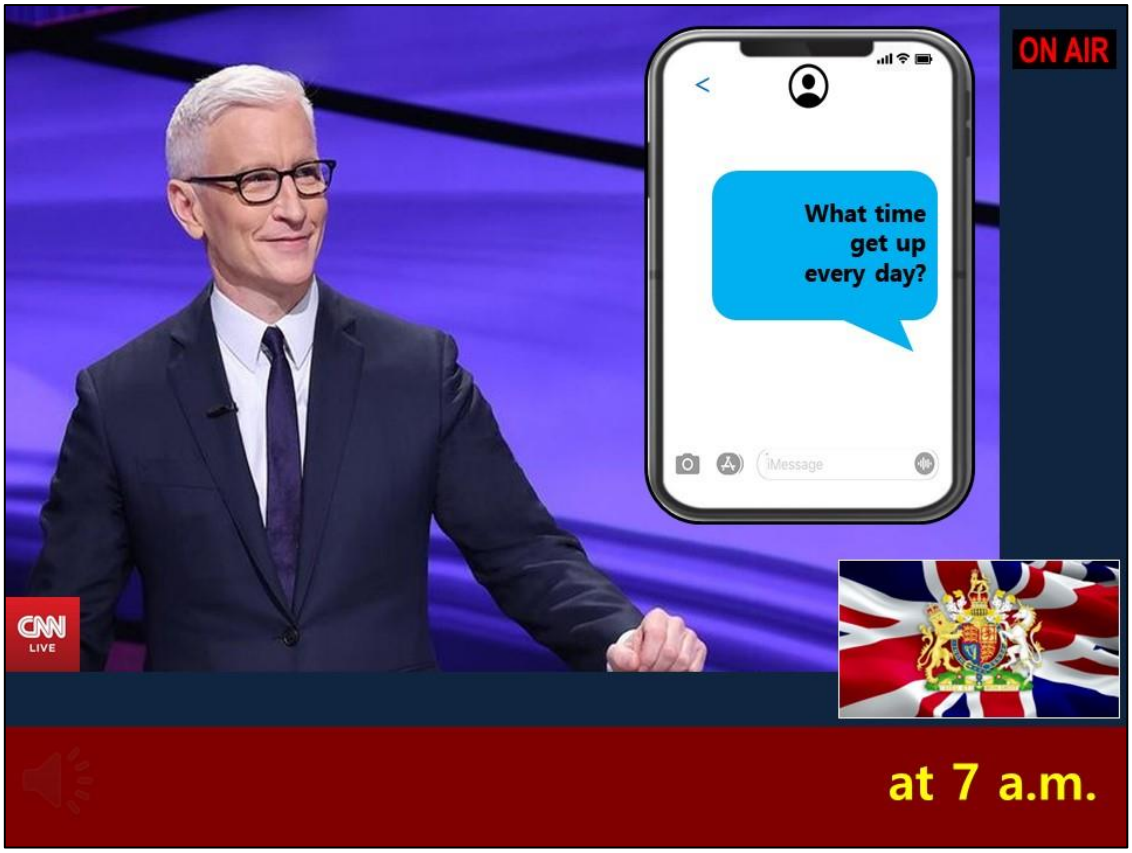


DIRECTIONS:

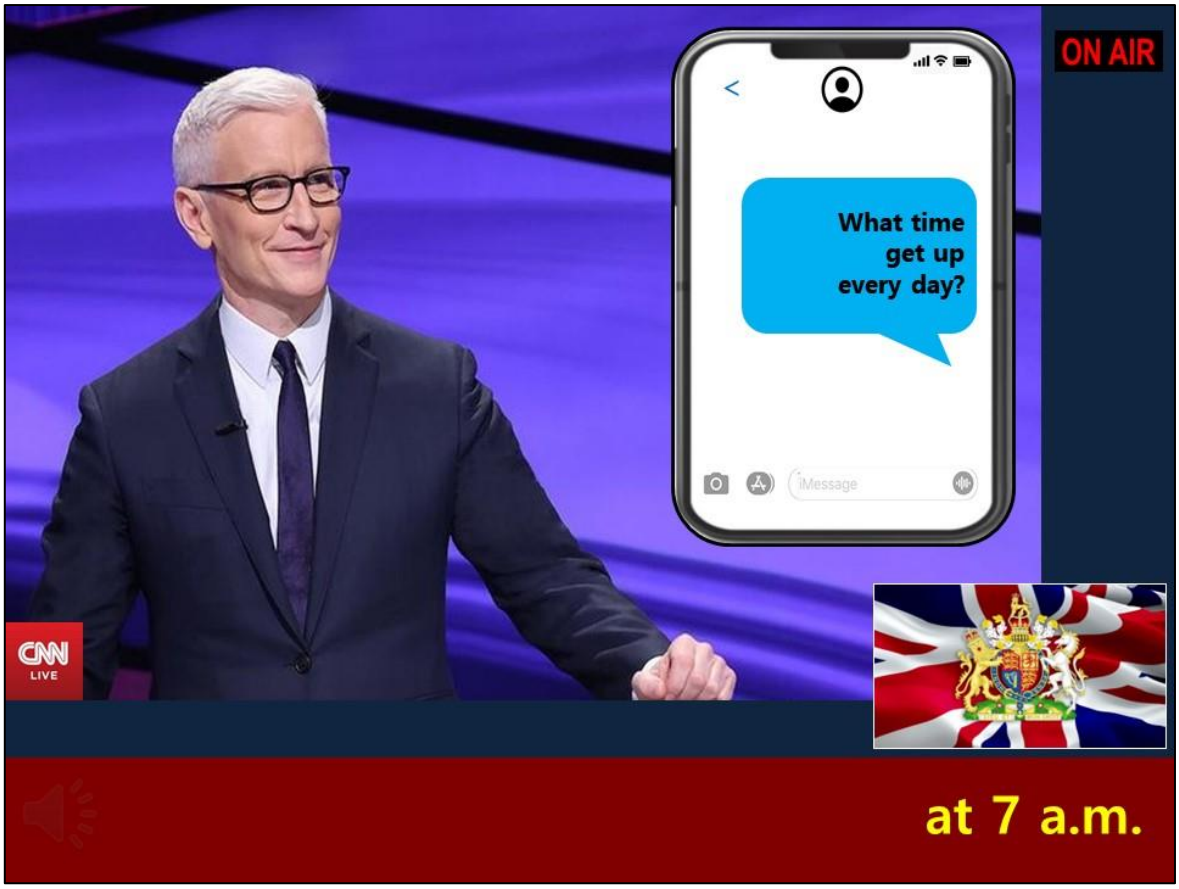
1) 시청자의 질문은 화면에 제시될 것입니다.
2) 제공된 정보를 바탕으로 질문에 대답하세요.
3) 공식 인터뷰이므로, <u>반드시 완성형 문장으로 답하세요.</u>

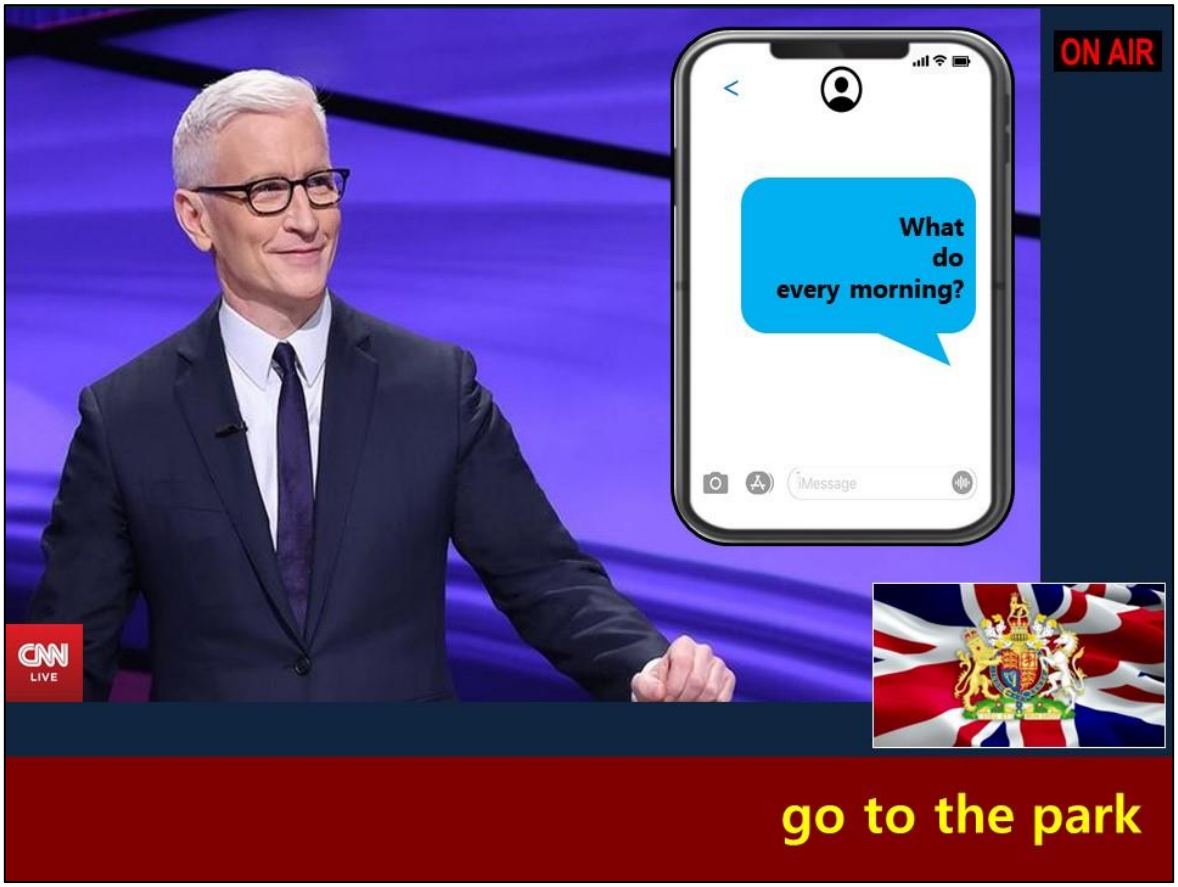


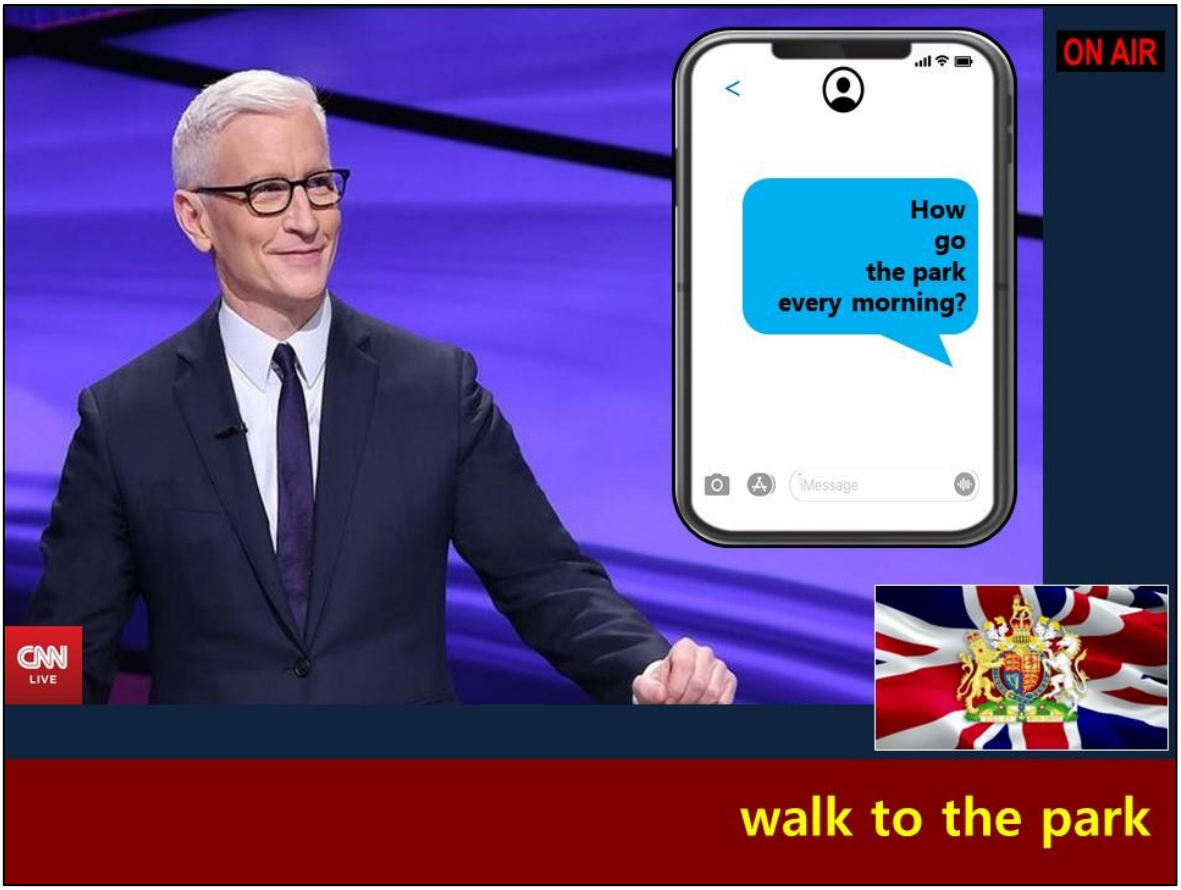
No.	✗
No, she can't.	✗
No, she can't speak German.	○

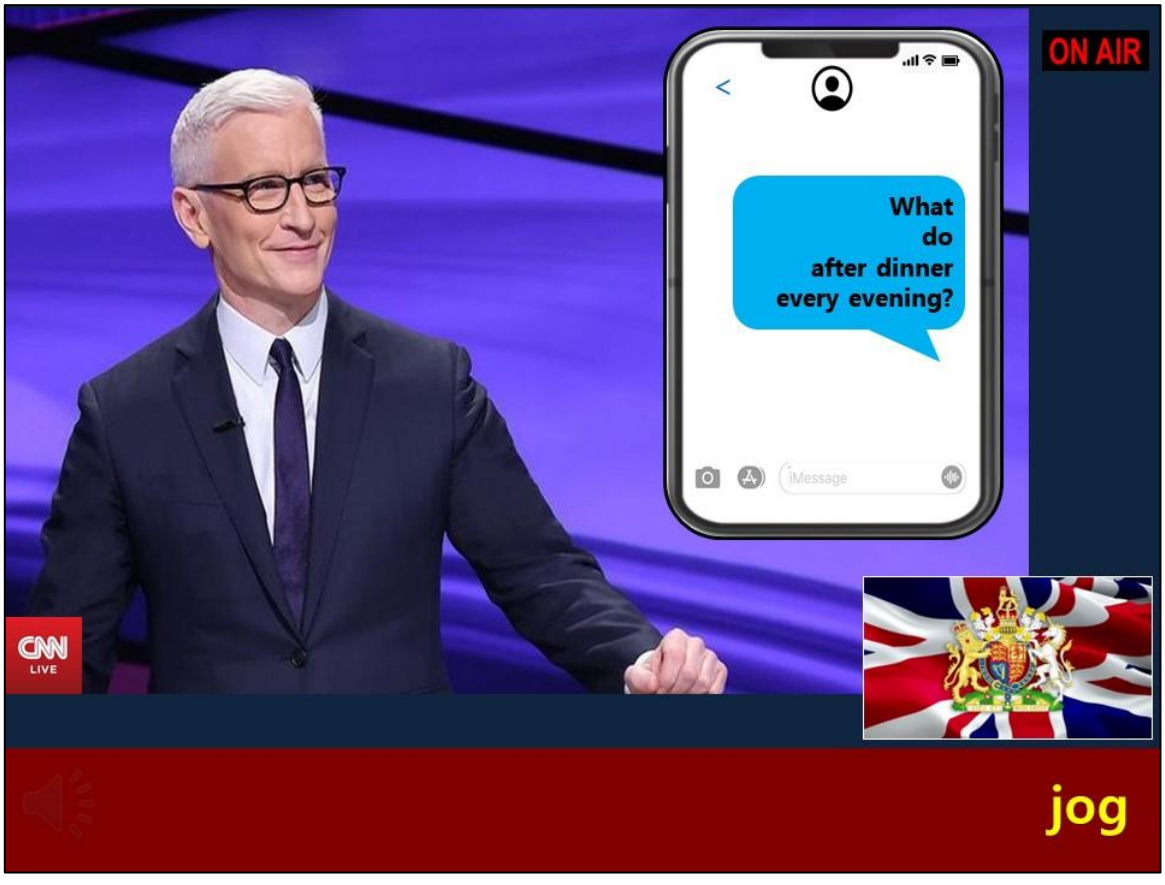

ON AIR
What time get up every day?
iMessage
CNN LIVE
at 7 a.m.

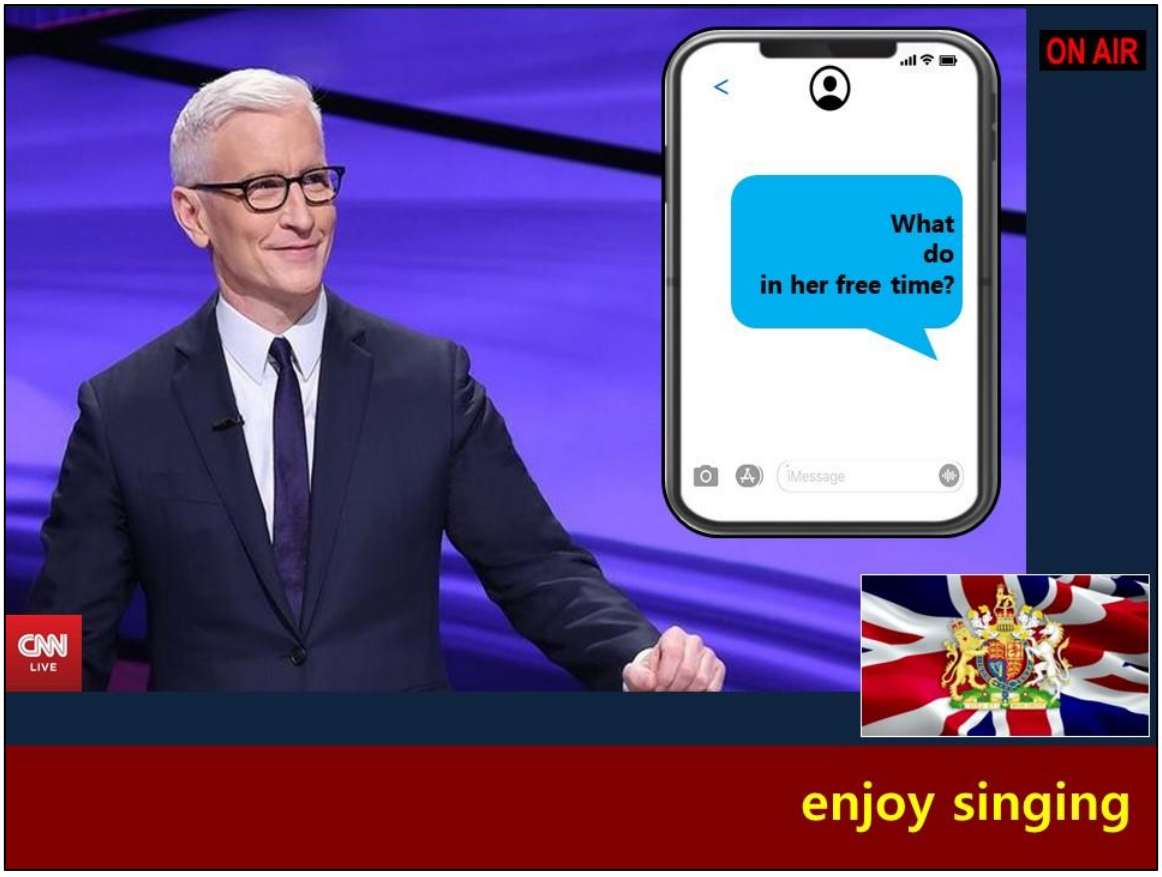

Let's begin!

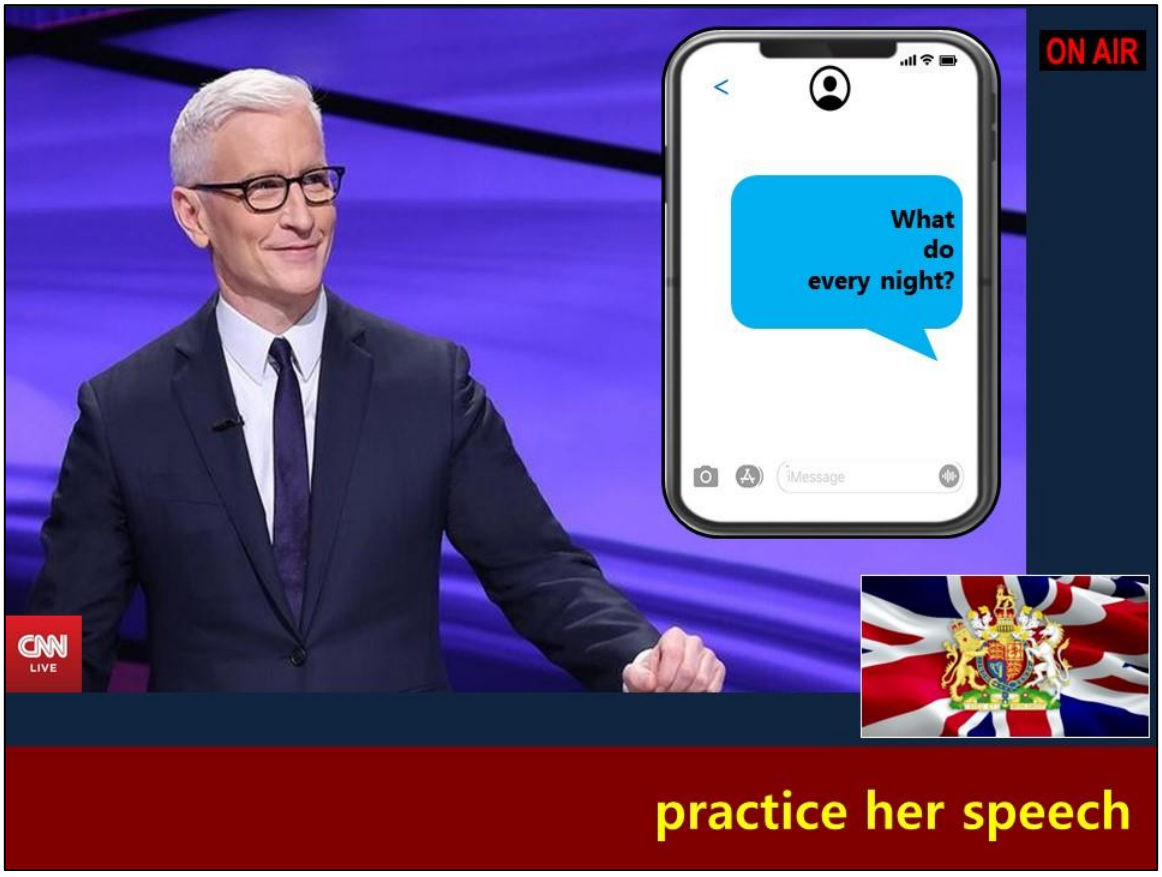

ON AIR
What time get up every day?
iMessage
CNN LIVE
at 7 a.m.


ON AIR
What do every morning?
iMessage
CNN LIVE
go to the park


ON AIR
How
go
the park
every morning?
iMessage
CNN
LIVE
walk to the park

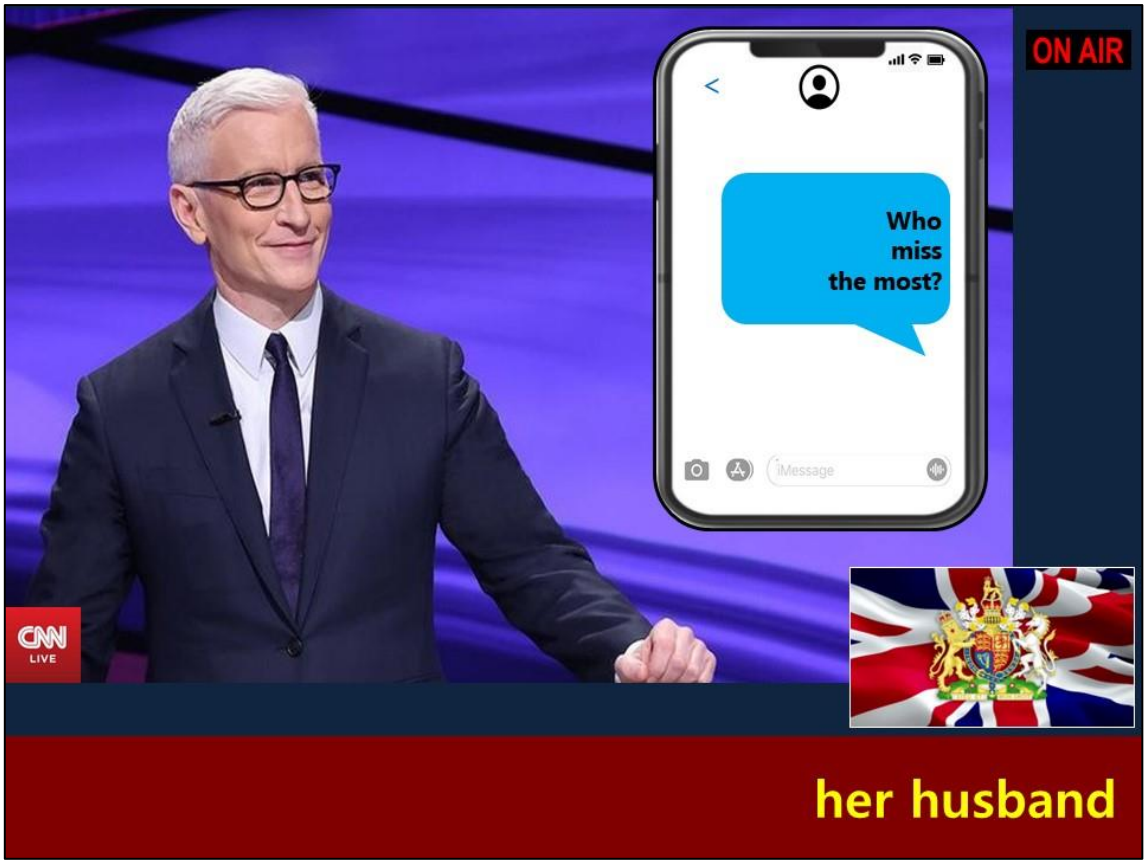

ON AIR
What
do
after dinner
every evening?
iMessage
CNN
LIVE
jog

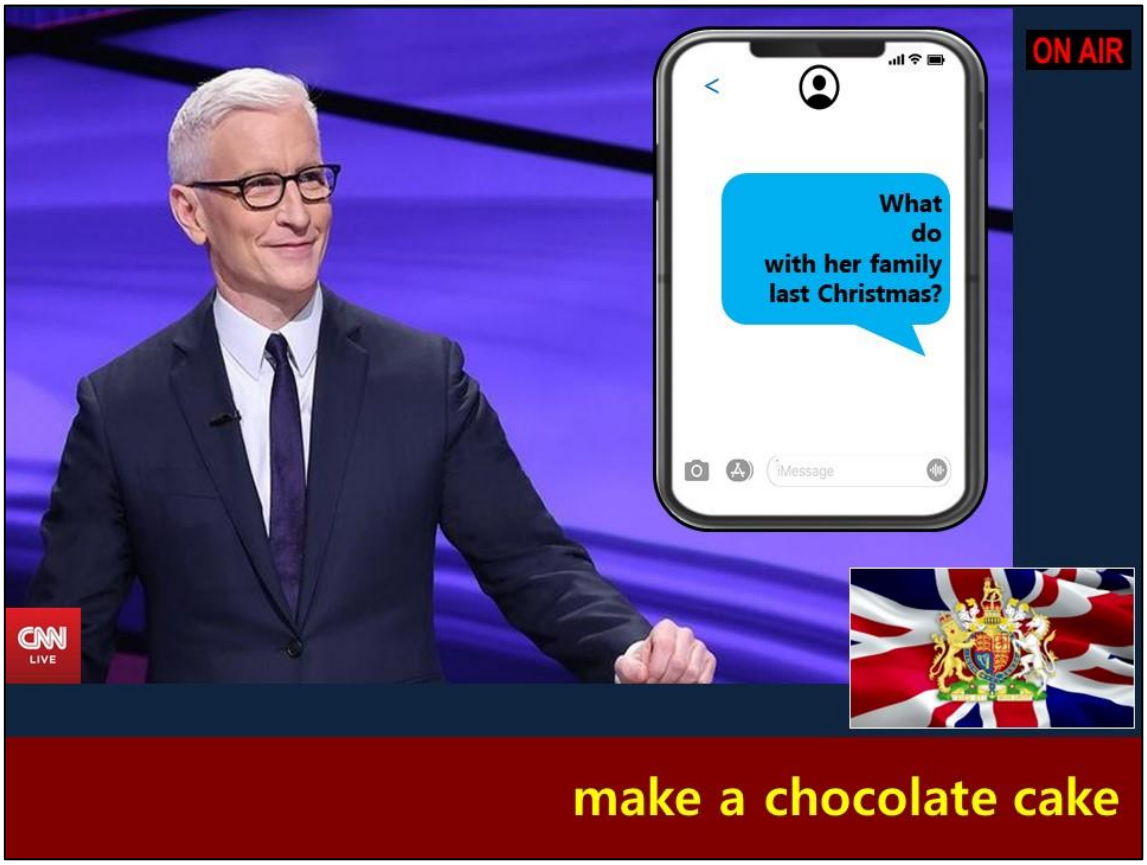

ON AIR
What
do
in her free time?
iMessage
CNN
LIVE
enjoy singing

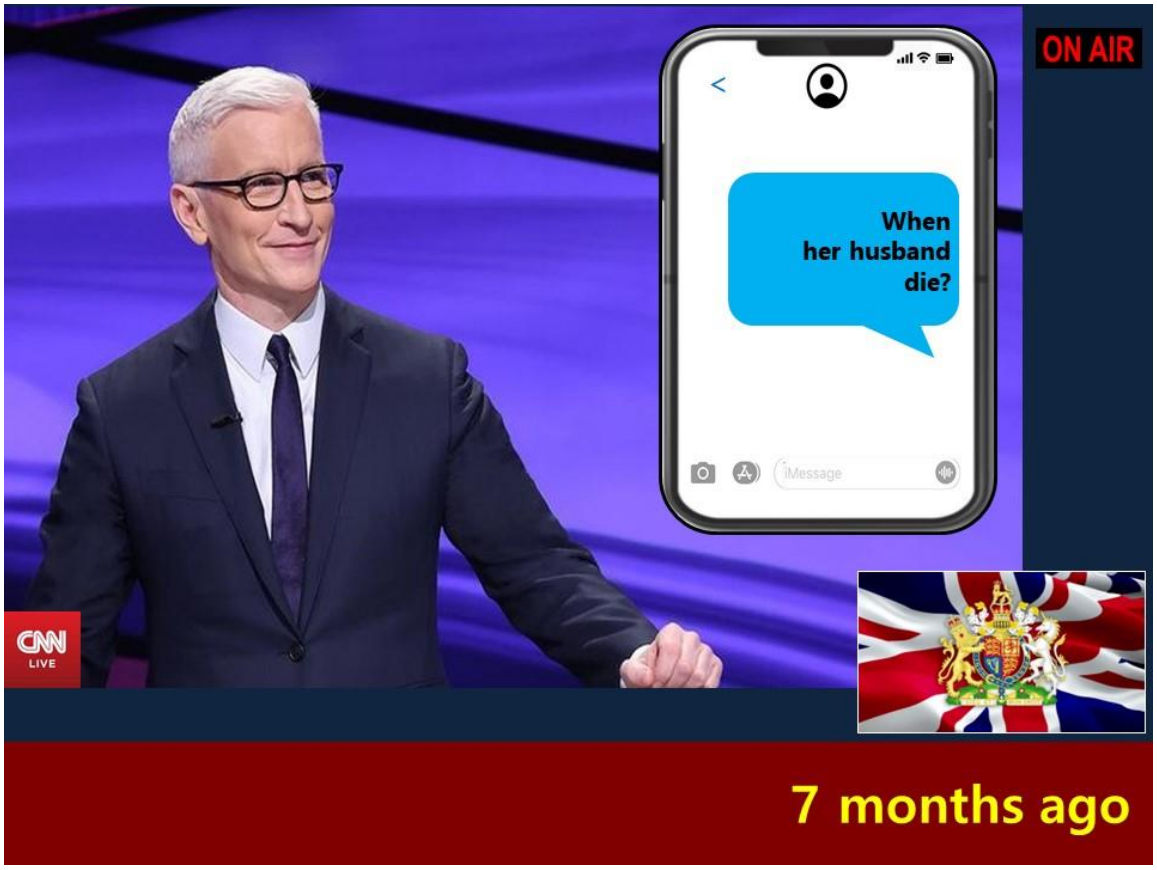

ON AIR
What
do
every night?
iMessage
CNN
LIVE
practice her speech


ON AIR
What do with her family last summer?
iMessage
CNN LIVE
plant roses

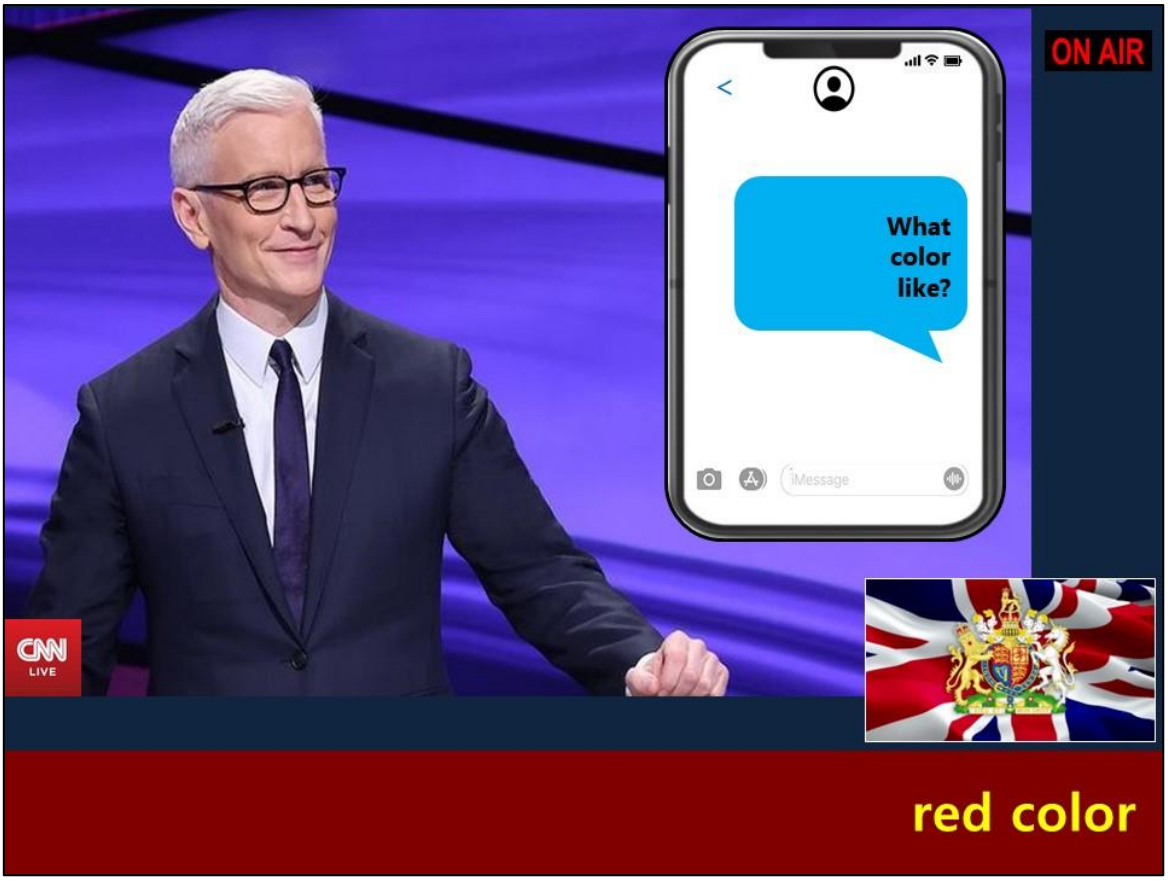

ON AIR
Where spend her holidays last year?
iMessage
CNN LIVE
in Scotland

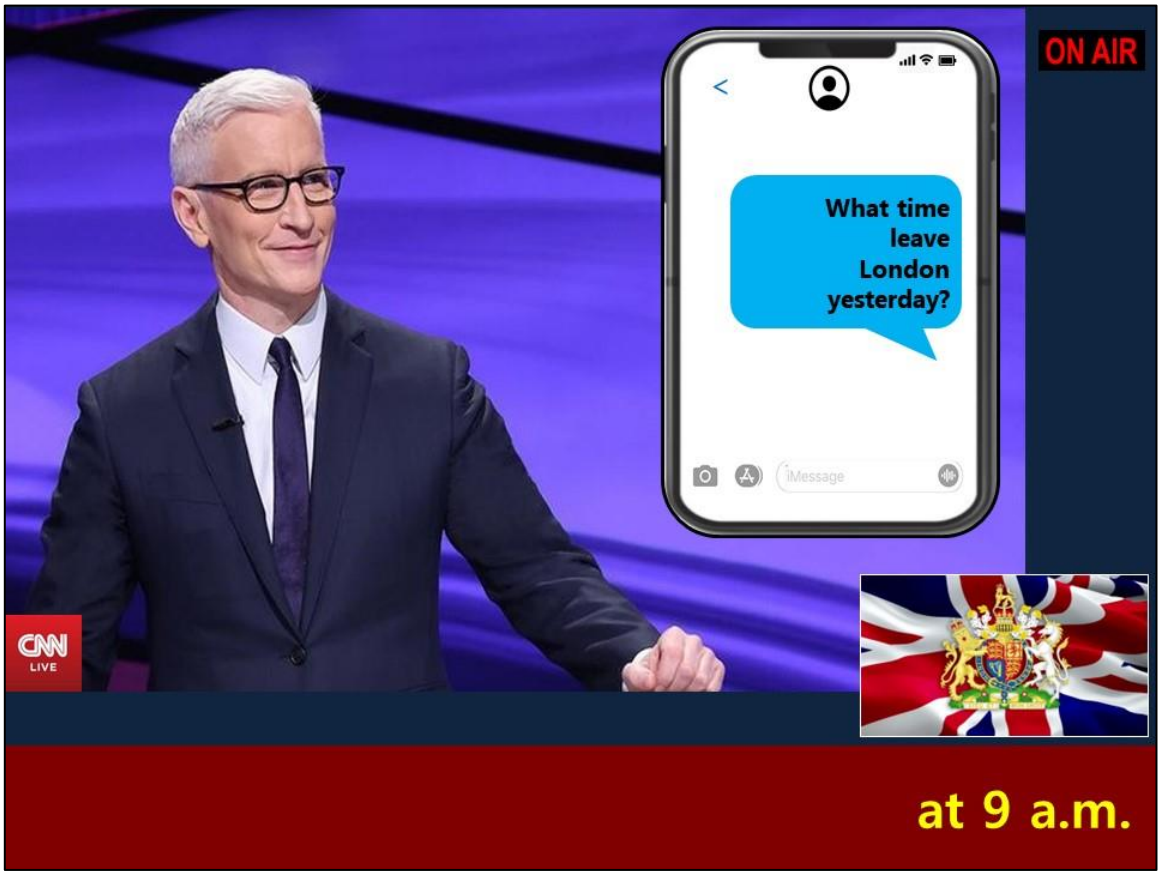

ON AIR
Who
miss
the most?
iMessage
CNN
LIVE
her husband


ON AIR
What
do
with her family
last Christmas?
iMessage
CNN
LIVE
make a chocolate cake

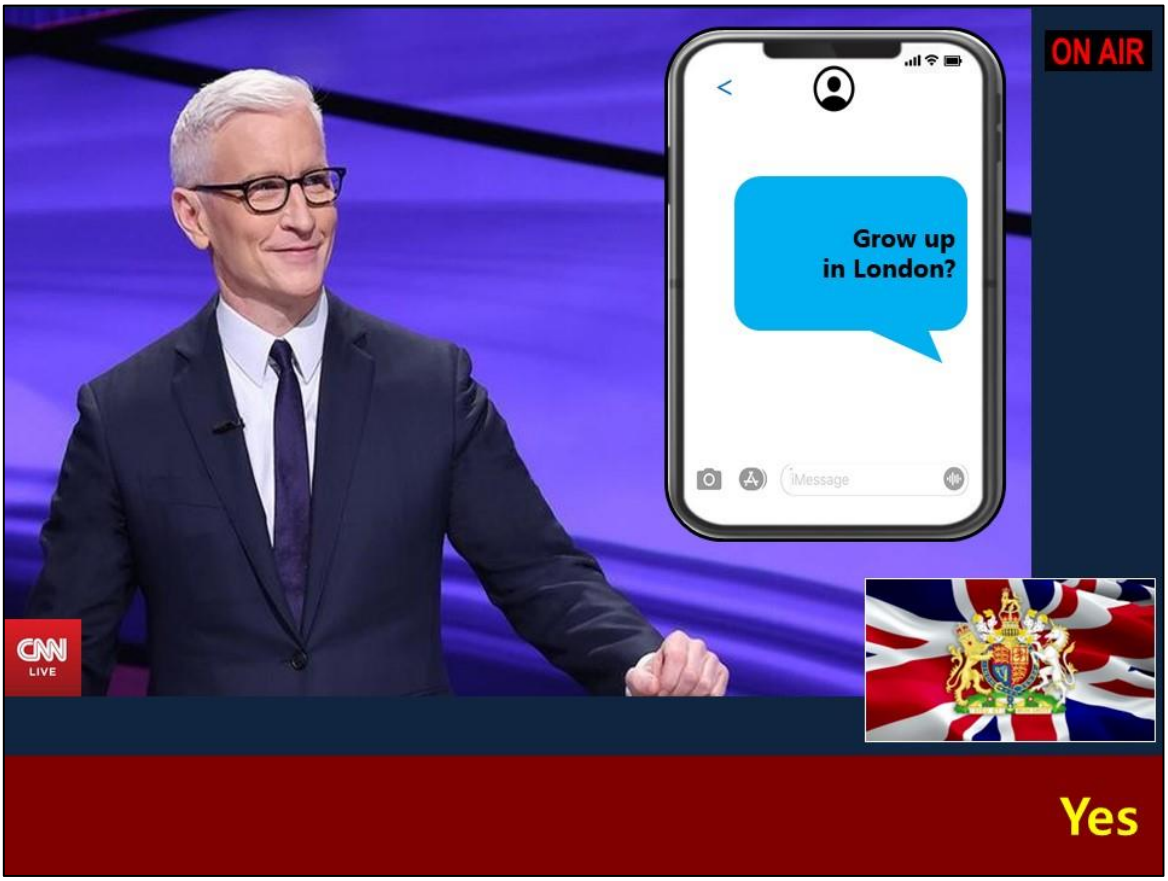

ON AIR
When her husband die?
iMessage
CNN LIVE
7 months ago


ON AIR
When get married?
iMessage
CNN LIVE
73 years ago

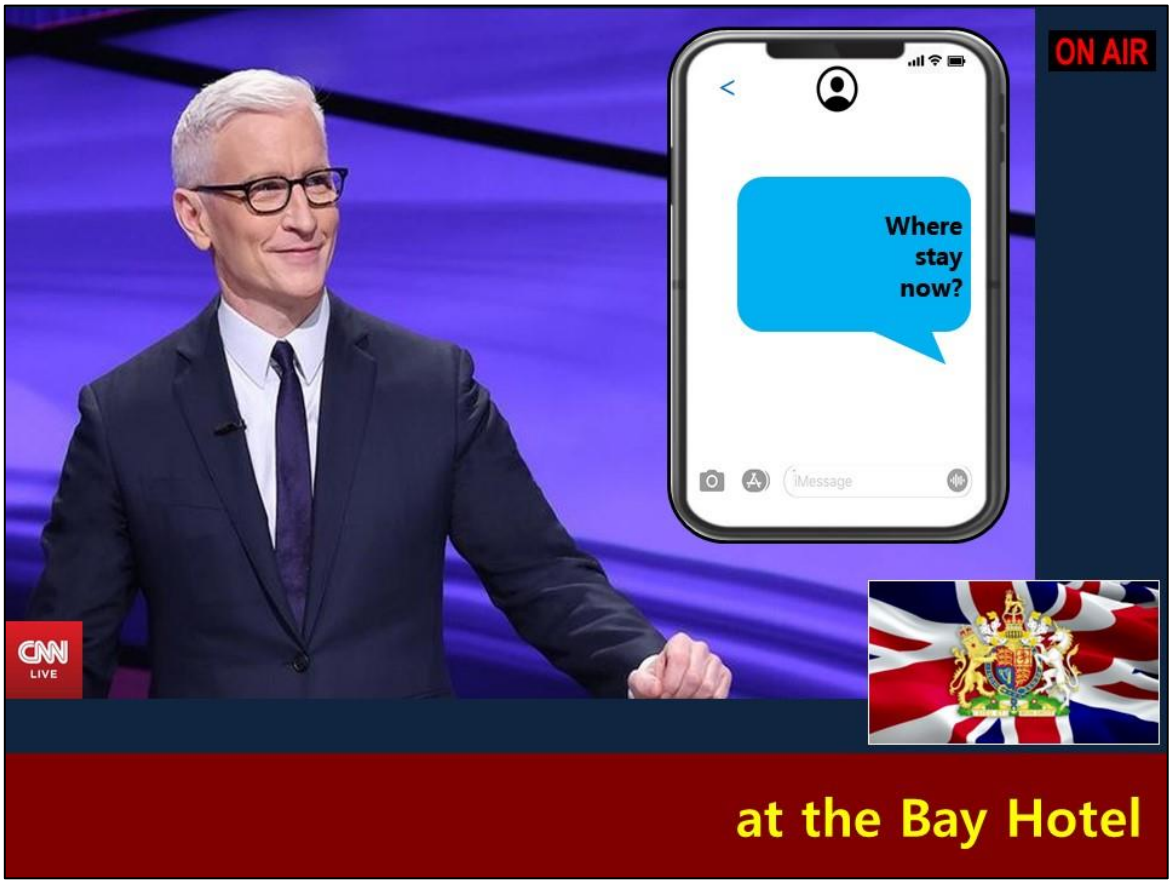

ON AIR
What color like?
iMessage
CNN LIVE
red color

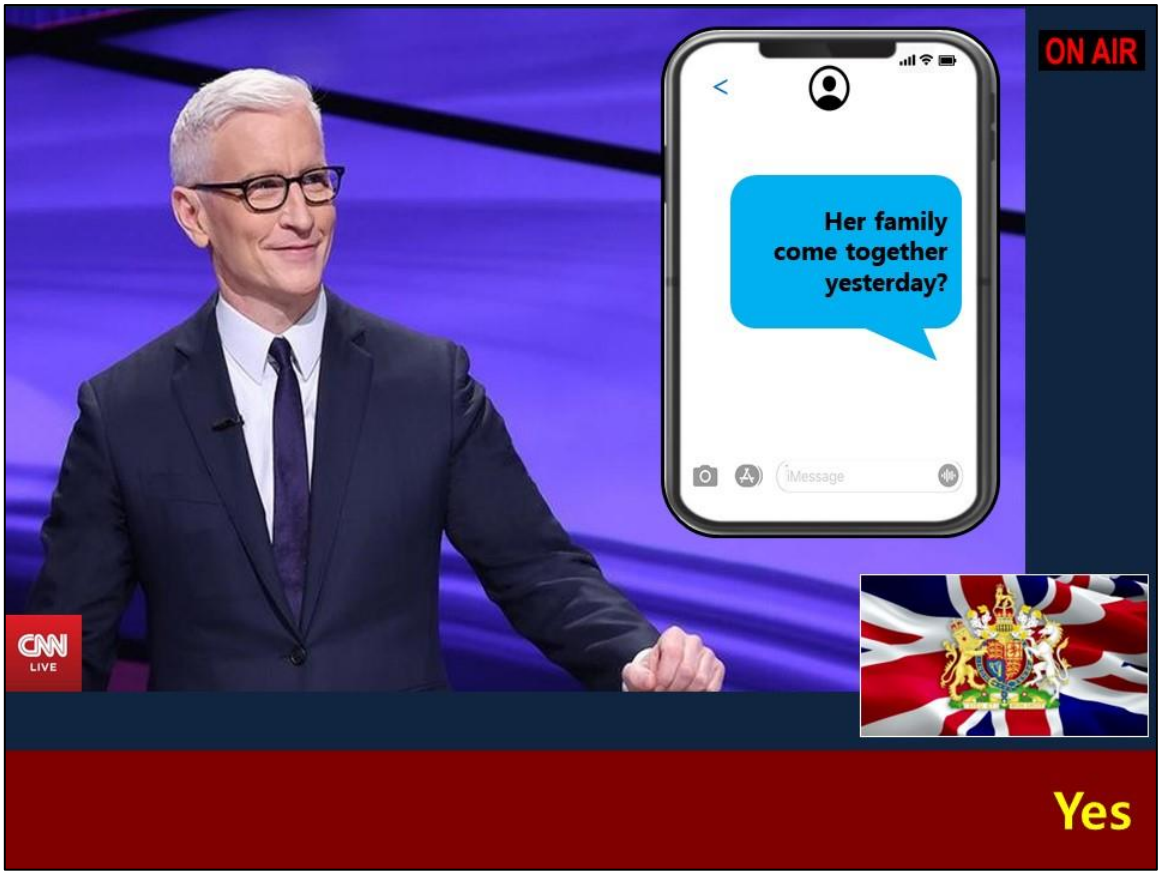

ON AIR
What time leave London yesterday?
iMessage
CNN LIVE
at 9 a.m.

