Peer review of Stage 1 Manuscript

Reading two languages simultaneously: the role of academic book language

The proposed cross-sectional research study seeks to examine the within subjects linguistic interdependence between word reading and vocabulary for bilinguals learning to read in English (written in the Roman alphabet) and Manipuri (written in an alphasyllabary called Meetei Mayek) in India. The authors propose to use parallel assessments in English and Meetei Mayek to examine word reading and vocabulary. They also include measures of phonological awareness, nonverbal IQ, and age as control variables. Demographic information will be collected from parents including information about income, educational level and literacy resources. The supplementary material also specifies measuring school level variables such as duration of Meetei Mayek and English classes.

This proposed study examines (1) the within-language zero-order correlation between word reading and vocabulary in English and in Meetei Mayek and (2a) the across-language association between word reading in Meetei Mayek while controlling for age, nonverbal IQ, phonological awareness in both languages, and English word reading, and (2b) the across-language association between vocabulary in Meetei Mayek while controlling for age, nonverbal IQ, phonological awareness in both languages, and English vocabulary. The study will refer to the triangle model of reading (for question 1), and the Linguistic Interdependence Hypothesis (for question 2). The authors propose to address these questions using both Frequentist and Bayesian analyses. I can comment only on the Frequentist analyses and rely on the editor and other reviewer(s) to comment on the Bayesian analyses.

These research questions are yet to be addressed for speakers of Manipuri, who first learn to read in English, which is used to teach them to read in Meetei Mayek. Thus, this study can add to the growing body of research on how aspects of the orthography (transparency and script type) and social and educational context (how reading is instructed, how often languages are used in the social and educational context) affect the correlation between first language and second language literacy and language skills. The study also contributes data from an under-represented sample (see discussion by Bylund et al., 2023).

Below, I provide comments on the Introduction, and suggestions to add some more contextual information. I then organize my review by the five categories indicated by PCI RR. I believe that the suggestions below can be feasibly addressed by the authors, so I recommend a revise and resubmit of the manuscript.

Suggestions for the Introduction/Literature Review section

The introduction/literature review introduces the key ideas of the paper, with reference to important terms in the literacy field including lexical quality hypothesis, orthographic transparency, triangle model of reading, lexical restructuring hypothesis, phonological processing. I felt that these terms were not explained with sufficient detail for a broader

audience and encourage the authors to expand these definitions. Specifically, Linguistic Interdependence, which is mentioned in research question 2, is not addressed sufficiently in the front matter of the manuscript. A detailed description of this hypothesis, with reference to some research examples would help the reader follow the rationale for research question 2.

Requesting more contextual information

The authors do well to present contextual information relevant to the study under the heading Biliteracy in India, such as addressing some aspects of the schooling environment (e.g., Meetei Mayek is taught with reference to English graphemes) and describing Meetei Mayek. Although this was a very detailed question, I had a few remaining questions that bear relevance for the study and ask the authors to add this information into the manuscript. I follow the recommendations, for example in the journal Infant and Child Development, that samples should be described in as much detail as possible.

For example, what educational activities are children in this sample exposed to before formal schooling? Is there a policy for children to attend early childhood education centres, or do they stay at home until Reception/Kindergarten or Grade 1? What languages are used in this early education? Please also comment on the extent of use of English and Meetei Mayek in the communities the children live in. The manuscript clearly indicates that English is dominant in schooling and that is the medium of instruction, but how familiar are children with Meetei Mayek? Is it their first language? I noted that Meetei Mayek is a lingua franca, suggesting that children may also be speakers of other languages. If so, please provide some information. How familiar are children with English? Is Meetei Mayek used in society, e.g., in road signs, i.e., do children get incidental exposure to Meetei Mayek or they only ever see it in school? What are the overall rates of literacy in the area (for example, in South Africa, literacy levels are generally very low with children in Grade 3 being unable to read in any language; this reduces the strength of correlations between word reading and other variables).

It is relevant for this study to specify children's familiarity with various languages, as the word reading, and vocabulary correlation is of interest. If children are not very familiar with the vocabulary of their language, but can decode, then one would expect a lower correlation between these two variables.

