Evaluating Multimodal Comprehension: Effects of Question Types in Documentary Viewing

Abstract

Successful listeners utilize both top-down and bottom-up strategies. Besides decoding phonemes, they rely on background information, gestures, and images to understand messages. However, the role of images in viewing comprehension has been largely overlooked. Only Durbahn et al. (2020) and Fievez (2020) have examined comprehension questions that require processing of visual messages alone or in combination with verbal messages, and their findings were mixed, indicating a need for further research.

Operationalizing visual-only questions as those that required attention to images, audio-only questions as those that depended solely on aural input, and audiovisual questions as those that could be answered using either visual or aural information, the following research question was formulated: *Is there a significant difference in participants' comprehension scores based on question types (visual-only, audio-only, audiovisual) across two documentary genres (wildlife, science and technology)?*

Based on Mayer's (2005) statement that "people can learn more deeply from words and pictures than from words alone" (p. 1), it was expected that participants would perform better on audiovisual questions compared to visual-only and audio-only questions.

This exploratory study used a Latin square design to control for the potential order effects in documentary viewing and test administration. Nineteen participants were divided into three groups. Each group watched the two documentaries and a third documentary that was not included in the analysis in a counterbalanced order over three consecutive days. Participants completed comprehension tests for each documentary each day. While the wildlife documentary included 8 image-based, 12 audio-based, and 19 audiovisual questions, the science and technology documentary had 7 image-based, 18 audio-based, and 15 audiovisual questions.

A two-way repeated measures ANOVA examined the effects of question type and documentary genre on comprehension scores. Results revealed a significant main effect of question type, $F(2, 36) = 12.59$, $p < .001$, partial $\eta^2 = .411$, indicating that comprehension scores differed significantly across the three question types. Post-hoc pairwise comparisons using Bonferroni correction indicated significantly higher scores for audiovisual questions ($EMM = 62.09$, $SE = 2.64$) compared to both audio-only ($EMM = 54.53$, $SE = 2.84$), $p = .004$, and visual-only questions ($EMM = 48.73$, $SE = 2.91$), $p < .001$. However, there was no significant difference between audio-only and visual-only questions, $p = .199$. The main effect of documentary type and the interaction effect between question type and documentary type were not statistically significant. These findings emphasize the importance of using questions that leverage the multimodality of audiovisual input, moving beyond traditional audio-only comprehension assessments.

Keywords: viewing comprehension, multimodal input, audiovisual, images

References

Durbahn, M., Rodgers, M., & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. *System, 88.* https://doi.org/10.1016/j.system.2019.102166

Fievez, I. (2020). *What you see is what you get? Use and effectiveness of multimodal input for second language learning* (Doctoral dissertation).

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning.* Cambridge University Press.

Author: Kadir Kaderoglu

Supervisors: Roger Gilabert & M. del Mar Suárez