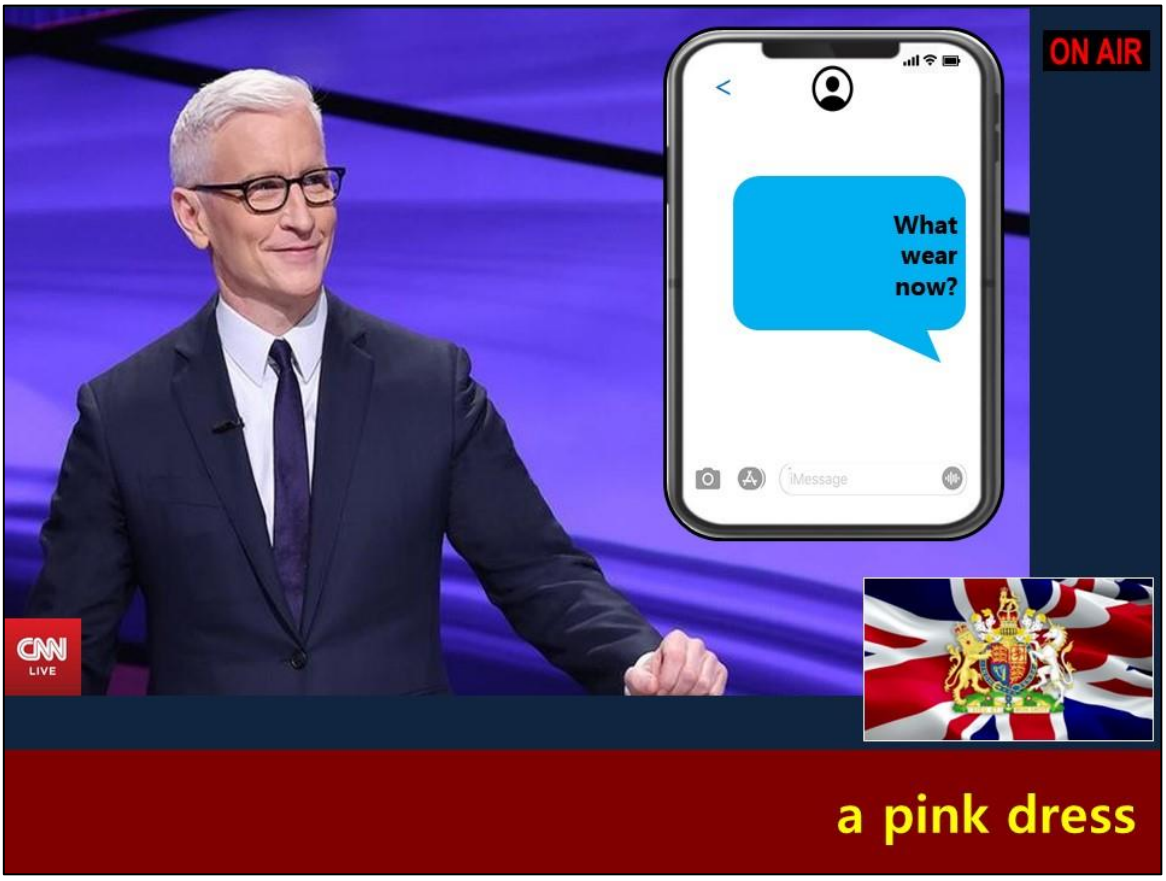

ON AIR
What do on the plane yesterday?
iMessage
CNN LIVE
listen to BBC news

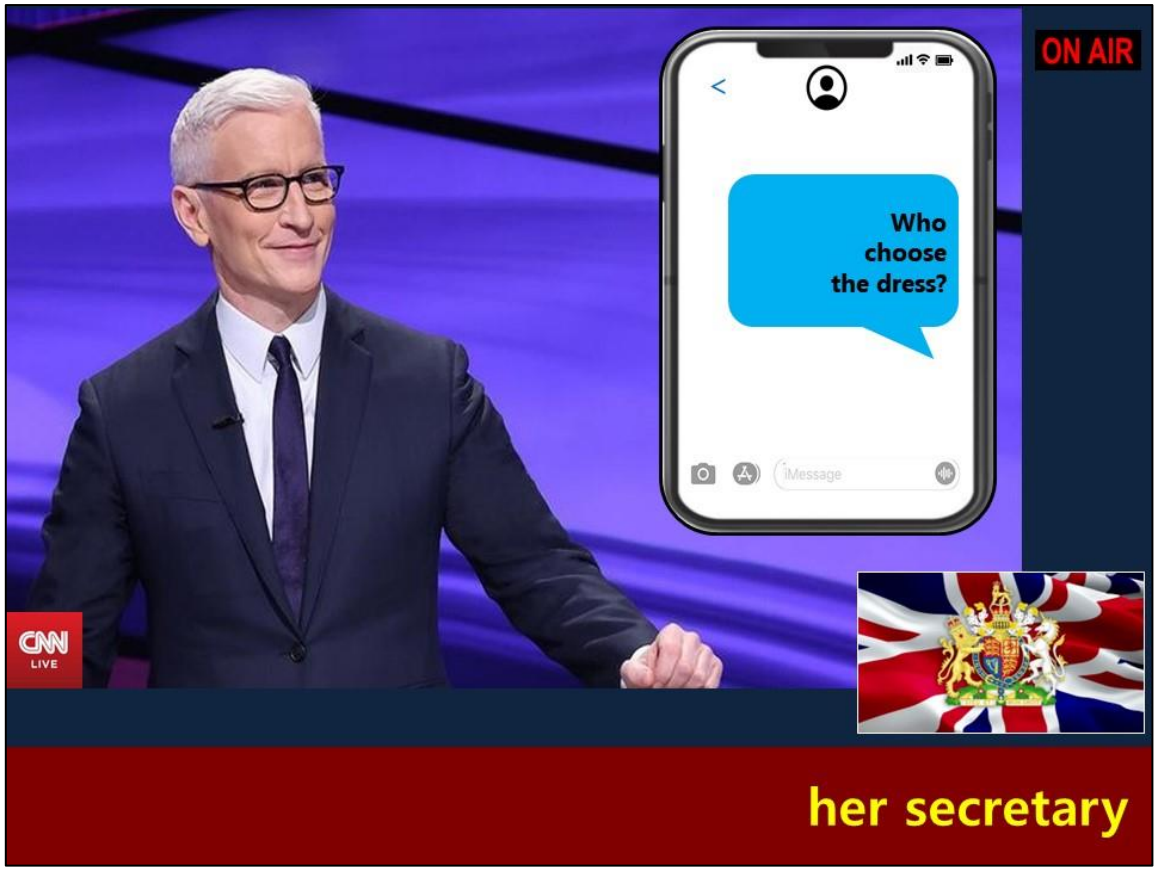

ON AIR
Grow up in London?
iMessage
CNN LIVE
Yes

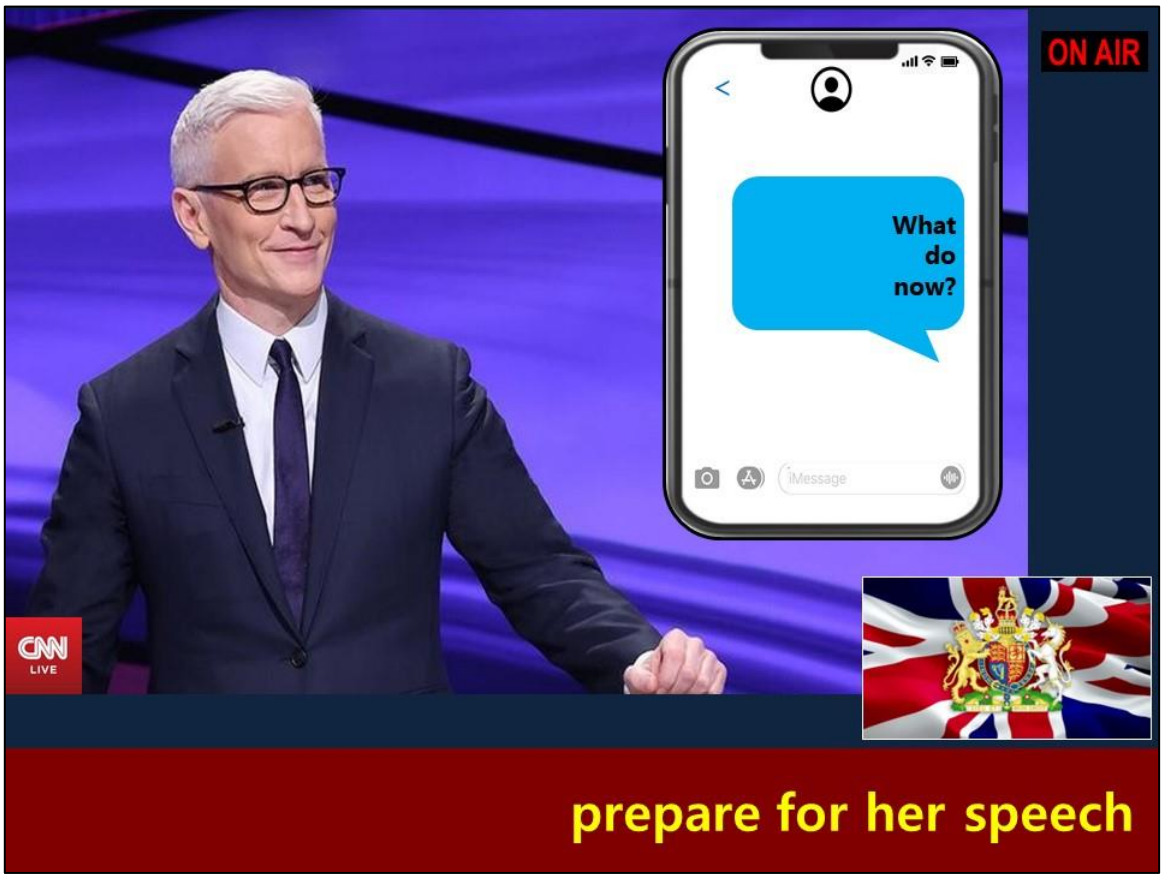

ON AIR
What time arrive in Cornwall yesterday?
iMessage
CNN LIVE
at 2 p.m.


ON AIR
Where stay now?
iMessage
CNN LIVE
at the Bay Hotel

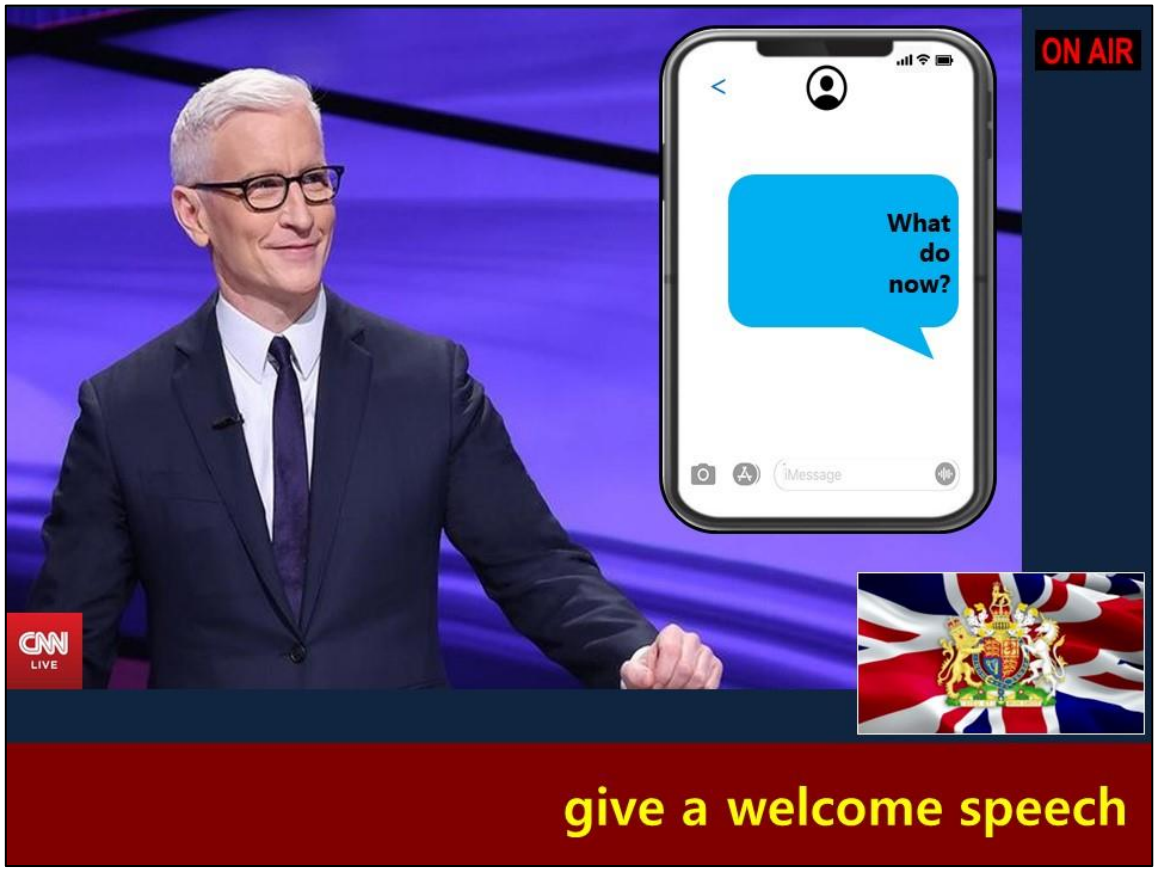

ON AIR
Her family come together yesterday?
iMessage
CNN LIVE
Yes

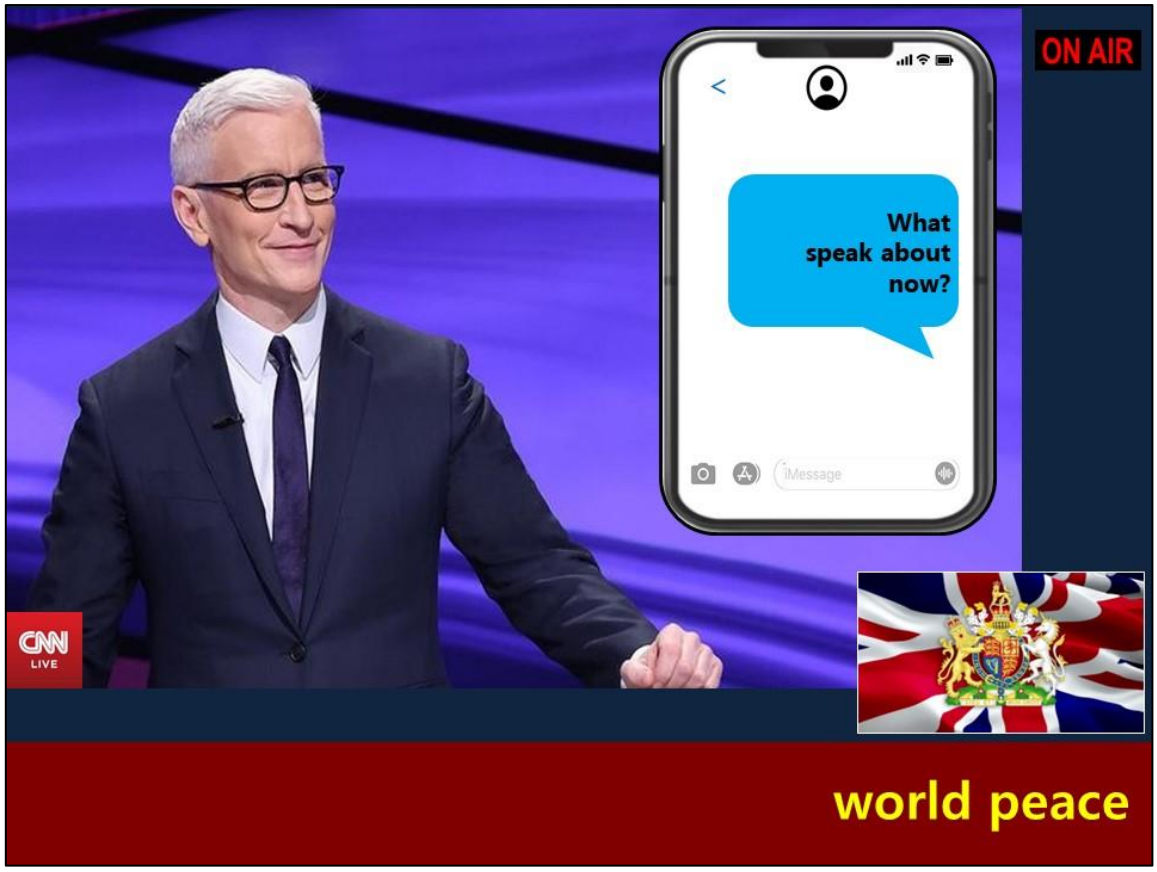

ON AIR
What wear now?
iMessage
CNN LIVE
a pink dress

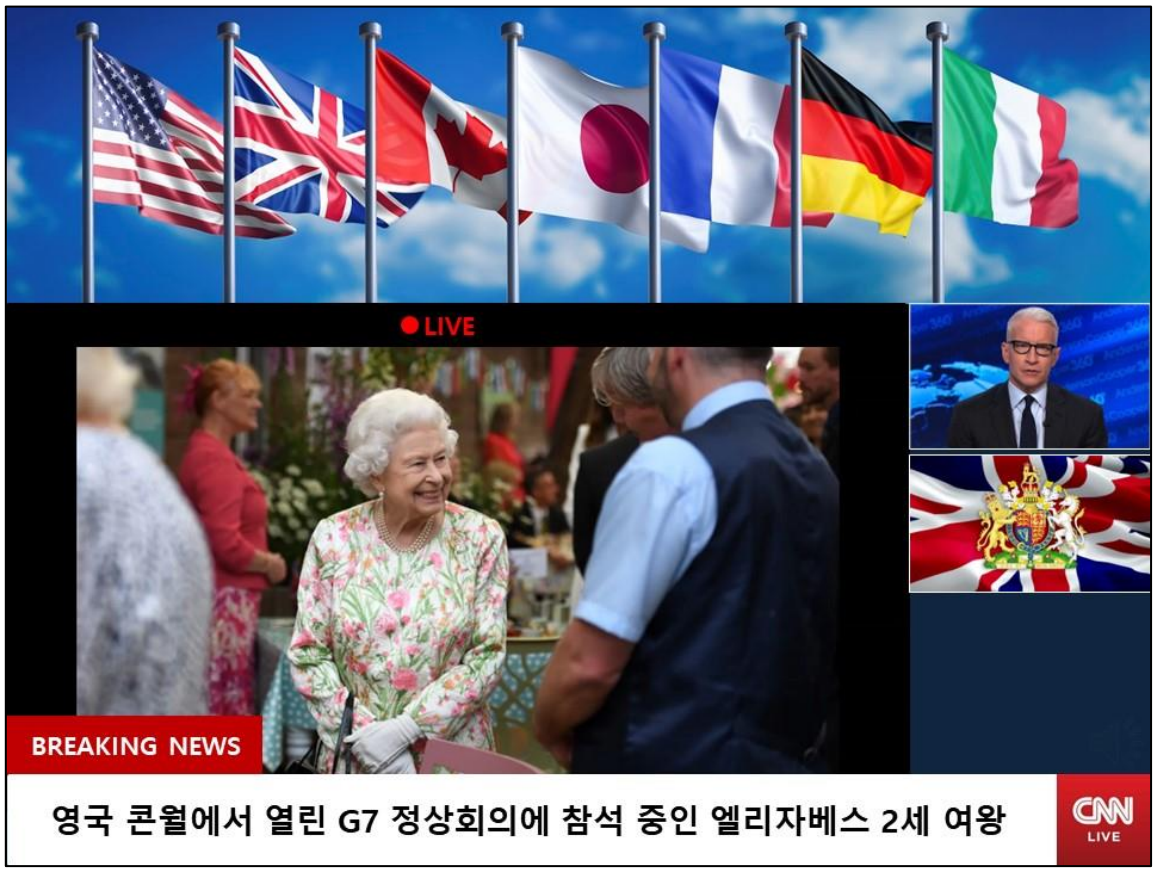

ON AIR
Who choose the dress?
iMessage
CNN LIVE
her secretary


ON AIR
What do now?
iMessage
CNN LIVE
prepare for her speech

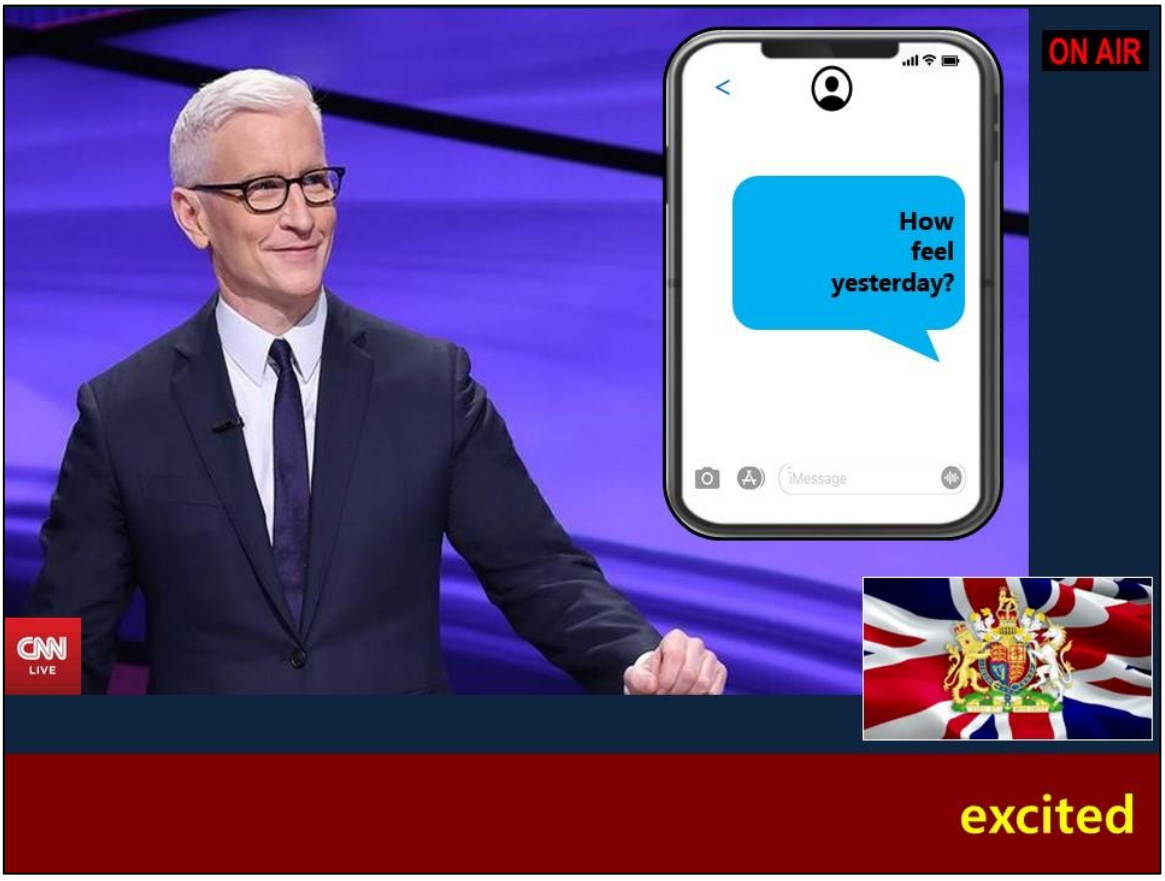

LIVE
BREAKING NEWS
G7 정상회의가 시작됐습니다.
CNN
LIVE

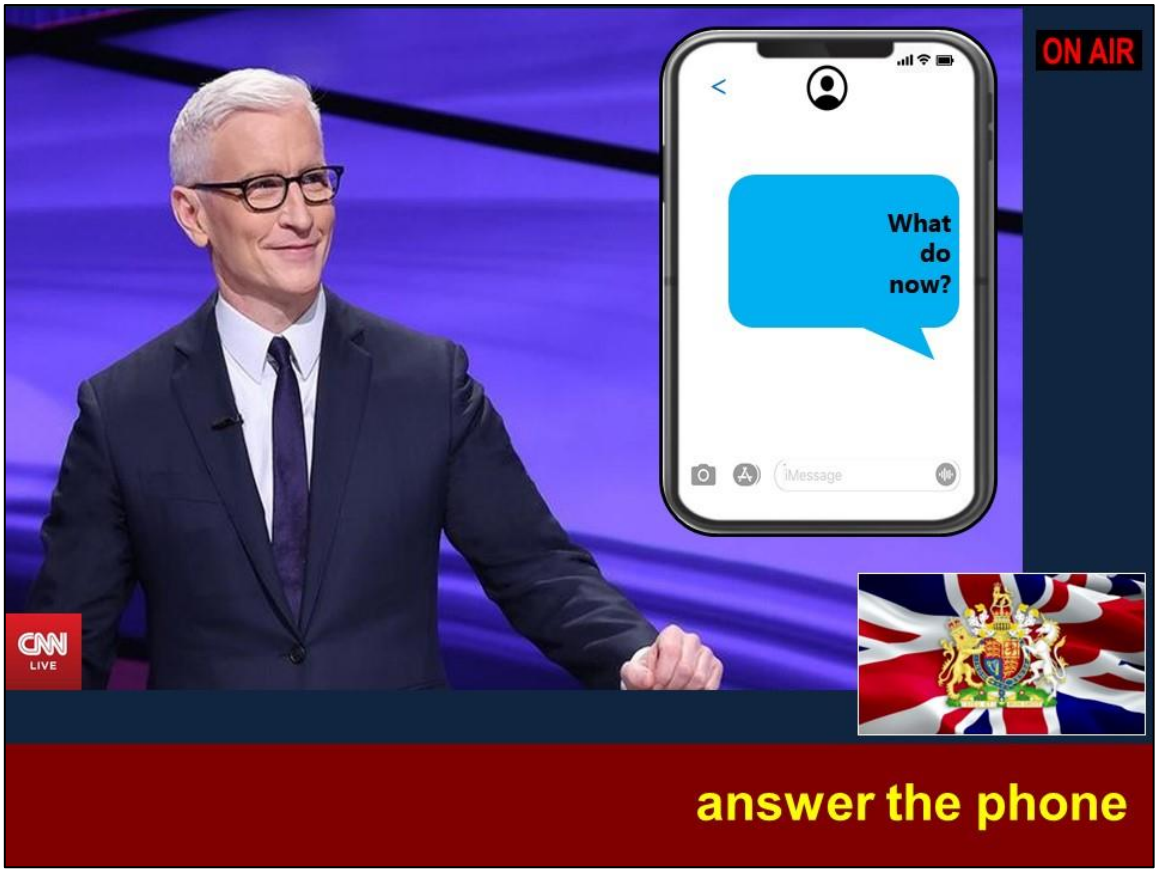

ON AIR
What do now?
iMessage
CNN
LIVE
give a welcome speech


ON AIR
What speak about now?
iMessage
CNN LIVE
world peace

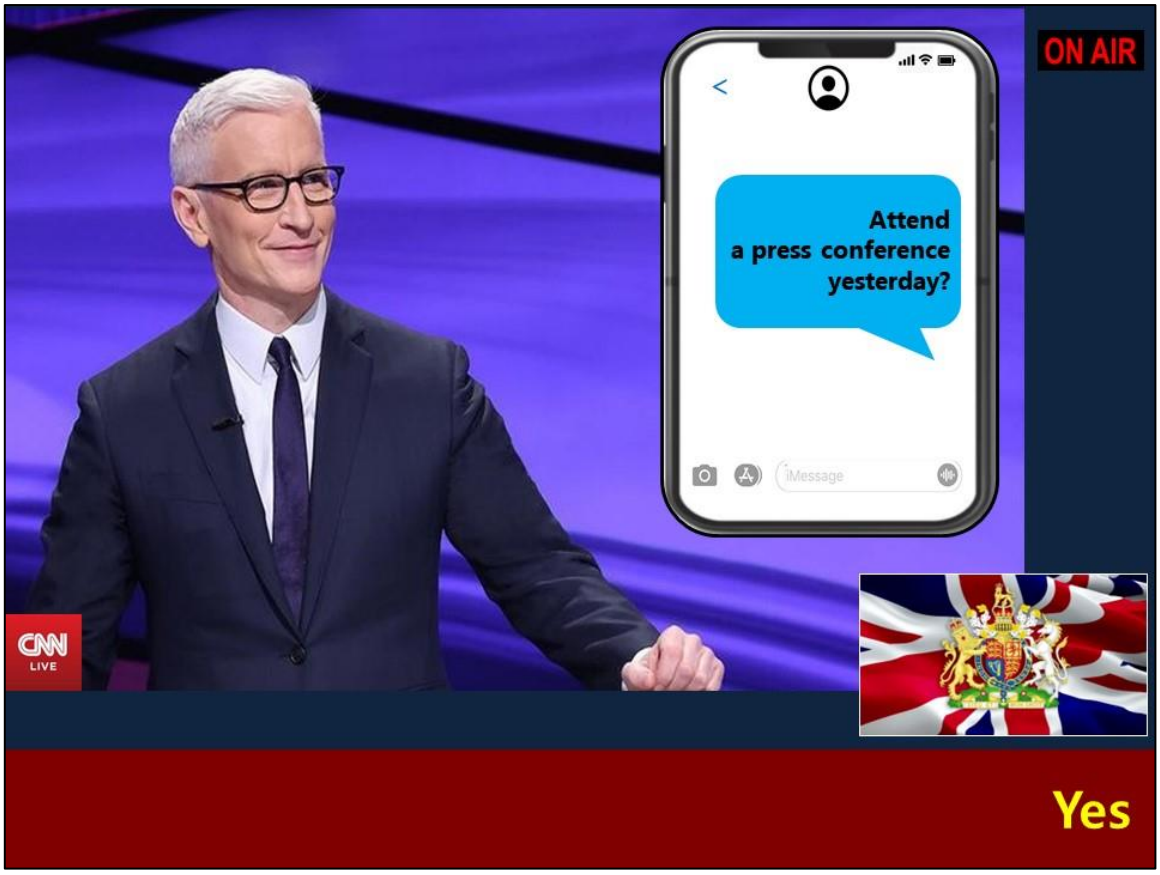

LIVE
BREAKING NEWS
영국 콘월에서 열린 G7 정상회의에 참석 중인 엘리자베스 2세 여왕
CNN LIVE

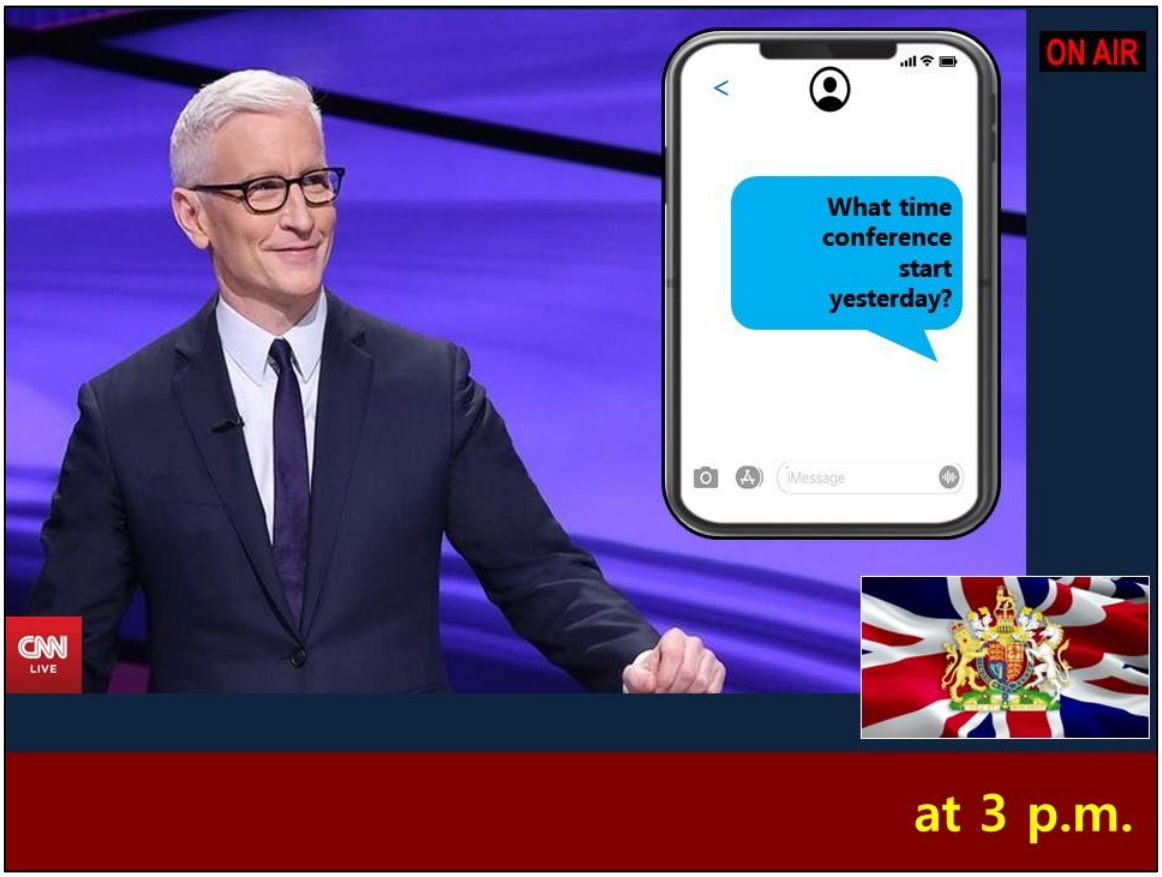

ON AIR
What do now?
iMessage
CNN LIVE
hold a tea party

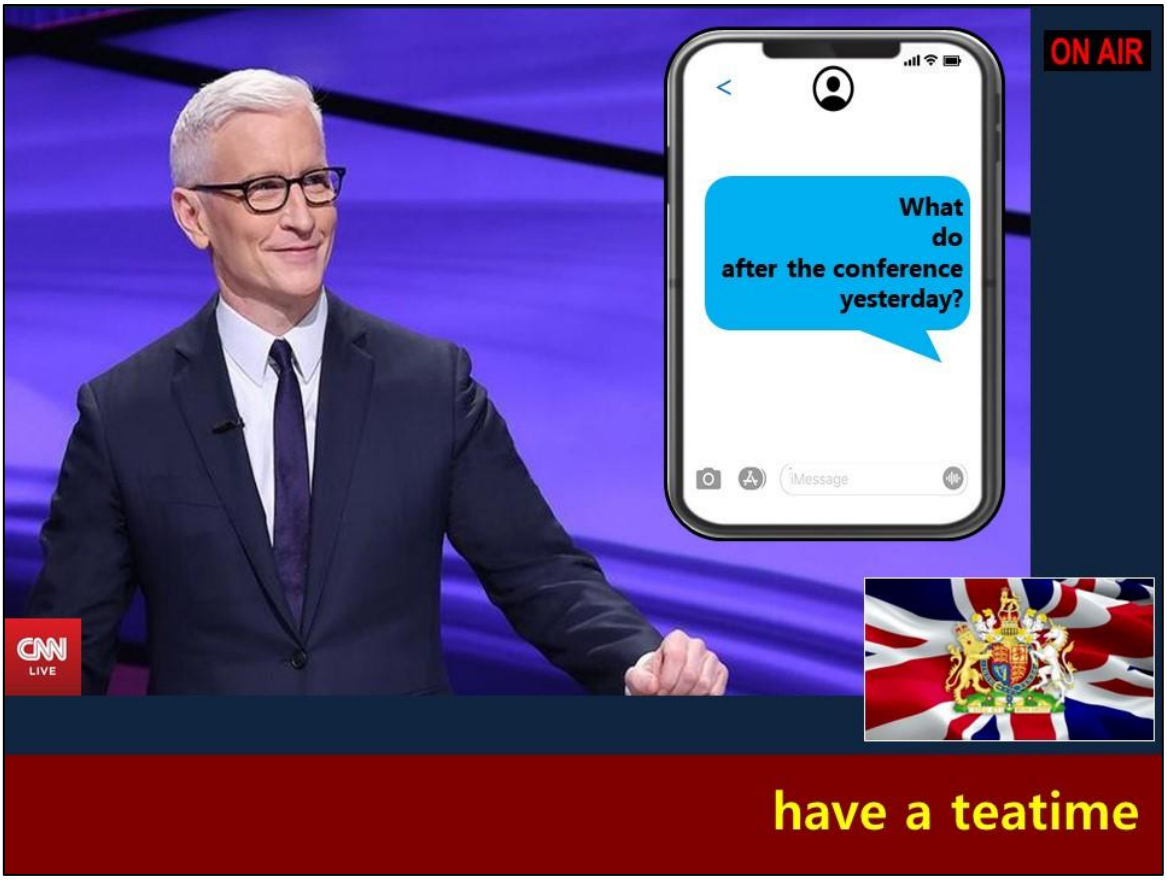

ON AIR
How feel yesterday?
iMessage
CNN LIVE
excited

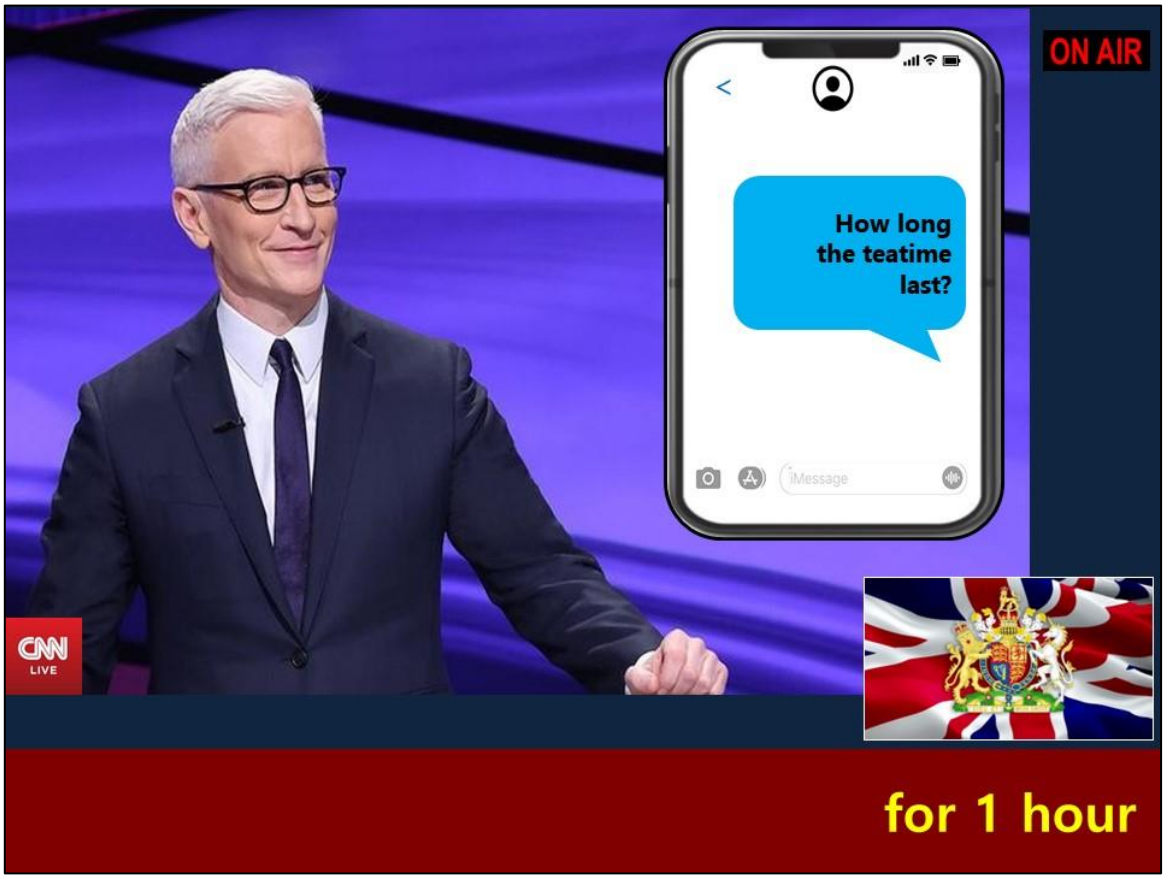

ON AIR
What do now?
iMessage
CNN LIVE
answer the phone


THE CORNISH ARMS
BREAKING NEWS
LIVE
G7 정상회의가 종료되었습니다.
CNN LIVE

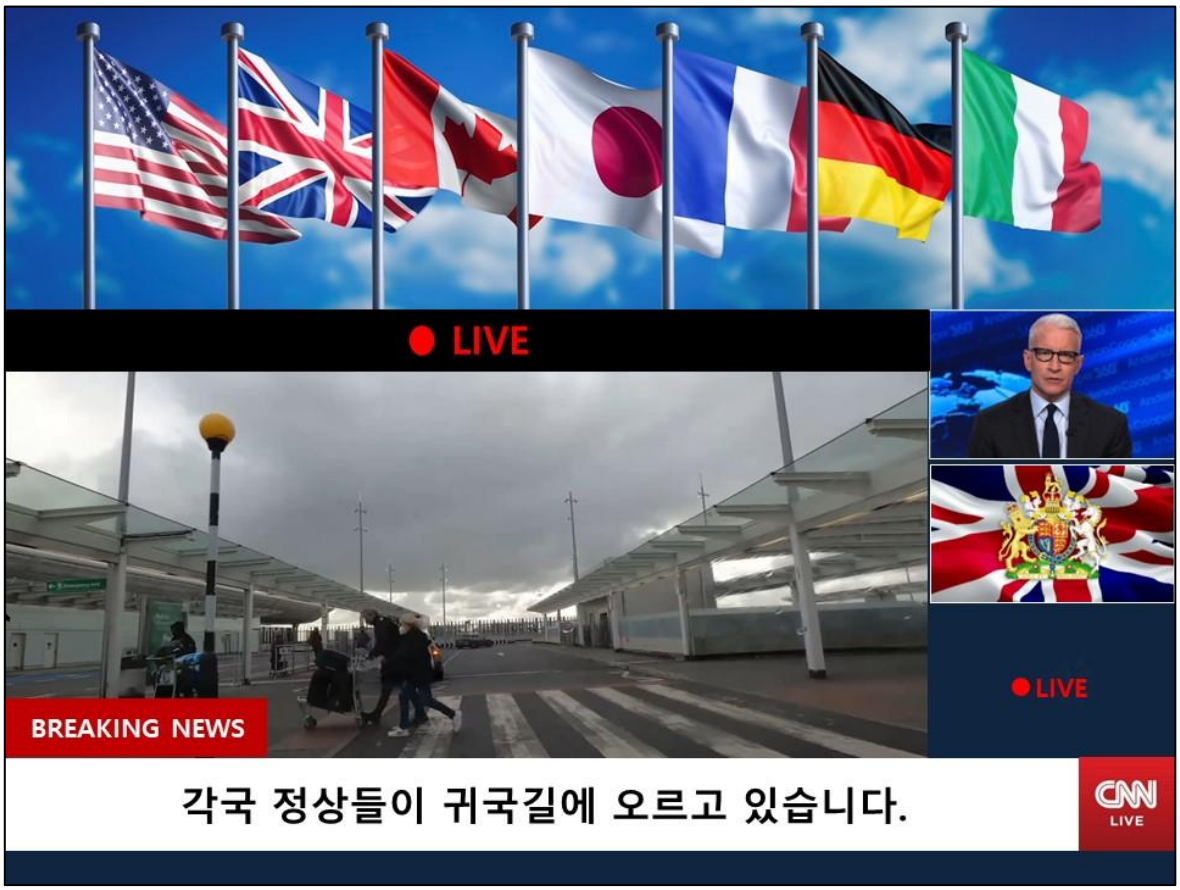

ON AIR
Attend a press conference yesterday?
iMessage
CNN LIVE
Yes


ON AIR
What time conference start yesterday?
iMessage
CNN LIVE
at 3 p.m.