1A. The scientific validity of the research question(s).

The research questions are valid questions asked in cross-linguistic research. The authors have received ethical approval for their study. I would recommend the authors provide more detail in the research questions, and to align the power analyses with the research questions.

For example, research question (RQ) 1 says *Are the relationships among measures of word reading and vocabulary knowledge similar or different in English and Meetei Mayek*? It is not explicitly stated whether the within language or across-language correlations are of interest here. I would think that the within-language correlation is the target of this question. Secondly,

the power analysis addresses a different question. The power analysis calculates the sample size needed to identify r = .50 between word-reading and vocabulary. However, the research question compares the magnitude of the within language correlation for these variables in English, compared to their within language correlation in Meetei Mayek. To address the first research question as proposed, the authors would need to specify what difference in the two correlations would be taken as evidence that they are sufficiently different. An equivalence test can be used to determine whether the difference is larger than this smallest effect. The first power analysis can still be included, but an additional research question should be included (e.g., Are vocabulary and word-reading correlated with one another in English and Meetei Mayek in this sample?). I also expect the vocabulary and word-reading correlation might be lower than r = .50, from experience in my context. With the intended sample size of 120 students, the study is powered to detect a correlation of r = .30.

The second research question asks, *is there linguistic interdependence (English to Meetei Mayek) in word reading and vocabulary knowledge?* I ask the authors to more explicitly specify what result they will refer to as evidence for linguistic interdependence (will you refer to variance explained, statistical significance of the model comparison of models 3 and 4, or the size of the beta coefficient of the English variable?). The power analysis addresses variance explained in the outcome. Other researchers will say there is evidence for interdependence if English word reading is a significant (p< .05) predictor of Meetei Mayek word reading after controlling for other variables. Other researchers may refer to the beta coefficient being at least a certain value.

Currently you only propose to examine the effect of English on Meetei Mayek, but you could also explore the same analyses with English variables as the outcomes. This result could be informative especially if language skills are stronger in Meetei Mayek.

1B. The logic, rationale, and plausibility of the proposed hypotheses, as applicable.

A pilot study with 116 children from grades 1 to 6 was conducted and is summarized in the supplemental material. It would be helpful to have more information about the summary statistics and Cronbach's alpha of each assessment, as well as the correlation matrix reported for the pilot. Having this information would allow the reader to better assess the applicability of the hypotheses.

The authors state that *The first hypothesis is that the association between respective measures of word reading and vocabulary knowledge will be stronger in English than in Meetei Mayek*. I recommend that the authors explicitly specify the difference in correlations that will be interpreted as a meaningful effect. I overall agree with the argument provided by the authors that a stronger correlation between word-reading and vocabulary may be expected in English because (1) Meetei Mayek is more transparent than English, so vocabulary knowledge is more weakly associated with word reading, and (2) children learn English vocabulary from their English word reading. More detail on the language skills of students (as requested in the context section above) will further support this interpretation of the data.

The second hypothesis states *word reading and vocabulary knowledge in English would positively contribute to their equivalent performance in to Meetei Mayek (MM)*. I agree that this is a plausible hypothesis given the literature on cross-linguistic transfer. I would ask the authors to more specifically state the interpretation of possible outcomes. The interpretation of data is specified: *If English to Meetei Mayek additive patterns are not observed, then it might suggest a subtractive biliterate and/or bilingual environment.* I would interpret a positive beta coefficient of the English variable in the final model AND a significant model comparison as evidence of positive transfer of English to Meetei Mayek. I would interpret a negative beta coefficient of the English variable in the final model AND a significant model comparison as evidence of negative transfer of English to Meetei Mayek. I am not sure how I would interpret a null effect, where adding English to the model results in a non-significant p value, and the model comparison is not significant. I look forward to reading the authors' more detailed views on this hypothesis.

1C. The soundness and feasibility of the methodology and analysis pipeline (including statistical power analysis or alternative sampling plans where applicable).

