In my first review I observed three relevant aspects that deserved more attention from the authors. The first one was the excessively shallow revision of the theoretical models concerning visual word recognition. The second one was related to the need for more emphasis on how the previously presented models and experimental results could impact on educational and clinical praxis. The last query was related to the text assessing the mixed linear models. In my view, the first query has been properly addressed and I am satisfied with the way the authors have worked on the theoretical part of the paper. Some minor comments in relation to that are:

- the text in which the convolutional neural networks are explained deserves more detail. For those not familiar with this type of models the paragraph will be of little help to understand them. Maybe by describing the paper by Bowers et al. (2022) the potential readers would have a clearer idea about the rationale behind these complex models.
- To me, it is unclear the way the authors claim that models of visual-word recognition should have a more dynamic character. If I understand correctly, the authors relate that with the "contextual diversity", but I think this variable is as dynamic or static as many others. At some point in the paragraph, the study of orthographic processing in children is mentioned. I really think that the best approach to justify the need for a more dynamic understanding of visual word recognition is through the issue of how children learn to read (relating more explicitly contextual diversity and learning to read would be fine, see for instance Hsiao and Nation, 2018, JoM&L). The "contextual diversity" as the latent semantic is a very much interesting issue, but I think it does not reflect the dynamic processes involved in word processing by itself.

When the authors explain the relevance of the eye movements in sentence reading, they cite several papers. I suggest including this, as it is related to one of the previous variables mentioned (semantic transparency)

Peterson et al (2011). Morphological priming during reading: Evidence from eye movements. L&CP

With respect to the second issue raised in the first revision, I think that it is still possible to substantially improve it. The epigraph of the educational implications is less than one page long despite the relevance for many potential readers. It highlights the font and color as interesting issues (and I agree with that), but there are many others as, for example, the role of morphological processing (and awareness). There are many papers, and even

systematic reviews, in which it is recommended the training in morphological knowledge.
From graphemes to morphemes: An alternative way to improve skills in children with dyslexia. D. Traficante. 2012.
A meta-analysis of morphological interventions: effects on literacy achievement of children with literacy difficulties, Goodwin & Ahn, 2010

Lastly, I find the text improved in relation to how researchers approach statistical analyses. I would suggest commenting, once they are mentioned aspects such as the transformation of the RT, the relevance of the random slopes. I believe that the random slopes are critical, and many papers report no random slopes in their models. As far as I think they are key to generalize the results and it is unconservative not to use them, I think is worth mentioning it. In relation to that, I would recommend the explicit mention to the model run when reporting the data analyses in the papers.

In summary, I am more positive regarding the second version of the paper although I would like to suggest a third version in which previous queries are considered. Importantly, I think that all queries raised are relatively easy to address.