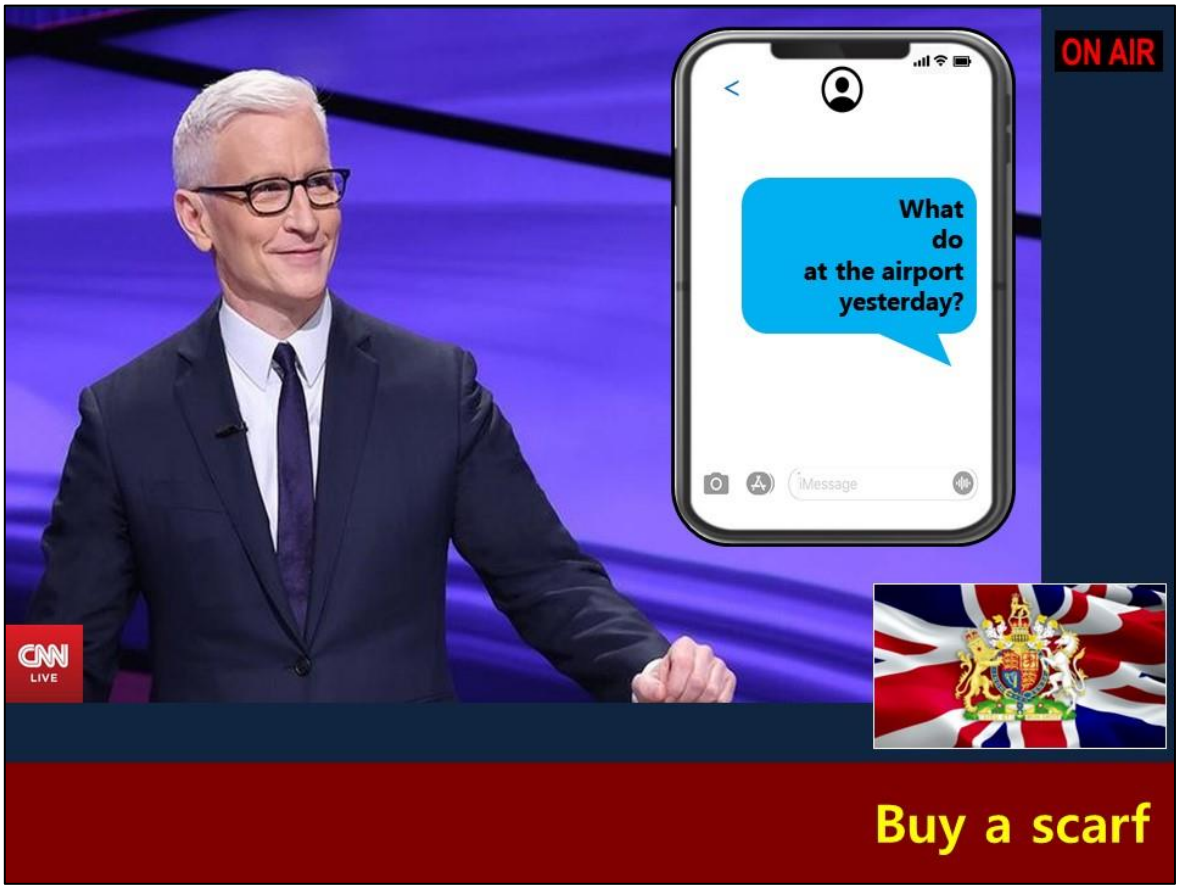

ON AIR
What do after the conference yesterday?
iMessage
CNN LIVE
have a teatime

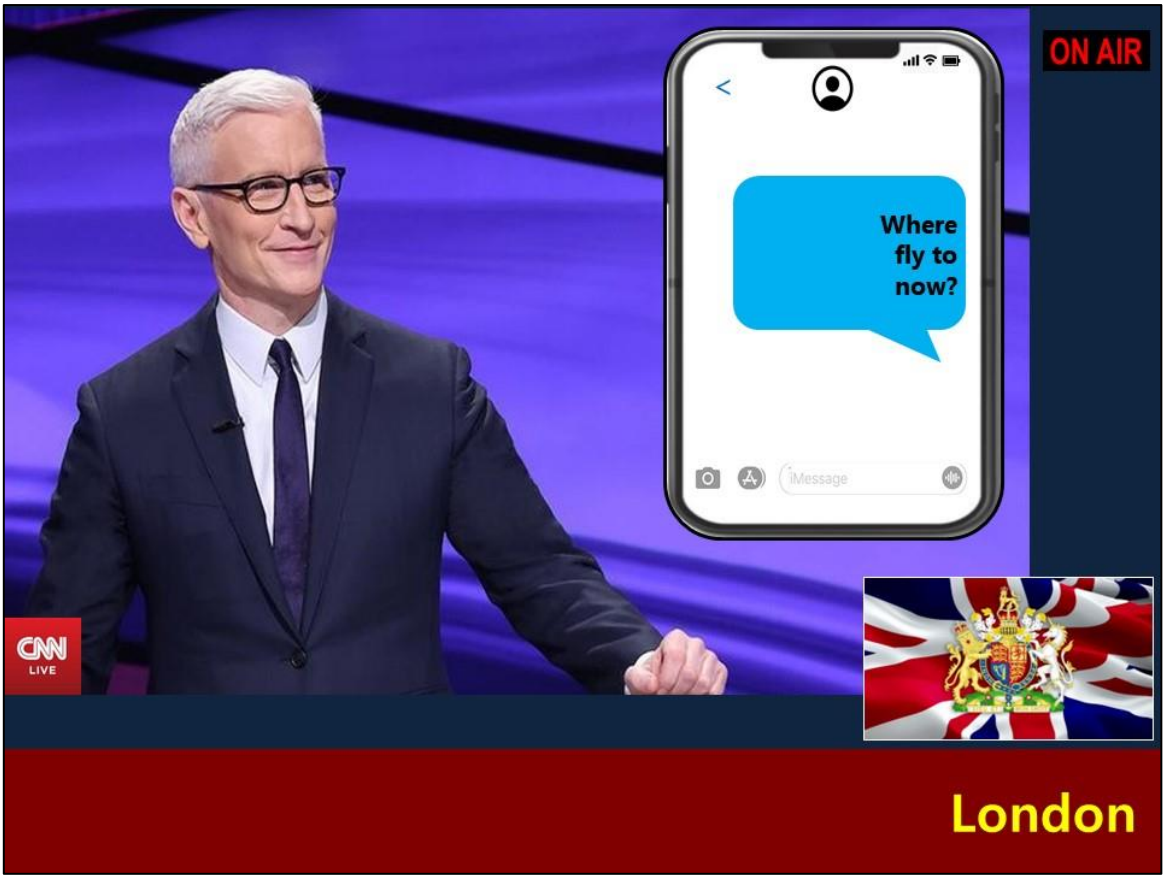

ON AIR
How long the teatime last?
iMessage
CNN LIVE
for 1 hour

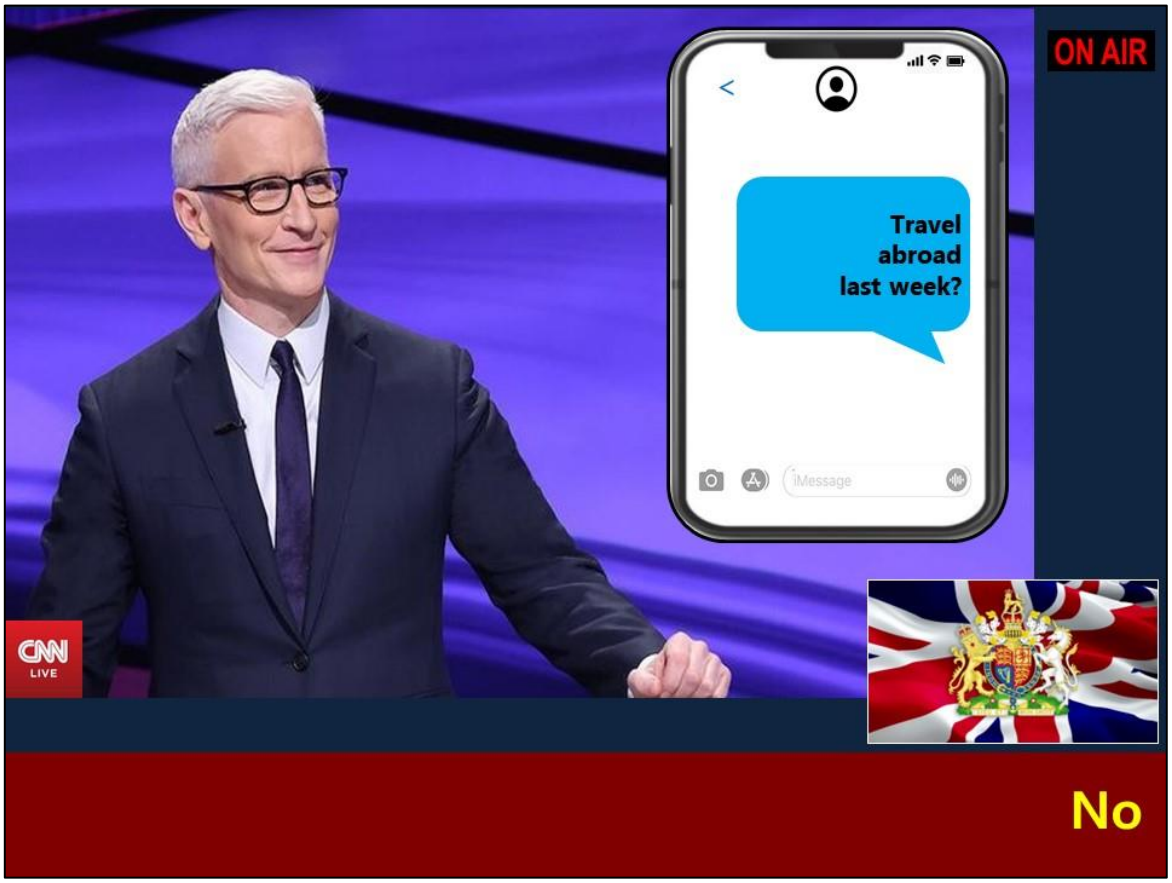

ON AIR
What time finish work yesterday?
iMessage
CNN LIVE
at 11 p.m.

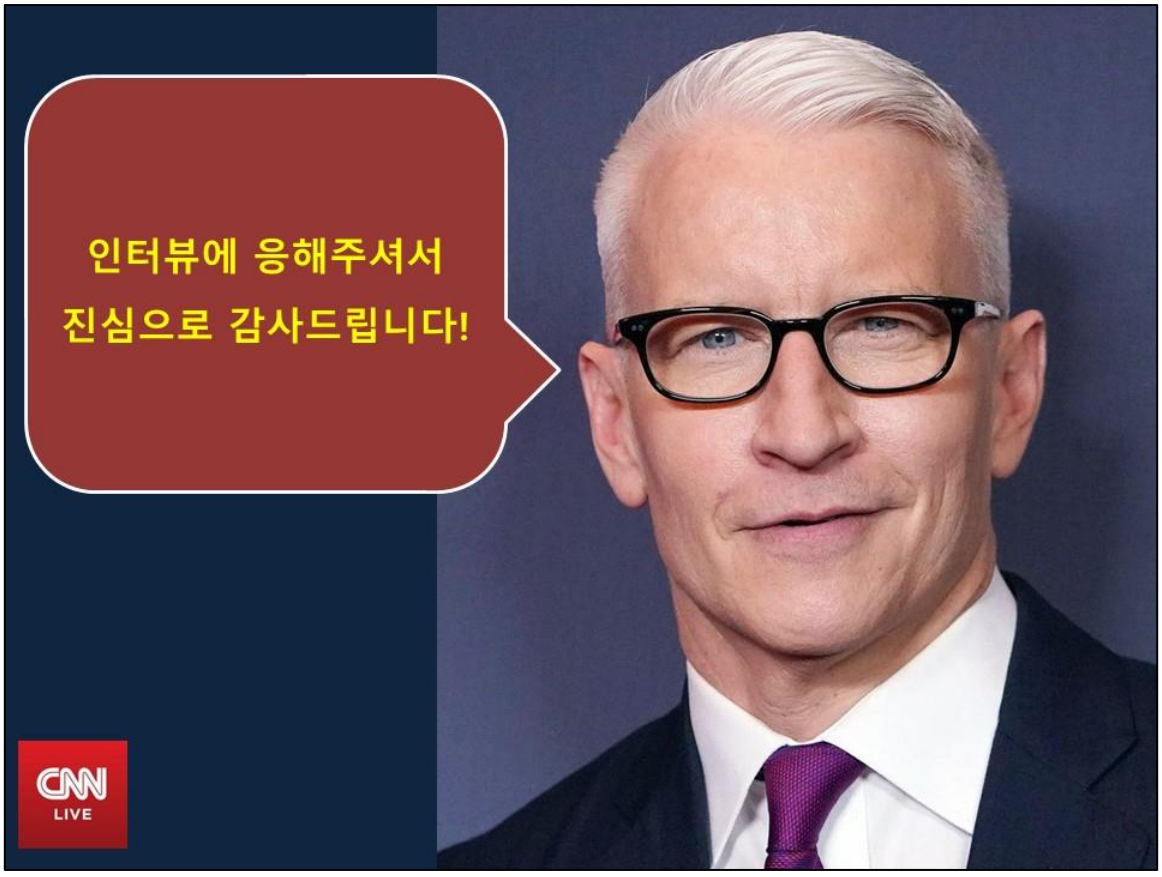

LIVE
LIVE
BREAKING NEWS
각국 정상들이 귀국길에 오르고 있습니다.
CNN LIVE


ON AIR
What do now?
iMessage
CNN LIVE
head to the airport


ON AIR
What do at the airport yesterday?
iMessage
CNN LIVE
Buy a scarf


ON AIR
Where fly to now?
iMessage
CNN LIVE
London


ON AIR
Travel abroad last week?
iMessage
CNN LIVE
No


인터뷰에 응해주셔서
진심으로 감사드립니다!
CNN
LIVE


BERLIN
LONDON
CNN
BREAKING
NEWS
JOHANNESBURG

Appendix A17: A Sample of WPT in Study 2






BBC Live Chat
BBC NEWS POLITICS
politics live


BREAKING NEWS
독일 메르켈 총리가 G7 정상회의 참석차 영국을 방문 중입니다.
BBC NEWS BREAKING


BBC
Live Chat
BREAKING NEWS
총리에 대해 궁금한 점이 있으신가요?
오늘 BBC 라이브 채팅에 참여하세요!
BBC NEWS BREAKING


ON AIR
환영합니다!
당신은
앙겔라 메르켈 독일 총리의
대변인입니다.
BBC NEWS BREAKING



DIRECTIONS:

당신은 앙겔라 메르켈 독일 총리의 대변인입니다.
총리는 G7 정상회의 참석차 영국을 공식 방문 중입니다.
오늘 당신은 총리를 대신해 BBC 라이브 채팅에 참여할 예정입니다.



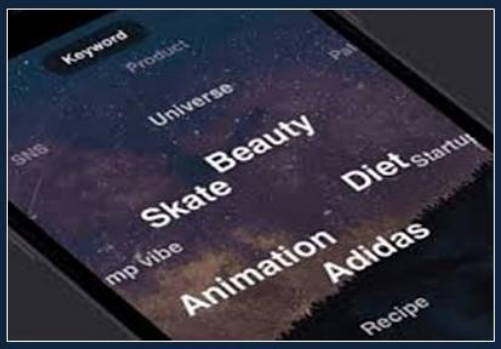


DIRECTIONS:

오늘 시청자들은 총리에 대해 궁금한 점을 실시간 채팅을 통해 질문할 수 있습니다.
시간 관계상 키워드로 된 짧은 질문만 가능합니다.
당신은 대변인으로서 시청자들의 질문에 답해야 합니다.

당신이 질문에 적절히 답변할 수 있도록 총리의 보좌관이 화면 하단에 정보를 제공해 줄 것입니다.
제공된 정보는 대변인인 당신에게만 보입니다.

BBC Live Chat



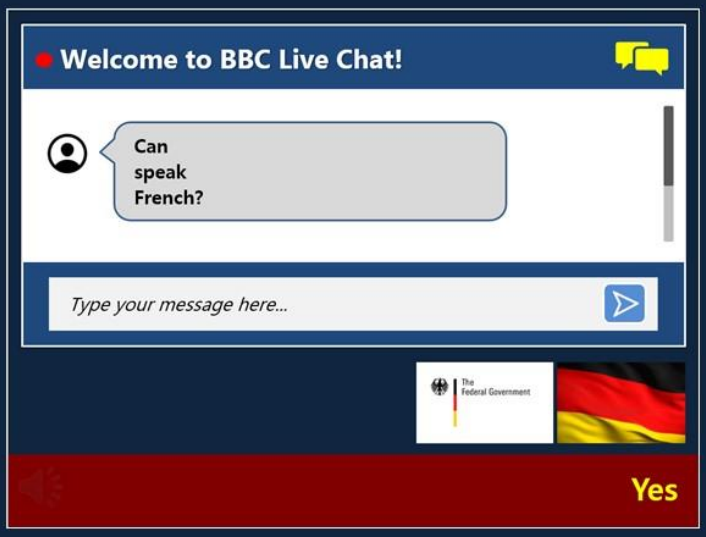
DIRECTIONS:

1) 총 30개의 질문이 주어집니다. 질문은 화면에 제시될 것입니다.



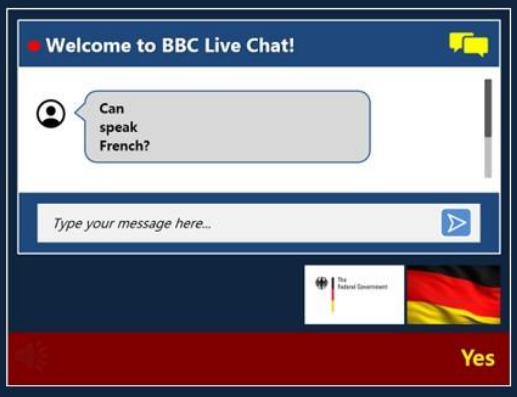
DIRECTIONS:

1) 총 30개의 질문이 주어집니다. 질문은 화면에 제시될 것입니다.
2) 주어진 정보를 바탕으로 채팅창에 답변을 입력하세요.



DIRECTIONS:

1) 시청자들의 질문은 화면에 제시될 것입니다.
2) 주어진 정보를 바탕으로 채팅창에 답변을 입력하세요.
3) 공식 석상이므로, 반드시 주어진 단어를 사용해 완성형 문장으로 답하세요.



Yes. ✗

Yes, she can. ✗

Yes, she can speak French. ○

Let's begin!

BBC Live Chat



Welcome to BBC Live Chat!

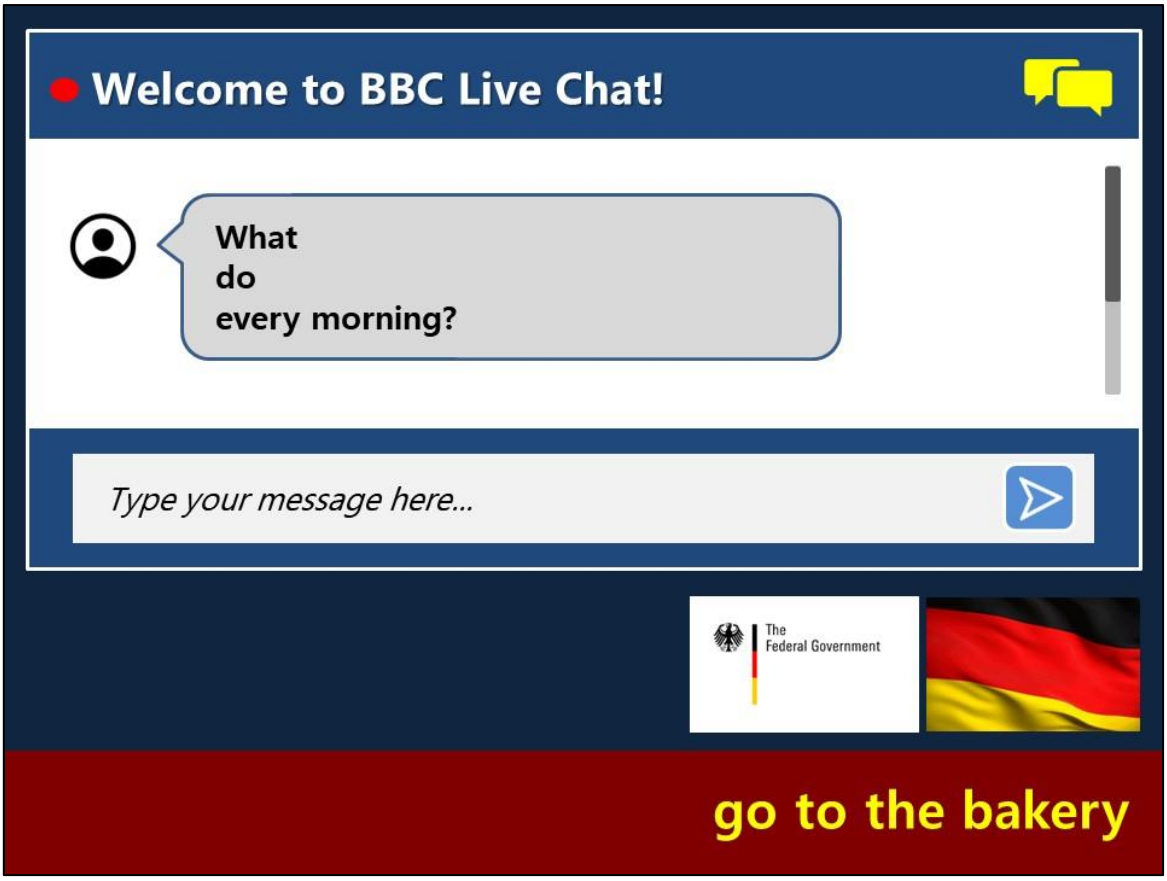
What time
get up
every day?

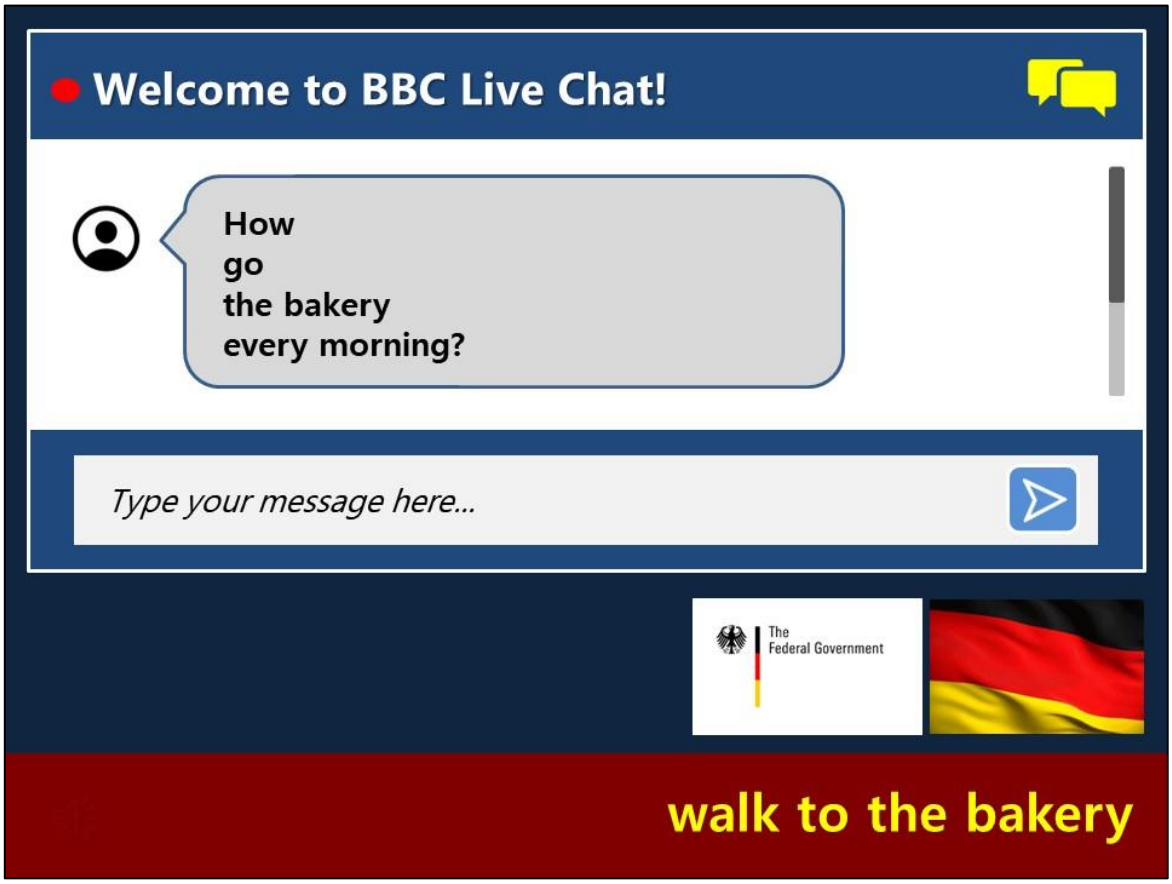
Type your message here...

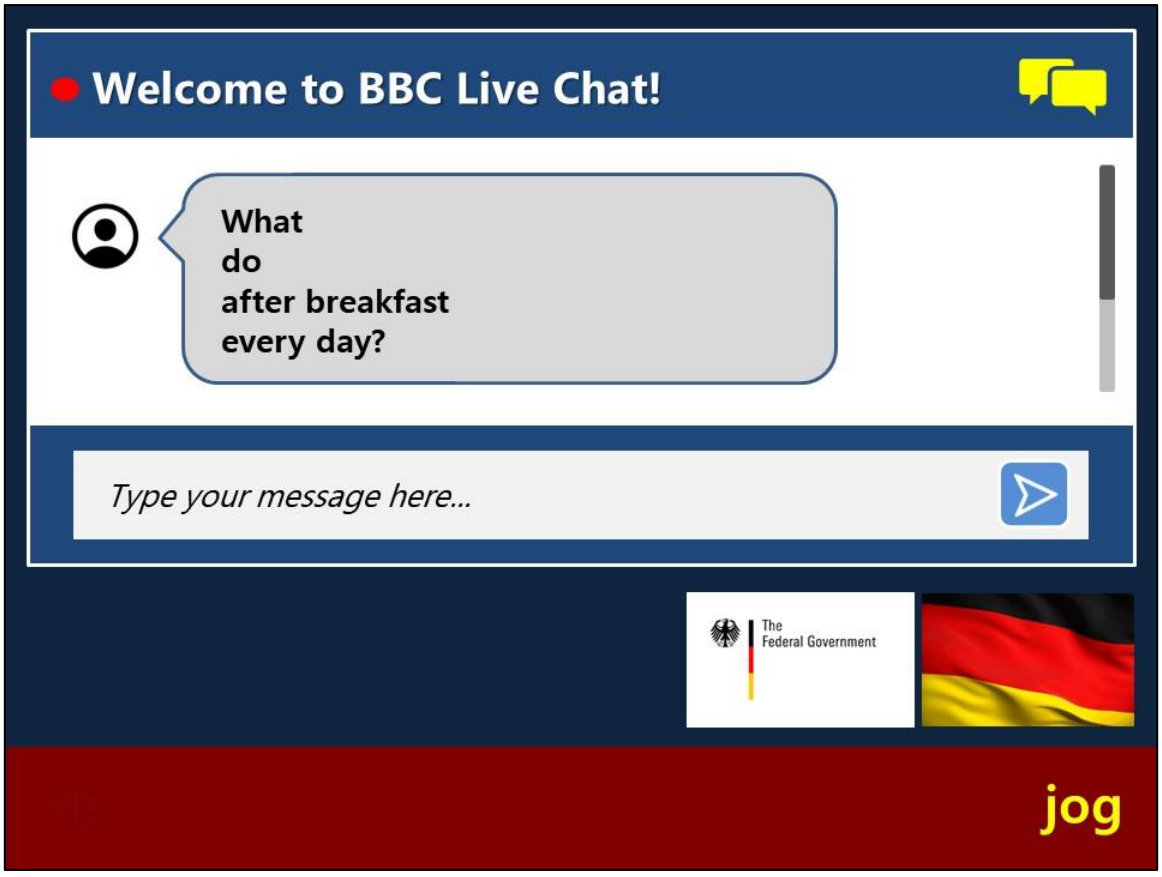
The Federal Government

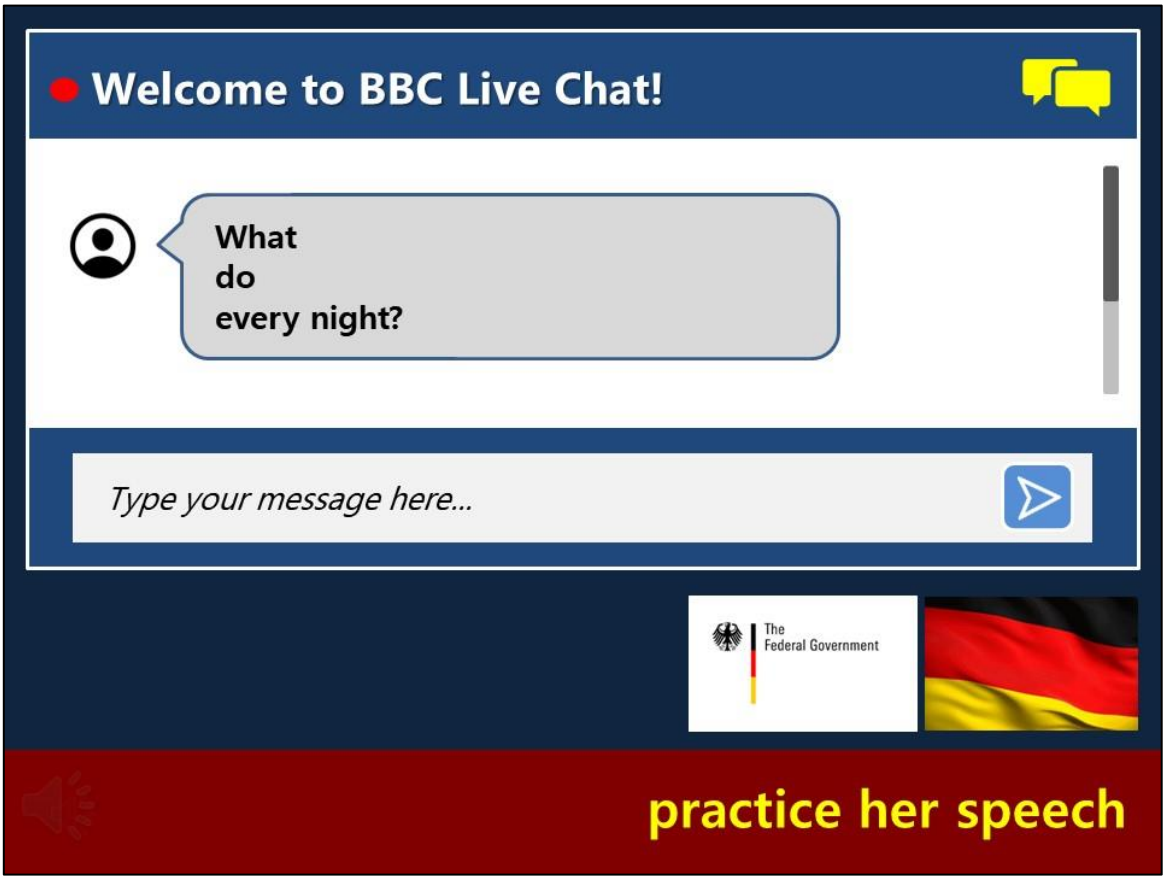


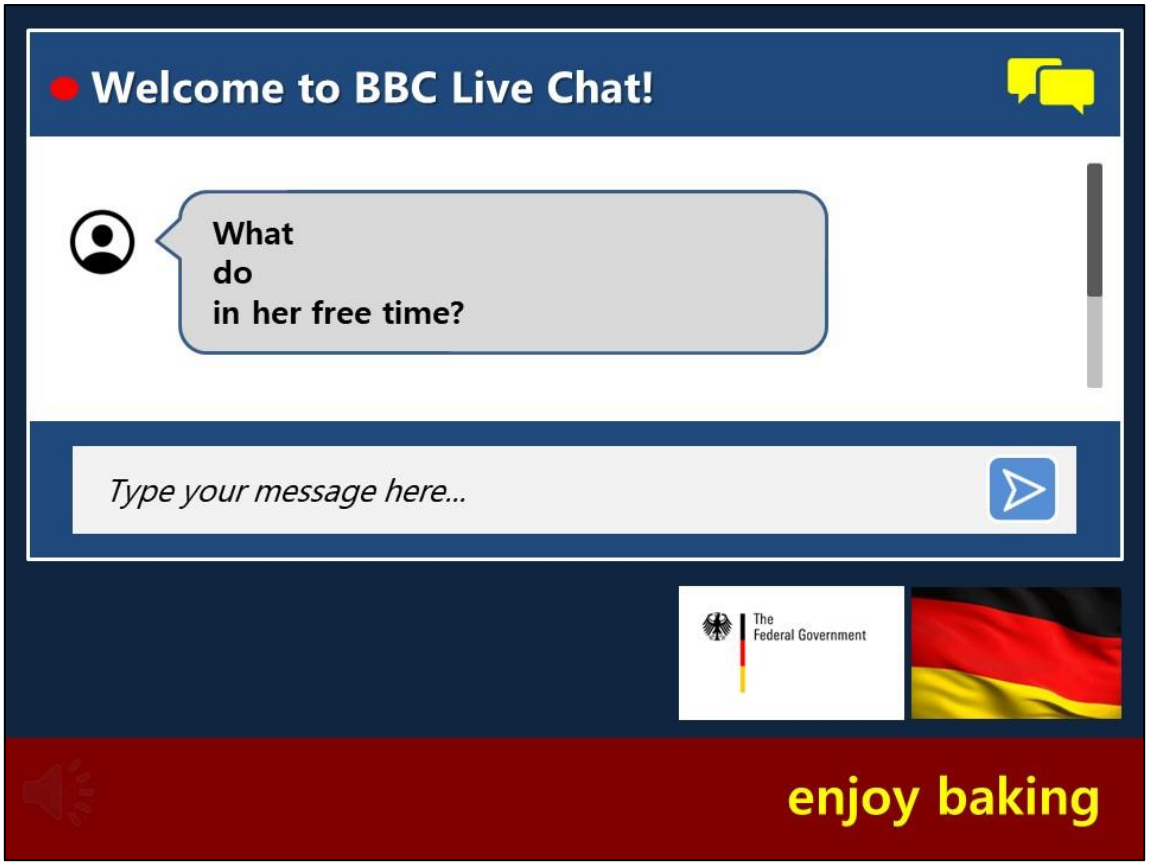
at 6 a.m.

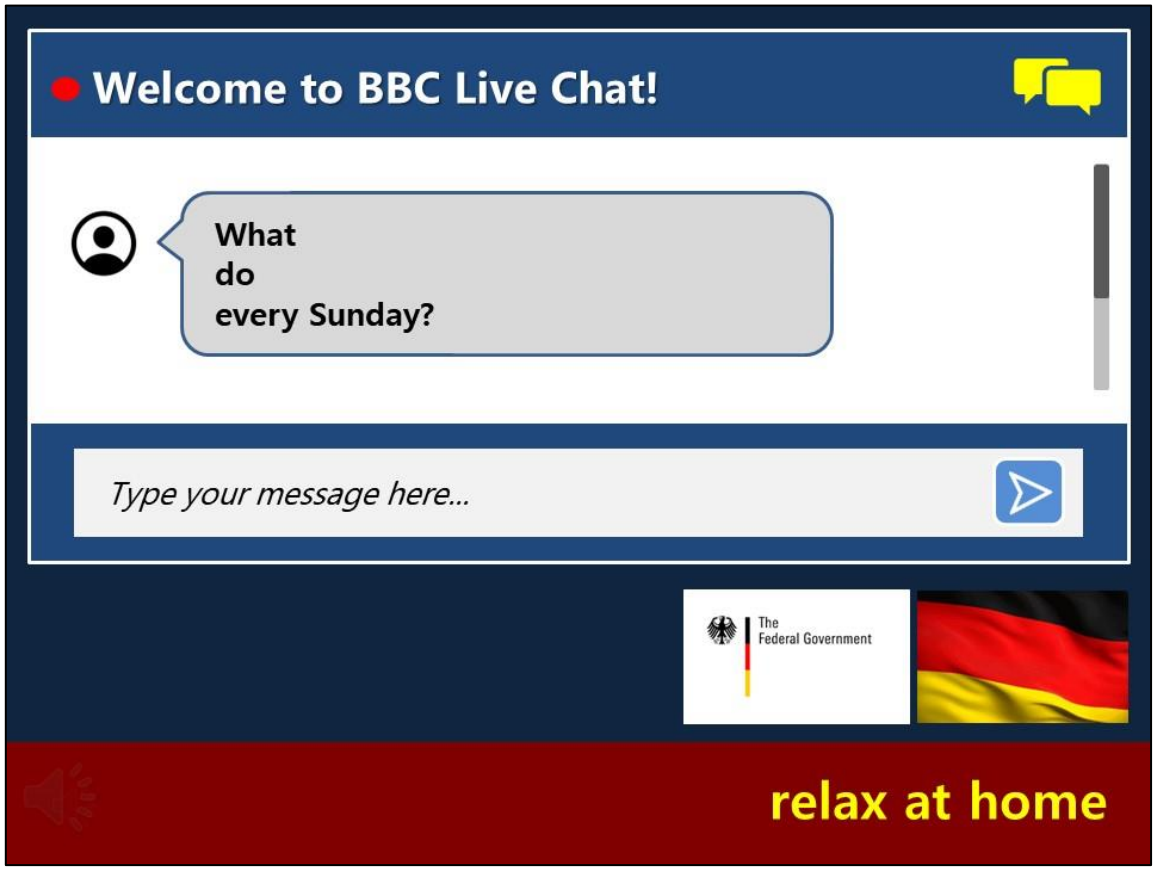

Welcome to BBC Live Chat!
What
do
every morning?
Type your message here...
The
Federal Government
go to the bakery

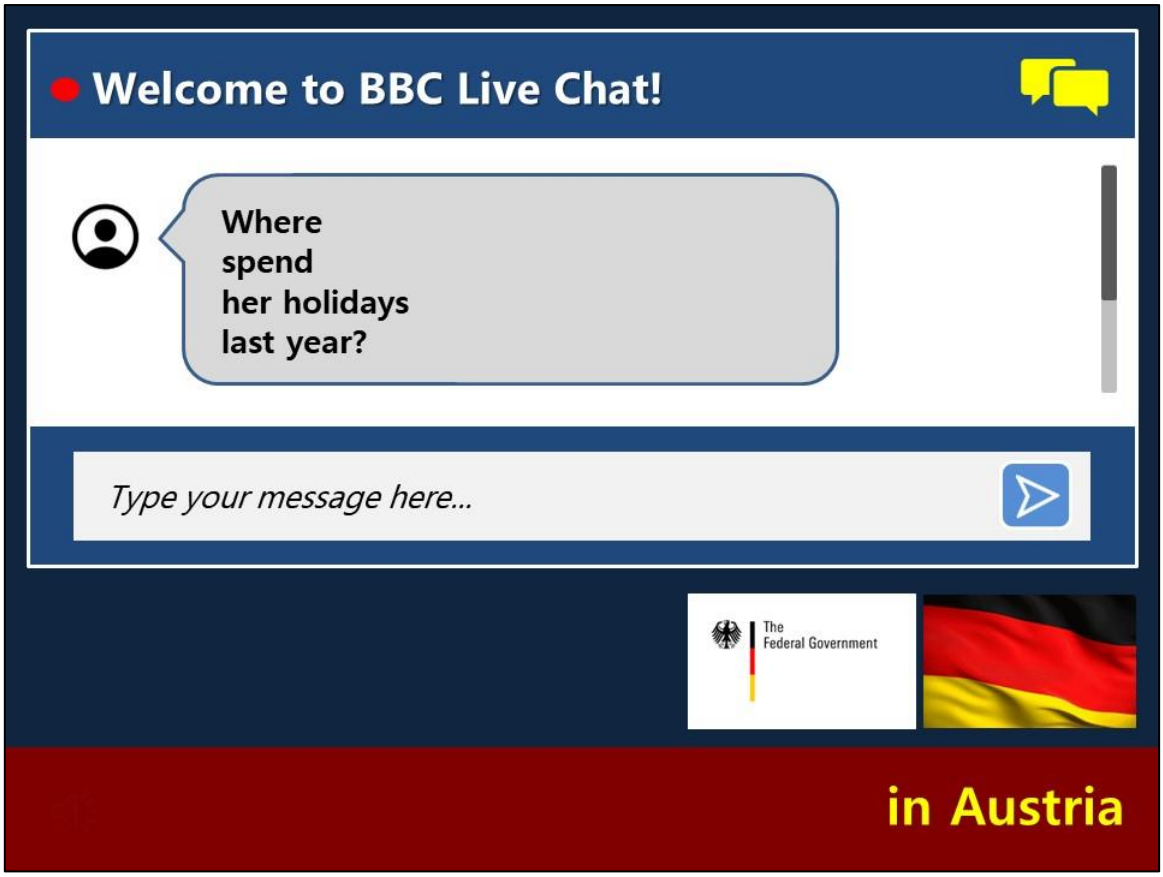

Welcome to BBC Live Chat!
How
go
the bakery
every morning?
Type your message here...
The
Federal Government
walk to the bakery

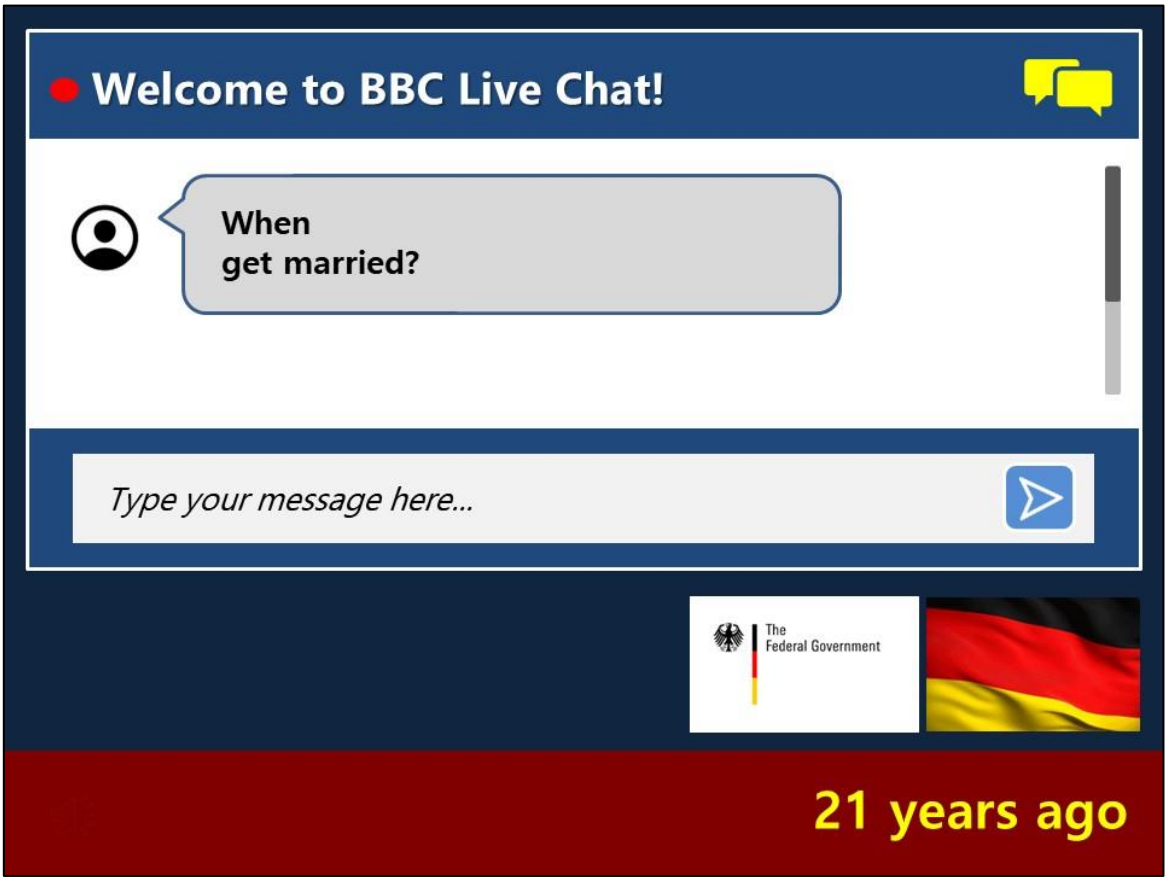

Welcome to BBC Live Chat!
What
do
after breakfast
every day?
Type your message here...
The
Federal Government
jog

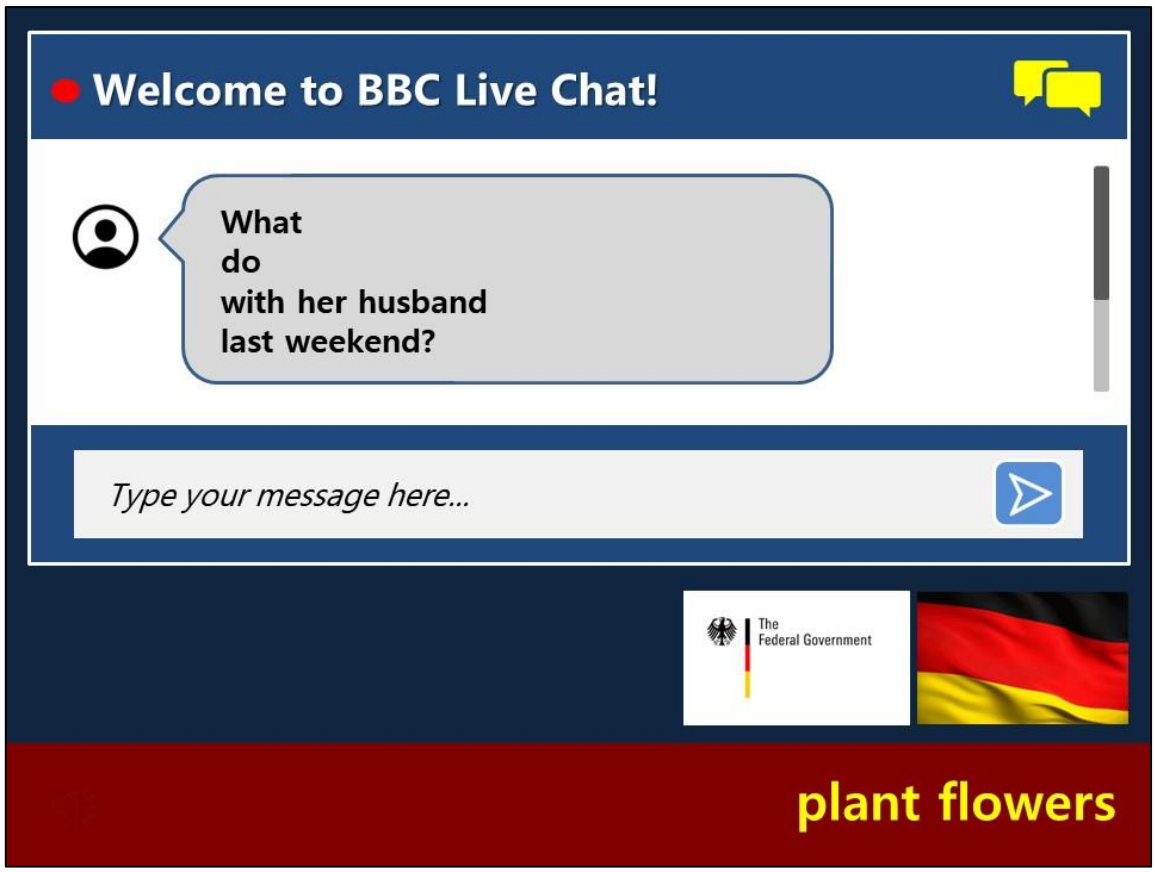

Welcome to BBC Live Chat!
What
do
every night?
Type your message here...
The
Federal Government
practice her speech

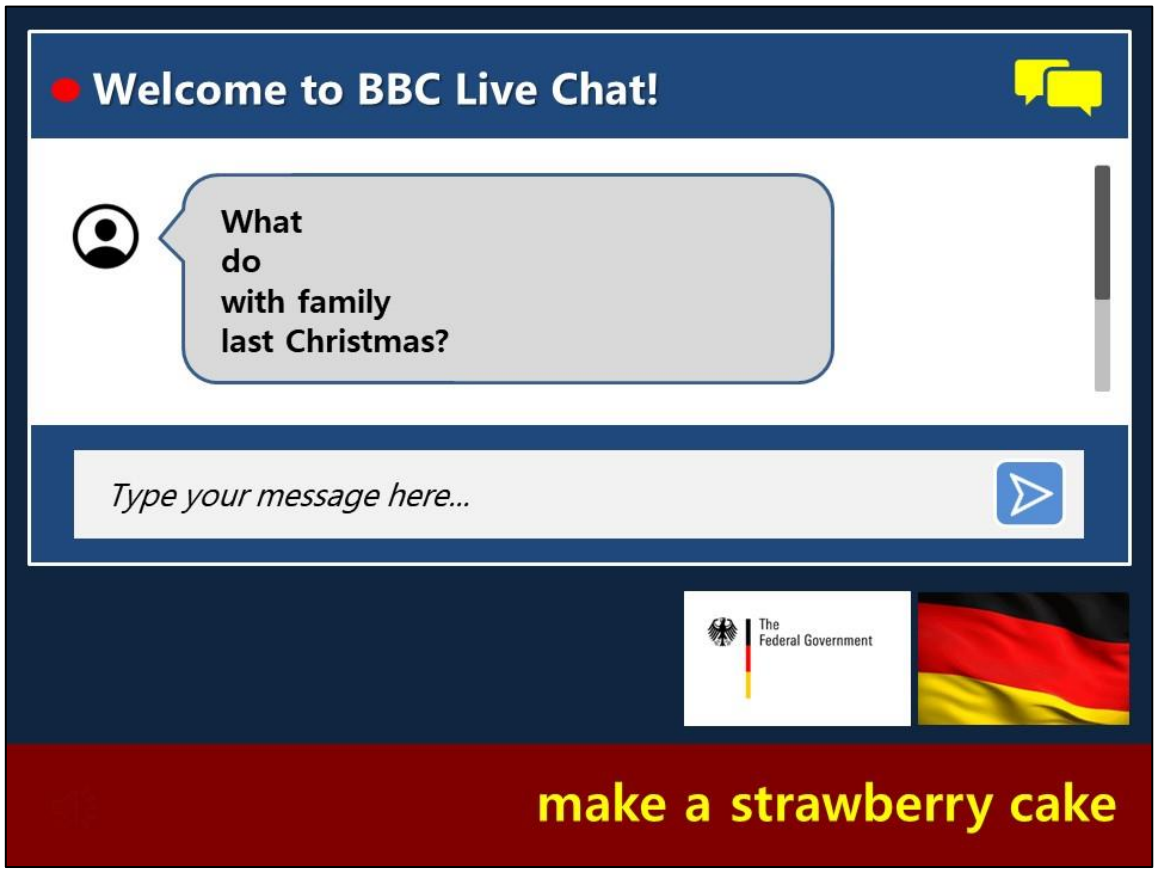

Welcome to BBC Live Chat!
What
do
in her free time?
Type your message here...
The
Federal Government
enjoy baking