The authors propose using a parallel test battery in English and Meetei Mayek. I commend the authors for their effort in creating assessments for the study. I would encourage the authors to share the Meetei Mayek tests for review. The adaptation of word reading tasks is straightforward. Because the pilot study summary statistics were not included in the supplemental materials, I am unable to evaluate whether 30 items in the receptive vocabulary test, and 30 items in the expressive vocabulary test adaptations will be sufficient or not to avoid ceiling or floor effects. The authors indicate they will use sets $1-15$ for the receptive vocabulary measurement which are the easier items in English. This could be enough items if children are not very familiar with English, but I would expect ceiling effects if children are very familiar with English. Similarly, 15 items may be too few for Meetei Mayek if children are very familiar with the language. More information on how vocabulary items were selected across languages is needed for the reader to evaluate test equivalence (the supplement states: *From this pool of 108, 30 items each are selected that are best suitable for English and Meetei Mayek based on the content covered in the school curriculum for the target population.*) I would also recommend against interpreting differences in the mean scores between English and Meetei Mayek tasks, because the tests may not be equivalent in difficulty level. For example, a score of 15 in English receptive vocabulary, and a score of 10 in Meetei Mayek should not be interpreted as children being more familiar with English.

Please see comments on the research questions and associated power analyses in section 1A. I also encourage the authors to more explicitly justify their selected effect sizes of interest. The sampling plan is described well based on the existing power analyses, but the authors may need to update the sampling plan pending a revised power analysis for research question 1 which examines the difference in correlation coefficients between the within-langauge correlations in English, and Meetei Mayek for word reading and vocabulary.

The authors propose ways to handle missing data via imputation. Please explicitly state which dataset (imputed or original with missing data) will be used in your analysis.

1D. Whether the clarity and degree of methodological detail is sufficient to closely replicate the proposed study procedures and analysis pipeline and to prevent undisclosed flexibility in the procedures and analyses.

The manuscript provides some detail of the methods to be used, but I would suggest more detail be included regarding the test development (probably best in the supplemental material), and the proposed analytic plan (in the manuscript). The strength of this study is that there are multiple measures of the main constructs of word reading and vocabulary, but this also complicates the analyses, and associated interpretation of results. For example, will the number of variables be reduced before analysis, e.g., by summing z scores, or will the analyses be run on the different indicator variables? If the former, justify the choice of data reduction technique; if the latter, how will you interpret the results if only some analyses support the hypotheses? Some reviewers may also point out that it seems data will be collected from multiple schools, so children are nested within schools. Will this level of clustering be accounted for, and does it need to be accounted for?

I would suggest being more explicit about the sensitivity analyses in the main manuscript, for example, exactly state that you will run the analysis with and without outliers/missing data being addressed and compare the results, for example. Which results do you anticipate including in the main manuscript?

I do not know Bayesian analysis so I could not review the proposed Bayesian analyses for this study.

1E. Whether the authors have considered sufficient outcome-neutral conditions (e.g., absence of floor or ceiling effects; positive controls; other quality checks) for ensuring that the obtained results are able to test the stated hypotheses or answer the stated research question(s).

The interpretations of the potential results could be made more explicit, especially for neutral outcomes (see comment in section 1B).

I also suggest that the authors specify what interpretation they will have if there are ceiling or floor effects in the data. I expect that ceiling or floor effects would decrease the strength of any potential correlation between variables.

Terminology

I include below some suggestions of alternate phrasing of terms to bring them more in line with other research on the topic.

I understand blending and elision tasks as phonological awareness tasks, which fall under the umbrella term phonological processing. Phonological processing also includes measures of rapid automatised naming and phonological working memory (as measured in the Comprehensive Test of Phonological Processing (CTOPP)). Since only phonological awareness will

be measured out of the three phonological processing measures in the CTOPP, I recommend using the term "phonological awareness".

In the description of Meetei Mayek, it says *Meetei Mayek has 27* **alphabets**. I would have used the term "graphemes" instead of "alphabets". An alphabet refers to a writing system where consonants and vowels are represented separately by graphemes (letters, in this case). If the correct term is "alphabets", please include a footnote explaining this to readers.

~~~

I thank the authors for the opportunity to review this manuscript and apologize for the delay in responding!