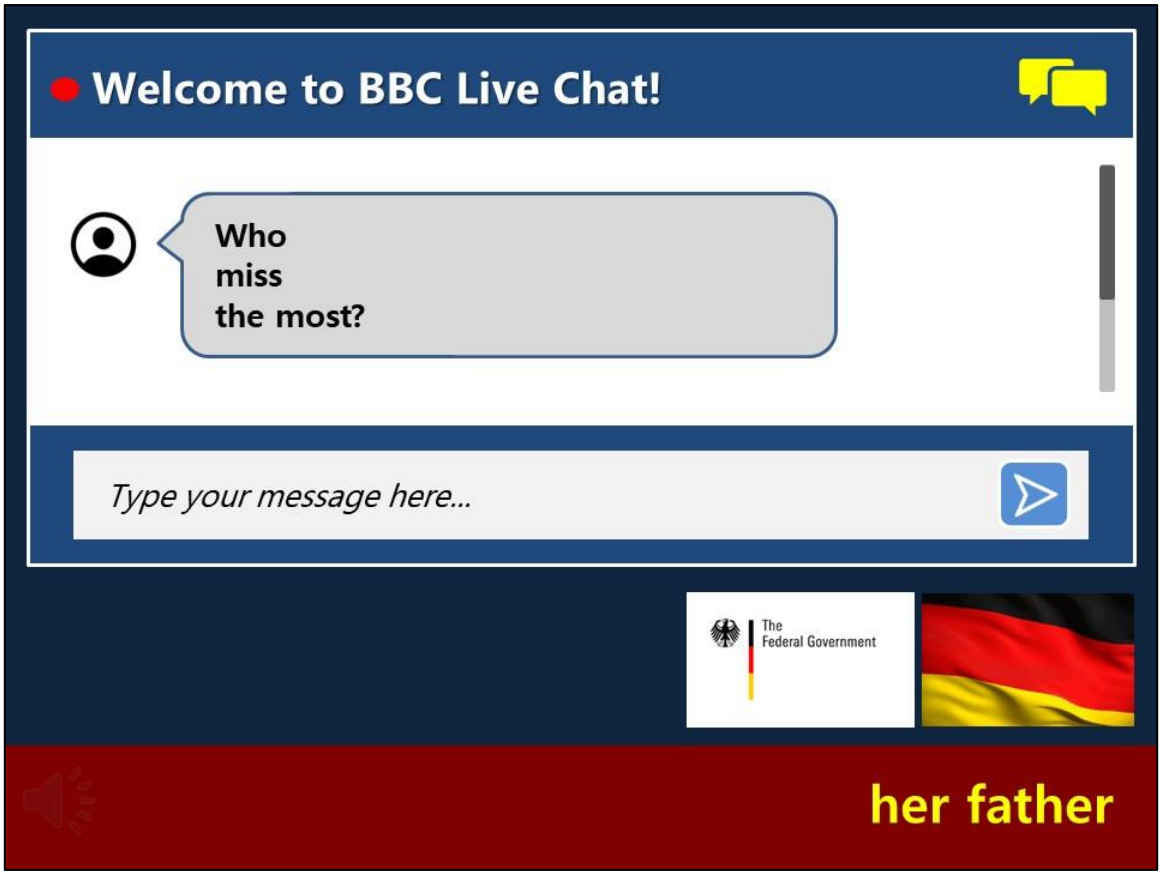

Welcome to BBC Live Chat!
What
do
every Sunday?
Type your message here...
The
Federal Government
relax at home

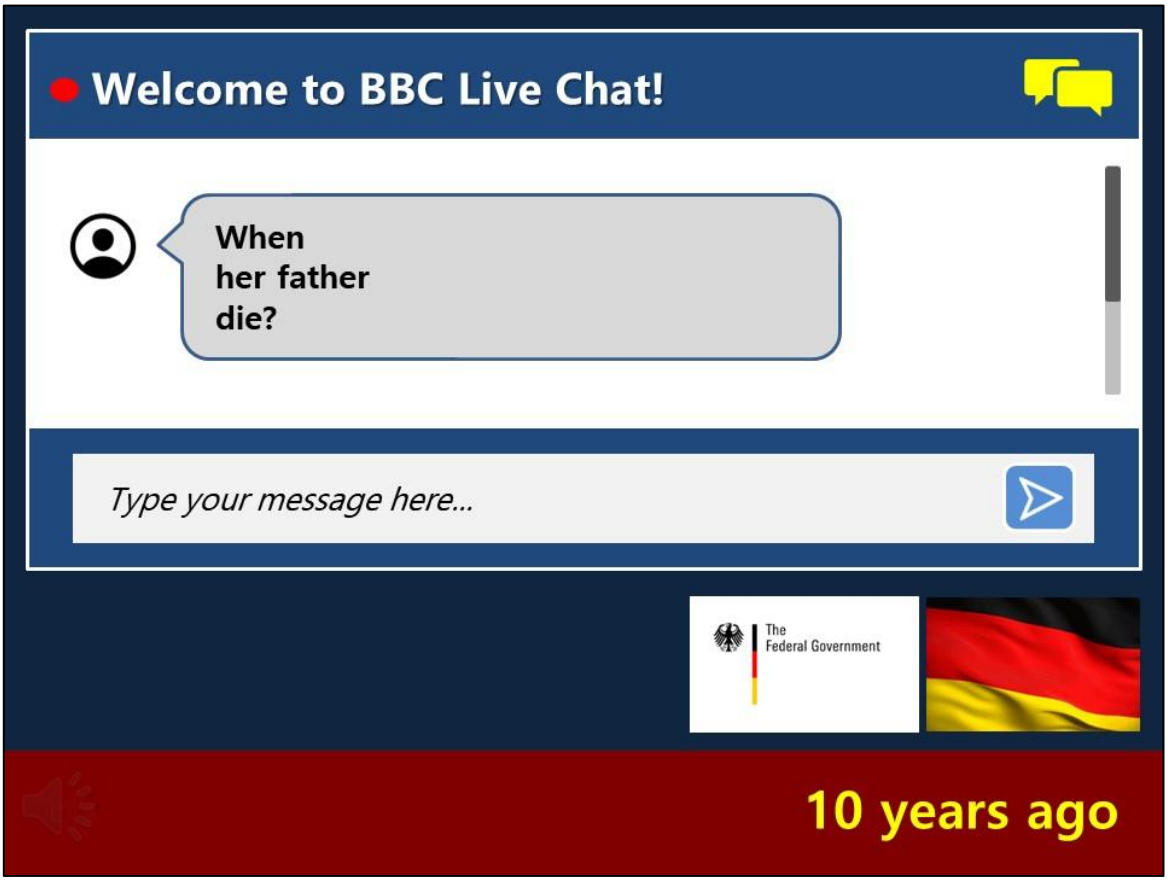

Welcome to BBC Live Chat!
Where
spend
her holidays
last year?
Type your message here...
The
Federal Government
in Austria

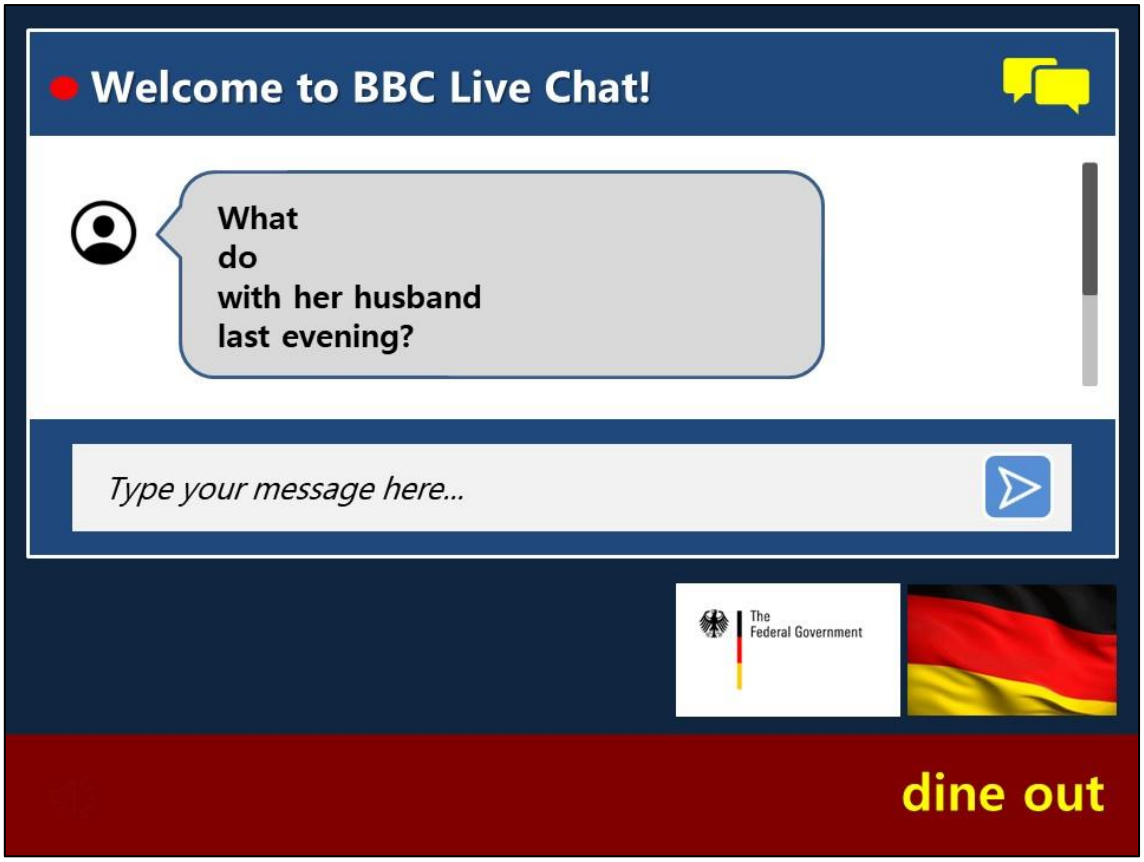

Welcome to BBC Live Chat!
When
get married?
Type your message here...
The
Federal Government
21 years ago

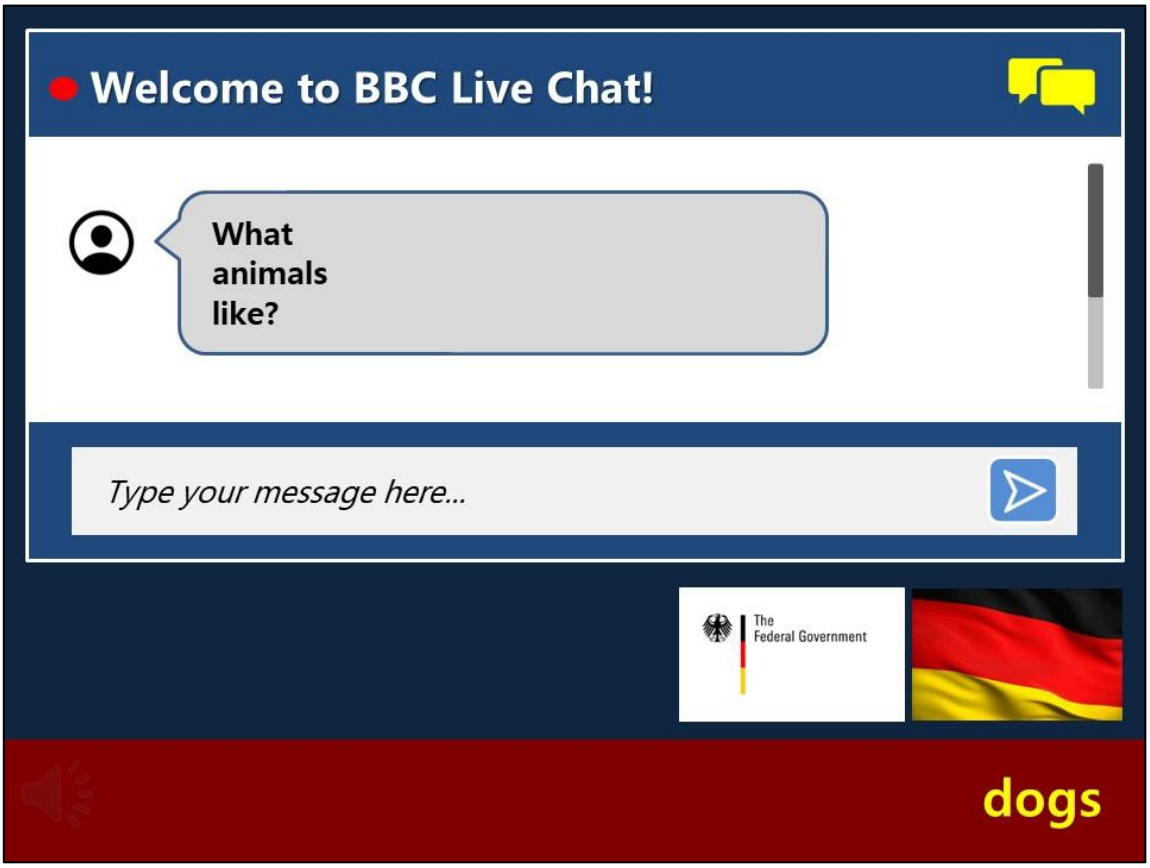

Welcome to BBC Live Chat!
What
do
with her husband
last weekend?
Type your message here...
The
Federal Government
plant flowers

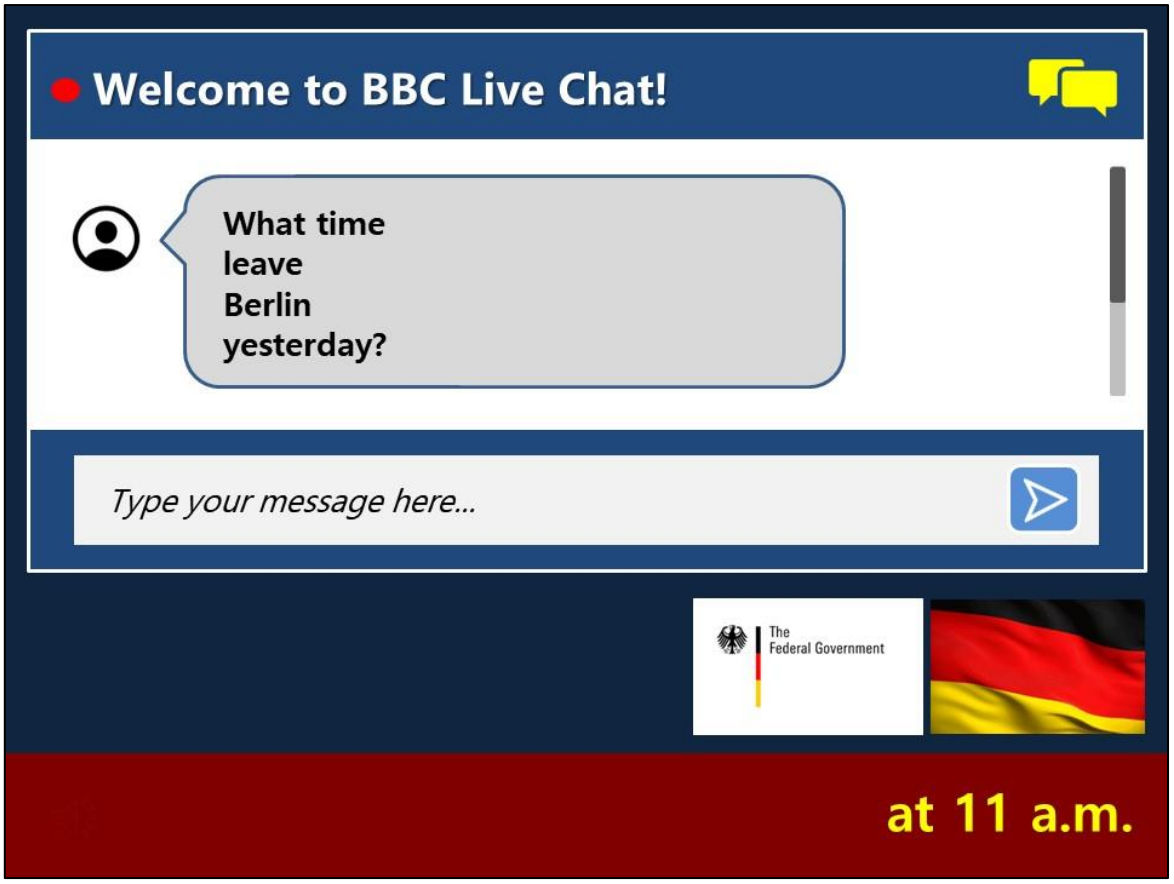

Welcome to BBC Live Chat!
What
do
with family
last Christmas?
Type your message here...
The
Federal Government
make a strawberry cake

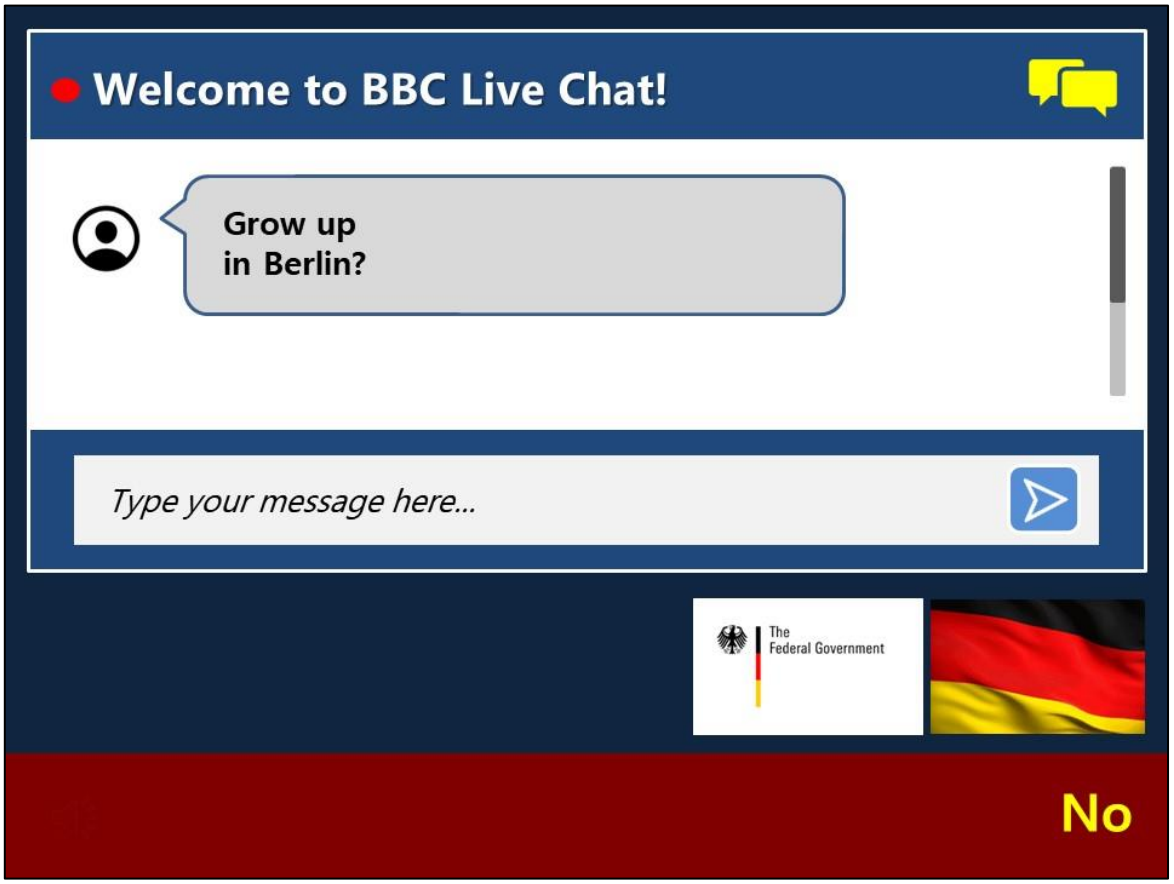

Welcome to BBC Live Chat!
Who
miss
the most?
Type your message here...
The
Federal Government
her father

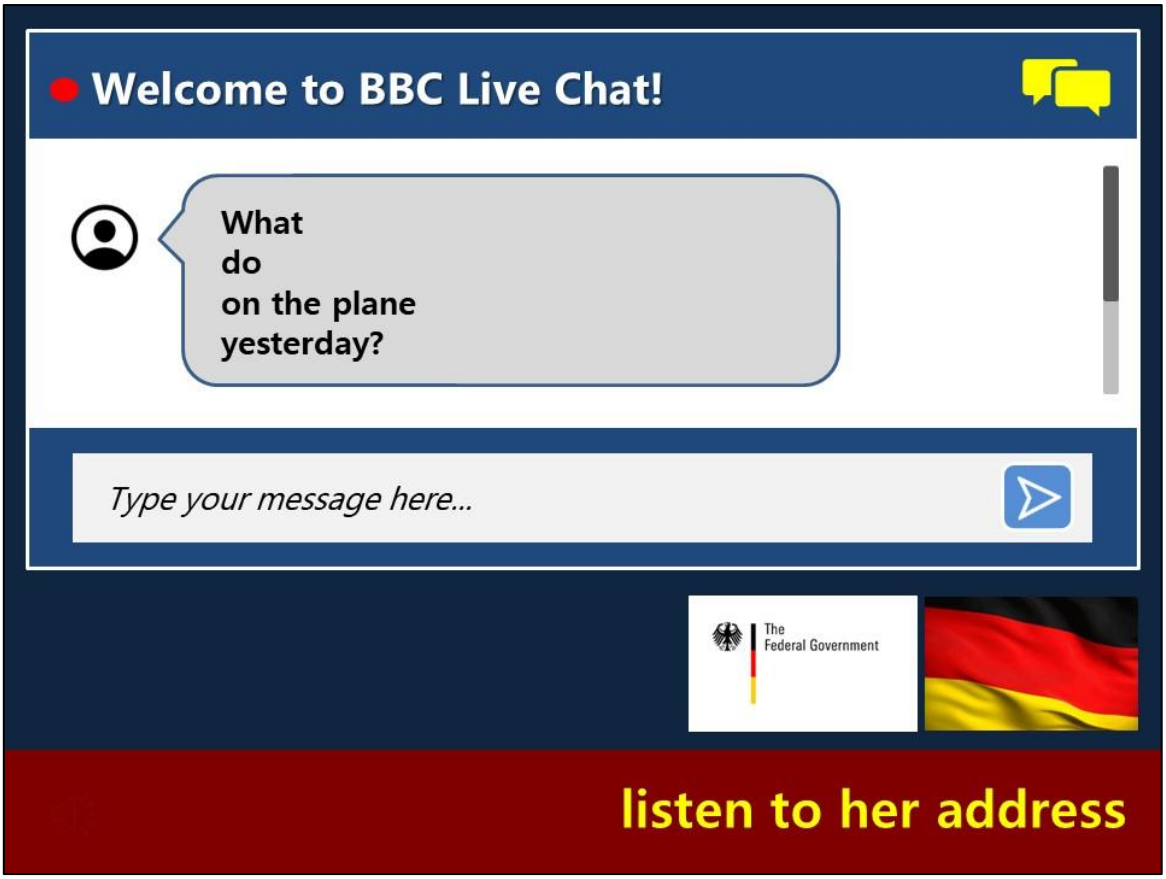

Welcome to BBC Live Chat!
When
her father
die?
Type your message here...
The
Federal Government
10 years ago

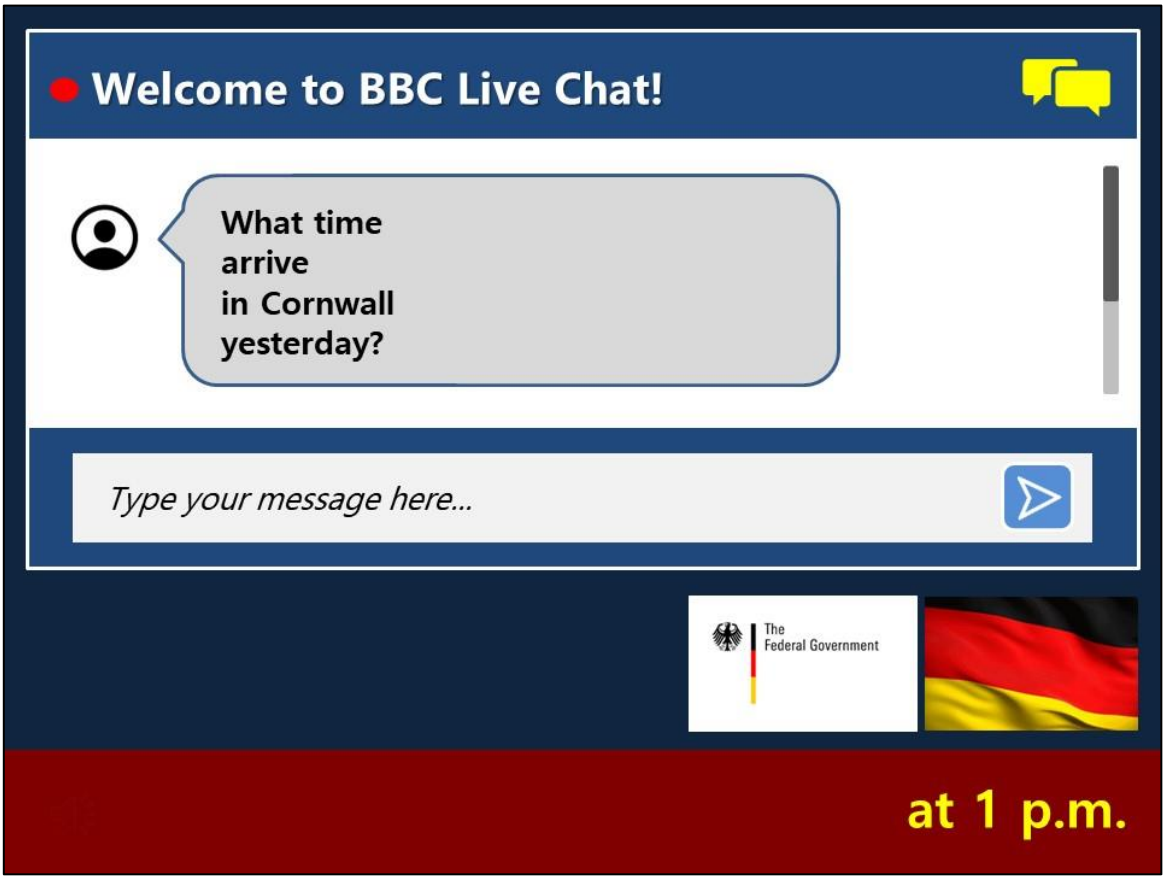

Welcome to BBC Live Chat!
What
do
with her husband
last evening?
Type your message here...
The
Federal Government
dine out

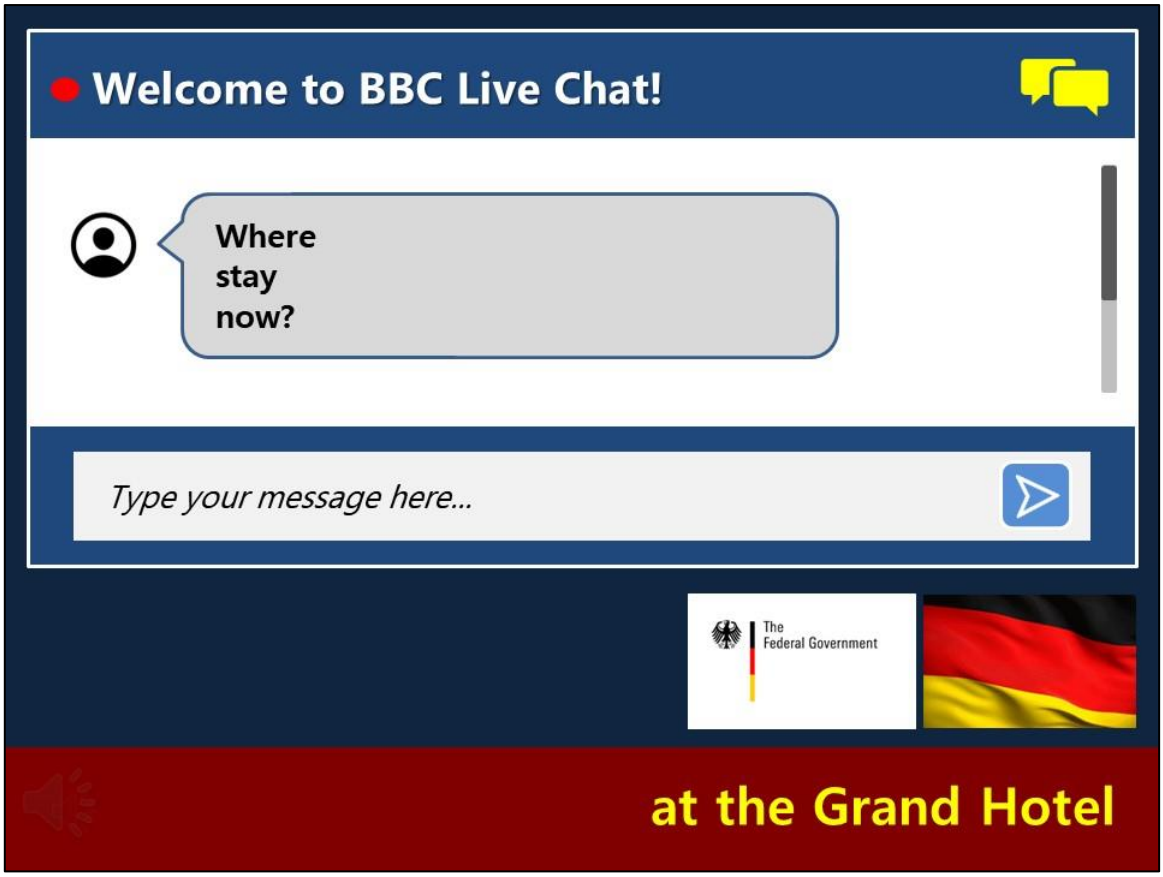

Welcome to BBC Live Chat!
What
animals
like?
Type your message here...
The
Federal Government
dogs

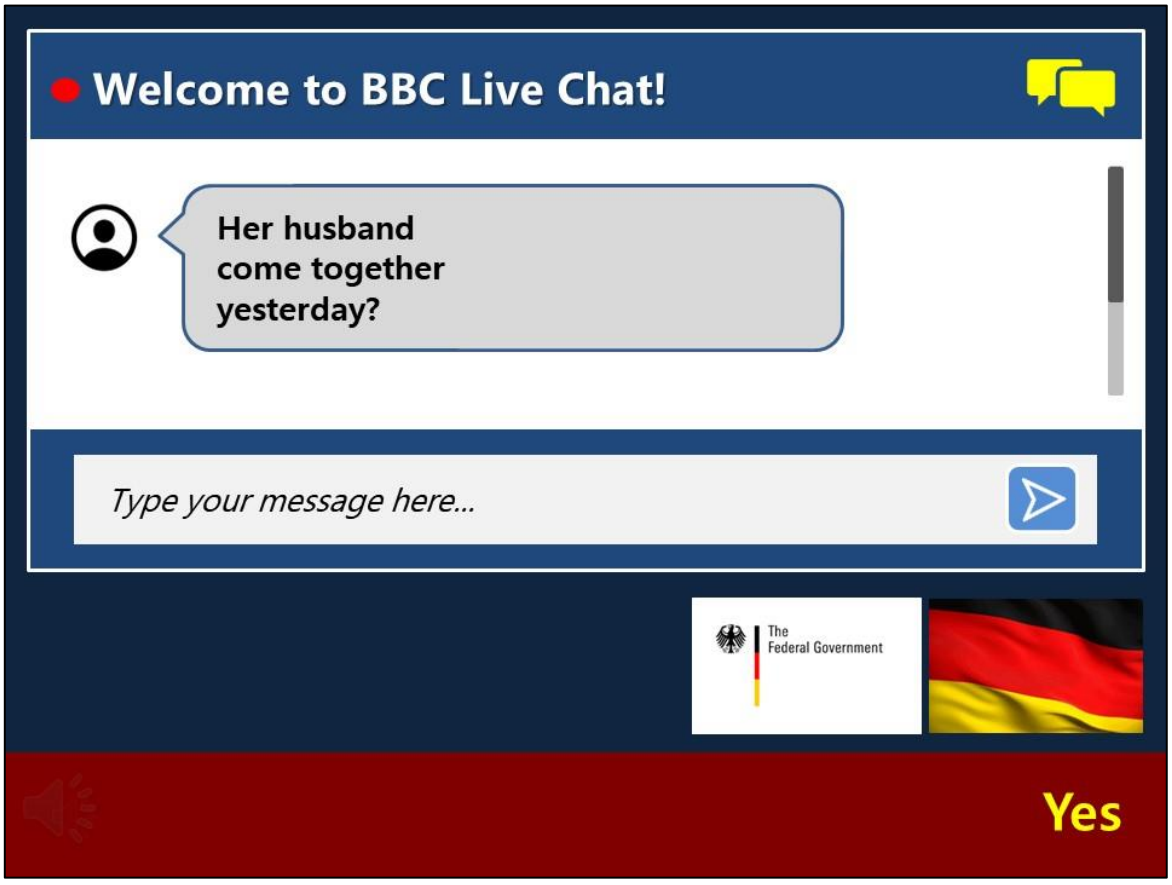

Welcome to BBC Live Chat!
What time leave Berlin yesterday?
Type your message here...
The Federal Government
at 11 a.m.


Welcome to BBC Live Chat!
Grow up in Berlin?
Type your message here...
The Federal Government
No

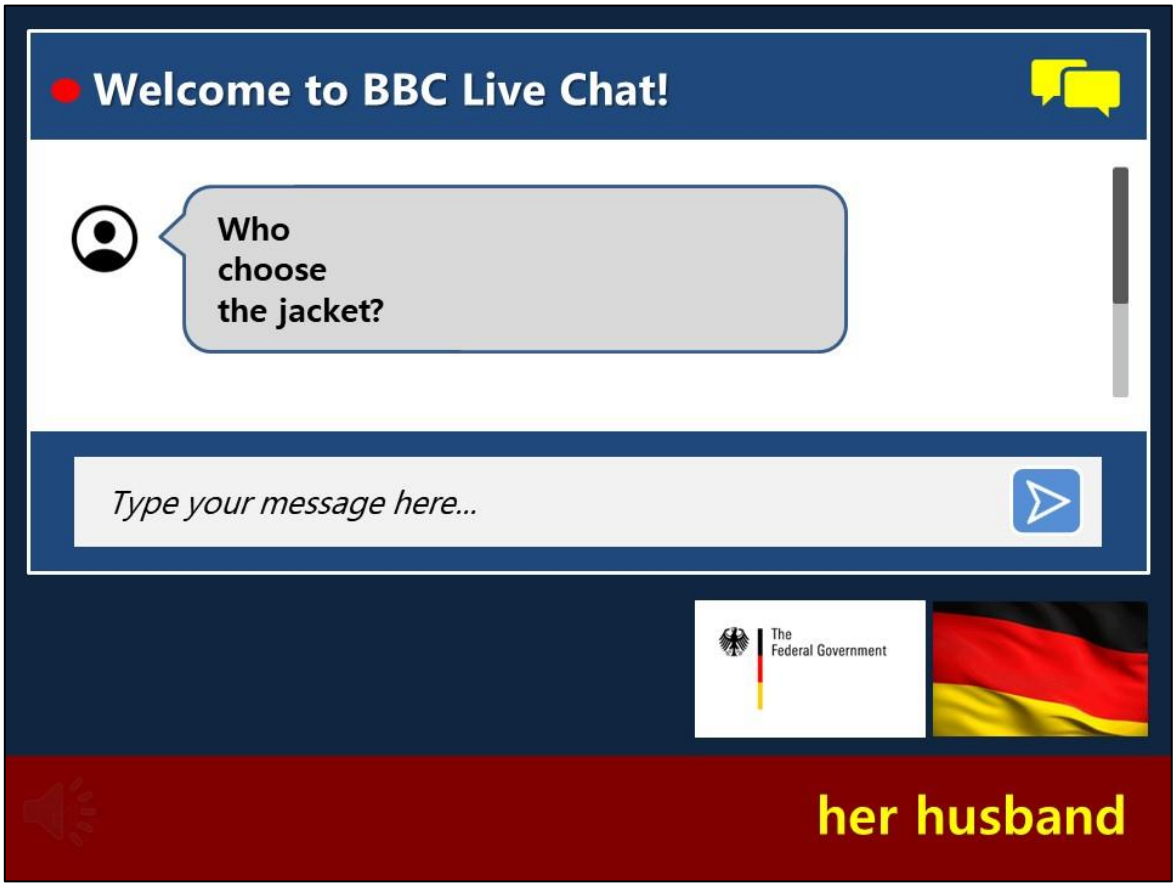

Welcome to BBC Live Chat!
What
do
on the plane
yesterday?
Type your message here...
The
Federal Government
listen to her address


Welcome to BBC Live Chat!
What time
arrive
in Cornwall
yesterday?
Type your message here...
The
Federal Government
at 1 p.m.


Welcome to BBC Live Chat!
Where
stay
now?
Type your message here...
The
Federal Government
at the Grand Hotel


Welcome to BBC Live Chat!
Her husband
come together
yesterday?
Type your message here...
The
Federal Government
Yes


Welcome to BBC Live Chat!
What wear now?
Type your message here...
The Federal Government
a blue jacket


Welcome to BBC Live Chat!
Who choose the jacket?
Type your message here...
The Federal Government
her husband

Welcome to BBC Live Chat!

What do now?

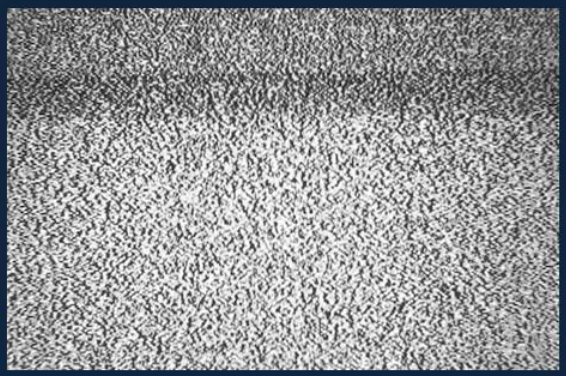
Type your message here...

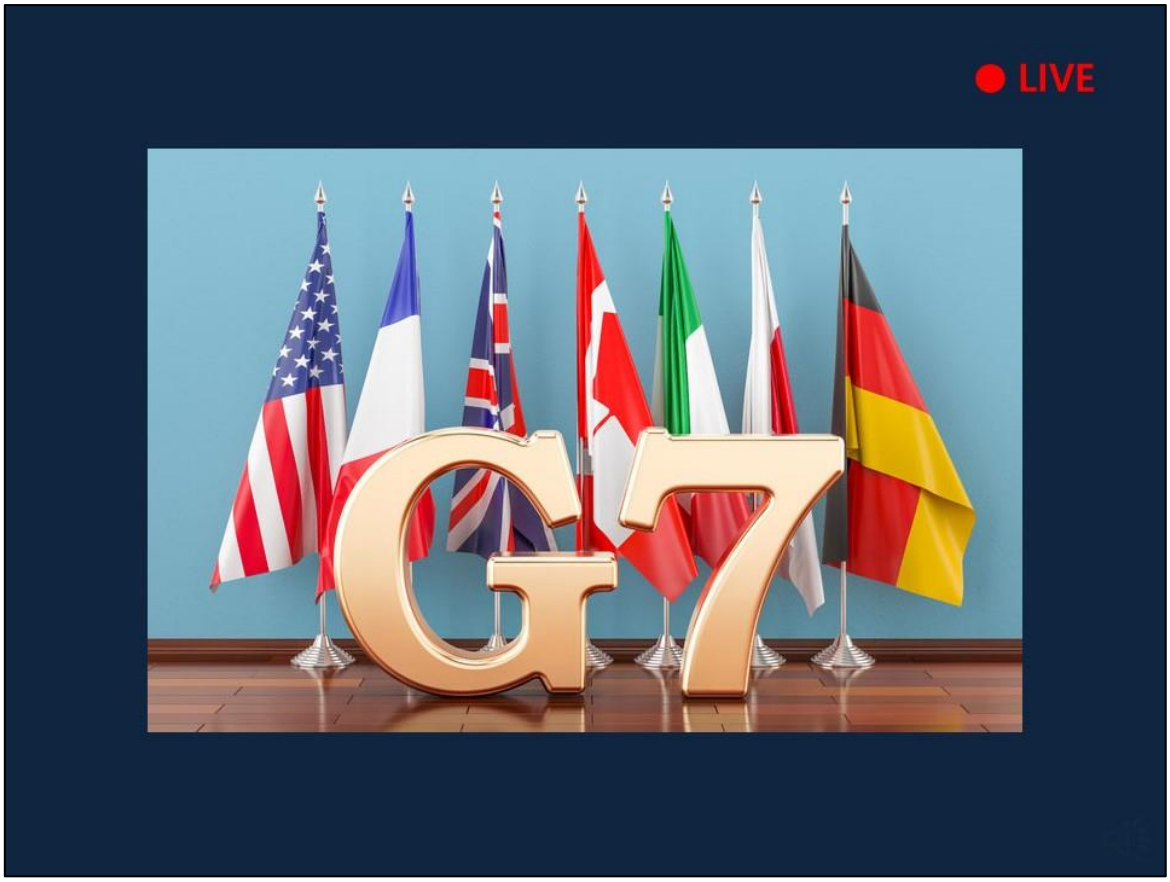
The Federal Government

prepare for her interview

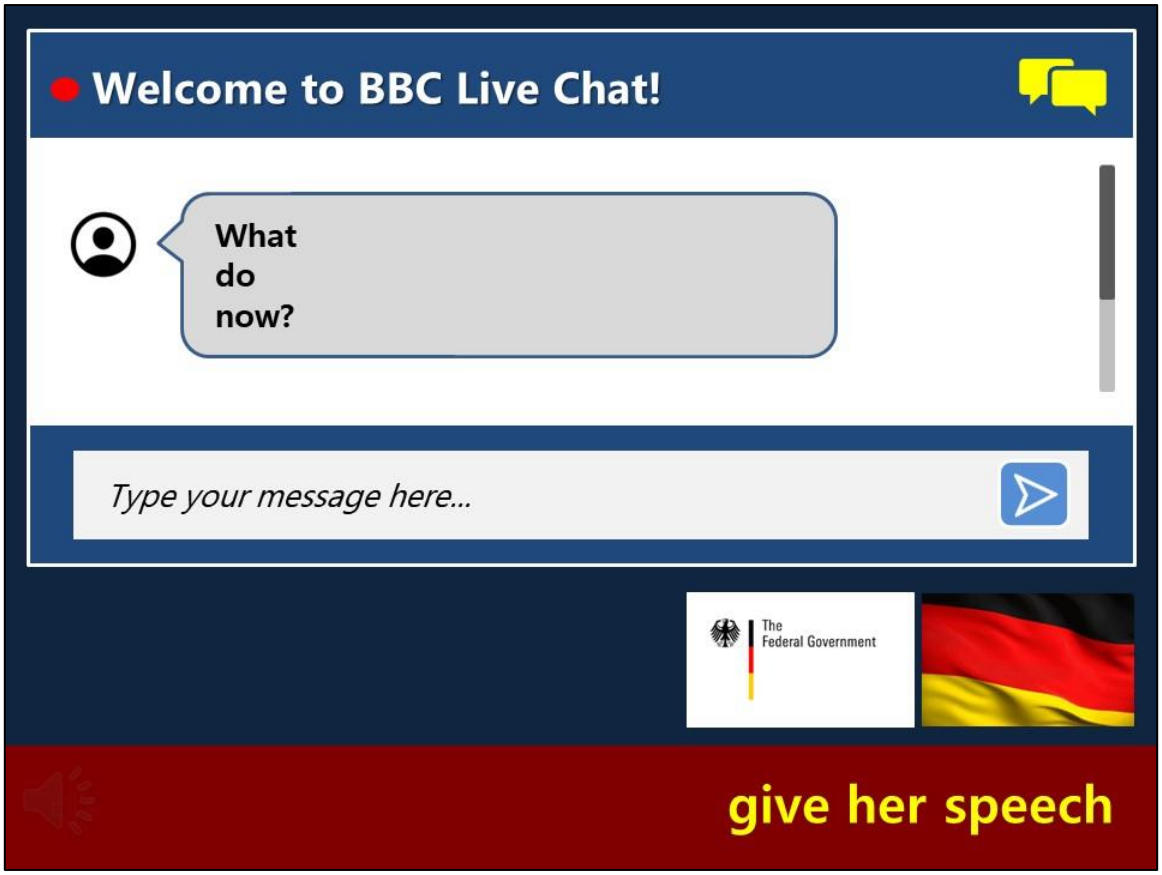
LIVE

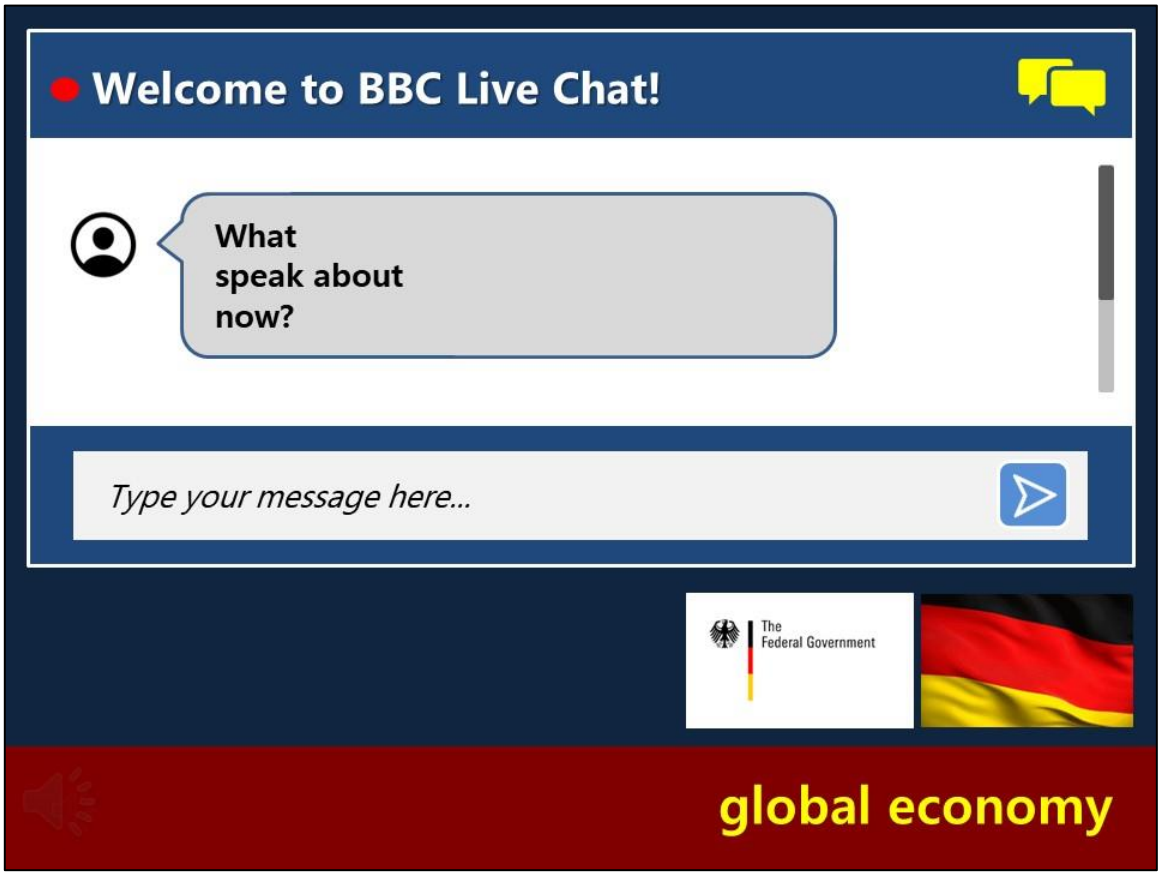


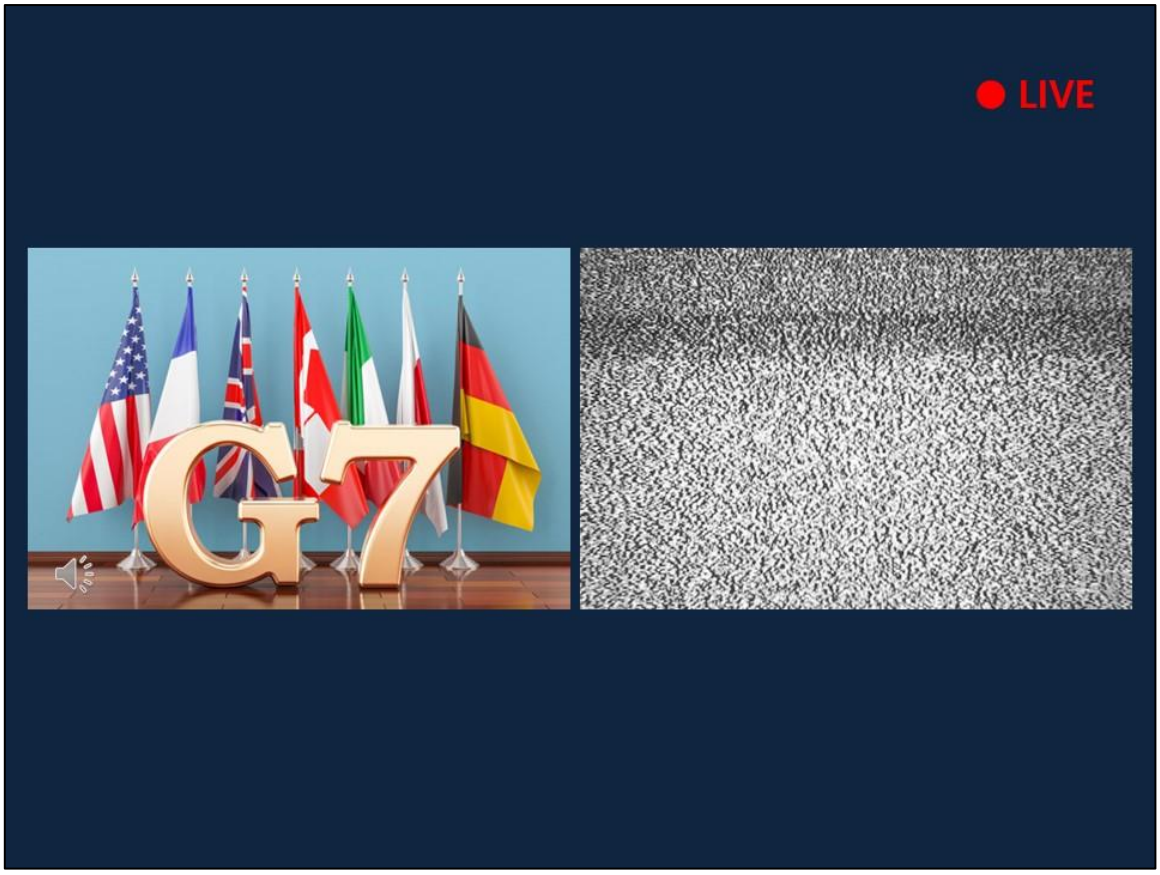



LIVE
G7

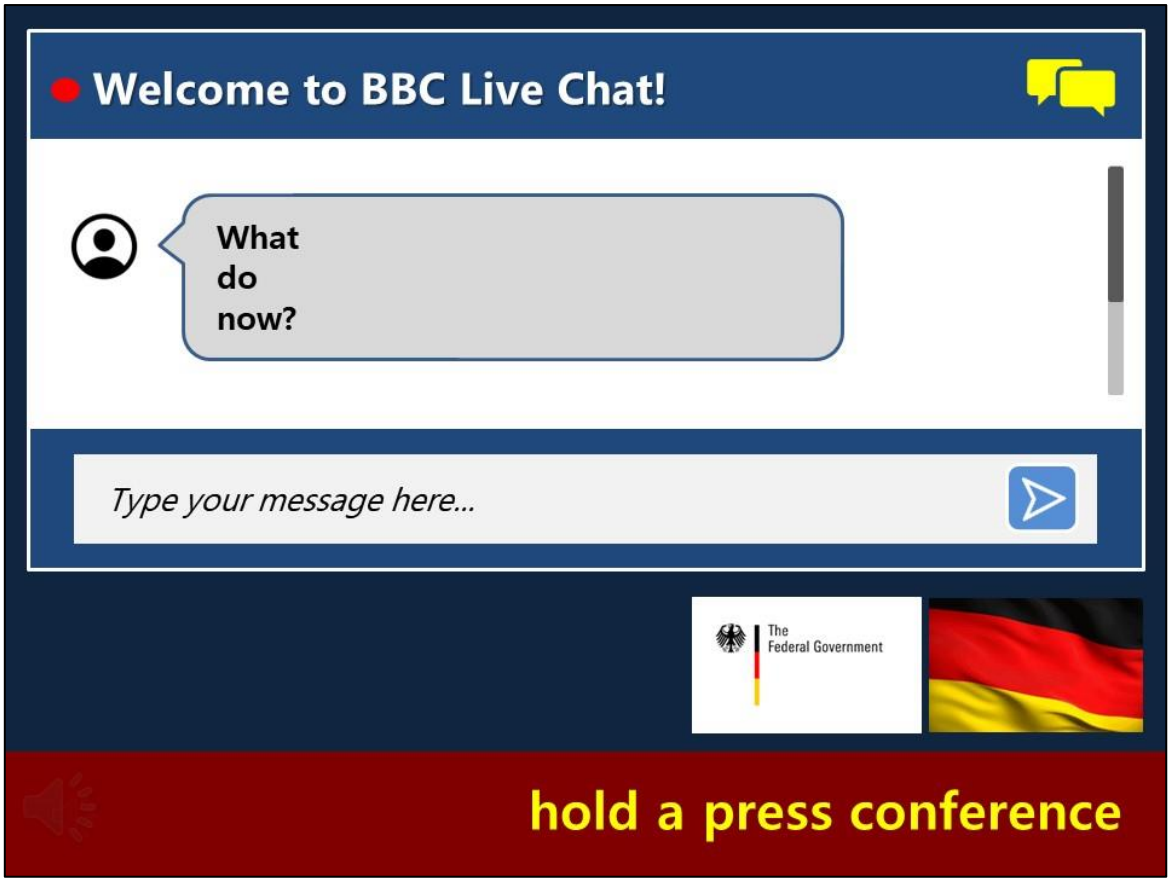

politics live
BBC
Live
Chat
LIVE
CGTN
G7
Cornwall
United Kingdom 2021
BREAKING NEWS
G7 정상회의가 시작됐습니다!
BBC
NEWS
BREAKING

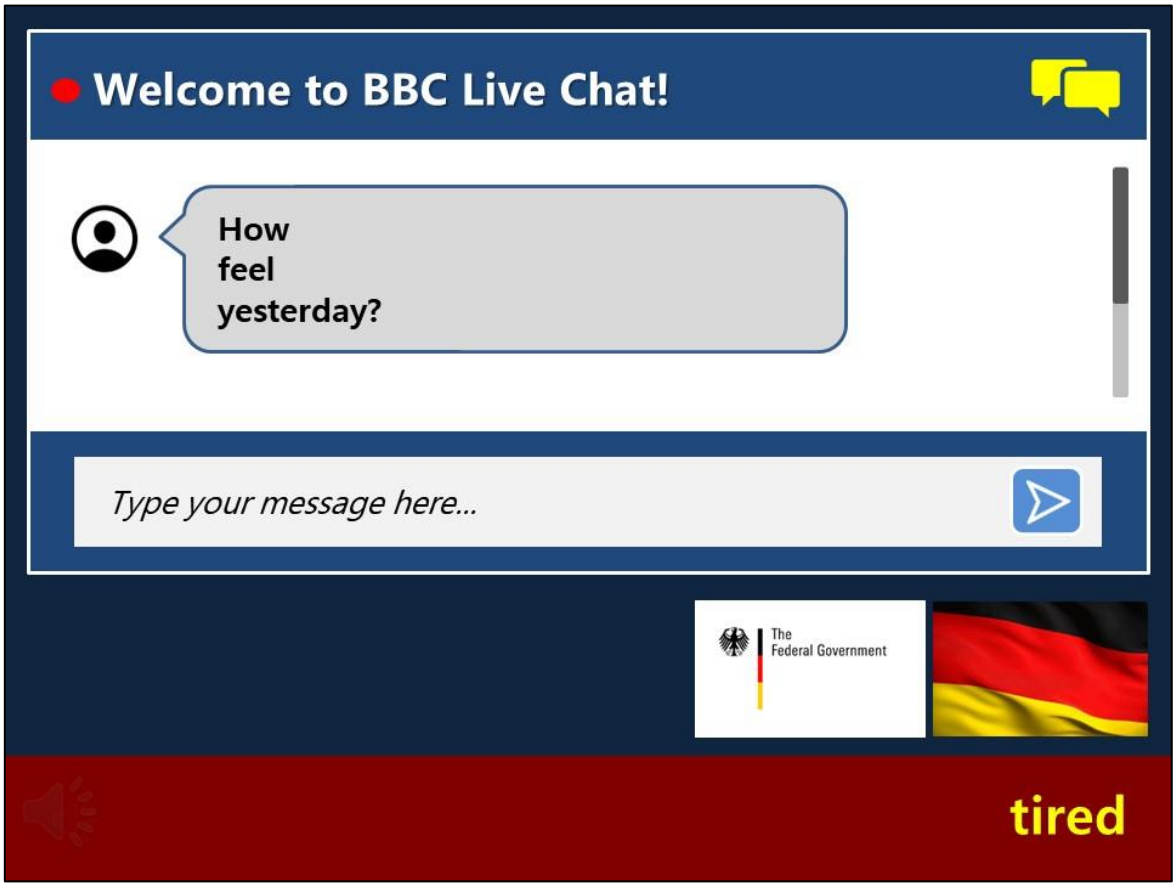

Welcome to BBC Live Chat!
What
do
now?
Type your message here...
The
Federal Government
give her speech


Welcome to BBC Live Chat!
What
speak about
now?
Type your message here...
The
Federal Government
global economy

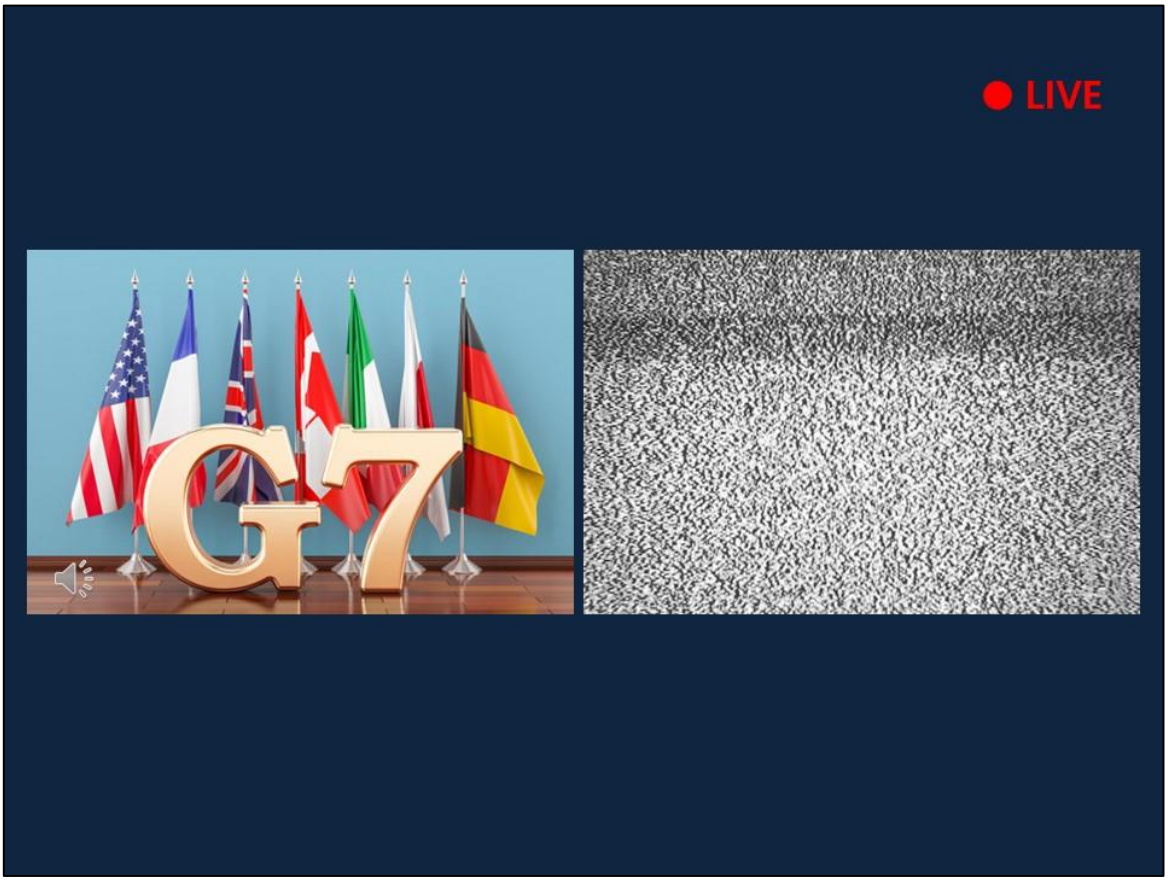

LIVE
G7


politics live
BBC
Live
Chat
LIVE
BREAKING NEWS
G7 정상회의가 영국 콘월에서 진행 중입니다.
BBC
NEWS
BREAKING

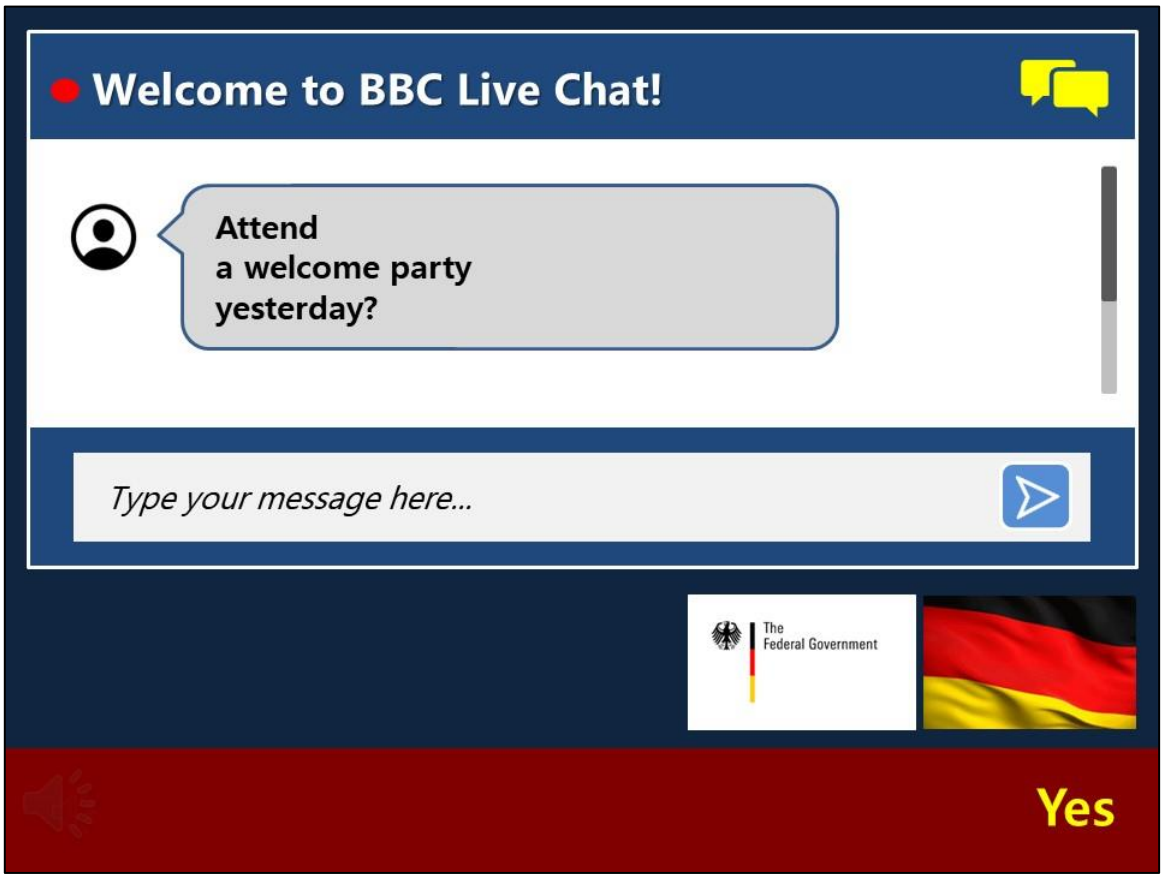

Welcome to BBC Live Chat!
What do now?
Type your message here...
The Federal Government
hold a press conference

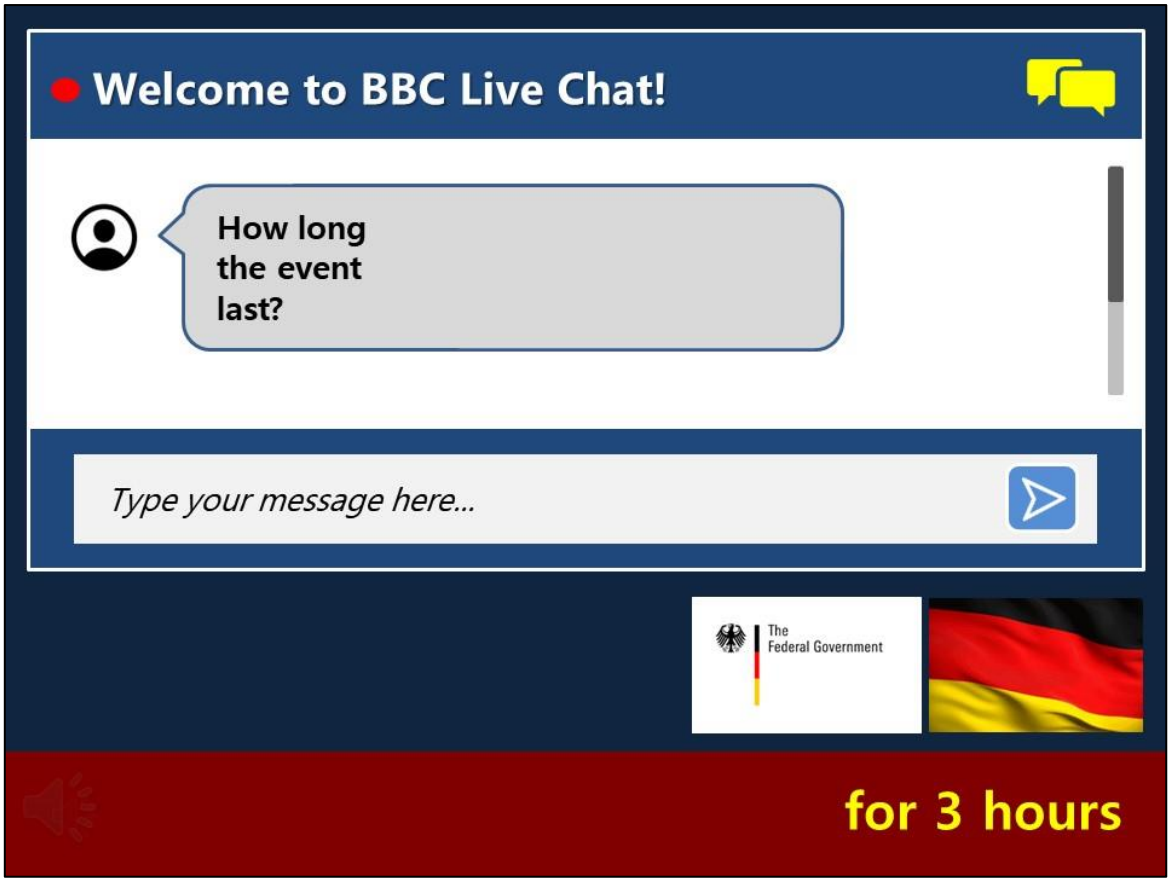

Welcome to BBC Live Chat!
How feel yesterday?
Type your message here...
The Federal Government
tired

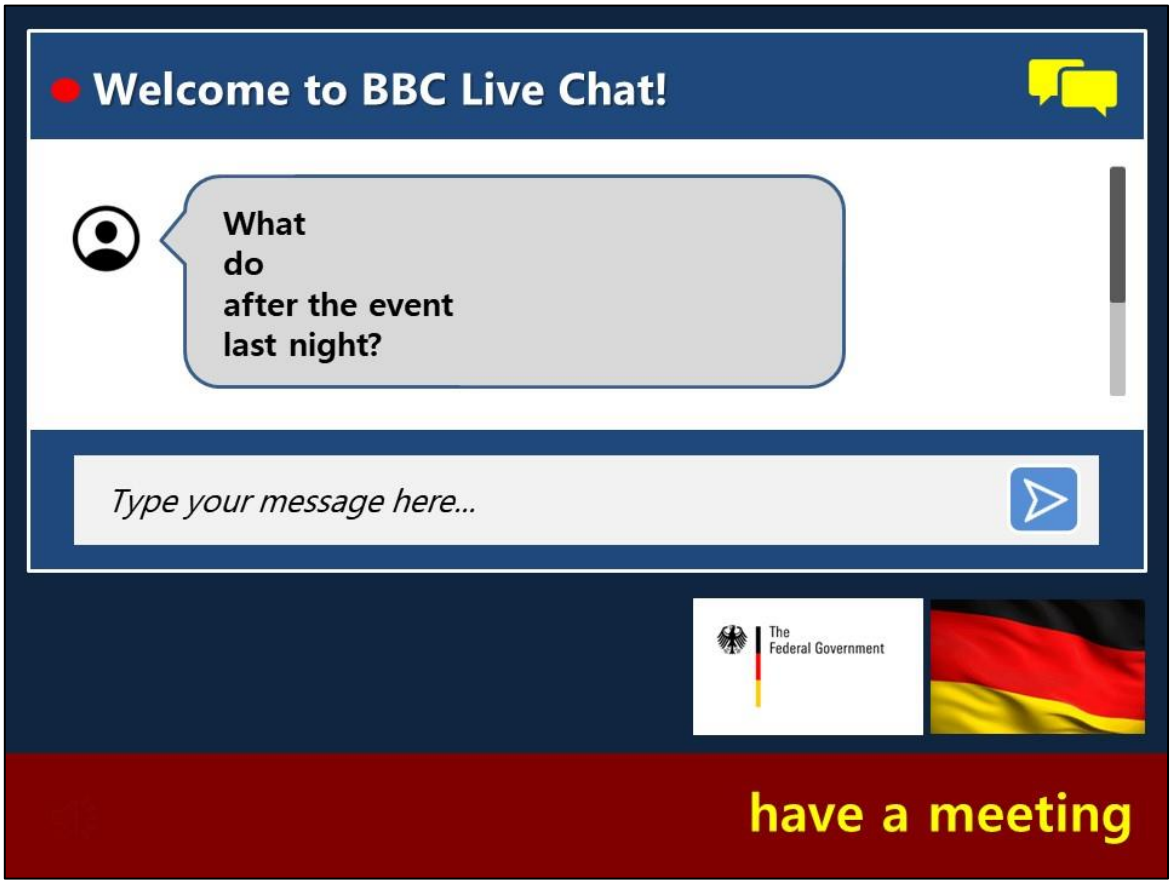

Welcome to BBC Live Chat!
What do now?
Type your message here...
The Federal Government
answer reporters' questions

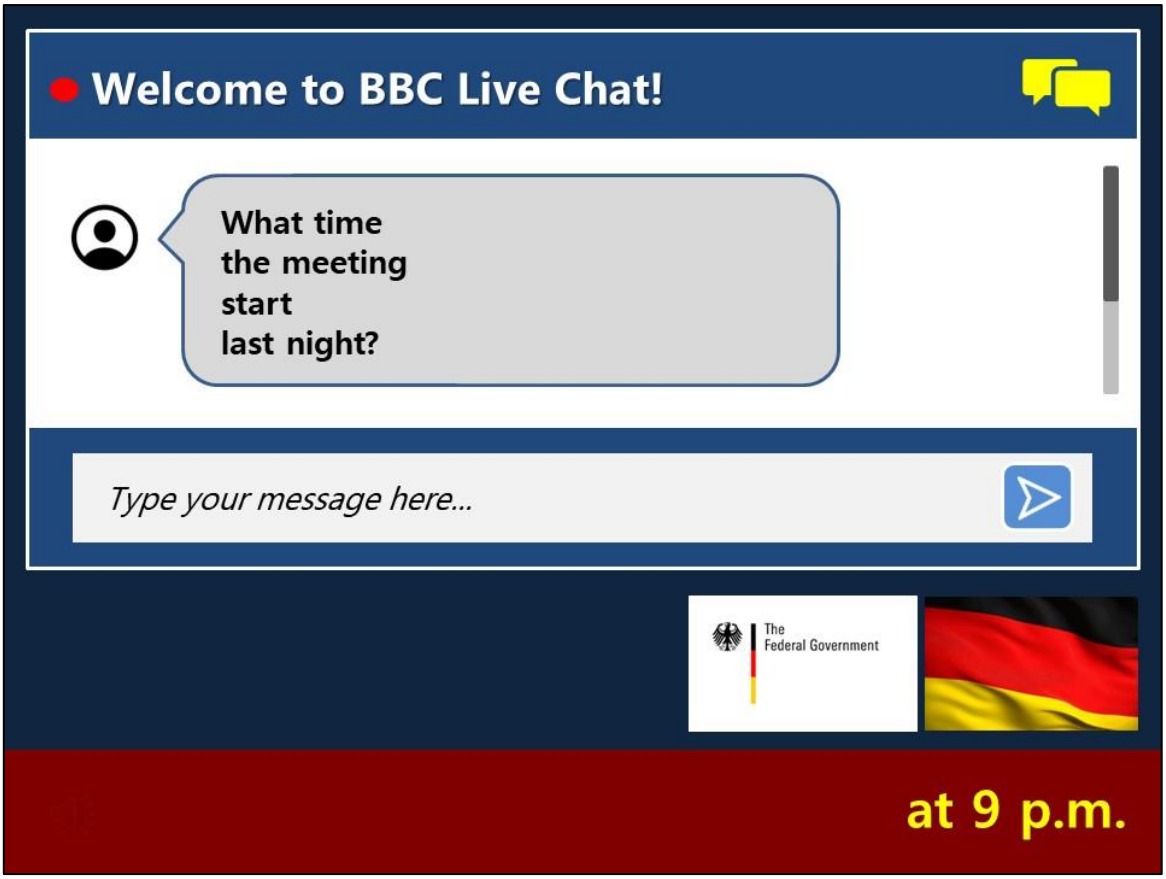

LIVE
G7

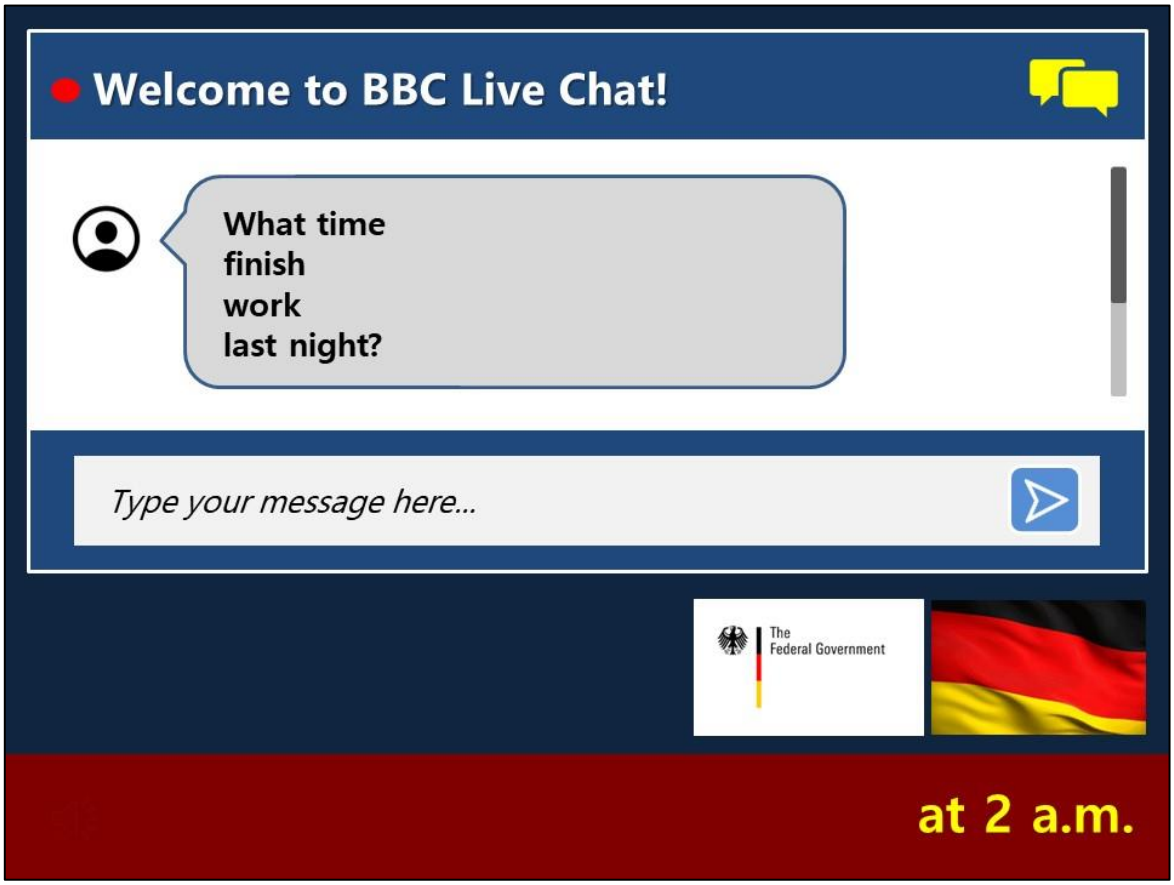

politics live
BBC
Live Chat
LIVE
THE CORNISH ARMS
BREAKING NEWS
G7 정상회의가 종료되었습니다.
BBC NEWS BREAKING

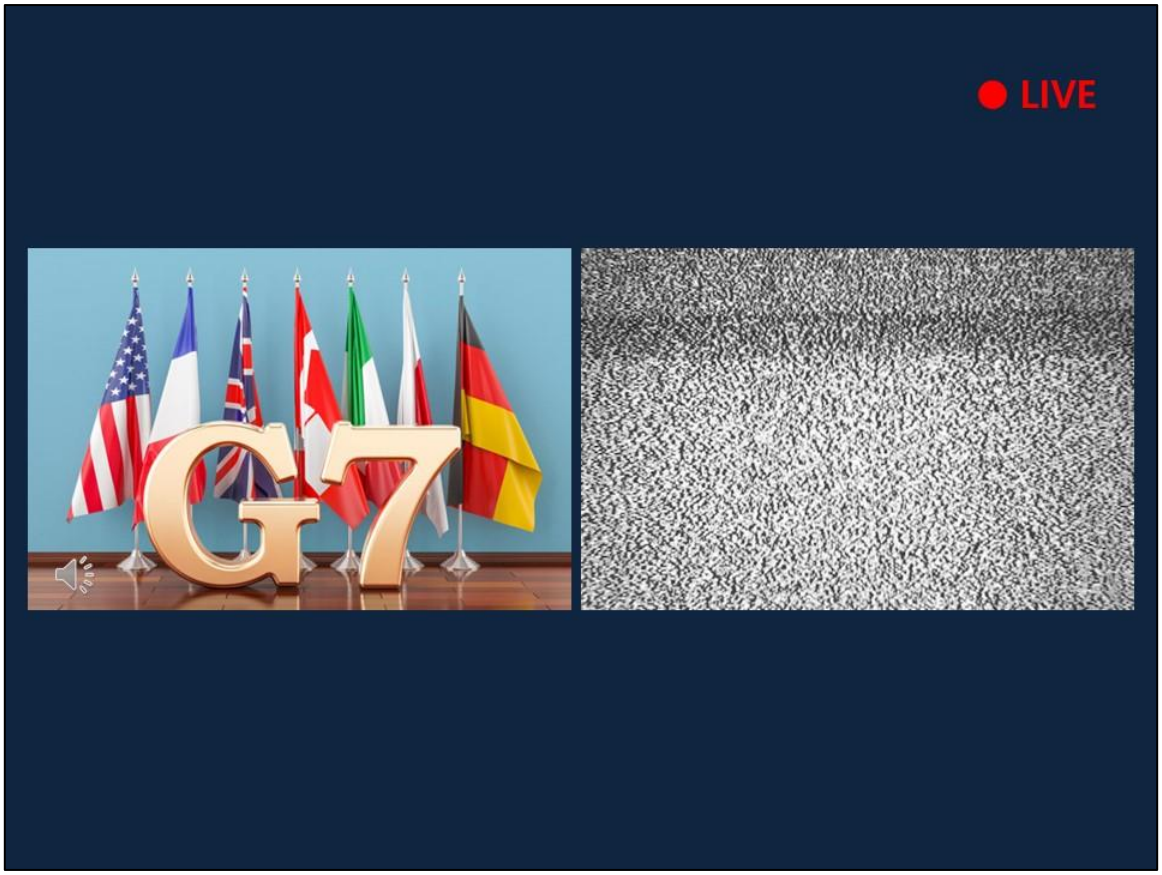

Welcome to BBC Live Chat!
Attend a welcome party yesterday?
Type your message here...
The Federal Government
Yes

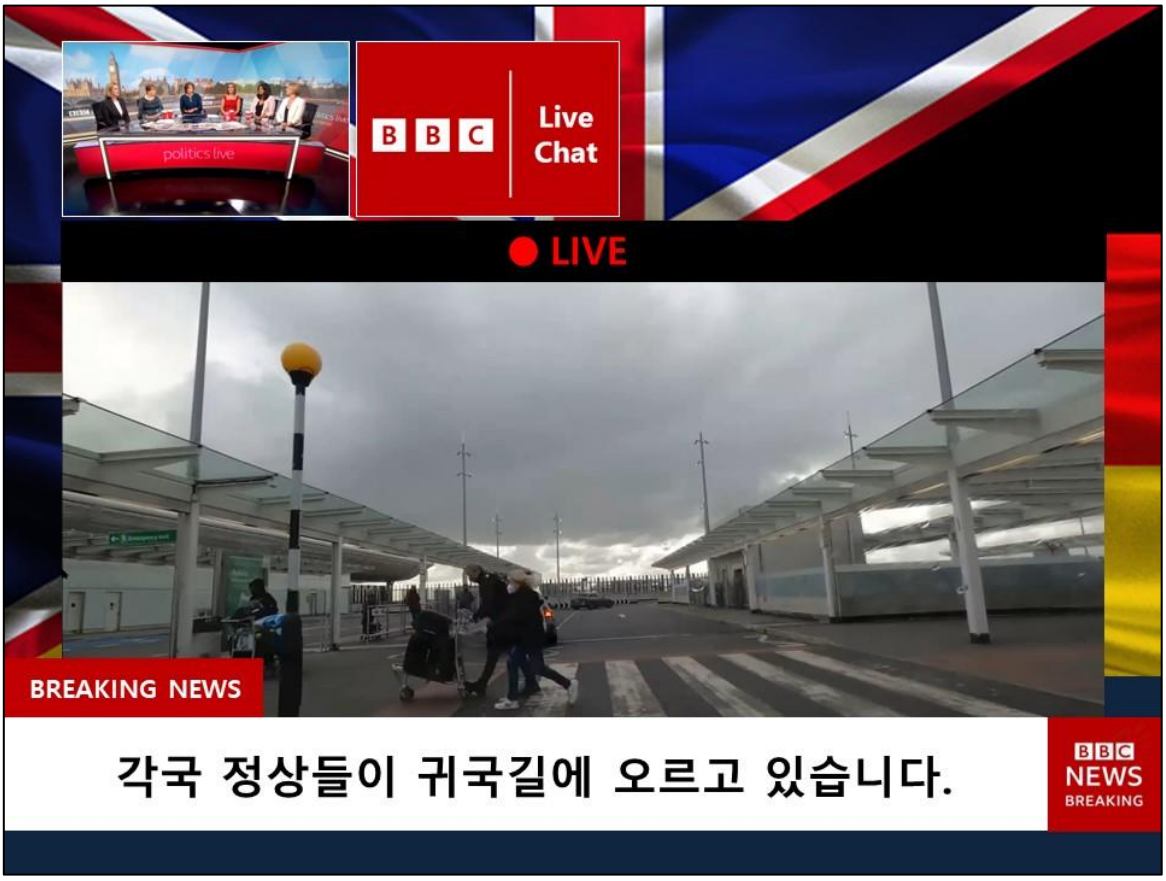

Welcome to BBC Live Chat!
How long the event last?
Type your message here...
The Federal Government
for 3 hours

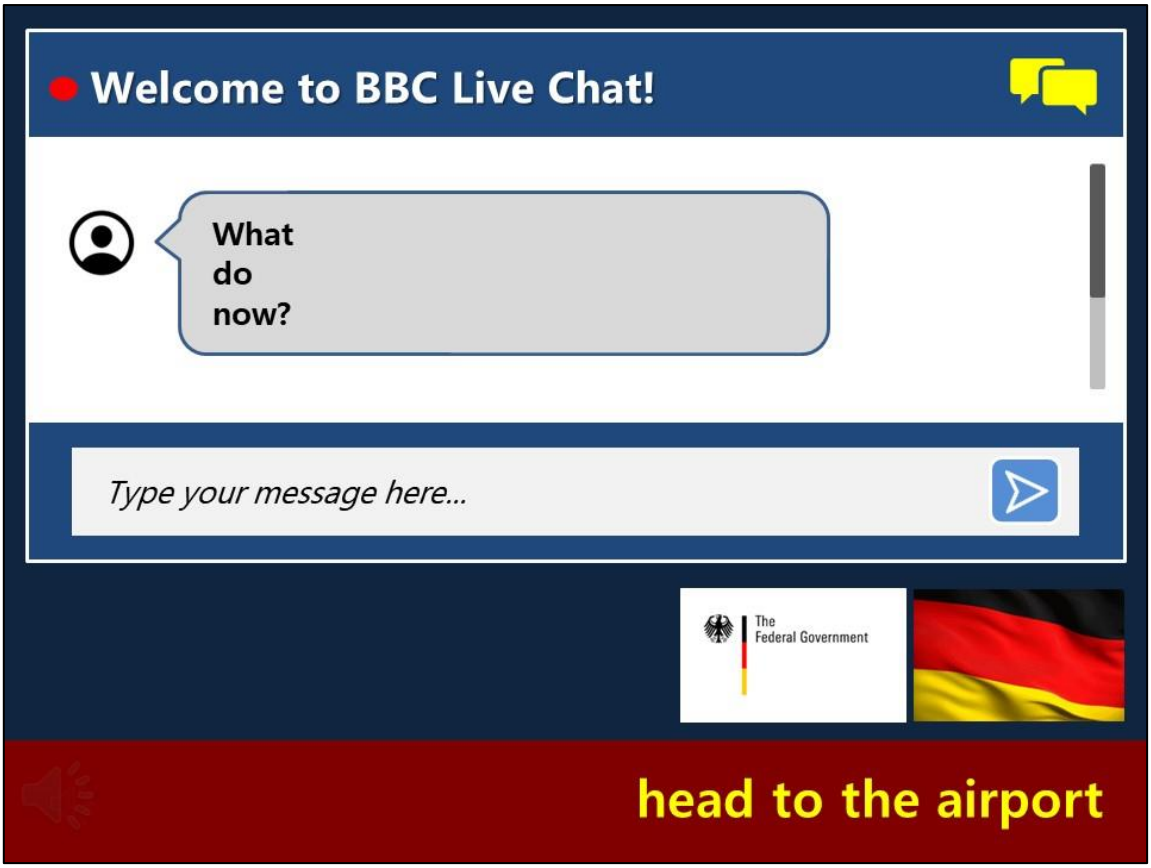

Welcome to BBC Live Chat!
What do after the event last night?
Type your message here...
The Federal Government
have a meeting

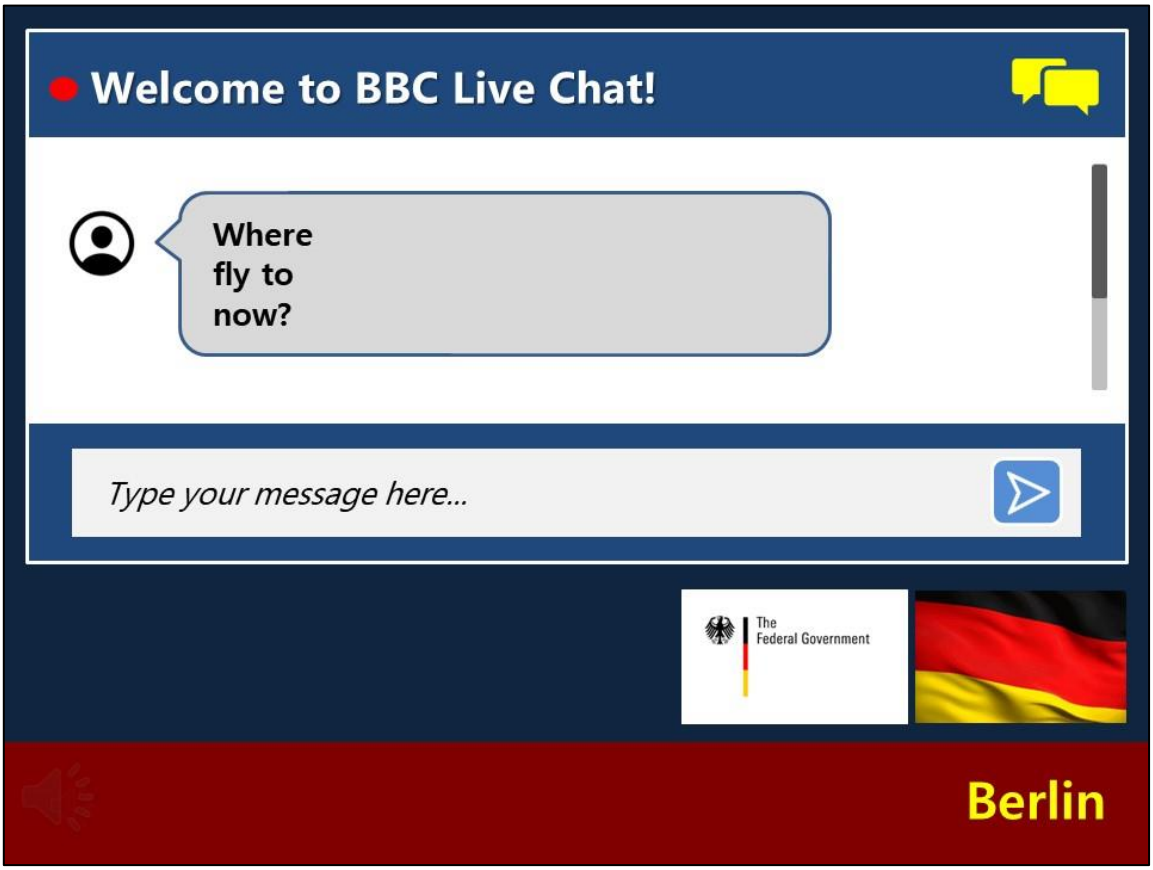

Welcome to BBC Live Chat!
What time
the meeting
start
last night?
Type your message here...
The
Federal Government
at 9 p.m.


Welcome to BBC Live Chat!
What time
finish
work
last night?
Type your message here...
The
Federal Government
at 2 a.m.

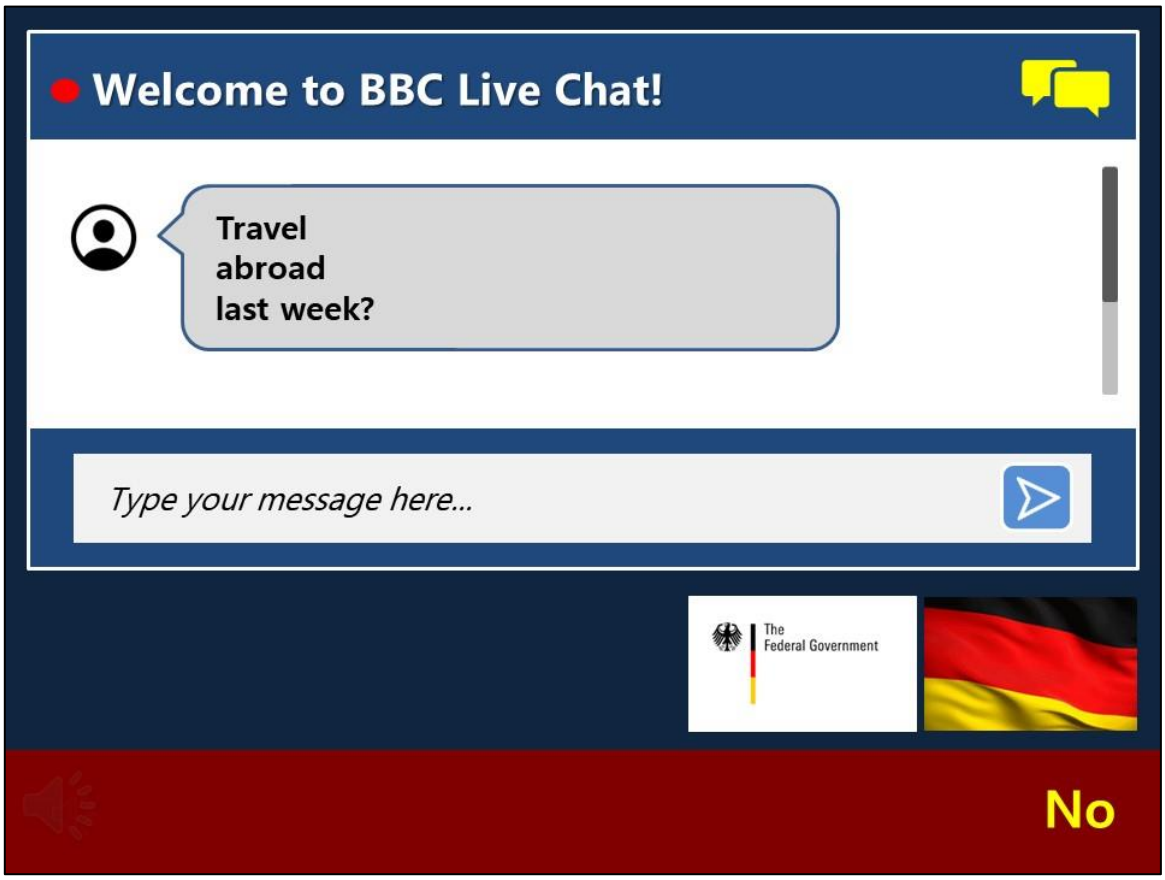

LIVE
G7


politics live
BBC
Live
Chat
LIVE
BREAKING NEWS
각국 정상들이 귀국길에 오르고 있습니다.
BBC
NEWS
BREAKING


Welcome to BBC Live Chat!
What do now?
Type your message here...
The Federal Government
head to the airport


Welcome to BBC Live Chat!
Where fly to now?
Type your message here...
The Federal Government
Berlin


Welcome to BBC Live Chat!
What
do
at duty free shops
yesterday?
Type your message here...
The
Federal Government
buy a lipstick


Welcome to BBC Live Chat!
Travel
abroad
last week?
Type your message here...
The
Federal Government
No


인터뷰에 응해 주셔서
진심으로 감사드립니다!
BBC
NEWS


BERLIN
LONDON
BBC
WORLD
NEWS
JOHANNESBURG

Appendix A18: A Sample of SOPT in Study 2

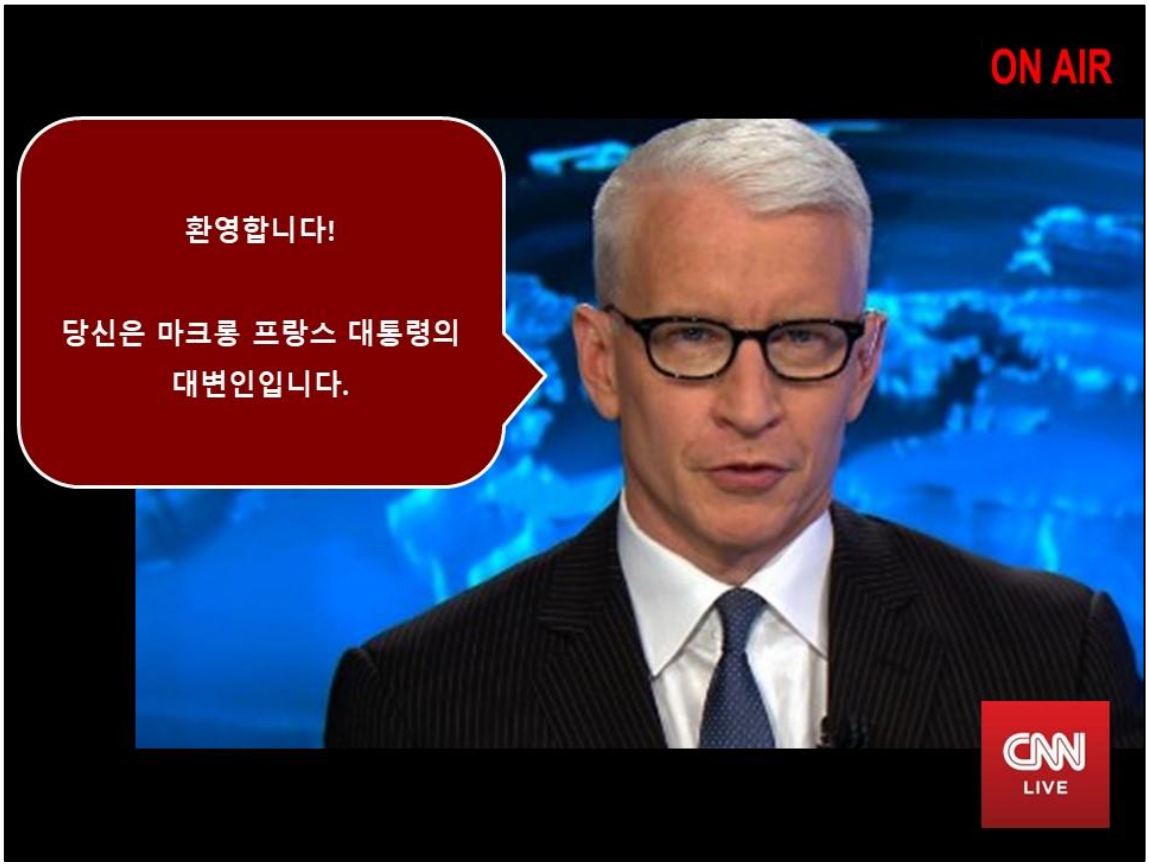





INTERVIEW WITH CNN
CNN
BREAKING
NEWS


G7
BREAKING NEWS
오늘 마크롱 프랑스 대통령은 G7 회의에서
각국 정상들과 회담을 가질 예정입니다.
CNN
LIVE


SMS
sending...
BREAKING NEWS
마크롱 대통령에 대해 궁금한 점이 있나요?
당신의 질문을 CNN 뉴스에 보내세요!
CNN
LIVE


ON AIR
환영합니다!
당신은 마크롱 프랑스 대통령의
대변인입니다.
CNN
LIVE

DIRECTIONS:

당신은 마크롱 프랑스 대통령의 대변인입니다.

마크롱 대통령은 영국 콘월에서 열리는 G7 회의에서 각국 정상들과 회담을 가질 예정입니다. 오늘 당신은 대통령을 대신해 CNN과 인터뷰를 할 예정입니다.



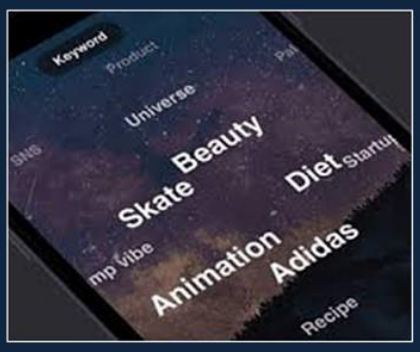


With the Macron's spokesperson

DIRECTIONS:

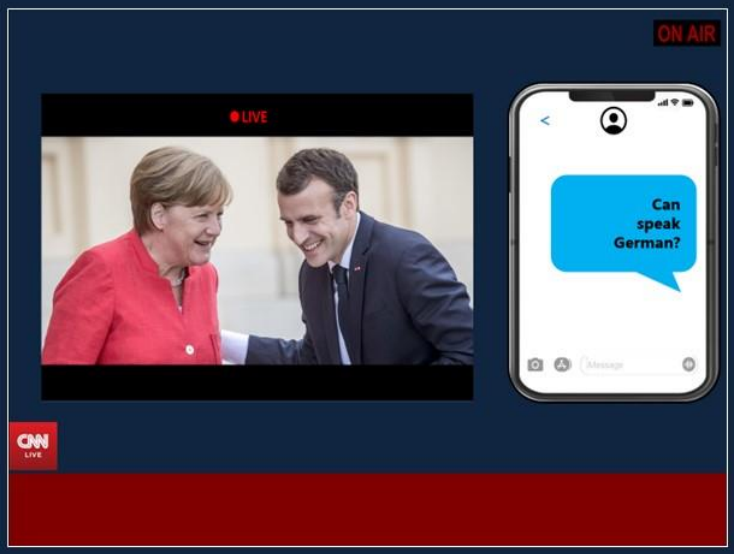
시청자들은 마크롱 대통령에 대해 궁금한 점을 문자로 질문할 수 있습니다.
시간 관계상 키워드로 된 짧은 질문만 가능합니다.
당신은 대변인으로서 시청자들의 질문에 답해야 합니다.





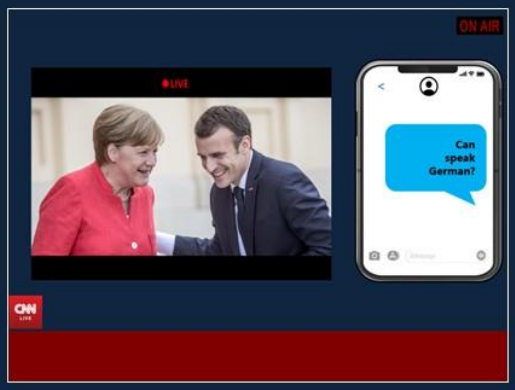
DIRECTIONS:

1) 시청자의 질문은 화면에 제시될 것입니다.



DIRECTIONS:

1) 시청자의 질문은 화면에 제시될 것입니다.
2) 제공된 화면을 바탕으로 질문에 대답하세요.
3) 공식 석상이므로, 반드시 완성형 문장으로 답하세요.



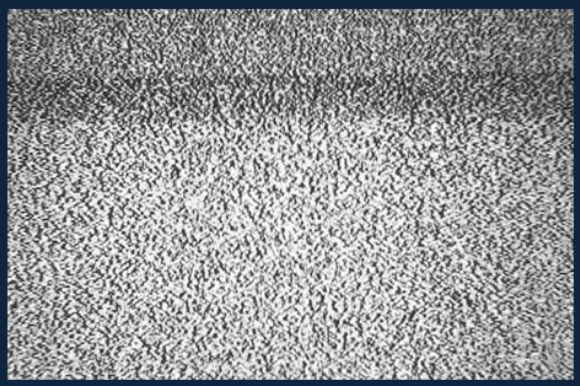
Yes.	✗
Yes, he can.	✗
Yes, he can speak German.	○

Let's begin!



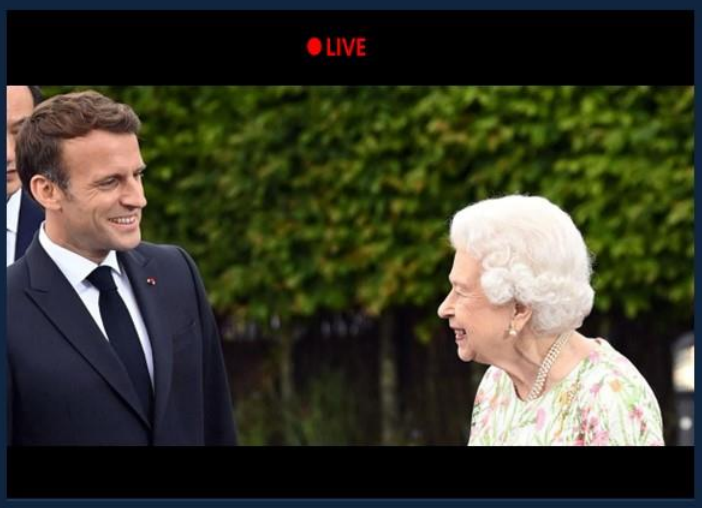
● LIVE





Now...

ON AIR



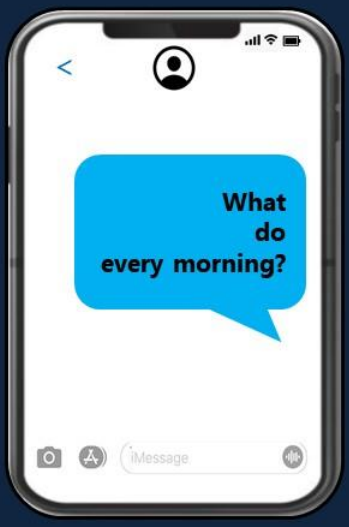


CNN LIVE

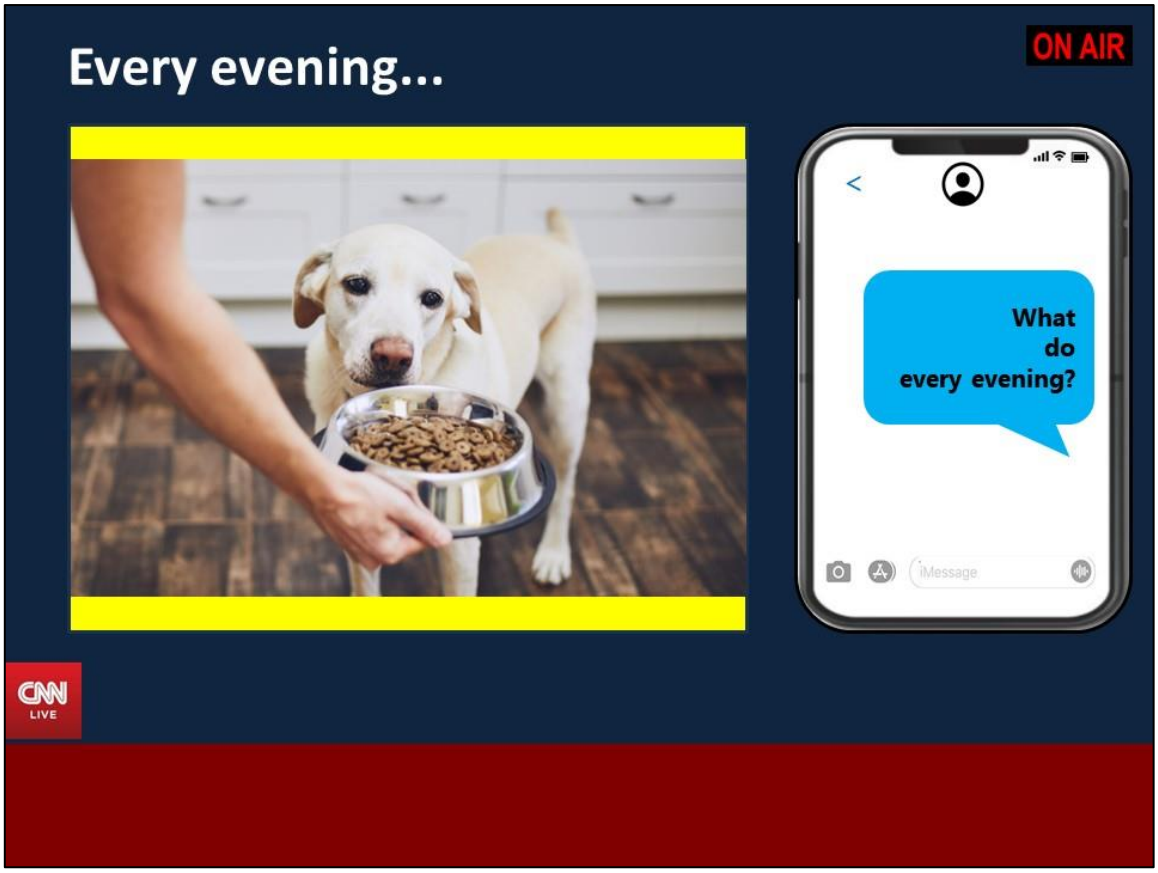
Every morning...

ON AIR





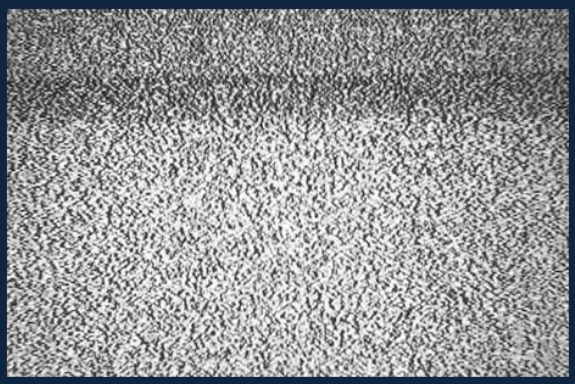
CNN LIVE


Every evening...
ON AIR
What do every evening?
iMessage
CNN LIVE


LIVE
BREAKING NEWS
G7 정상회의가 시작됐습니다.
CNN LIVE

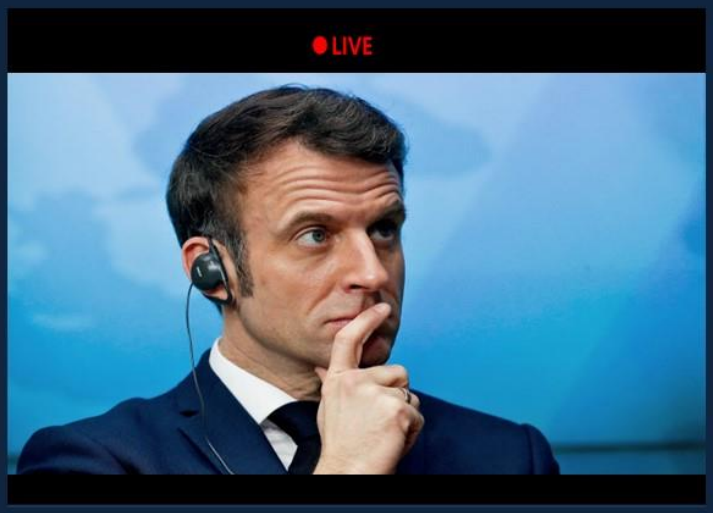
● LIVE





Now...

ON AIR





CNN LIVE

Every Monday...

ON AIR





CNN LIVE

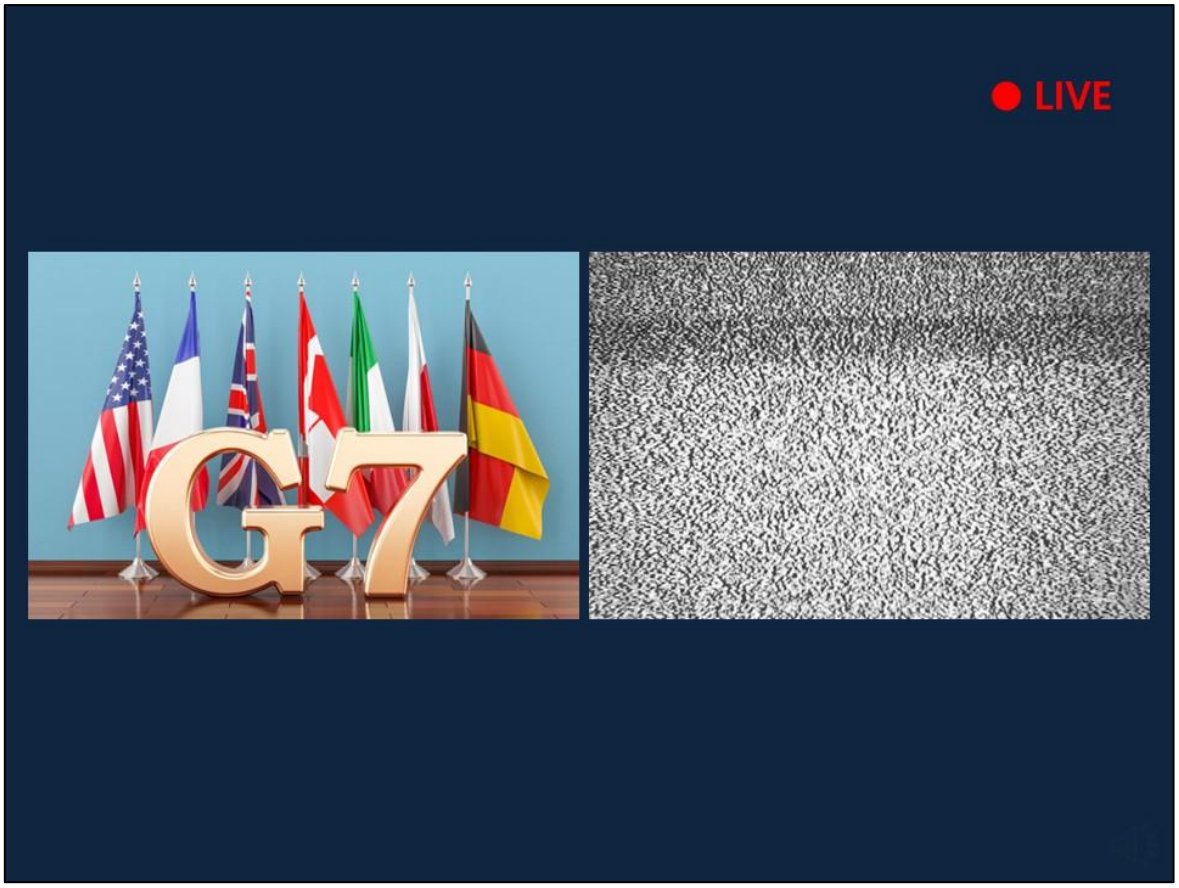
Every Tuesday...

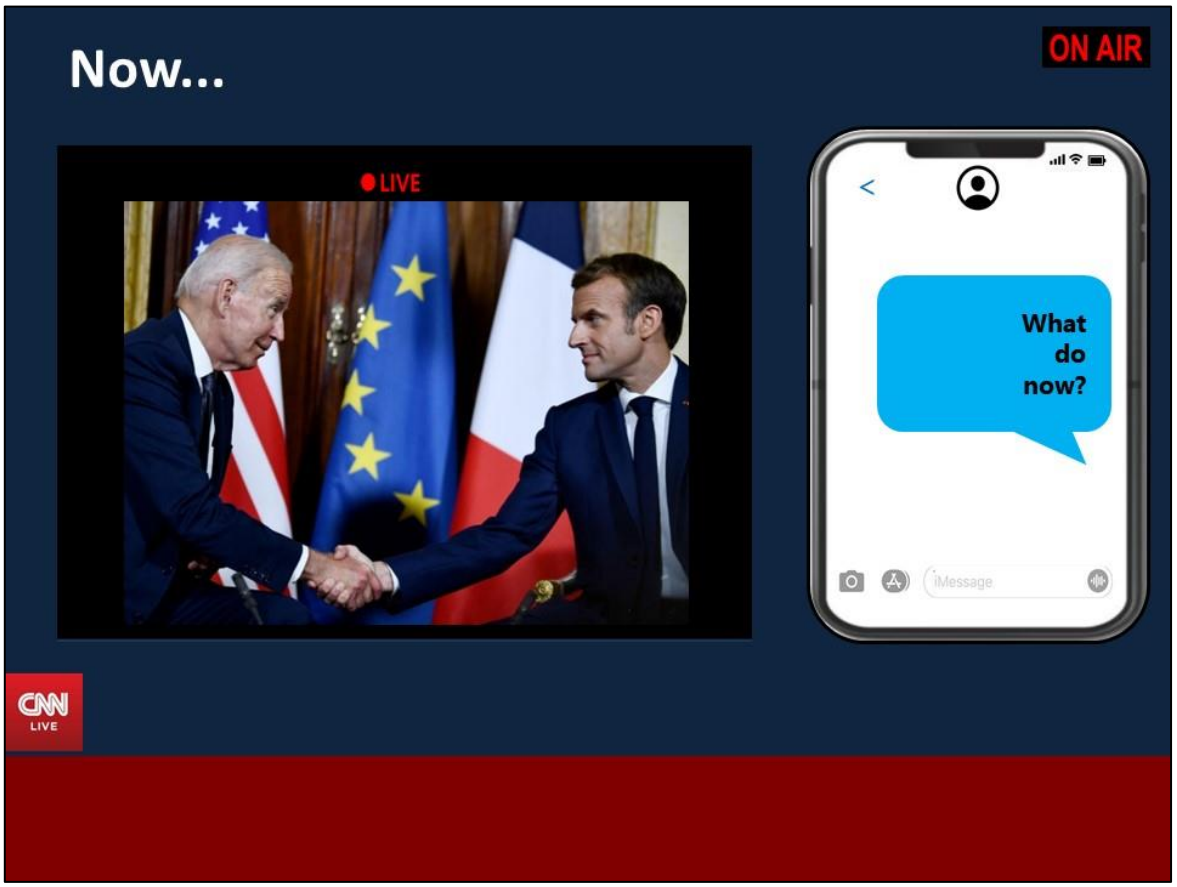
ON AIR





CNN LIVE


LIVE
G7


Now...
ON AIR
LIVE
What do now?
iMessage
CNN LIVE

Every Wednesday...

ON AIR





CNN LIVE

Every Thursday...

ON AIR





CNN LIVE

Every Friday...

ON AIR

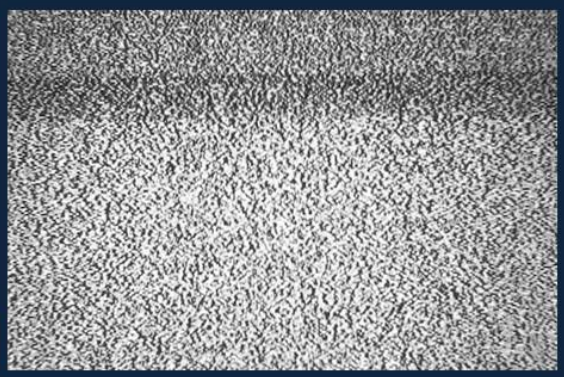




CNN LIVE

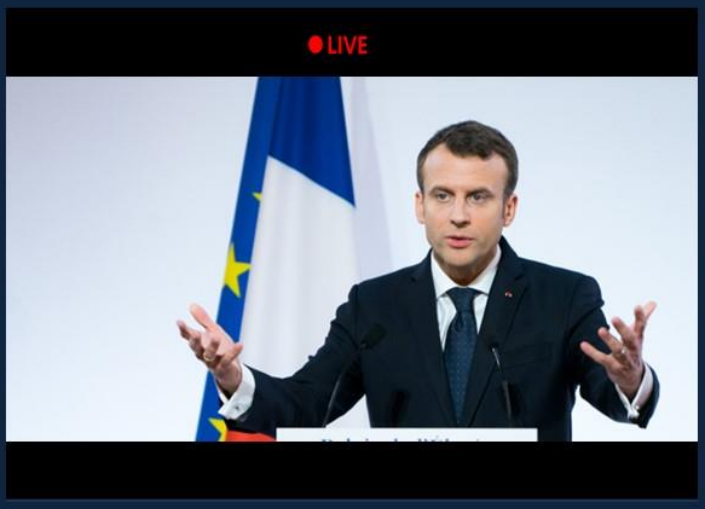
● LIVE





Now...

ON AIR





CNN LIVE

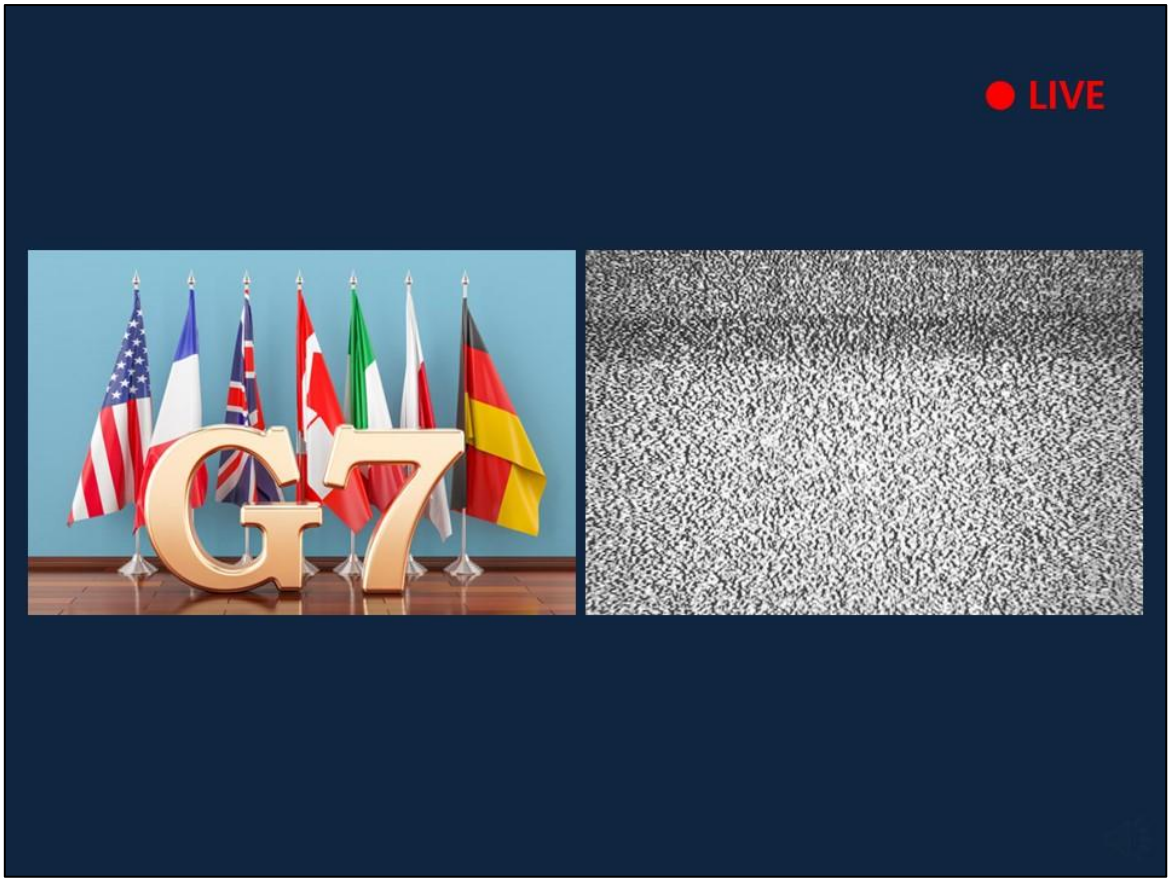
Every Saturday...

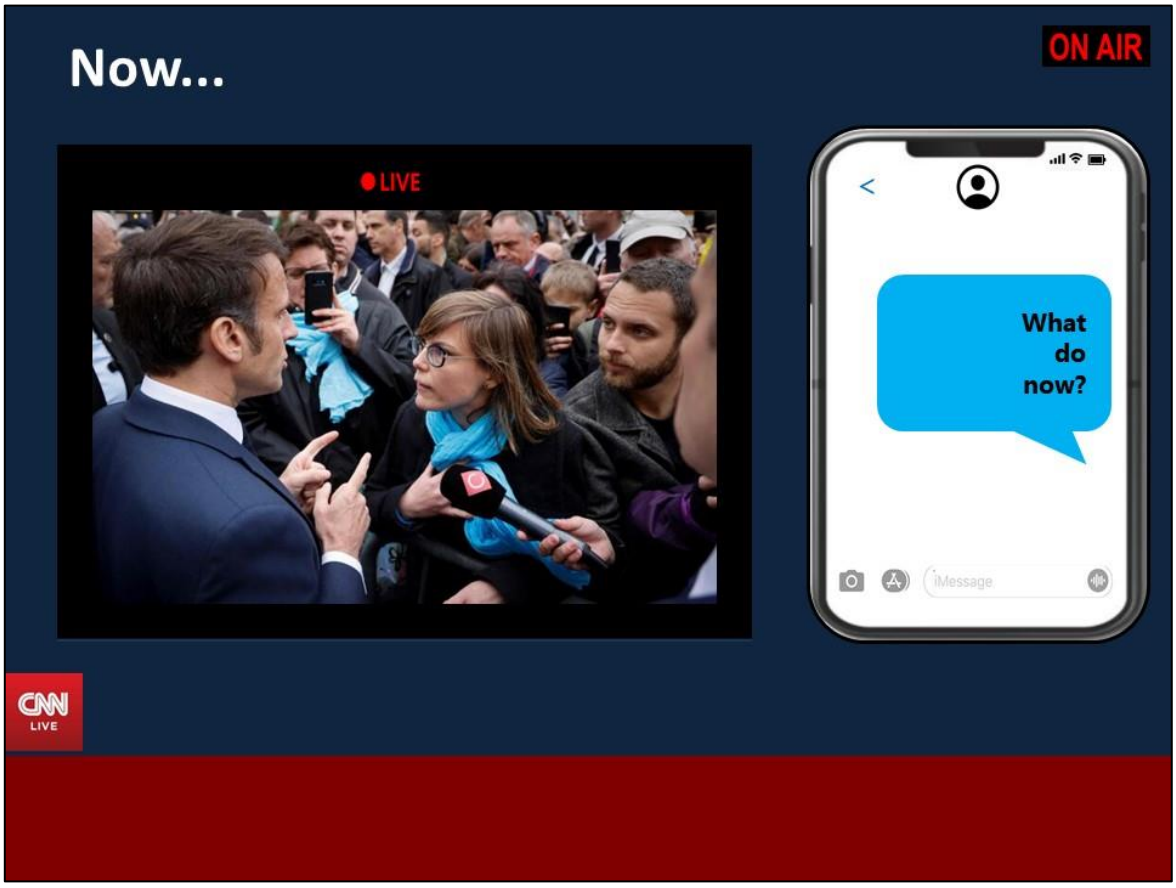
ON AIR

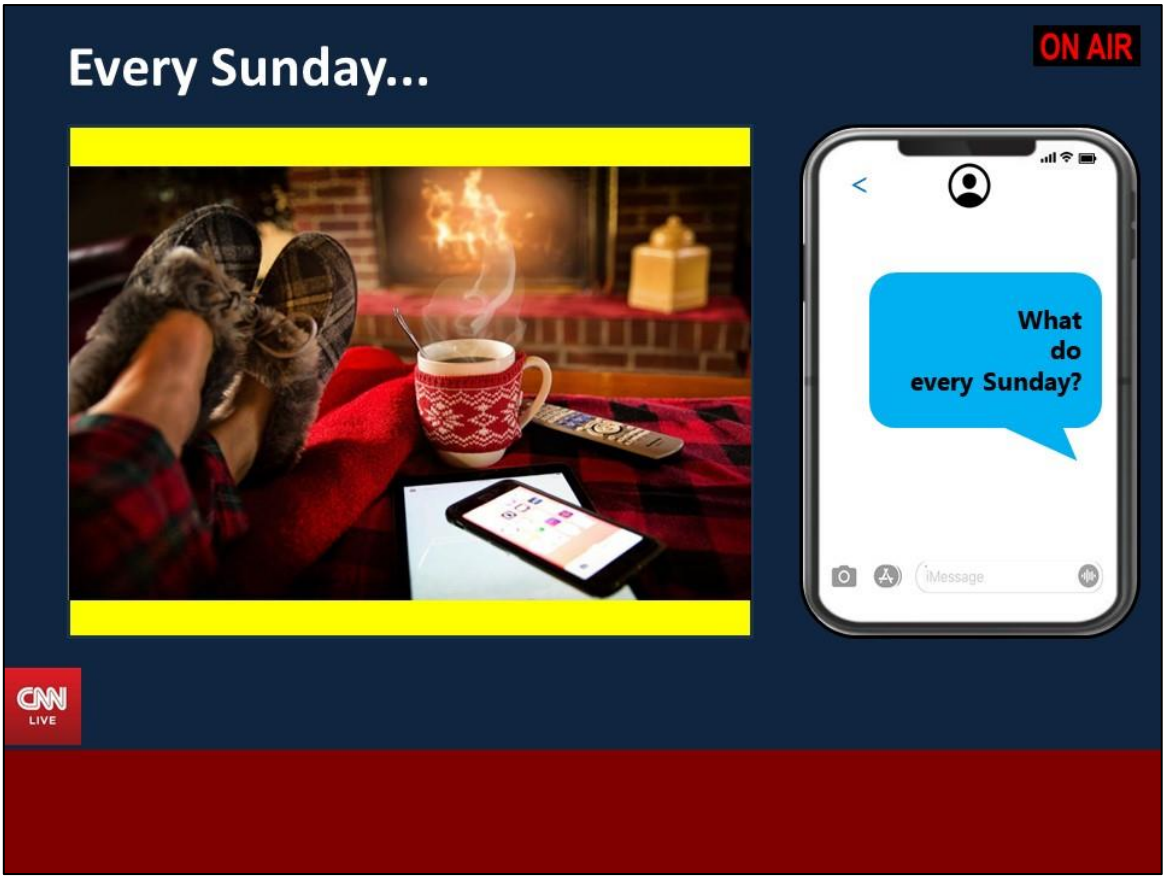


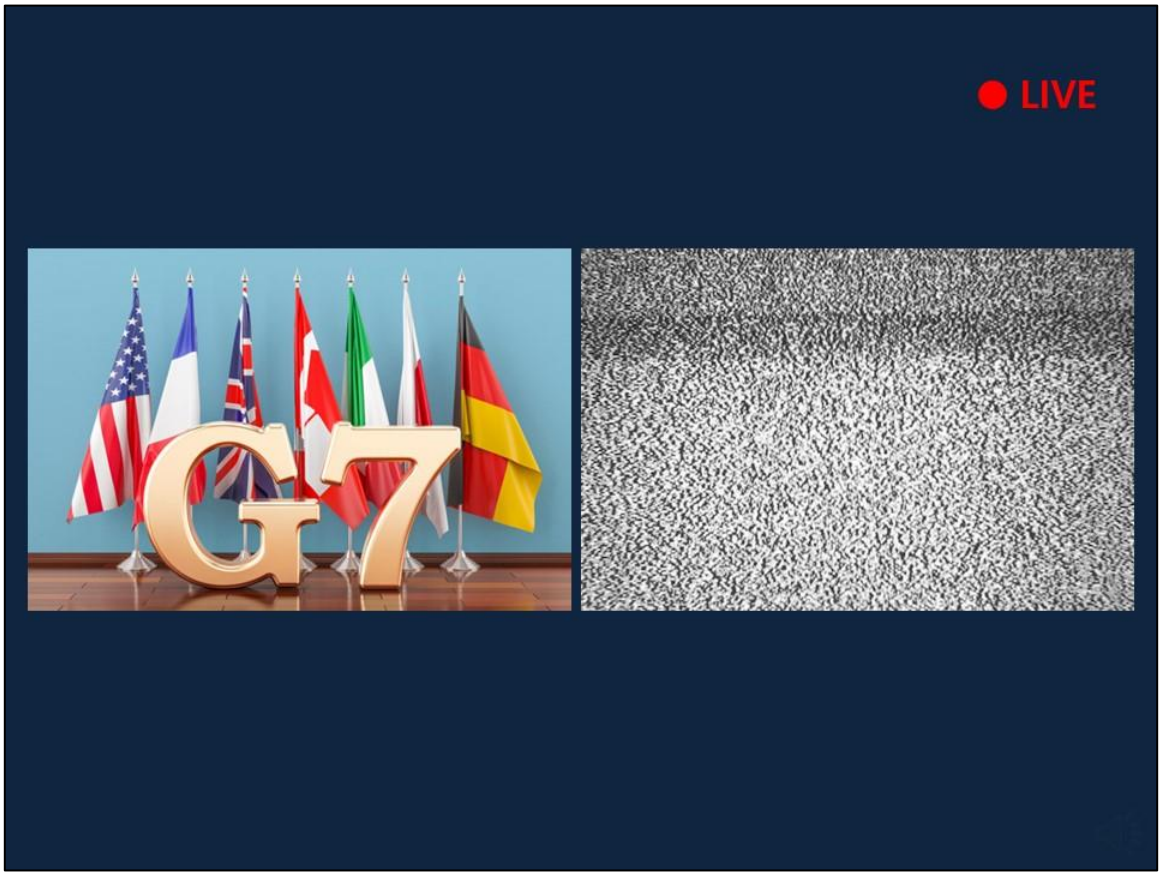


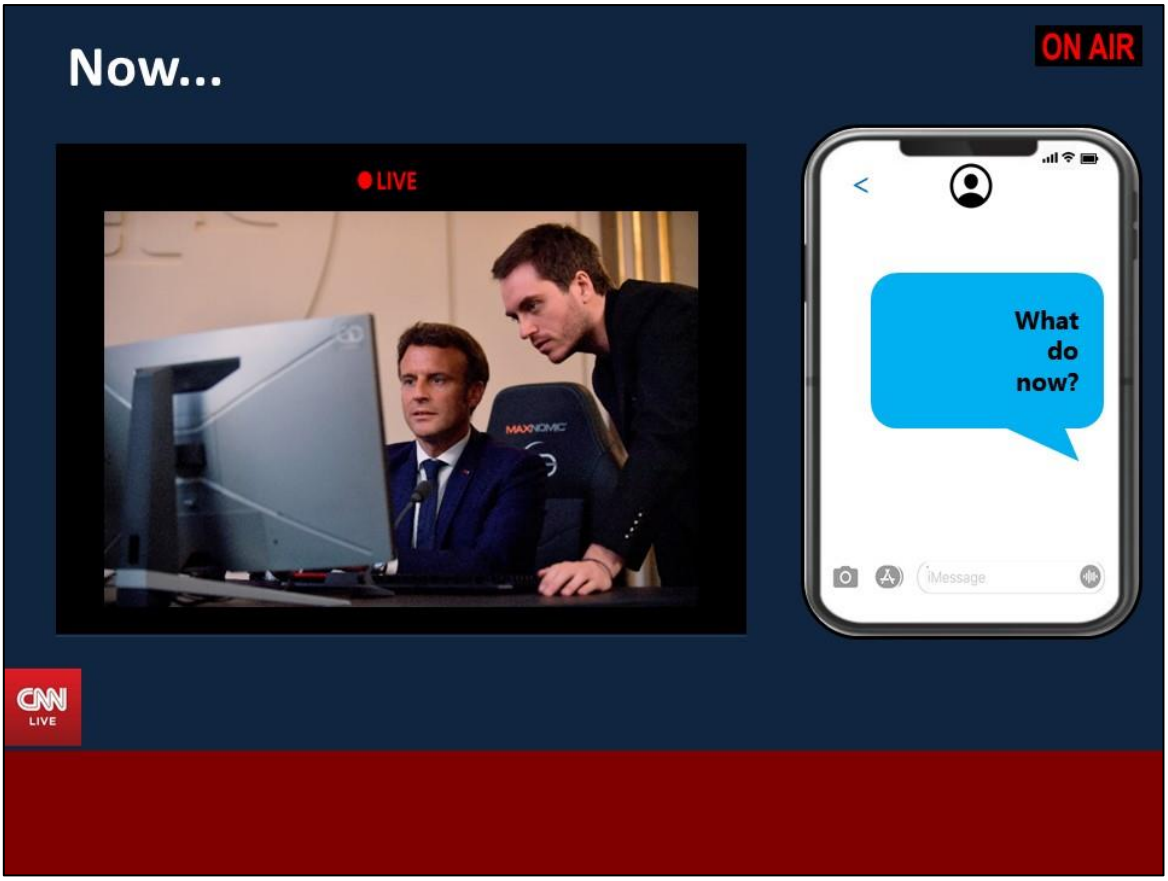
CNN LIVE

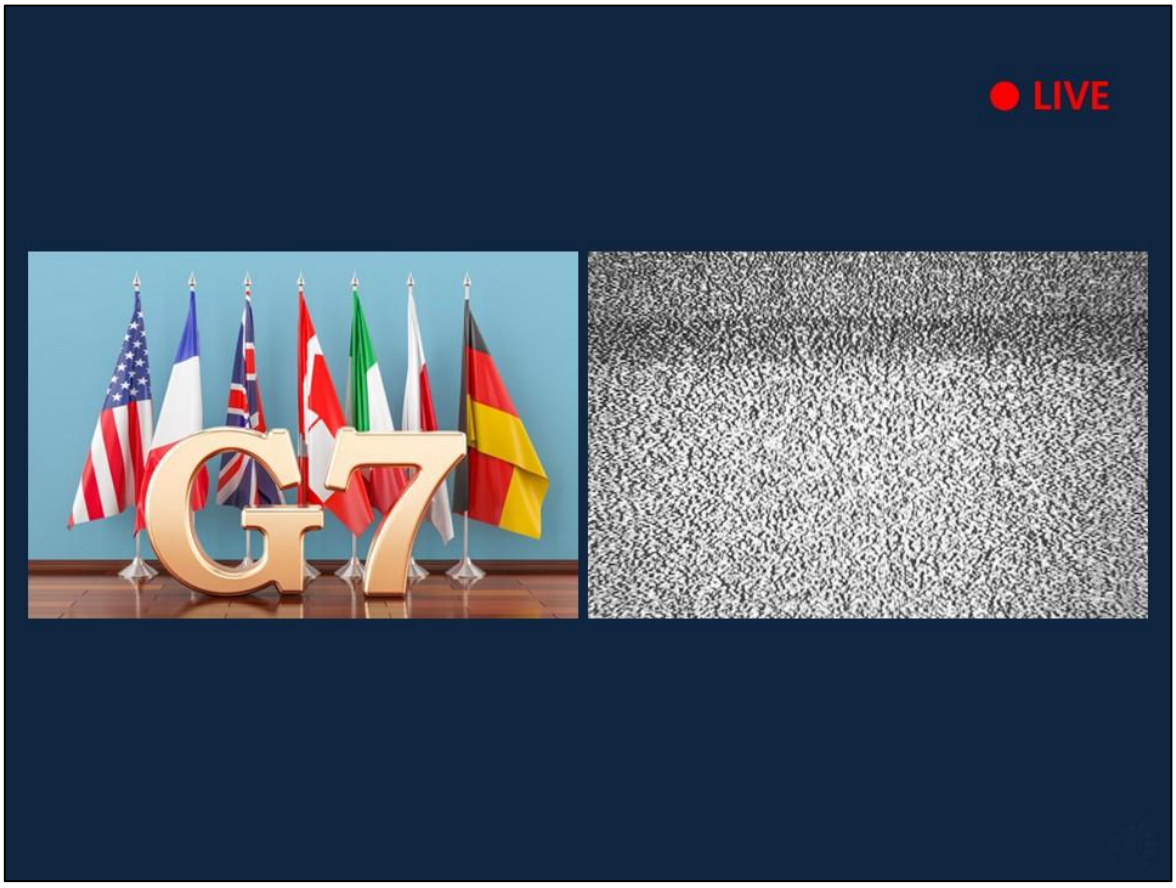

LIVE
G7


Now...
ON AIR
LIVE
What do now?
iMessage
CNN LIVE

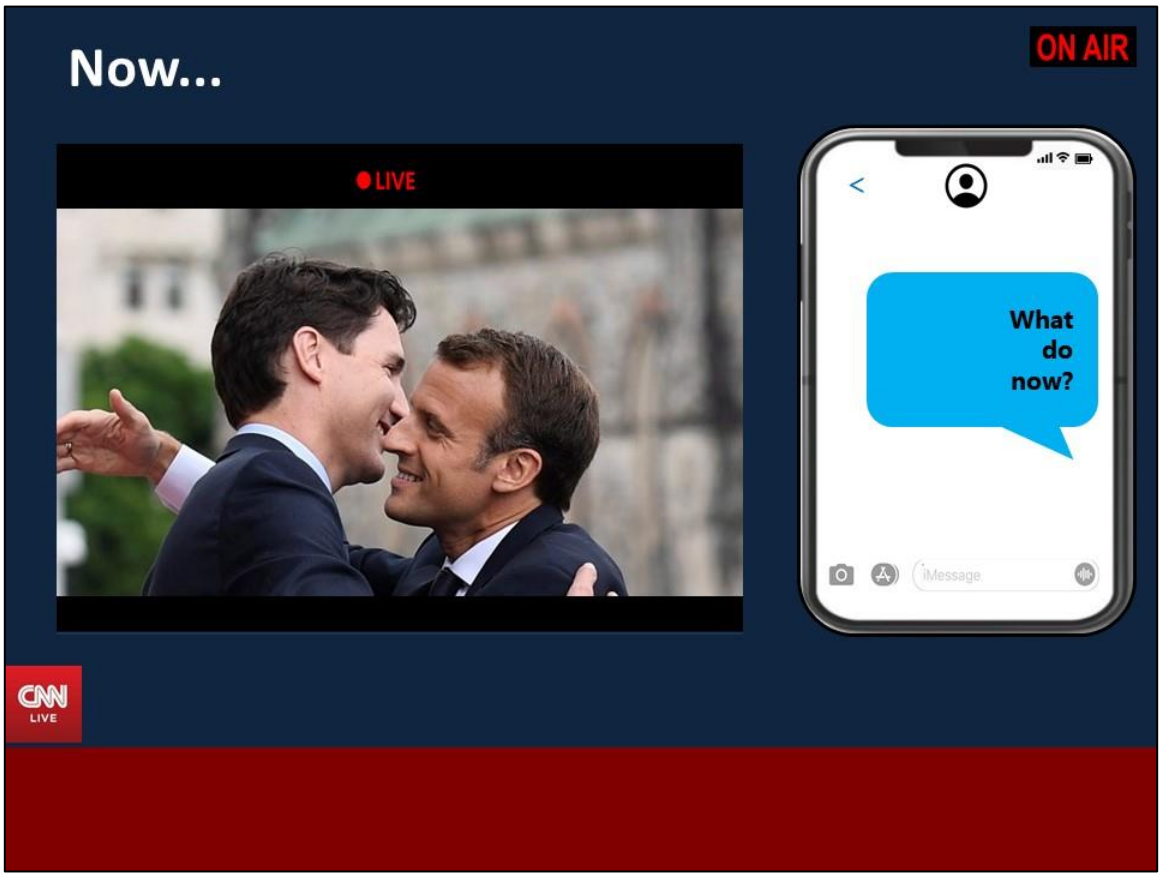

Every Sunday...
ON AIR
What do every Sunday?
iMessage
CNN LIVE

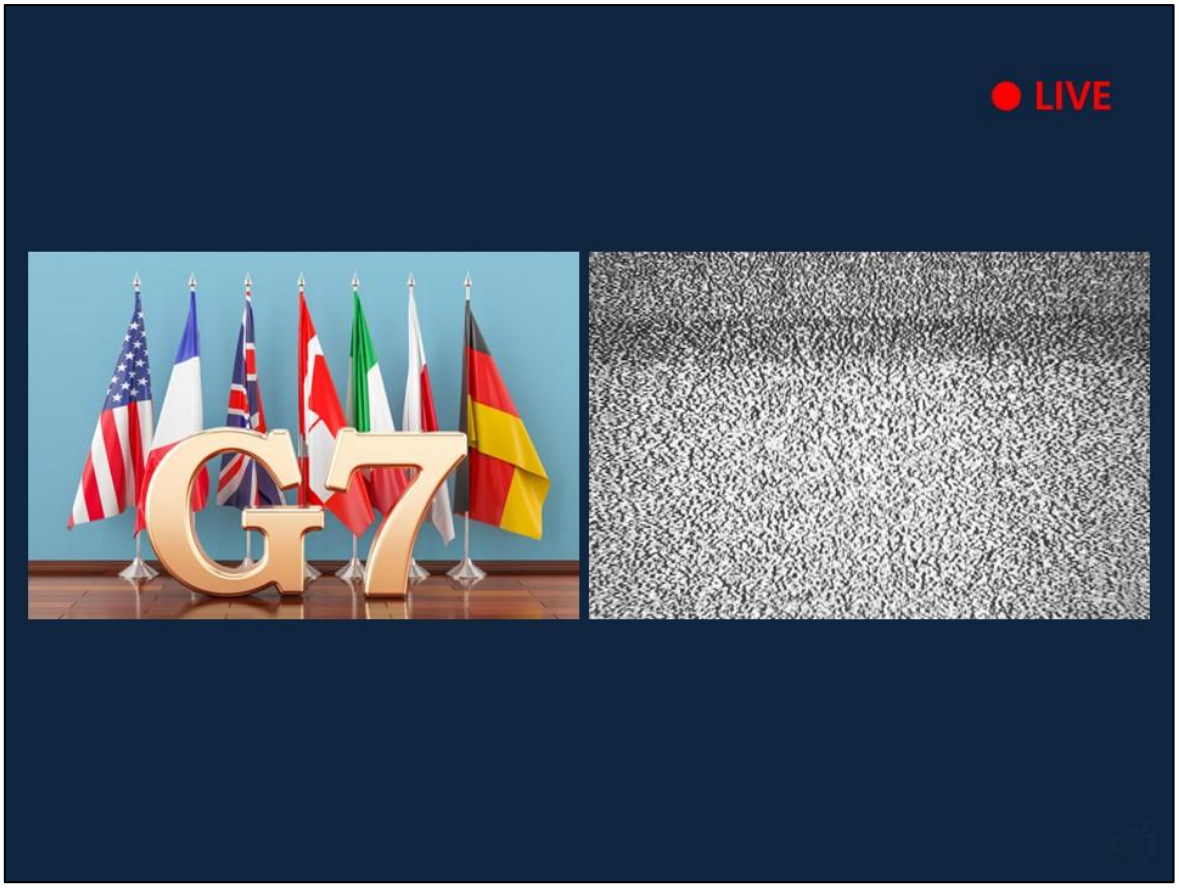

LIVE
G7

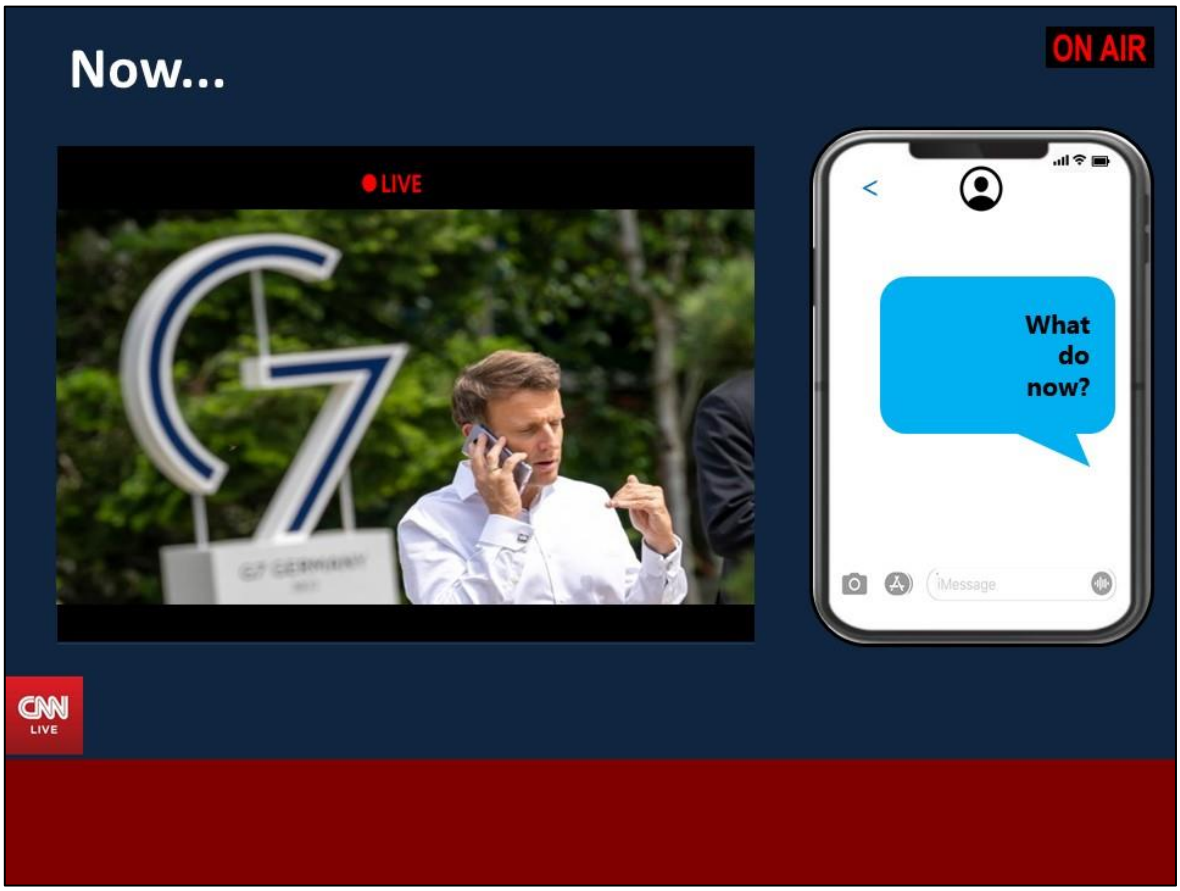

Now...
ON AIR
LIVE
What do now?
iMessage
CNN LIVE


LIVE
G7

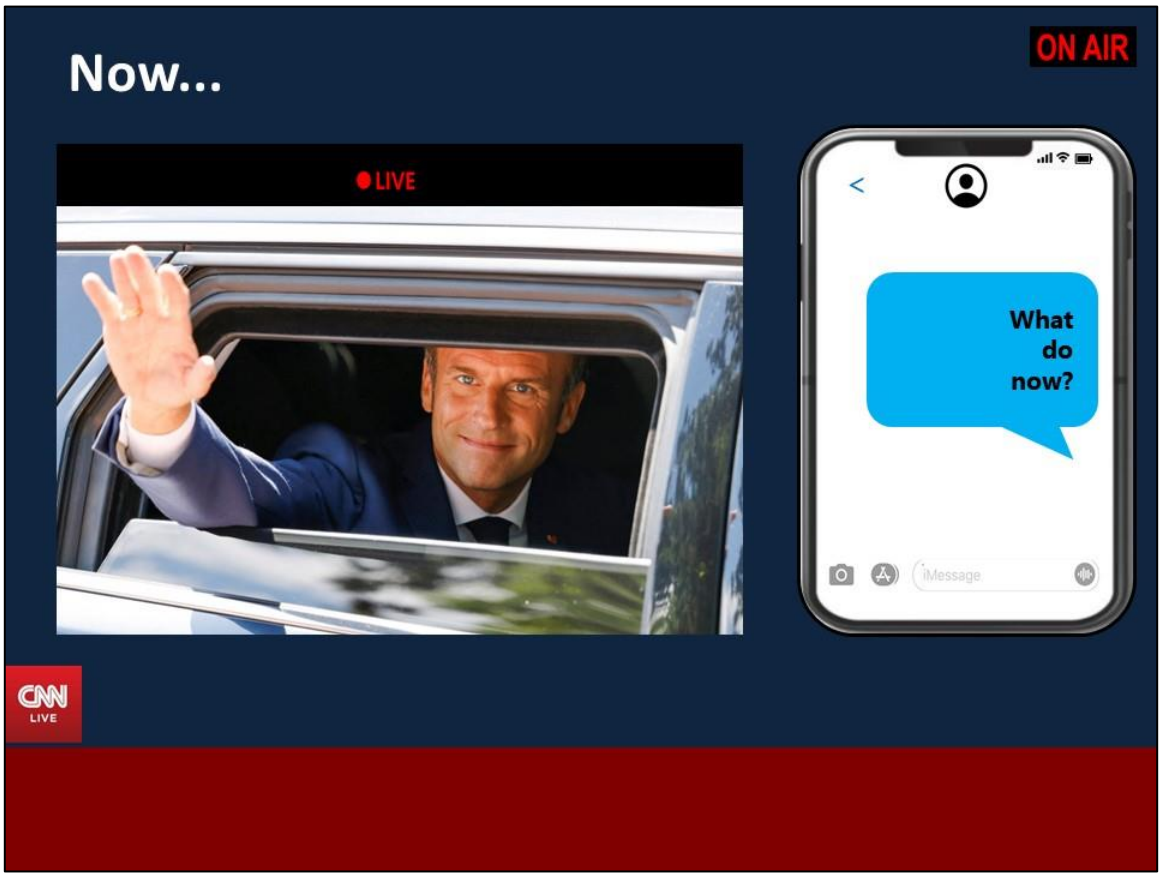

THE CORNISH ARMS
LIVE
BREAKING NEWS
G7 정상회의가 종료되었습니다.
CNN
LIVE

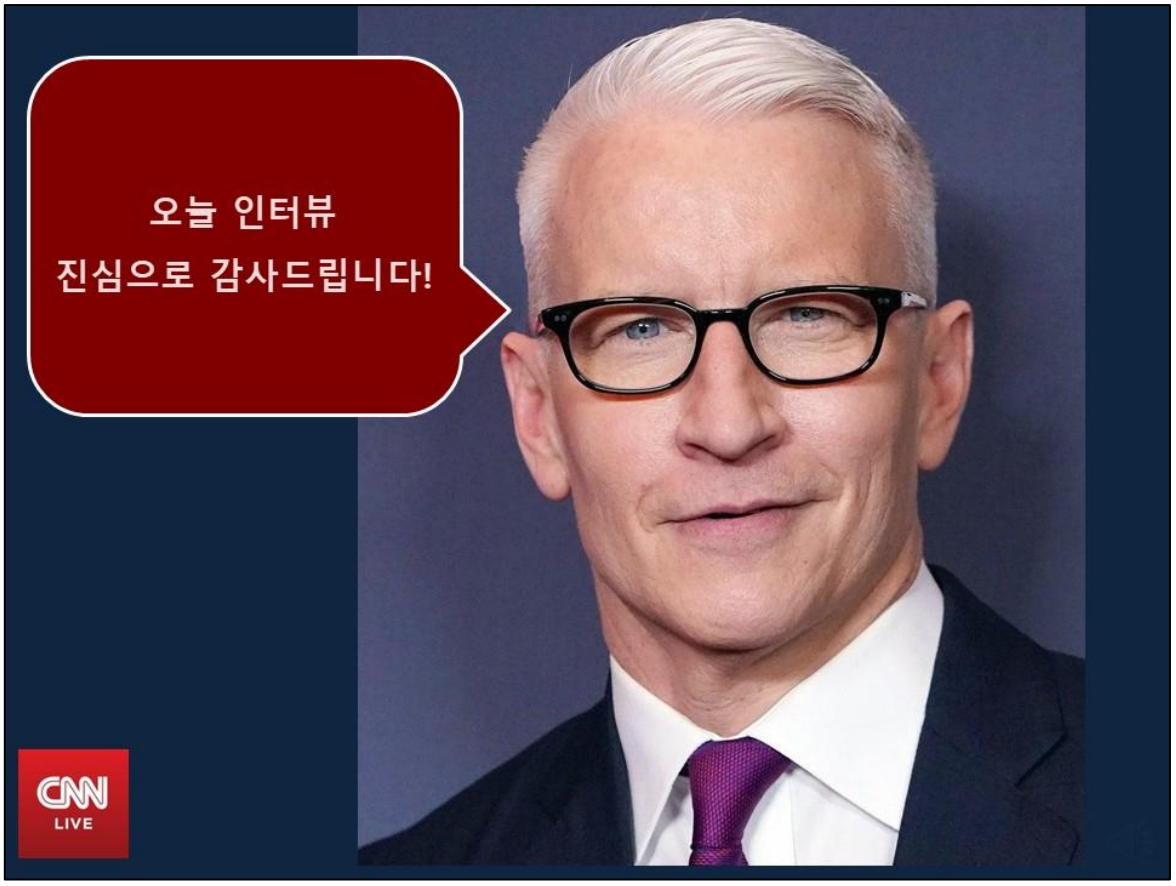

Now...
ON AIR
LIVE
What do now?
iMessage
CNN
LIVE


LIVE
G7


Now...
ON AIR
LIVE
What do now?
iMessage
CNN LIVE


LIVE
LIVE
BREAKING NEWS
각국 정상들이 귀국길에 오르고 있습니다.
CNN
LIVE


Now...
ON AIR
LIVE
What do now?
iMessage
CNN
LIVE


오늘 인터뷰
진심으로 감사드립니다!
CNN
LIVE


BERLIN
LONDON
CNN
BREAKING
NEWS
JOHANNESBURG

Appendix A19: Background Questionnaire for Study 2

설 문 지 (Questionnaire)

참가자 번호: _____

귀하께서는 각 항목별로 어느 부분에 해당이 되는지 **각 해당란에 체크**해 주시기 바랍니다.

1. 귀하의 성별은 무엇입니까?	**남성**	**여성**	**기타**		
	☐	☐	☐		
2. 귀하의 연령은 어떻게 되십니까? (　　세)	**10 대**	**20 대**	**30 대**	**40 대**	**50 대**
	☐	☐	☐	☐	☐
3. 귀하의 시력은 얼마입니까?	**2.0 이상**	**2.0-1.5**	**1.5-1.0**	**1.0 이하**	
	☐	☐	☐	☐	
4. 안경이나 콘택트렌즈를 사용 중이신가요?	**예**	**아니요**			
	☐	☐			
5. 안경이나 콘택트렌즈의 교정시력은 어떻게 되시나요?	**2.0 이상**	**2.0-1.5**	**1.5-1.0**	**1.0 이하**	
	☐	☐	☐	☐	
6. 글을 읽을 때 어떠한 매체를 선호하시나요?	**종이책**	**전자책**	**핸드폰**	**테블릿 PC**	**컴퓨터 (노트북)**
	☐	☐	☐	☐	☐
7. 표준 영어 능력 시험 점수가 있습니까??	**토익**	**토플**	**아이엘츠**	**기타**	
	☐	☐	☐	☐	
8. 영어를 언제 배우기 시작하셨습니까?	**10 대**	**20 대**	**30 대**	**40 대**	**50 대**
	☐	☐	☐	☐	☐
9. 총 영어 교육 기간을 적어주세요.	_____년_____ 개월				
10. 영어권 국가에서 생활한 경험이 있나요?	**예**	**아니요**			
	☐	☐			

11. 있다면 국가 이름과 체류 기간 및 영어권 국가에서의 교육 기간을 적어주세요.	
● 국가: ① ② ③	● 체류기간/교육 기간 ① ② ③

12. 1(초보)부터 5(고급)까지의 척도에서 각각의 영어 능력을 평가해보세요					
	1	**2**	**3**	**4**	**5**
① 말하기	☐	☐	☐	☐	☐
② 듣기	☐	☐	☐	☐	☐
③ 읽기	☐	☐	☐	☐	☐
④ 쓰기	☐	☐	☐	☐	☐
13. 한국어와 영어 외에 사용하는 언어가 있나요?	**예**	**아니요**			
	☐	☐			
14. 있다면 사용하는 언어를 적어 주세요. 그리고 각 언어에 대해 본인의 레벨을 체크해 주세요.					
	초급	**중급**	**고급**		
① 언어:	☐	☐	☐		
② 언어:	☐	☐	☐		
③ 언어:	☐	☐	☐		

15. 하루에 소셜 네트워킹 사이트는 (SNS) 얼마나 사용하시나요?

① 1 시간 미만

② 1–2 시간

③ 2–3 시간

④ 4 시간 이상

16. 어떤 SNS 애플리케이션이나 플랫폼을 선호하시나요? (예: Facebook, Instagram, Twitter)

① Facebook

② Instagram

③ Twitter

④ 카카오톡

⑤ 기타:

17. 어떠한 유형의 SNS 커뮤니케이션을 선호하시나요?

① 텍스트 기반 (예: 문자, 이메일, 실시간 문자 채팅)

② 음성 기반 (예: 전화)

③ 비디오 기반 (예: 화상 채팅)

④ 혼합 형태 (텍스트, 음성, 비디오)

⑤ 기타:

18. **실시간 커뮤니케이션 플랫폼 중에 어떤 것을 선호하시나요?**

① Zoom

② Microsoft Teams

③ Skype

④ 기타_____

		매우 편함	**편함**	**중립**	**불편함**	**매우 불편함**
19.	SNS 나 인터넷을 사용하는 것에 대해 얼마나 편안한가요?	☐	☐	☐	☐	☐
20.	하루에 인터넷은 얼마나 사용하시나요?	**1 시간 미만**	**2 시간**	**3 시간**	**4 시간 이상**	**기타**
		☐	☐	☐	☐	☐
21.	언어 습득을 위해 언어 학습 앱이나 온라인 리소스를 사용한 경험이 있나요?	**예**	**아니요**			
		☐	☐			

22. 있다면 학습 경험에 대해 간략하게 적어 주세요.

그 밖에 하고 싶은 말은 무엇인가요?

(자유롭게 서술하셔도 좋습니다)

소중한 의견 감사합니다!

Appendix A20: Exit Questionnaire for Study 2

출구 설문조사 (Exit Questionnaire)

귀하께서는 각 항목별로 어느 부분에 해당이 되는지 **각 해당란에 체크**해 주시기 바랍니다.

1. 나는 이 과업이 어려웠다고 생각한다.	**매우 그렇다**	**그렇다**	**보통이다**	**그렇지않다**	**매우 그렇지않다**
	☐	☐	☐	☐	☐
2. 나는 이 과업이 흥미롭다고 느꼈다	**매우 그렇다**	**그렇다**	**보통이다**	**그렇지않다**	**매우 그렇지않다**
	☐	☐	☐	☐	☐
3. 나는 이 과업이 내 영어실력 향상에 도움이 되었다고 생각한다.	**매우 그렇다**	**그렇다**	**보통이다**	**그렇지않다**	**매우 그렇지않다**
	☐	☐	☐	☐	☐
4. 나는 이런 과업을 또 해보고 싶다.	**매우 그렇다**	**그렇다**	**보통이다**	**그렇지않다**	**매우 그렇지않다**
	☐	☐	☐	☐	☐

1. 이 연구의 목적이 무엇이라고 생각합니까?

_____

2. 이 연구를 통해 영어 문법이나 단어를 배운 것이 있습니까?

_____

3. 이 연구에 참여하는 동안 특정한 영어 문법이나 단어에 주의를 기울였습니까?

_____

4. 주의를 기울였다면 그 이유는 무엇입니까?

_____

5. 연구 밖에서 영어 문법이나 단어를 찾아보거나 공부한 적이 있습니까?

_____

6. 귀하의 영어 답변에 대해 피드백을 받는다면 어떤 형태의 피드백을 받고 싶습니까?

① 말로 받는 피드백

② 글로 받는 피드백

③ 말과 글로 동시에 받는 피드백

④ 기타 _____

7. 고르신 피드백을 선호하는 이유는 무엇입니까?

_____

8. 화상회의를 통한 영어 학습이 효과가 있다고 생각하십니까?

_____

9. 그렇게 생각하는 이유는 무엇입니까?

_____

10. 실험 중에 화면의 어느 부분을 중점적으로 보셨나요? 그 이유는 무엇인가요?

_____

11. 피드백을 받으셨나요? (Yes → 12 번 질문 / No → 15 번 질문)

_____

12. 피드백이 특정한 영어 문법이나 단어에 주의를 기울이게 했습니까?

_____

13. 피드백이 귀하의 영어 학습에 도움이 됐다고 생각하십니까?

_____

14. 피드백이 귀하의 시선 처리에 영향을 미쳤나요? 그랬다면 이유는 무엇인가요?

_____

15. 전반적인 eye-tracking 실험과정에 대해 어떻게 생각하십니까?

_____

16. 전반적인 working memory 측정 과정에 대해 어떻게 생각하십니까?

_____

17. 이 연구에 대해 남기시고 싶은 말씀이나 제안 사항이 있다면 적어주세요.

_____

_____

_____

_____

Appendix A21. Results for the Linear Mixed-effects Model Examining Performance on the OPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Est	*SE*	*t*	*P*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control group as reference level										
(Intercept)	−1.56	0.07	−21.47	<.001***	.28	.30	<.01	<.01	<.01	.06
Oral	0.02	0.13	−0.12	.908						
Written	0.01	0.12	−0.05	.960						
Multi	0.01	0.10	0.15	.881						
Time2	0.13	0.10	1.34	.182						
Time3	0.09	0.10	0.88	.382						
Oral:Time2	0.95	0.14	6.71	<.001***						
Written:Time2	1.05	0.14	7.42	<.001***						
Multi:Time2	1.86	0.14	13.18	<.001***						
Oral:Time3	0.85	0.14	6.04	<.001***						
Written:Time3	0.98	0.14	7.01	<.001***						
Multi:Time3	1.82	0.14	12.93	<.001***						
Oral group as reference level										
(Intercept)	−1.58	0.11	−13.78	<.001***	.28	.30	<.01	<.01	.07	.27
Control	0.02	0.13	0.12	.908						
Written	0.01	0.11	0.08	.935						
Multi	0.03	0.13	0.24	.815						
Time2	1.08	0.10	10.85	<.001***						
Time3	0.93	0.10	9.40	<.001***						
Control:Time2	−0.95	0.14	−6.71	<.001***						
Written:Time2	0.10	0.14	0.71	.481						
Multi:Time2	0.91	0.14	6.48	<.001***						
Control:Time3	−0.85	0.14	−6.04	<.001***						
Written:Time3	0.14	0.14	0.97	.333						
Multi:Time3	0.97	0.14	6.88	<.001***						
Written group as reference level										
(Intercept)	−1.57	0.11	−14.05	<.001***	.28	.30	<.01	<.01	.07	.26
Control	0.01	0.12	0.05	.960						
Oral	−0.01	0.11	−0.08	.935						
Multi	0.02	0.12	0.17	.865						
Time2	1.18	0.10	11.84	<.001***						
Time3	1.07	0.10	10.77	<.001***						
Control:Time2	−1.05	0.14	−7.42	<.001***						
Oral:Time2	−0.10	0.14	−0.70	.481						
Multi:Time2	0.81	0.14	5.77	<.001***						
Control:Time3	−0.98	0.14	−7.01	<.001***						
Oral:Time3	−0.14	0.14	−0.97	.333						
Multi:Time3	0.83	0.14	5.91	<.001***						
Formula: *lmer(log_OPT_PP ~ 1 + Group*Time + (1|Participant) + (Group|Item))*										

Note. Time1 = pretest (reference level); Time2 = immediate posttest; Time3 = delayed posttest.

Appendix A22. Results for the Logistic Mixed-effects Models Examining Performance on the OPT – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
Control group as reference level						
(Intercept)	–1.76	0.20	–8.66	<.001***	0.17	[0.12, 0.26]
Oral	–0.09	0.24	–0.37	.713	0.91	[0.57, 1.47]
Written	–0.03	0.24	–0.12	.903	0.97	[0.60, 1.56]
Multi	–0.06	0.24	–0.24	.807	0.94	[0.59, 1.52]
Time2	0.03	0.24	0.12	.904	1.03	[0.64, 1.64]
Time3	0.03	0.24	0.12	.904	1.03	[0.64, 1.64]
Oral:Time2	1.26	0.32	3.88	<.001***	3.52	[1.87, 6.62]
Written:Time2	1.33	0.32	4.12	<.001***	3.77	[2.01, 7.05]
Multi:Time2	3.09	0.33	9.30	<.001***	22.03	[11.53, 42.12]
Oral:Time3	0.68	0.33	2.07	.038*	1.98	[1.04, 3.77]
Written:Time3	0.70	0.33	2.15	.032*	2.02	[1.07, 3.82]
Multi:Time3	2.46	0.33	7.60	<.001***	11.76	[6.25, 22.13]
Oral group as reference level						
(Intercept)	–1.85	0.21	–8.90	<.001***	0.16	[0.11, 0.24]
Control	0.09	0.25	0.37	.713	1.09	[0.68, 1.77]
Written	0.06	0.25	0.25	.805	1.06	[0.66, 1.72]
Multi	0.03	0.25	0.12	.901	1.03	[0.64, 1.67]
Time2	1.29	0.22	5.91	<.001***	3.62	[2.37, 5.53]
Time3	0.71	0.23	3.15	.002**	2.04	[1.31, 3.17]
Control:Time2	–1.26	0.32	–3.88	<.001***	0.28	[0.15, 0.53]
Written:Time2	0.07	0.30	0.22	.823	1.07	[0.59, 1.94]
Multi:Time2	1.83	0.32	5.81	<.001***	6.26	[3.38, 11.60]
Control:Time3	–0.68	0.33	–2.07	.038*	0.51	[0.27, 0.96]
Written:Time3	0.02	0.32	0.06	.953	1.02	[0.55, 1.90]
Multi:Time3	1.78	0.31	5.68	<.001***	5.94	[3.22, 10.96]
Written group as reference level						
(Intercept)	–1.79	0.20	–8.74	<.001***	0.17	[0.11, 0.25]
Control	0.03	0.24	0.12	.903	1.03	[0.64, 1.65]
Oral	–0.06	0.25	–0.25	.805	0.94	[0.58, 1.52]
Multi	–0.03	0.24	–0.12	.902	0.97	[0.60, 1.56]
Time2	1.35	0.21	6.34	<.001***	3.88	[2.56, 5.88]
Time3	0.73	0.22	3.30	<.001***	2.08	[1.35, 3.20]
Control:Time2	–1.33	0.32	–4.13	<.001***	0.27	[0.14, 0.50]
Oral:Time2	–0.07	0.30	–0.22	.823	0.93	[0.52, 1.69]
Multi:Time2	1.77	0.31	5.64	<.001***	5.85	[3.18, 10.78]
Control:Time3	–0.70	0.33	–2.15	.032*	0.50	[0.26, 0.94]
Oral:Time3	–0.02	0.32	–0.06	.953	0.98	[0.53, 1.82]
Multi:Time3	1.76	0.31	5.68	<.001***	5.83	[3.18, 10.69]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.11	.33			
*glmer(OPT_PS ~ 1+ Group*Time + (1|Participant) + (Group|Item))*						

Appendix A23. Results for the Linear Mixed-effects Model Examining Performance on the WPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control group as reference level										
(Intercept)	−1.56	0.07	−21.72	<.001***	.29	.31	<.01	<.01	<.01	.05
Oral	−0.01	0.12	−0.12	.904						
Written	−0.01	0.12	−0.05	.960						
Multi	0.01	0.10	0.15	.881						
Time2	0.15	0.10	1.50	.126						
Time3	0.08	0.10	0.82	.415						
Oral:Time2	0.98	0.14	6.98	<.001***						
Written:Time2	1.05	0.14	7.52	<.001***						
Multi:Time2	1.91	0.14	13.63	<.001***						
Oral:Time3	0.85	0.14	6.13	<.001***						
Written:Time3	0.99	0.14	7.12	<.001***						
Multi:Time3	1.87	0.14	13.37	<.001***						
Oral group as reference level										
Intercept	−1.57	0.12	−13.52	<.001***	.29	.31	<.01	<.01	.08	.28
Control	0.01	0.13	0.12	.907						
Written	0.01	0.11	0.08	.938						
Multi	0.03	0.13	0.23	.817						
Time2	1.13	0.10	11.42	<.001***						
Time3	0.94	0.10	9.47	<.001***						
Control:Time2	−0.98	0.14	−6.98	<.001***						
Written:Time2	0.07	0.14	0.54	.592						
Multi:Time2	0.93	0.14	6.66	<.001***						
Control:Time3	−0.86	0.14	−6.13	<.001***						
Written:Time3	0.14	0.14	0.99	.323						
Multi:Time3	1.01	0.14	7.23	<.001***						
Written group as reference level										
Intercept	−1.56	0.11	−14.21	<.001***	.29	.31	<.01	<.01	.06	.25
Control	0.01	0.12	0.05	.960						
Oral	−0.01	0.11	−0.08	.938						
Multi	0.02	0.12	0.18	.860						
Time2	1.20	0.10	12.18	<.001***						
Time3	1.07	0.10	10.87	<.001***						
Control:Time2	−1.05	0.14	−7.51	<.001***						
Oral:Time2	−0.07	0.14	−0.54	.592						
Multi:Time2	0.86	0.14	6.12	<.001***						
Control:Time3	−1.00	0.14	−7.12	<.001***						
Oral:Time3	−0.14	0.14	−0.99	.323						
Multi:Time3	0.87	0.14	6.24	<.001***						
Formula: *lmer(log_WPT_PP ~ 1 + Group*Time + (1|Participant) + (Group|Item))*										

Appendix A24. Results for the Generalized Linear Mixed-effects Models Examining Performance on the WPT – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
Control group as reference level						
(Intercept)	–1.71	0.19	–9.06	<.001***	0.18	[0.12, 0.26]
Oral	–0.06	0.24	0.24	.810	0.94	[0.59, 1.51]
Written	–0.03	0.24	0.12	.905	0.97	[0.61, 1.55]
Multi	–0.09	0.24	0.36	.716	0.92	[0.57, 1.47]
Time2	0.03	0.24	0.12	.905	1.03	[0.65, 1.63]
Time3	0.03	0.24	–0.12	.905	1.03	[0.65, 1.63]
Oral:Time2	1.24	0.32	3.88	<.001***	3.45	[1.85, 6.42]
Written:Time2	1.46	0.32	4.60	<.001***	4.30	[2.32, 7.97]
Multi:Time2	3.58	0.34	10.46	<.001***	35.74	[18.35, 69.62]
Oral:Time3	0.64	0.33	1.97	.049*	1.89	[1.01, 3.57]
Written:Time3	0.71	0.32	2.19	.028*	2.03	[1.08, 3.81]
Multi:Time3	2.53	0.32	7.85	<.001***	12.55	[6.69, 23.52]
Oral group as reference level						
(Intercept)	–1.77	0.19	–9.22	<.001***	0.17	[0.12, 0.25]
Control	0.06	0.24	0.24	.810	1.06	[0.66, 1.69]
Written	0.03	0.24	0.12	.904	1.03	[0.64, 1.65]
Multi	–0.03	0.24	–0.12	.903	0.97	[0.60, 1.56]
Time2	1.27	0.21	5.93	<.001***	3.54	[2.34, 5.37]
Time3	0.67	0.22	3.00	.003**	1.95	[1.26, 3.01]
Control:Time2	–1.24	0.32	–3.88	<.001***	0.29	[0.16, 0.54]
Written:Time2	0.22	0.30	0.74	.462	1.25	[0.70, 2.24]
Multi:Time2	2.34	0.33	7.18	<.001***	10.37	[5.50, 19.57]
Control:Time3	–0.64	0.33	–1.97	.049*	0.53	[0.28, 0.99]
Written:Time3	0.07	0.31	0.22	.828	1.07	[0.58, 1.97]
Multi:Time3	1.89	0.31	6.08	<.001***	6.62	[3.61, 12.15]
Written group as reference level						
(Intercept)	–1.74	0.19	–9.14	<.001***	0.17	[0.12, 0.25]
Control	0.03	0.24	0.12	.905	1.03	[0.65, 1.64]
Oral	–0.03	0.24	–0.12	.904	0.97	[0.61, 1.56]
Multi	–0.06	0.24	–0.24	.808	0.94	[0.59, 1.51]
Time2	1.49	0.21	7.06	<.001***	4.42	[2.93, 6.66]
Time3	0.74	0.21	3.36	<.001***	2.09	[1.36, 3.20]
Control:Time2	–1.46	0.32	–4.60	<.001***	0.23	[0.13, 0.43]
Oral:Time2	–0.22	0.30	–0.74	.462	0.80	[0.45, 1.44]
Multi:Time2	2.12	0.32	6.55	<.001***	8.32	[4.43, 15.63]
Control:Time3	–0.71	0.32	–2.19	.028*	0.49	[0.26, 0.93]
Oral:Time3	–0.07	0.31	–0.22	.828	0.93	[0.51, 1.72]
Multi:Time3	1.82	0.31	5.90	<.001***	6.19	[3.39, 11.30]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.07	.26				
*glmer(WPT_PS ~ 1+ Group*Time + (1|Participant) + (1|Item))*						

Appendix A25. Results for the Linear Mixed-effects Model Examining Performance on the SOPT – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Est	*SE*	*t*	*P*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Control group as reference level										
(Intercept)	−1.58	0.07	−22.91	<.001***	.30	.32	<.01	<.01	<.01	<.01
Oral	−0.03	0.13	−0.23	.822						
Written	−0.01	0.13	−0.10	.918						
Multi	−0.00	0.10	0.03	.975						
Time2	0.17	0.10	1.70	.090						
Time3	0.12	0.10	1.18	.238						
Oral:Time2	0.98	0.14	7.07	<.001***						
Written:Time2	1.07	0.14	7.77	<.001***						
Multi:Time2	1.92	0.14	13.91	<.001***						
Oral:Time3	0.87	0.14	6.29	<.001***						
Written:Time3	0.99	0.14	7.17	<.001***						
Multi:Time3	1.90	0.14	13.72	<.001***						
Oral group as reference level										
Intercept	−1.61	0.12	−13.92	<.001***	.30	.32	<.01	<.01	.08	.28
Control	0.03	0.13	0.23	.822						
Written	0.02	0.12	0.15	.885						
Multi	0.03	0.13	0.22	.829						
Time2	1.14	0.10	11.71	<.001***						
Time3	0.98	0.10	10.09	<.001***						
Control:Time2	−0.97	0.14	−7.08	<.001***						
Written:Time2	0.10	0.14	0.69	.488						
Multi:Time2	0.95	0.14	6.85	<.001***						
Control:Time3	−0.87	0.14	−6.30	<.001***						
Written:Time3	0.12	0.14	0.88	.379						
Multi:Time3	1.03	0.14	7.43	<.001***						
Written group as reference level										
Intercept	−1.60	0.11	−14.26	<.001***	.30	.32	<.01	<.01	.07	.26
Control	0.01	0.12	0.11	.910						
Oral	−0.02	0.12	−0.15	.886						
Multi	0.01	0.12	0.09	.931						
Time2	1.24	0.10	12.68	<.001***						
Time3	1.11	0.10	11.33	<.001***						
Control:Time2	−1.07	0.14	−7.77	<.001***						
Oral:Time2	−0.10	0.14	−0.69	.489						
Multi:Time2	0.85	0.14	6.15	<.001***						
Control:Time3	−0.99	0.14	−7.18	<.001***						
Oral:Time3	−0.12	0.14	−0.88	.379						
Multi:Time3	0.91	0.14	6.55	<.001***						
Formula: *lmer(log_SOPT_PP ~ 1 + Group*Time + (1|Participant) + (Group|Item))*										

Appendix A26. Results for the Generalized Linear Mixed-effects Models Examining Performance on the SOPT – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
Control group as reference level						
(Intercept)	–1.83	0.21	–8.51	<.001***	0.16	[0.11, 0.24]
Oral	–0.09	0.25	–0.38	.706	0.91	[0.56, 1.48]
Written	–0.00	0.25	–0.00	.999	1.00	[0.62, 1.62]
Multi	–3.08	0.25	–0.12	.901	0.97	[0.60, 1.57]
Time2	0.05	0.24	0.25	.807	1.06	[0.66, 1.71]
Time3	–0.03	0.24	0.12	.902	1.03	[0.64, 1.66]
Oral:Time2	1.33	0.33	4.04	<.001***	3.78	[1.99, 7.19]
Written:Time2	1.40	0.32	4.29	<.001***	4.04	[2.14, 7.63]
Multi:Time2	3.18	0.38	9.41	<.001***	24.06	[12.45, 46.52]
Oral:Time3	0.85	0.33	2.53	.011*	2.34	[1.22, 4.49]
Written:Time3	0.81	0.33	2.45	.014*	2.25	[1.18, 4.29]
Multi:Time3	2.58	0.33	7.81	<.001***	13.20	[6.93, 25.83]
Oral group as reference level						
(Intercept)	–1.92	0.22	–8.76	<.001***	0.15	[0.10, 0.22]
Control	0.09	0.25	0.38	.706	1.10	[0.67, 1.79]
Written	0.09	0.25	0.38	.706	1.10	[0.67, 1.79]
Multi	0.06	0.25	0.25	.800	1.07	[0.65, 1.74]
Time2	1.39	0.22	6.29	<.001***	4.02	[2.61, 6.18]
Time3	0.88	0.23	3.86	<.001***	2.41	[1.54, 3.75]
Control:Time2	–1.33	0.33	–4.04	<.001***	0.26	[0.14, 0.50]
Written:Time2	0.07	0.31	0.21	.830	1.07	[0.59, 1.95]
Multi:Time2	1.85	0.32	5.76	<.001***	6.36	[3.40, 11.89]
Control:Time3	–0.85	0.33	–2.53	.011*	0.43	[0.22, 0.82]
Written:Time3	–0.04	0.32	–0.12	.908	0.96	[0.52, 1.79]
Multi:Time3	1.73	0.32	5.46	<.001***	5.65	[3.04, 10.48]
Written group as reference level						
(Intercept)	–1.83	0.21	–8.51	<.001***	0.16	[0.11, 0.24]
Control	0.00	0.25	0.00	.999	1.00	[0.62, 1.62]
Oral	–0.09	0.25	–0.38	.706	0.91	[0.56, 1.48]
Multi	–0.03	0.25	–0.12	.901	0.97	[0.60, 1.57]
Time2	1.46	0.22	6.77	<.001***	4.29	[2.82, 6.53]
Time3	0.82	0.22	3.79	<.001***	2.32	[1.50, 3.58]
Control:Time2	–1.40	0.33	–4.29	<.001***	0.25	[0.13, 0.47]
Oral:Time2	–0.07	0.31	–0.21	.830	0.94	[0.51, 1.71]
Multi:Time2	1.78	0.32	5.63	<.001***	5.95	[3.21, 11.04]
Control:Time3	–0.81	0.33	–2.45	.014*	0.44	[0.23, 0.85]
Oral:Time3	0.04	0.32	0.12	.908	1.04	[0.56, 1.93]
Multi:Time3	1.77	0.31	5.65	<.001***	5.86	[3.18, 10.79]
Random effects	*variance*	*SD*				
Participant (Int.)	<.01	<.01				
Item (Int.)	.14	.37				
*glmer(WPT_PS ~ 1+ Group*Time + (1|Participant) + (1|Item))*						

Appendix A27. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (NWS) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Est	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
				NWS						
Control group as reference level										
(Intercept)	–3.40	1.21	–2.82	.005**	.31	.32	<.01	<.01	.02	.15
Oral:Time2:NWS	2.37	1.11	2.13	.033*						
Written:Time2:NWS	0.01	1.08	0.01	.990						
Multi:Time2:NWS	1.34	1.12	1.20	.232						
Oral:Time3:NWS	2.62	1.11	2.36	.018*						
Written:Time3:NWS	–0.22	1.13	–0.20	.842						
Multi:Time3:NWS	1.86	1.12	1.66	.097						
Oral group as reference level										
(Intercept)	–0.93	1.16	–0.80	.426	.31	.32	<.01	<.01	.02	.15
Control:Time2:NWS	–2.37	1.11	–2.13	.033*						
Written:Time2:NWS	–2.36	1.07	–2.20	.028*						
Multi:Time2:NWS	–1.03	1.11	–0.93	.350						
Control:Time3:NWS	–2.62	1.11	–2.36	.018*						
Written:Time3:NWS	–2.85	1.11	–2.55	.011*						
Multi:Time3:NWS	–0.76	1.11	–0.69	.492						
Written group as reference level										
(Intercept)	–2.01	1.11	–1.81	.071	.31	.32	<.01	<.01	.07	.26
Control:Time2:NWS	–0.01	1.08	–0.01	.990						
Oral:Time2:NWS	2.36	1.07	2.20	.028*						
Multi:Time2:NWS	1.33	1.08	1.23	.219						
Control:Time3:NWS	0.22	1.13	0.20	.842						
Oral:Time3:NWS	2.85	1.11	2.55	.011*						
Multi:Time3:NWS	2.09	1.12	1.86	.063						
*Formula: lmer(log_SOPT ~ Group*Time*log_NWS + (1|Participant) + (Group|Item))*										

Appendix A28. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (AOSPAN) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
AOSPAN										
Control group as reference level										
(Intercept)	–7.92	4.12	–1.92	.055	.28	.30	<.01	<.01	<.01	<.01
Oral:Time2:AOS	2.32	1.95	1.19	.230						
Written:Time2:AOS	1.59	2.00	0.80	.425						
Multi:Time2:AOS	1.94	2.12	0.92	.360						
Oral:Time3:AOS	4.36	1.88	2.32	.021*						
Written:Time3:AOS	3.80	1.89	2.01	.044*						
Multi:Time3:AOS	2.99	2.06	1.45	.146						
Oral group as reference level										
(Intercept)	4.36	3.54	1.23	.218	.28	.30	<.01	<.01	.07	.27
Control:Time2:AOS	–2.32	1.95	–1.19	.230						
Written:Time2:AOS	–0.73	1.80	–0.40	.686						
Multi:Time2:AOS	–0.38	1.94	–0.19	.846						
Control:Time3:AOS	–4.36	1.88	–2.32	.021*						
Written:Time3:AOS	–0.56	1.76	–0.32	.752						
Multi:Time3:AOS	–1.37	1.94	–0.71	.478						
Written group as reference level										
(Intercept)	3.66	3.65	1.01	.315	.28	.30	<.01	<.01	.07	.26
Control:Time2:AOS	–1.59	2.00	–0.80	.425						
Oral:Time2:AOS	0.73	1.80	0.40	.686						
Multi:Time2:AOS	0.35	1.99	0.18	.860						
Control:Time3:AOS	–3.80	1.89	–2.01	.044*						
Oral:Time3:AOS	0.56	1.76	0.32	.752						
Multi:Time3:AOS	–0.82	1.95	–0.42	.675						
*Formula: lmer(log_SOPT ~ Group*Time*log_AOSPAN + (1|Participant) + (Group|Item))*										

Note. AOS = AOSPAN

Appendix 29. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (AOSPAN) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
AOSPAN										
Control group as reference level										
(Intercept)	–8.62	4.09	–2.11	.035*	.29	.31	<.01	<.01	<.01	<.01
Oral:Time2:AOS	1.78	1.93	0.92	.357						
Written:Time2:AOS	1.08	1.98	0.54	.586						
Multi:Time2:AOS	2.29	2.11	1.14	.256						
Oral:Time3:AOS	4.08	1.87	2.19	.029*						
Written:Time3:AOS	3.83	1.88	2.04	.041*						
Multi:Time3:AOS	3.26	2.04	1.60	.111						
Oral group as reference level										
(Intercept)	3.24	3.51	0.92	.357	.29	.31	<.01	<.01	.08	.28
Control:Time2:AOS	–1.78	1.93	–0.92	.357						
Written:Time2:AOS	–0.71	1.79	–0.39	.693						
Multi:Time2:AOS	0.61	1.93	0.32	.752						
Control:Time3:AOS	–4.08	1.87	–2.19	.029*						
Written:Time3:AOS	–0.25	1.75	–0.15	.885						
Multi:Time3:AOS	–0.83	1.92	–0.43	.668						
Written group as reference level										
(Intercept)	4.38	3.62	1.21	.226	.29	.31	<.01	<.01	.06	.25
Control:Time2:AOS	–1.08	1.98	–0.54	.586						
Oral:Time2:AOS	0.71	1.79	0.39	.693						
Multi:Time2:AOS	1.31	1.97	0.67	.505						
Control:Time3:AOS	–3.83	1.88	–2.04	.041*						
Oral:Time3:AOS	0.25	1.75	0.15	.885						
Multi:Time2:AOS	–0.57	1.93	–0.30	.767						

Appendix A30. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (AOSPAN) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
AOSPAN										
Control group as reference level										
(Intercept)	–7.19	4.03	–1.78	.075	.31	.33	<.01	<.01	<.01	<.01
Oral:Time2:AOS	4.16	1.91	2.18	.029*						
Written:Time2:AOS	1.58	1.96	0.81	.419						
Multi:Time2:AOS	1.96	2.08	0.94	.345						
Oral:Time3:AOS	6.78	1.84	3.68	<.001***						
Written:Time3:AOS	3.82	1.85	2.07	.039*						
Multi:Time3:AOS	2.81	2.01	1.40	.163						
Oral group as reference level										
(Intercept)	3.88	3.46	1.12	.263	.31	.33	<.01	<.01	.08	.28
Control:Time2:AOS	–4.16	1.91	–2.18	.029*						
Written:Time2:AOS	–2.58	1.76	–1.46	.144						
Multi:Time2:AOS	–2.20	1.90	–1.16	.247						
Control:Time3:AOS	–6.78	1.84	–3.68	<.001***						
Written:Time3:AOS	–2.99	1.72	–1.72	.077						
Multi:Time3:AOS	–3.97	1.90	–2.09	.036*						
Written group as reference level										
(Intercept)	2.97	3.57	0.83	.406	.31	.33	<.01	<.01	.07	.27
Control:Time2:AOS	–1.60	1.95	–0.82	.413						
Oral:Time2:AOS	2.58	1.76	1.46	.144						
Multi:Time2:AOS	0.38	1.95	0.20	.846						
Control:Time3:AOS	–3.83	1.85	–2.02	.039*						
Oral:Time3:AOS	2.95	1.72	1.72	.086						
Multi:Time3:AOS	–1.01	1.91	–0.53	.594						

Appendix A31. Results for the Linear Mixed-effects Models Examining Moderating Effects of Working Memory Measures on Gains of SOPT (Stop Signal) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			Stop Signal			
Control group as reference level						
(Intercept)	–7.97	14.14	–0.56	.573	0.00	[0.00, 3.28]
Oral:Time2:Stop	–0.46	4.82	–0.10	.924	0.63	[0.00, 1.15]
Written:Time2:Stop	0.08	5.09	0.02	.988	1.08	[0.00, 2.22]
Multi:Time2:Stop	–2.15	5.23	–0.41	.681	0.16	[0.00, 3.10]
Oral:Time3:Stop	4.49	4.91	0.91	.361	0.89	[0.01, 1.29]
Written:Time3:Stop	–0.03	5.23	–0.01	.995	0.97	[0.00, 2.61]
Multi:Time3:Stop	–3.52	5.12	–0.69	.492	0.03	[0.00, 2.40]
Oral group as reference level						
(Intercept)	–2.13	10.29	–0.21	.836	0.12	[0.00, 6.12]
Control:Time2:Stop	0.46	3.62	0.13	.899	1.58	[0.00, 1.83]
Written:Time2:Stop	0.54	3.59	0.15	.881	1.71	[0.00, 1.89]
Multi:Time2:Stop	–1.69	3.09	–0.55	.584	0.18	[0.00, 7.68]
Control:Time3:Stop	–4.49	3.51	–1.28	.202	0.01	[0.00, 1.06]
Written:Time3:Stop	–4.52	3.73	–1.21	.226	0.01	[0.00, 1.58]
Multi:Time3:Stop	–8.00	3.08	–2.60	.009**	0.00	[0.00, 0.13]
Written group as reference level						
(Intercept)	–16.31	16.55	–0.99	.324	0.00	[0.00, 2.50]
Control:Time2:Stop	–0.08	5.09	–0.02	.986	0.92	[0.00, 5.81]
Oral:Time2:Stop	–0.54	4.92	–0.11	.889	0.58	[0.00, 1.09]
Multi:Time2:Stop	–2.23	5.32	–0.42	.549	0.11	[0.00, 1.54]
Conrol:Time3:Stop	0.03	5.23	0.01	.990	1.06	[0.00, 7.17]
Oral:Time3:Stop	4.52	5.07	0.89	.567	1.40	[0.00, 1.12]
Multi:Time3:Stop	–3.49	5.27	–0.66	.508	0.03	[0.00, 1.34]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.14	.38			
*glmer(SOPT_PS ~ 1+ Group * Time *log_Stop+ (1|Participant) + (1|Item))*						

Note. Stop = stop signal

Appendix A32. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (NWS) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
				NWS						
Control group as reference level										
(Intercept)	–3.52	1.22	–2.88	.004**	.28	.30	<.01	<.01	<.01	.06
Oral:Time2:NWS	1.11	1.13	0.99	.325						
Written:Time2:NWS	0.21	1.10	0.19	.852						
Multi:Time2:NWS	1.03	1.14	0.91	.364						
Oral:Time3:NWS	1.06	1.13	0.94	.346						
Written:Time3:NWS	–0.40	1.14	–0.35	.729						
Multi:Time3:NWS	1.44	1.14	1.27	.205						
Oral group as reference level										
(Intercept)	–0.70	1.18	–0.60	.552	.28	.30	<.01	<.01	.07	.27
Control:Time2:NWS	–1.11	1.13	–0.99	.325						
Written:Time2:NWS	–0.91	1.09	–0.83	.405						
Multi:Time2:NWS	–0.08	1.13	–0.07	.943						
Control:Time3:NWS	–1.06	1.13	–0.94	.346						
Written:Time3:NWS	–1.46	1.13	–1.29	.197						
Multi:Time3:NWS	0.38	1.13	0.34	.738						
Written group as reference level										
(Intercept)	–1.72	1.13	0.80	.128	.28	.30	<.01	<.01	.07	.26
Control:Time2:NWS	–0.21	1.10	–0.19	.851						
Oral:Time2:NWS	0.91	1.09	0.83	.405						
Multi:Time2:NWS	0.83	1.10	0.75	.451						
Control:Time3:NWS	0.40	1.14	0.35	.729						
Oral:Time3:NWS	1.46	1.03	1.29	.197						
Multi:Time3:NWS	1.84	1.14	1.61	.107						

Appendix A33. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (NWS) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
				NWS						
Control group as reference level										
(Intercept)	–3.74	1.22	–3.08	.002**	.29	.31	<.01	<.01	<.01	.07
Oral:Time2:NWS	0.59	1.12	052	.601						
Written:Time2:NWS	0.15	1.09	0.14	.887						
Multi:Time2:NWS	1.53	1.13	1.36	.175						
Oral:Time3:NWS	1.19	1,12	1.07	.287						
Written:Time3:NWS	0.11	1.13	0.09	.926						
Multi:Time3:NWS	1.75	1.11	1.58	.114						
Oral group as reference level										
(Intercept)	–1.34	1.17	–1.14	.253	.29	.31	<.01	<.01	.08	.28
Control:Time2:NWS	–0.58	1.12	–0.52	.601						
Written:Time2:NWS	–0.43	1.08	–0.40	.690						
Multi:Time2:NWS	0.95	1.12	–0.85	.397						
Control:Time3:NWS	–1.19	1.12	–1.07	.287						
Written:Time3:NWS	–1.09	1.12	–0.97	.333						
Multi:Time3:NWS	0.56	1.09	0.51	.611						
Written group as reference level										
(Intercept)	–1.52	1.12	–1.36	.175	.29	.31	<.01	<.01	.06	.25
Control:Time2:NWS	–0.15	1.09	–0.14	.887						
Oral:Time2:NWS	0.43	1.08	0.40	.690						
Multi:Time2:NWS	1.38	1.08	1.27	.206						
Control:Time3:NWS	–0.11	1.14	–0.09	.926						
Oral:Time3:NWS	1.09	1.12	0.97	.333						
Multi:Time3:NWS	1.64	1.11	1.48	.139						

Appendix A34. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Forward Corsi Block) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Forward Corsi Block										
Control group as reference level										
(Intercept)	–1.47	4.38	–0.34	.737	.28	.30	<.01	<.01	<.01	<.01
Oral:Time2:NWS	0.67	2.03	0.33	.740						
Written:Time2:NWS	0.72	1.87	0.39	.700						
Multi:Time2:NWS	0.93	2.40	0.39	.700						
Oral:Time3:NWS	0.18	2.02	0.09	.928						
Written:Time3:NWS	0.56	1.86	0.30	.763						
Multi:Time3:NWS	1.79	2.39	0.75	.455						
Oral group as reference level										
(Intercept)	–2.97	3.81	–0.78	.436	.28	.30	<.01	<.01	.07	.27
Control:Time2:NWS	–0.66	2.03	–0.33	.745						
Written:Time2:NWS	0.05	1.71	0.03	.978						
Multi:Time2:NWS	0.25	2.28	0.11	.912						
Control:Time3:NWS	–1.17	2.02	–0.09	.932						
Written:Time3:NWS	0.38	1.71	0.22	.824						
Multi:Time3:NWS	1.61	2.28	0.71	.481						
Written group as reference level										
(Intercept)	–1.86	3.07	–0.61	.545	.28	.30	<.01	<.01	.07	.26
Control:Time2:NWS	–0.70	1.87	–0.38	.706						
Oral:Time2:NWS	–0.05	1.71	–0.03	.978						
Multi:Time2:NWS	0.21	2.14	0.10	.923						
Control:Time3:NWS	–0.55	1.86	–0.30	.766						
Oral:Time3:NWS	–0.38	1.71	–0.22	.824						
Multi:Time3:NWS	1.23	2.14	0.58	.565						

Appendix A35. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Forward Corsi Block) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Forward Corsi Block										
Control group as reference level										
(Intercept)	–1.83	4.35	–0.42	.673	.29	.31	<.01	<.01	<.01	.07
Oral:Time2:NWS	0.27	2.01	0.14	.891						
Written:Time2:NWS	0.47	1.85	0.26	.798						
Multi:Time2:NWS	0.02	2.38	0.01	.993						
Oral:Time3:NWS	0.05	2.00	0.03	.980						
Written:Time3:NWS	0.59	1.84	0.32	.751						
Multi:Time3:NWS	1.27	2.38	0.53	.594						
Oral group as reference level										
(Intercept)	–2.47	3.78	–0.65	.514	.29	.31	<.01	<.01	.08	.28
Control:Time2:NWS	–0.27	2.01	–0.14	.891						
Written:Time2:NWS	0.20	1.69	0.12	.906						
Multi:Time2:NWS	–0.25	2.26	–0.11	.911						
Control:Time3:NWS	–0.05	2.00	–0.03	.980						
Written:Time3:NWS	0.54	1.69	0.32	.752						
Multi:Time3:NWS	1.22	2.26	0.54	.590						
Written group as reference level										
(Intercept)	–1.54	3.05	–0.51	.613	.29	.31	<.01	<.01	.06	.25
Control:Time2:NWS	–0.47	1.85	–0.26	.798						
Oral:Time2:NWS	–0.20	1.70	–0.12	.906						
Multi:Time2:NWS	–0.45	2.12	–0.21	.831						
Control:Time3:NWS	–0.59	1.84	–0.32	.751						
Oral:Time3:NWS	–0.54	1.70	–0.32	.752						
Multi:Time3:NWS	0.68	2.12	0.32	.748						

Appendix A36. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Forward Corsi Block) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Forward Corsi Block										
Control group as reference level										
(Intercept)	–0.70	4.30	–0.16	.871	.30	.32	<.01	<.01	<.01	<.01
Oral:Time2:NWS	0.56	1.99	0.28	.777						
Written:Time2:NWS	0.23	1.83	0.12	.901						
Multi:Time2:NWS	–0.59	2.36	–0.25	.802						
Oral:Time3:NWS	–0.23	1.98	–0.11	.909						
Written:Time3:NWS	0.77	1.82	0.42	.671						
Multi:Time3:NWS	0.46	2.35	0.20	.844						
Oral group as reference level										
(Intercept)	–2.56	3.74	–0.69	.493	.30	.33	<.01	<.01	.08	.28
Control:Time2:NWS	–0.55	1.99	–0.28	.783						
Written:Time2:NWS	–0.34	1.67	–0.20	.841						
Multi:Time2:NWS	–1.15	2.23	–0.52	.606						
Control:Time3:NWS	0.24	1.98	–0.12	.905						
Written:Time3:NWS	1.00	1.67	0.60	.550						
Multi:Time3:NWS	0.69	2.23	0.31	.758						
Written group as reference level										
(Intercept)	–2.02	3.01	–0.67	.502	.30	.33	<.01	<.01	.07	.27
Control:Time2:NWS	–0.21	1.83	–0.12	.908						
Oral:Time2:NWS	0.34	1.67	0.20	.841						
Multi:Time2:NWS	–0.82	2.10	–0.39	.697						
Control:Time3:NWS	–0.76	1.82	–0.42	.675						
Oral:Time3:NWS	–1.00	1.67	–0.60	.550						
Multi:Time3:NWS	–0.31	2.10	–0.15	.882						

Appendix A37. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Color Shape Task) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Color Shape										
Control group as reference level										
(Intercept)	–1.62	3.73	–0.43	.664	.28	.30	<.01	<.01	<.01	.06
Oral:Time2:NWS	0.38	1.25	0.31	.759						
Written:Time2:NWS	–0.51	1.20	–0.42	.673						
Multi:Time2:NWS	0.21	1.18	0.18	.860						
Oral:Time3:NWS	0.10	1.25	0.08	.939						
Written:Time3:NWS	0.07	1.20	0.06	.952						
Multi:Time3:NWS	0.42	1.18	0.36	.720						
Oral group as reference level										
(Intercept)	–0.19	3.68	–0.05	.958	.28	.30	<.01	<.01	.07	.27
Control:Time2:NWS	–0.38	1.25	–0.31	.759						
Written:Time2:NWS	–0.89	1.19	–0.75	.454						
Multi:Time2:NWS	–0.17	1.17	–0.15	.882						
Control:Time3:NWS	–0.10	1.25	–0.08	.939						
Written:Time3:NWS	–0.02	1.19	–0.02	.984						
Multi:Time3:NWS	0.33	1.17	0.28	.779						
Written group as reference level										
(Intercept)	–3.53	3.41	–1.04	.301	.28	.30	<.01	<.01	.07	.26
Control:Time2:NWS	0.51	1.20	0.42	.673						
Oral:Time2:NWS	0.89	1.19	0.75	.454						
Multi:Time2:NWS	0.72	1.12	0.64	.522						
Control:Time3:NWS	–0.07	1.20	–0.06	.952						
Oral:Time3:NWS	0.02	1.19	0.02	.984						
Multi:Time3:NWS	0.35	1.12	0.32	.753						

Appendix A38. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Color Shape Task) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Color Shape										
Control group as reference level										
(Intercept)	–1.47	3.71	–0.40	.692	.29	.31	<.01	<.01	<.01	.07
Oral:Time2:NWS	0.68	1.24	0.55	.583						
Written:Time2:NWS	–0.24	1.19	–0.20	.839						
Multi:Time2:NWS	0.54	1.17	0.46	.648						
Oral:Time3:NWS	0.22	1.24	0.18	.857						
Written:Time3:NWS	–0.01	1.19	–0.01	.995						
Multi:Time3:NWS	0.50	1.17	0.43	.669						
Oral group as reference level										
(Intercept)	0.35	3.66	0.10	.925	.29	.31	<.01	<.01	.08	.28
Control:Time2:NWS	–0.68	1.24	–0.55	.582						
Written:Time2:NWS	–0.92	1.18	–0.78	.434						
Multi:Time2:NWS	–0.15	1.16	–0.13	.900						
Control:Time3:NWS	–0.22	1.24	–0.18	.857						
Written:Time3:NWS	–0.23	1.18	–0.20	.845						
Multi:Time3:NWS	0.28	1.16	0.24	.811						
Written group as reference level										
(Intercept)	–4.07	3.38	–1.20	.229	.29	.31	<.01	<.01	.06	.25
Control:Time2:NWS	0.24	1.19	0.20	.839						
Oral:Time2:NWS	0.92	1.18	0.78	.434						
Multi:Time2:NWS	0.78	1.11	0.70	.483						
Control:Time3:NWS	0.01	1.19	0.01	.995						
Oral:Time3:NWS	0.23	1.18	0.20	.845						
Multi:Time3:NWS	0.51	1.11	0.46	.647						

Appendix A39. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Color Shape Task) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Color Shape Task										
Control group as reference level										
(Intercept)	–2.64	3.66	–0.72	.472	.30	.32	<.01	<.01	<.01	<.01
Oral:Time2:NWS	0.75	1.23	0.61	.540						
Written:Time2:NWS	–0.27	1.18	–0.23	.817						
Multi:Time2:NWS	0.47	1.16	0.41	.684						
Oral:Time3:NWS	0.41	1.23	0.33	.739						
Written:Time3:NWS	0.32	1.18	0.27	.784						
Multi:Time3:NWS	0.91	1.16	0.79	.431						
Oral group as reference level										
(Intercept)	–1.21	3.61	–0.33	.738	.30	.33	<.01	<.01	.08	.28
Control:Time2:NWS	–0.75	1.23	–0.61	.540						
Written:Time2:NWS	–1.03	1.17	–0.88	.380						
Multi:Time2:NWS	–0.28	1.15	–0.24	.808						
Control:Time3:NWS	–0.41	1.23	–0.33	.739						
Written:Time3:NWS	–0.09	1.17	–0.07	.942						
Multi:Time3:NWS	0.50	1.15	0.44	.660						
Written group as reference level										
(Intercept)	–2.65	3.34	–0.79	.429	.30	.33	<.01	<.01	.07	.27
Control:Time2:NWS	0.27	1.18	0.23	.817						
Oral:Time2:NWS	1.03	1.17	0.88	.380						
Multi:Time2:NWS	0.75	1.10	0.68	.496						
Control:Time3:NWS	–0.32	1.18	–0.27	.784						
Oral:Time3:NWS	0.09	1.17	0.07	.942						
Multi:Time3:NWS	0.59	1.10	0.54	.591						

Appendix A40. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Stop Signal Task) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Stop Signal Task										
Control group as reference level										
(Intercept)	0.46	5.73	0.08	.936	.28	.30	<.01	<.01	<.01	.06
Oral:Time2:NWS	–0.59	2.07	–0.28	.778						
Written:Time2:NWS	0.73	2.27	0.32	.749						
Multi:Time2:NWS	–0.20	2.21	–0.09	.926						
Oral:Time3:NWS	–0.25	2.06	–0.12	.902						
Written:Time3:NWS	–0.32	2.26	–0.14	.887						
Multi:Time3:NWS	–0.77	2.20	–0.35	.727						
Oral group as reference level										
(Intercept)	–0.06	5.88	–0.01	.992	.28	.30	<.01	<.01	.07	.27
Control:Time2:NWS	0.59	2.07	0.28	.778						
Written:Time2:NWS	1.31	2.28	0.58	.565						
Multi:Time2:NWS	0.38	2.22	0.17	.864						
Control:Time3:NWS	0.25	2.06	0.12	.902						
Written:Time3:NWS	–0.07	2.28	–0.03	.977						
Multi:Time3:NWS	–0.52	2.22	–0.23	.816						
Written group as reference level										
(Intercept)	–1.30	6.90	–0.19	.851	.28	.30	<.01	<.01	.07	.26
Control:Time2:NWS	–0.73	2.27	–0.32	.749						
Oral:Time2:NWS	–1.31	2.28	–0.58	.565						
Multi:Time2:NWS	–0.93	2.40	–0.39	.699						
Control:Time3:NWS	0.32	2.26	0.14	.887						
Oral:Time3:NWS	0.07	2.28	0.03	.977						
Multi:Time3:NWS	–0.45	2.40	–0.19	.852						

Appendix A41. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Stop Signal Task) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Stop Signal Task										
Control group as reference level										
(Intercept)	1.20	5.69	0.02	.983	.29	.31	<.01	<.01	<.01	.07
Oral:Time2:NWS	0.19	2.06	0.09	.928						
Written:Time2:NWS	0.58	2.25	0.26	.799						
Multi:Time2:NWS	–1.10	2.19	–0.50	.617						
Oral:Time3:NWS	–0.29	2.05	–0.14	.887						
Written:Time3:NWS	–0.01	2.24	–0.00	.997						
Multi:Time3:NWS	–1.54	2.19	–0.70	.482						
Oral group as reference level										
(Intercept)	–0.65	5.84	–0.11	.911	.29	.31	<.01	<.01	.08	.28
Control:Time2:NWS	–0.19	2.06	–0.09	.928						
Written:Time2:NWS	0.39	2.26	0.17	.863						
Multi:Time2:NWS	–1.28	2.21	–0.58	.561						
Control:Time3:NWS	0.29	2.05	0.14	.887						
Written:Time3:NWS	0.28	2.26	0.13	.900						
Multi:Time3:NWS	–1.25	2.21	–0.57	.572						
Written group as reference level										
(Intercept)	–1.89	6.85	–0.28	.783	.29	.31	<.01	<.01	.06	.25
Control:Time2:NWS	–0.58	2.25	–0.26	.799						
Oral:Time2:NWS	–0.39	2.26	–0.17	.863						
Multi:Time2:NWS	–1.67	2.39	–0.70	.483						
Control:Time3:NWS	0.01	2.24	0.00	.997						
Oral:Time3:NWS	–0.28	2.26	–0.13	.900						
Multi:Time3:NWS	–1.53	2.39	–0.64	.521						

Appendix A42. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Stop Signal Task) – Present Progressive

	Fixed effects						Random effects			
							by participant		*by item*	
	Estimate	*SE*	*t*	*p*	*R2m*	*R2c*	*variance*	*SD*	*variance*	*SD*
Stop Signal Task										
Control group as reference level										
(Intercept)	0.95	5.62	0.17	.865	.30	.33	<.01	<.01	.01	.08
Oral:Time2:NWS	–0.43	2.03	–0.21	.831						
Written:Time2:NWS	–0.20	2.22	–0.09	.927						
Multi:Time2:NWS	–1.59	2.17	–0.73	.463						
Oral:Time3:NWS	–0.11	2.02	–0.06	.956						
Written:Time3:NWS	–1.04	2.22	–0.47	.639						
Multi:Time3:NWS	–2.33	2.16	–1.08	.280						
Oral group as reference level										
(Intercept)	2.52	5.76	0.44	.662	.30	.33	<.01	<.01	.08	.28
Control:Time2:NWS	0.43	2.03	0.21	.831						
Written:Time2:NWS	0.23	2.24	0.10	.918						
Multi:Time2:NWS	–1.16	2.18	–0.53	.596						
Control:Time3:NWS	0.11	2.02	0.06	.956						
Written:Time3:NWS	–0.93	2.24	–0.42	.678						
Multi:Time3:NWS	–2.22	2.18	–1.02	.308						
Written group as reference level										
(Intercept)	–1.59	6.76	–0.24	.814	.30	.33	<.01	<.01	.07	.27
Control:Time2:NWS	0.20	2.23	0.09	.927						
Oral:Time2:NWS	–0.23	2.24	–0.10	.918						
Multi:Time2:NWS	–1.39	2.36	–0.59	.557						
Control:Time3:NWS	1.04	2.22	0.47	.639						
Oral:Time3:NWS	0.92	2.24	0.42	.678						
Multi:Time3:NWS	–1.29	2.36	–0.55	.584						

Appendix A43. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (NWS) – Third Person Singular

Fixed effects	*Estimate*	*SE*	*z*	*p*	*OR*	95% CI for OR
			NWS			
Control group as reference level						
(Intercept)	–1.33	2.96	–0.45	.654	0.27	[0.00, 8.56]
Oral:Time2:NWS	–0.52	2.56	–0.21	.838	0.59	[0.00, 8.73]
Written:Time2:NWS	0.15	2.52	0.06	.953	1.16	[0.01, 1.57]
Multi:Time2:NWS	–0.19	2.64	–0.07	.943	0.83	[0.00, 1.44]
Oral:Time3:NWS	–1.18	2.71	–0.44	.663	0.31	[0.01, 6.10]
Written:Time3:NWS	–1.27	2.76	–0.46	.646	0.28	[0.00, 6.07]
Multi:Time3:NWS	–0.32	2.71	–0.12	.907	0.73	[0.00, 1.43]
Oral group as reference level						
(Intercept)	–3.57	2.91	–1.23	.220	0.03	[0.00, 8.24]
Control:Time2:NWS	0.52	2.55	0.21	.837	1.69	[0.01, 2.43]
Written:Time2:NWS	0.67	2.34	0.29	.774	1.96	[0.02, 1.87]
Multi:Time2:NWS	0.33	2.60	0.13	.897	1.40	[0.01, 2.22]
Control:Time3:NWS	1.18	2.56	0.46	.644	3.26	[0.02, 4.82]
Written:Time3:NWS	–0.09	2.45	–0.04	.972	0.92	[0.01, 1.09]
Multi:Time3:NWS	0.86	2.50	0.35	.730	2.37	[0.02, 3.11]
Written group as reference level						
(Intercept)	0.05	2.70	0.02	.985	1.05	[0.01, 2.03]
Control:Time2:NWS	–0.15	2.46	–0.06	.952	0.86	[0.01, 1.03]
Oral:Time2:NWS	–0.67	2.31	–0.29	.772	0.51	[0.01, 4.64]
Multi:Time2:NWS	–0.34	2.41	–0.14	.889	0.71	[0.01, 7.92]
Conrol:Time3:NWS	1.27	2.61	0.49	.628	3.55	[0.02, 5.76]
Oral:Time3:NWS	0.08	2.52	0.03	.973	1.09	[0.01, 1.47]
Multi:Time3:NWS	0.95	2.48	0.38	.702	2.58	[0.02, 3.23]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.11	.33			
*glmer(OPT_PS ~ 1+ Group*Time*log_NWS + (1|Participant) + (1|Item))*						

Appendix A44. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (NWS) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			NWS			
Control group as reference level						
(Intercept)	–3.05	2.83	–1.08	.281	0.05	[0.00, 1.18]
Oral:Time2:NWS	–0.27	2.43	–0.11	.912	0.76	[0.01, 8.78]
Written:Time2:NWS	0.86	2.37	0.36	.717	2.36	[0.02, 2.42]
Multi:Time2:NWS	0.18	2.59	0.07	.944	1.20	[0.01, 1.87]
Oral:Time3:NWS	1.18	2.55	0.07	.945	1.19	[0.01, 1.73]
Written:Time3:NWS	–1.12	2.49	–0.05	.962	0.89	[0.01, 1.15]
Multi:Time3:NWS	0.46	2.51	0.18	.855	1.58	[0.01, 2.12]
Oral group as reference level						
(Intercept)	–2.16	2.61	–0.83	.408	0.12	[0.00, 1.88]
Control:Time2:NWS	0.27	2.36	0.11	.909	1.31	[0.01, 1.31]
Written:Time2:NWS	1.13	2.16	0.52	.603	3.08	[0.05, 2.09]
Multi:Time2:NWS	0.45	2.42	0.19	.852	1.57	[0.01, 1.75]
Control:Time3:NWS	–1.18	2.40	–0.07	.942	0.84	[0.01, 9.10]
Written:Time3:NWS	–0.30	2.32	–0.13	.899	0.74	[0.01, 6.86]
Multi:Time3:NWS	0.28	2.28	0.12	.902	1.32	[0.01, 1.12]
Written group as reference level						
(Intercept)	0.20	3.02	0.07	.947	1.22	[0.00, 4.41]
Control:Time2:NWS	–0.86	2.70	–0.32	.750	0.42	[0.00, 8.16]
Oral:Time2:NWS	–1.13	2.57	–0.44	.662	0.32	[0.00, 4.89]
Multi:Time2:NWS	–0.67	2.83	–0.24	.812	0.51	[0.00, 1.27]
Conrol:Time3:NWS	0.12	2.86	0.04	.967	1.13	[0.00, 2.98]
Oral:Time3:NWS	0.30	2.74	0.11	.914	1.34	[0.01, 2.83]
Multi:Time3:NWS	0.58	2.77	0.21	.835	1.78	[0.01, 3.95]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.07	.26			
*glmer(WPT_PS ~ 1+ Group*Time*log_NWS + (1|Participant) + (1|Item))*						

Appendix A45. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (NWS) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			NWS			
Control group as reference level						
(Intercept)	–2.95	3.12	–0.94	.345	0.05	[0.00, 2.31]
Oral:Time2:NWS	0.09	2.64	–0.03	.973	1.09	[0.01, 1.88]
Written:Time2:NWS	0.98	2.60	0.38	.706	2.66	[0.02, 4.21]
Multi:Time2:NWS	1.10	2.78	0.40	.693	3.00	[0.01, 6.77]
Oral:Time3:NWS	–1.39	2.69	–0.52	.605	0.25	[0.00, 4.70]
Written:Time3:NWS	–1.09	2.72	–0.40	.689	0.34	[0.00, 6.75]
Multi:Time3:NWS	0.47	2.68	0.18	.861	1.60	[0.01, 3.01]
Oral group as reference level						
(Intercept)	–4.54	2.97	–1.53	.127	0.01	[0.00, 3.51]
Control:Time2:NWS	–0.09	2.64	–0.03	.973	1.09	[0.01, 1.88]
Written:Time2:NWS	0.89	2.31	0.38	.702	2.43	[0.02, 3.22]
Multi:Time2:NWS	1.01	2.54	0.40	.691	2.74	[0.01, 3.90]
Control:Time3:NWS	1.39	2.69	0.52	.605	0.25	[0.00, 4.70]
Written:Time3:NWS	0.31	2.49	0.12	.902	1.36	[0.01, 1.76]
Multi:Time3:NWS	1.86	2.55	0.73	.465	6.44	[0.04, 9.32]
Written group as reference level						
(Intercept)	–0.11	2.52	–0.04	.966	0.90	[0.01, 4.41]
Control:Time2:NWS	–0.98	2.60	–0.38	.706	2.66	[0.02, 4.21]
Oral:Time2:NWS	–0.89	2.31	–0.38	.702	2.43	[0.03, 3.22]
Multi:Time2:NWS	0.12	2.27	0.05	.957	1.13	[0.01, 1.27]
Conrol:Time3:NWS	1.09	2.72	0.40	.689	0.34	[0.00, 6.75]
Oral:Time3:NWS	–0.31	2.49	–0.12	.902	1.36	[0.01, 1.76]
Multi:Time3:NWS	1.55	2.27	0.68	.494	4.73	[0.05, 5.95]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.14	.38			
*glmer(SOPT_PS ~ 1+ Group*Time*log_NWS + (1|Participant) + (1|Item))*						

Appendix A46. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Forward Corsi Block) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			Forward Corsi Block			
Control group as reference level						
(Intercept)	6.45	10.25	0.63	.529	6.30	[0.00, 7.99]
Oral:Time2:NWS	0.03	4.05	0.01	.995	0.97	[0.01, 2.60]
Written:Time2:NWS	–0.82	4.10	–0.20	.842	2.66	[0.00, 6.74]
Multi:Time2:NWS	0.95	5.53	0.17	.864	2.58	[0.00, 6.70]
Oral:Time3:NWS	–0.55	3.85	–0.14	.887	1.73	[0.00, 3.13]
Written:Time3:NWS	–1.24	4.13	–0.30	.763	0.34	[0.00, 1.09]
Multi:Time3:NWS	–1.45	5.36	–0.27	.787	2.34	[0.00, 3.01]
Oral group as reference level						
(Intercept)	0.11	8.66	0.01	.990	1.12	[0.00, 2.39]
Control:Time2:NWS	–0.03	4.05	–0.01	.995	0.97	[0.01, 2.60]
Written:Time2:NWS	–0.84	3.58	–0.24	.814	0.43	[0.02, 4.60]
Multi:Time2:NWS	0.92	4.90	0.19	.851	2.51	[0.01, 3.57]
Control:Time3:NWS	0.55	3.85	0.14	.887	1.73	[0.00, 3.13]
Written:Time3:NWS	–0.70	3.25	–0.22	.830	0.50	[0.01, 2.81]
Multi:Time3:NWS	–0.91	4.63	–0.20	.845	0.40	[0.04, 3.37]
Written group as reference level						
(Intercept)	–4.47	7.92	–0.56	.573	0.01	[0.00, 5.85]
Control:Time2:NWS	0.82	4.10	0.20	.842	2.66	[0.00, 6.74]
Oral:Time2:NWS	0.84	3.74	0.23	.822	2.43	[0.00, 3.43]
Multi:Time2:NWS	1.76	4.95	0.36	.722	1.13	[0.00, 9.14]
Conrol:Time3:NWS	1.24	4.13	0.30	.763	0.34	[0.00, 1.09]
Oral:Time3:NWS	0.70	3.76	0.19	.853	1.36	[0.01, 3.05]
Multi:Time3:NWS	–0.21	5.67	–0.04	.971	4.73	[0.00, 5.17]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.10	.33			
*glmer(OPT_PS ~ 1+ Group*Time*log_Forward + (1|Participant) + (1|Item))*						

Appendix A47. Results for the Linear Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Forward Corsi Block) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			Forward Corsi Block			
Control group as reference level						
(Intercept)	5.99	8.86	0.68	.499	4.01	[0.00, 1.27]
Oral:Time2:NWS	–0.05	4.21	–0.01	.991	0.95	[0.00, 3.53]
Written:Time2:NWS	–1.58	3.50	–0.45	.652	0.21	[0.00, 1.90]
Multi:Time2:NWS	1.14	4.96	0.23	.819	3.11	[0.00, 4.97]
Oral:Time3:NWS	–0.31	4.32	–0.07	.943	0.74	[0.00, 3.32]
Written:Time3:NWS	–1.51	3.75	–0.40	.688	0.22	[0.00, 3.32]
Multi:Time3:NWS	–0.98	4.51	–0.22	.827	0.37	[0.00, 2.48]
Oral group as reference level						
(Intercept)	0.28	9.29	0.03	.976	1.33	[0.00, 9.83]
Control:Time2:NWS	0.05	4.21	0.01	.991	0.95	[0.00, 3.53]
Written:Time2:NWS	–1.53	3.64	–0.42	.675	0.22	[0.00, 2.63]
Multi:Time2:NWS	1.18	5.40	0.22	.826	3.27	[0.00, 1.22]
Control:Time3:NWS	0.31	4.32	0.07	.943	0.74	[0.00, 3.32]
Written:Time3:NWS	–1.20	3.79	–0.32	.752	0.30	[0.00, 4.92]
Multi:Time3:NWS	–0.68	5.02	–0.14	.893	0.51	[0.00, 9.01]
Written group as reference level						
(Intercept)	–5.45	6.04	–0.90	.367	0.00	[0.00, 5.58]
Control:Time2:NWS	1.58	3.50	0.45	.652	0.21	[0.00, 1.90]
Oral:Time2:NWS	1.53	3.64	0.42	.675	0.22	[0.00, 2.63]
Multi:Time2:NWS	2.71	4.35	0.62	.533	0.00	[0.00, 7.23]
Conrol:Time3:NWS	1.51	3.75	0.40	.688	0.22	[0.00, 3.32]
Oral:Time3:NWS	1.20	3.79	0.32	.752	0.30	[0.00, 4.92]
Multi:Time3:NWS	0.52	3.79	0.14	.890	0.01	[0.00, 2.72]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.07	.26			
*glmer(WPT_PS ~ 1+ Group*Time*log_Forward + (1|Participant) + (1|Item))*						

Appendix A48. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Forward Corsi Block) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			Forward Corsi Block			
Control group as reference level						
(Intercept)	5.33	8.13	0.66	.512	2.72	[0.00, 1.58]
Oral:Time2:NWS	0.93	3.84	0.24	.808	2.55	[0.00, 4.59]
Written:Time2:NWS	–0.32	3.30	–0.10	.922	0.73	[0.00, 4.51]
Multi:Time2:NWS	–1.17	4.17	–0.28	.779	0.31	[0.00, 1.05]
Oral:Time3:NWS	–1.82	3.87	–0.47	.639	0.16	[0.00, 3.11]
Written:Time3:NWS	–1.32	3.33	–0.40	.691	0.27	[0.00, 1.76]
Multi:Time3:NWS	–4.33	3.87	–1.12	.264	0.01	[0.00, 2.52]
Oral group as reference level						
(Intercept)	3.16	9.75	0.32	.746	2.36	[0.00, 4.26]
Control:Time2:NWS	–0.93	4.65	–0.20	.841	0.39	[0.00, 3.41]
Written:Time2:NWS	–1.26	3.75	–0.36	.738	0.28	[0.00, 4.27]
Multi:Time2:NWS	–2.10	5.24	–0.40	.688	0.12	[0.00, 3.34]
Control:Time3:NWS	1.82	4.73	0.38	.701	6.14	[0.00, 6.27]
Written:Time3:NWS	0.49	3.87	0.13	.899	1.63	[0.00, 3.12]
Multi:Time3:NWS	–2.51	5.10	–0.49	.622	0.06	[0.00, 1.69]
Written group as reference level						
(Intercept)	–2.98	12.87	–0.23	.817	0.05	[0.00, 4.02]
Control:Time2:NWS	0.32	5.14	0.06	.950	1.38	[0.00, 3.13]
Oral:Time2:NWS	1.26	5.73	0.22	.826	3.51	[0.00, 2.50]
Multi:Time2:NWS	–0.85	9.78	–0.09	.931	0.43	[0.00, 8.23]
Conrol:Time3:NWS	1.32	6.01	0.22	.826	3.76	[0.00, 4.65]
Oral:Time3:NWS	–0.49	6.23	–0.08	.937	0.61	[0.01, 1.15]
Multi:Time3:NWS	–2.00	10.25	–0.29	.770	0.05	[0.00, 2.36]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.14	.38			
*glmer(SOPT_PS ~ 1+ Group*Time*log_Forward + (1|Participant) + (1|Item))*						

Appendix A49. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Color Shape) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
		Color Shape				
Control group as reference level						
(Intercept)	2.31	10.67	0.22	.829	1.01	[0.00, 1.10]
Oral:Time2:NWS	1.36	3.20	0.43	.670	3.90	[0.01, 1.99]
Written:Time2:NWS	1.35	3.02	0.45	.654	3.87	[0.01, 1.39]
Multi:Time2:NWS	0.88	3.04	0.29	.772	2.42	[0.01, 9.07]
Oral:Time3:NWS	–0.13	3.57	–0.04	.970	0.87	[0.00, 9.24]
Written:Time3:NWS	1.55	3.50	0.44	.658	4.71	[0.01, 4.34]
Multi:Time3:NWS	1.14	3.40	0.36	.738	3.12	[0.00, 2.38]
Oral group as reference level						
(Intercept)	1.09	9.20	0.12	.905	2.99	[0.00, 1.86]
Control:Time2:NWS	–1.36	2.88	–0.47	.636	0.26	[0.01, 7.02]
Written:Time2:NWS	–0.01	2.58	–0.00	.997	0.99	[0.00, 1.51]
Multi:Time2:NWS	–0.48	2.62	–0.18	.855	0.62	[0.00, 1.02]
Control:Time3:NWS	0.13	2.93	0.05	.963	1.14	[0.00, 3.47]
Written:Time3:NWS	1.68	2.68	0.63	.530	5.38	[0.03, 1.01]
Multi:Time3:NWS	1.27	2.62	0.49	.627	3.57	[0.02, 5.87]
Written group as reference level						
(Intercept)	3.43	10.76	0.32	.750	3.10	[0.00, 4.01]
Control:Time2:NWS	–1.35	3.52	–0.39	.700	0.26	[0.00, 2.45]
Oral:Time2:NWS	0.01	3.04	0.00	.998	1.01	[0.00, 3.76]
Multi:Time2:NWS	–0.47	3.25	–0.15	.885	0.62	[0.00, 3.56]
Conrol:Time3:NWS	–1.55	3.39	–0.46	.648	0.21	[0.00, 1.58]
Oral:Time3:NWS	–1.68	2.90	–0.58	.561	0.19	[0.01, 5.26]
Multi:Time3:NWS	–0.31	3.08	–0.13	.894	0.66	[0.00, 2.71]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.11	.33			
*glmer(OPT_PS ~ 1+ Group*Time*log_Color + (1|Participant) + (1|Item))*						

Appendix A50. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Color Shape) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
		Color Shape				
Control group as reference level						
(Intercept)	–0.20	8.49	–0.02	.981	0.82	[0.00, 1.27]
Oral:Time2:NWS	1.63	2.73	0.60	.551	5.10	[0.02, 1.05]
Written:Time2:NWS	1.80	2.85	0.63	.526	6.08	[0.02, 1.56]
Multi:Time2:NWS	0.69	3.02	0.23	.820	1.99	[0.01, 7.17]
Oral:Time3:NWS	0.32	2.44	0.13	.895	1.38	[0.01, 1.61]
Written:Time3:NWS	2.11	2.39	0.89	.376	8.32	[0.07, 8.82]
Multi:Time3:NWS	0.97	2.39	0.40	.686	2.63	[0.02, 2.77]
Oral group as reference level						
(Intercept)	1.93	7.26	0.27	.791	6.88	[0.00, 9.75]
Control:Time2:NWS	–1.63	2.37	–0.66	.510	0.20	[0.00, 2.43]
Written:Time2:NWS	0.18	2.01	0.09	.930	1.19	[0.02, 5.96]
Multi:Time2:NWS	–0.94	2.26	–0.42	.676	0.39	[0.00, 3.18]
Control:Time3:NWS	–0.32	2.51	–0.13	.897	0.72	[0.01, 9.58]
Written:Time3:NWS	1.80	2.26	0.80	.427	6.02	[0.07, 4.94]
Multi:Time3:NWS	0.64	2.16	0.30	.766	1.90	[0.03, 1.30]
Written group as reference level						
(Intercept)	4.31	9.38	0.46	.646	7.43	[0.00, 6.53]
Control:Time2:NWS	–1.80	3.08	–0.59	.558	0.16	[0.00, 6.72]
Oral:Time2:NWS	–0.18	3.01	–0.06	.954	0.84	[0.00, 2.96]
Multi:Time2:NWS	–1.12	2.55	–0.44	.661	0.33	[0.00, 4.75]
Conrol:Time3:NWS	–2.12	3.02	–0.70	.483	0.12	[0.00, 4.35]
Oral:Time3:NWS	–1.80	3.10	–0.58	.563	0.17	[0.00, 7.06]
Multi:Time3:NWS	–1.15	2.52	–0.46	.647	0.32	[0.00, 4.31]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.07	.26			

*glmer(WPT_PS ~ 1+ Group*Time*log_Color + (1|Participant) + (1|Item))*

Appendix A51. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (Color Shape) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
		Color Shape				
Control group as reference level						
(Intercept)	1.43	9.28	0.15	.878	4.16	[0.00, 2.98]
Oral:Time2:NWS	1.67	2.93	0.57	.568	5.32	[0.02, 1.60]
Written:Time2:NWS	1.37	2.79	0.49	.624	3.93	[0.02, 9.12]
Multi:Time2:NWS	1.60	2.84	0.56	.575	4.93	[0.02, 1.26]
Oral:Time3:NWS	0.59	2.97	0.20	.843	1.80	[0.01, 5.96]
Written:Time3:NWS	1.45	2.84	0.51	.610	4.26	[0.02, 1.08]
Multi:Time3:NWS	1.63	2.79	0.58	.559	5.11	[0.02, 1.18]
Oral group as reference level						
(Intercept)	0.91	8.02	0.11	.910	2.47	[0.00, 1.54]
Control:Time2:NWS	–1.67	2.74	–0.61	.542	0.19	[0.01, 3.92]
Written:Time2:NWS	–0.30	2.22	–0.14	.892	0.74	[0.01, 5.60]
Multi:Time2:NWS	–0.07	2.41	–0.03	.975	0.93	[0.01, 1.01]
Control:Time3:NWS	–0.59	2.73	–0.22	.829	0.55	[0.00, 1.14]
Written:Time3:NWS	0.86	2.34	0.37	.714	2.36	[0.02, 2.26]
Multi:Time3:NWS	1.04	2.49	0.42	.676	2.83	[0.02, 3.63]
Written group as reference level						
(Intercept)	–16.31	16.55	–0.99	.324	0.00	[0.00, 8.59]
Control:Time2:NWS	–0.08	5.09	–0.02	.988	0.92	[0.00, 1.90]
Oral:Time2:NWS	–0.54	4.92	–0.11	.913	0.58	[0.00, 8.64]
Multi:Time2:NWS	–2.23	5.32	–0.42	.675	0.11	[0.00, 3.43]
Conrol:Time3:NWS	0.03	5.23	0.01	.995	1.03	[0.00, 2.77]
Oral:Time3:NWS	4.52	5.07	0.89	.373	9.15	[0.01, 1.79]
Multi:Time3:NWS	–3.49	5.27	–0.66	.508	0.03	[0.00, 8.79]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.14	.38			

*glmer(SOPT_PS ~ 1+ Group*Time*log_Color + (1|Participant) + (1|Item))*

Appendix A52. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (Stop Signal) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			Stop Signal			
Control group as reference level						
(Intercept)	–9.64	10.87	–0.89	.375	0.00	[0.00, 1.04]
Oral:Time2:NWS	–1.20	3.60	–0.33	.740	0.30	[0.00, 3.41]
Written:Time2:NWS	–0.65	4.32	–0.15	.880	0.52	[0.00, 2.36]
Multi:Time2:NWS	–2.32	4.77	–0.49	.626	0.09	[0.00, 1.08]
Oral:Time3:NWS	0.77	3.69	0.21	.834	2.17	[0.00, 2.87]
Written:Time3:NWS	–1.20	3.82	–0.31	.753	0.31	[0.00, 5.19]
Multi:Time3:NWS	–3.24	4.29	–0.76	.450	0.03	[0.00, 1.69]
Oral group as reference level						
(Intercept)	–3.19	14.34	–0.22	.824	0.04	[0.00, 5.72]
Control:Time2:NWS	1.20	5.58	0.21	.830	3.31	[0.00, 1.75]
Written:Time2:NWS	0.55	5.37	0.10	.919	1.73	[0.00, 6.14]
Multi:Time2:NWS	–1.12	7.26	–0.16	.877	0.32	[0.00, 4.54]
Control:Time3:NWS	–0.77	5.22	–0.15	.882	0.46	[0.00, 1.21]
Written:Time3:NWS	–1.98	4.24	–0.47	.641	0.14	[0.00, 5.37]
Multi:Time3:NWS	–4.01	6.81	–0.59	.556	0.02	[0.00, 1.07]
Written group as reference level						
(Intercept)	–14.45	16.37	–0.94	.345	0.00	[0.00, 1.43]
Control:Time2:NWS	0.65	5.01	0.13	.897	1.92	[0.00, 3.38]
Oral:Time2:NWS	–0.55	4.87	–0.11	.911	0.58	[0.00, 7.66]
Multi:Time2:NWS	–1.67	5.26	–0.32	.750	0.19	[0.00, 5.34]
Conrol:Time3:NWS	1.20	5.17	0.23	.817	3.32	[0.00, 7.98]
Oral:Time3:NWS	1.97	5.09	0.39	.698	7.21	[0.00, 1.47]
Multi:Time3:NWS	–2.04	5.24	–0.39	.698	0.13	[0.00, 3.60]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.11	.33			
*glmer(OPT_PS ~ 1+ Group*Time*log_Stop + (1|Participant) + (1|Item))*						

Appendix A53. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (Stop Signal) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			Stop Signal			
Control group as reference level						
(Intercept)	–10.41	9.52	–1.09	.274	0.00	[0.00, 3.47]
Oral:Time2:NWS	–1.54	3.06	–0.50	.615	0.21	[0.00, 8.37]
Written:Time2:NWS	–3.11	3.28	–0.95	.342	0.04	[0.00, 2.65]
Multi:Time2:NWS	–0.22	3.49	–0.06	.950	0.80	[0.00, 7.21]
Oral:Time3:NWS	–0.45	3.65	–0.12	.902	0.64	[0.00, 7.93]
Written:Time3:NWS	–3.64	3.76	–0.97	.333	0.03	[0.00, 4.02]
Multi:Time3:NWS	–2.02	3.56	–0.57	.570	0.13	[0.00, 1.36]
Oral group as reference level						
(Intercept)	–9.61	16.02	–0.60	.549	0.00	[0.00, 2.47]
Control:Time2:NWS	1.54	5.62	0.27	.784	4.67	[0.00, 2.67]
Written:Time2:NWS	–1.57	7.43	–0.21	.833	0.21	[0.00, 4.06]
Multi:Time2:NWS	1.32	5.96	0.22	.824	3.76	[0.00, 4.18]
Control:Time3:NWS	0.45	3.96	0.11	.910	1.57	[0.00, 3.57]
Written:Time3:NWS	–3.19	4.75	–0.67	.502	0.04	[0.00, 4.35]
Multi:Time3:NWS	–1.57	4.29	–0.37	.715	0.21	[0.00, 9.01]
Written group as reference level						
(Intercept)	–22.52	10.38	–2.17	.030*	0.00	[0.00, 0.10]
Control:Time2:NWS	3.11	3.17	0.98	.326	2.25	[0.05, 1.08]
Oral:Time2:NWS	1.57	3.42	0.46	.646	4.81	[0.01, 3.78]
Multi:Time2:NWS	2.89	3.64	0.80	.426	1.81	[0.02, 2.17]
Conrol:Time3:NWS	3.64	3.38	1.08	.281	3.81	[0.05, 2.77]
Oral:Time3:NWS	3.19	3.71	0.86	.389	2.43	[0.02, 3.35]
Multi:Time3:NWS	1.62	3.72	0.44	.663	5.06	[0.00, 7.09]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.07	.26			
*glmer(WPT_PS ~ 1+ Group*Time*log_Stop + (1|Participant) + (1|Item))*						

Appendix A54. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of OPT (AOSPAN) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			AOSPAN			
Control group as reference level						
(Intercept)	2.39	8.29	0.29	.773	1.09	[0.00, 1.14]
Oral:Time2:NWS	–1.48	3.92	–0.38	.705	0.27	[0.00, 4.69]
Written:Time2:NWS	–3.37	3.82	–0.88	.377	0.03	[0.00, 5.84]
Multi:Time2:NWS	–1.64	4.81	–0.34	.732	0.19	[0.00, 2.26]
Oral:Time3:NWS	–0.47	4.32	–0.11	.914	0.62	[0.00, 2.88]
Written:Time3:NWS	–1.11	4.89	–0.23	.820	0.33	[0.00, 4.55]
Multi:Time3:NWS	–1.74	5.32	–0.33	.743	0.18	[0.00, 5.56]
Oral group as reference level						
(Intercept)	–3.56	9.71	–0.37	.714	0.03	[0.00, 4.75]
Control:Time2:NWS	1.48	4.02	0.37	.712	4.41	[0.00, 1.13]
Written:Time2:NWS	–1.89	4.38	–0.43	.666	0.15	[0.00, 7.71]
Multi:Time2:NWS	–0.16	5.35	–0.03	.976	0.85	[0.00, 2.89]
Control:Time3:NWS	0.47	4.40	0.11	.916	1.59	[0.00, 8.55]
Written:Time3:NWS	–0.64	4.69	–0.14	.891	0.52	[0.00, 4.90]
Multi:Time3:NWS	–1.27	5.64	–0.23	.821	0.28	[0.00, 1.68]
Written group as reference level						
(Intercept)	–5.67	6.80	–0.83	.405	0.00	[0.00, 1.99]
Control:Time2:NWS	3.38	3.74	0.90	.367	2.92	[0.01, 4.32]
Oral:Time2:NWS	1.89	3.23	0.59	.559	6.62	[0.01, 3.62]
Multi:Time2:NWS	1.73	3.50	0.49	.621	5.62	[0.00, 5.14]
Conrol:Time3:NWS	1.11	3.43	0.32	.746	3.04	[0.00, 2.46]
Oral:Time3:NWS	0.65	3.17	0.20	.838	1.91	[0.00, 9.17]
Multi:Time3:NWS	–0.63	3.10	–0.20	.839	0.53	[0.00, 2.24]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.11	.33			

*glmer(OPT_PS ~ 1+ Group*Time*log_AOSPAN + (1|Participant) + (1|Item))*

Appendix A55. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of WPT (AOSPAN) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			AOSPAN			
Control group as reference level						
(Intercept)	–2.57	10.02	–0.26	.798	0.08	[0.00, 2.35]
Oral:Time2:NWS	–1.52	3.77	–0.40	.687	0.22	[0.00, 3.39]
Written:Time2:NWS	–3.21	3.97	–0.81	.419	0.04	[0.00, 9.31]
Multi:Time2:NWS	–1.14	4.99	–0.23	.820	0.32	[0.00, 5.41]
Oral:Time3:NWS	1.69	4.38	0.39	.697	5.44	[0.00, 2.61]
Written:Time3:NWS	0.97	4.56	0.21	.831	2.64	[0.00, 1.92]
Multi:Time3:NWS	0.05	5.16	0.01	.992	1.05	[0.00, 2.46]
Oral group as reference level						
(Intercept)	–1.06	8.65	–0.12	.902	0.35	[0.00, 7.37]
Control:Time2:NWS	1.52	4.47	0.34	.734	4.57	[0.00, 2.77]
Written:Time2:NWS	–1.70	3.82	–0.44	.658	0.18	[0.00, 3.18]
Multi:Time2:NWS	0.38	4.55	0.08	.933	1.46	[0.00, 1.04]
Control:Time3:NWS	–1.69	4.38	–0.39	.699	0.18	[0.00, 9.38]
Written:Time3:NWS	–0.72	3.92	–0.18	.854	0.49	[0.00, 1.02]
Multi:Time3:NWS	–1.64	4.29	–0.38	.702	0.19	[0.00, 8.25]
Written group as reference level						
(Intercept)	–4.29	7.03	–0.61	.541	0.01	[0.00, 1.22]
Control:Time2:NWS	3.21	3.56	0.90	.366	2.49	[0.01, 2.55]
Oral:Time2:NWS	1.70	3.36	0.50	.614	5.45	[0.00, 3.83]
Multi:Time2:NWS	2.08	4.43	0.47	.639	7.98	[0.00, 4.48]
Conrol:Time3:NWS	–0.97	3.34	–0.29	.771	0.38	[0.00, 2.54]
Oral:Time3:NWS	0.72	3.59	0.20	.841	2.06	[0.00, 2.26]
Multi:Time3:NWS	–0.92	3.97	–0.23	.817	0.40	[0.00, 9.23]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.07	.26			
*glmer(WPT_PS ~ 1+ Group*Time*log_AOSPAN + (1|Participant) + (1|Item))*						

Appendix A56. Results for the Logistic Mixed-effects Models Examining Moderating Effects of WM Measures on Gains of SOPT (AOSPAN) – Third Person Singular

	Estimate	*SE*	*z*	*p*	*OR*	95% CI for OR
			AOSPAN			
Control group as reference level						
(Intercept)	–1.85	11.99	–0.16	.877	0.16	[0.00, 2.24]
Oral:Time2:NWS	–1.13	5.67	–0.20	.842	0.32	[0.00, 2.07]
Written:Time2:NWS	–2.01	5.58	–0.36	.718	0.13	[0.00, 7.09]
Multi:Time2:NWS	–2.01	5.61	–0.36	.721	0.13	[0.00, 7.60]
Oral:Time3:NWS	–0.21	4.94	–0.04	.966	0.81	[0.00, 1.24]
Written:Time3:NWS	–1.03	5.27	–0.20	.846	0.36	[0.00, 1.04]
Multi:Time3:NWS	–2.30	5.07	–0.45	.650	0.51	[0.00, 1.96]
Oral group as reference level						
(Intercept)	–6.68	6.93	–0.96	.335	0.00	[0.00, 9.35]
Control:Time2:NWS	1.13	3.36	0.34	.737	3.09	[0.00, 2.16]
Written:Time2:NWS	–0.89	3.13	–0.28	.777	0.41	[0.00, 1.83]
Multi:Time2:NWS	–0.88	3.45	–0.25	.799	0.42	[0.00, 3.50]
Control:Time3:NWS	0.21	3.80	0.06	.956	1.24	[0.00, 2.03]
Written:Time3:NWS	–0.82	3.10	–0.26	.793	0.44	[0.00, 1.88]
Multi:Time3:NWS	–2.08	3.56	–0.59	.558	0.12	[0.00, 1.29]
Written group as reference level						
(Intercept)	–4.96	7.50	–0.66	.508	0.01	[0.00, 1.59]
Control:Time2:NWS	2.01	3.82	0.53	.599	7.48	[0.00, 1.29]
Oral:Time2:NWS	0.88	3.59	0.25	.806	2.42	[0.00, 2.67]
Multi:Time2:NWS	0.01	3.42	0.00	.999	1.01	[0.00, 7.85]
Conrol:Time3:NWS	1.03	3.39	0.30	.762	2.79	[0.00, 2.06]
Oral:Time3:NWS	0.82	3.26	0.25	.803	2.26	[0.00, 1.31]
Multi:Time3:NWS	–1.27	3.46	–0.37	.714	0.28	[0.00, 2.38]
Random effects		*variance*	*SD*			
Participant (Intercept)		<.01	<.01			
Item (Intercept)		.14	.37			

*glmer(SOPT_PS ~ 1+ Group*Time*log_AOSPAN + (1|Participant) + (1|Item))*