ERP and Neural Decoding of L2 Morphosyntactic Processing: Classifying Learner Groups from Distributed EEG Patterns

Hoekeon Choi[1, 2], Haeun Ko[3], Sung-Eun Lee (cristlo5@snu.ac.kr)[1, 3, 4]
[1]Artificial Intelligence Institute of Seoul National University; [2]Department of English Language and Literature, Seoul National University; [3]Interdisciplinary Program in Cognitive Science, Seoul National University; [4]Department of German Language and Literature, Seoul National University

Background: Traditional ERP analyses of L2 morphosyntactic processing employ univariate comparisons within predefined time windows and ROIs, which may overlook distributed neural patterns distinguishing learner groups. Recent work has demonstrated that neural decoding can recover grammatical information from scalp EEG [1], but it remains unexplored whether grammaticality-related neural patterns can differentiate L2 learner groups. The present study integrates ERP analysis with neural decoding to examine whether Korean adults with preschool immersive L2 experience (English-only kindergarten, EK) and those without (non-EK) can be distinguished by their neural responses to subject-verb (S-V) and determiner-noun number (Det-N) agreement violations.

Method: Thirty-five Korean university students (EK: n = 16; non-EK: n = 19), with no participant having received further immersive English education after preschool, read English sentences with agreement violations word by word during 64-channel EEG recording. *ERP:* Difference waves (ungrammatical minus grammatical) were analyzed via linear mixed-effects models in the N400 (300-500 ms) and P600 (500-700 ms) windows across midline and lateral ROIs (Figure 1). *Neural Decoding:* Difference waves from 60 scalp channels were classified using sliding 200 ms windows (100 ms stride, 0-700 ms). Four ML (SVM, LR, LDA, Stacking Ensemble) and two DL classifiers (EEG Conformer, DB Conformer), included to test beyond linear separability, classified EK vs. non-EK per window via 5-fold stratified CV with graph-based and Gaussian augmentation.

Results: *ERP:* For Det-N, a Group × Feature interaction at midline revealed that the non-EK group showed greater negativity in the N400 window than the EK (p = .019, d = 0.59), while the EK group showed greater positivity in the P600 window than the non-EK (p = .020, d = 0.58; Figure 2). A marginal non-EK negativity was also observed in the left posterior area (p = .051). No group differences emerged for S-V. *Neural Decoding:* S-V yielded higher classification accuracy than Det-N, particularly in later time windows and in DL models (Table 1). The highest accuracy was 74.3% (EEG Conformer, S-V, 400–600 ms).

Discussion: ERP and neural decoding revealed complementary patterns. For Det-N, the two groups diverged in the direction of their ERP responses, suggesting distinct weightings of lexico-semantic vs. structure-based mechanisms [2]. This aligns with the usage-based view that frequent, interaction-rich L2 input during the preschool period supports proceduralization of structure-based routines [3], potentially mitigating L1 entrenchment before first-language patterns fully consolidate [4]. Neural decoding, however, found S-V more discriminative, with classification peaking in the 400–600 ms window, consistent with the prediction that S-V, absent in Korean, should maximally differentiate groups, whereas Det-N allows partial L1 transfer via *-tul* [5]. That S-V differences surfaced only in whole-brain multivariate analyses [1] demonstrates that conventional ERP analyses can miss distributed neural signatures of L2 processing differences. These contrasting patterns suggest that L1 similarity shapes the level at which group-level variation is detectable: Det-N differences map onto known ERP components, whereas S-V differences emerge only in distributed decoding patterns. Together, the two methods uncover complementary aspects of L2 processing differences that neither approach alone would reveal.

Figures

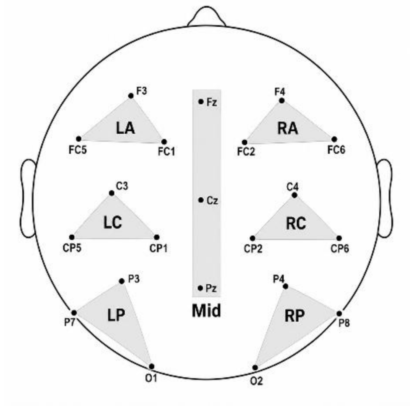

Figure 1. Regions of Interest for ERP analysis.

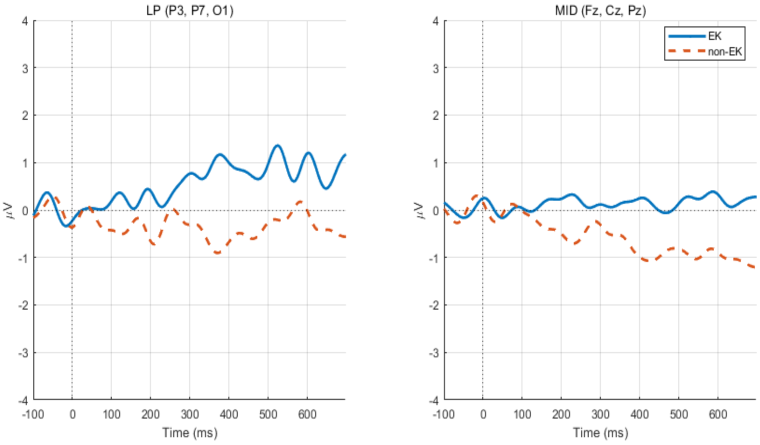

Figure 2. ERP difference waves (ungrammatical minus grammatical) by group for Det-N at left posterior and midline electrodes.

Tables

Table 1. Balanced classification accuracy by model and agreement type at 300-500 ms, 400-600 ms, and 500-700 ms.

	S-V			Det-N		
Model	**300-500 ms**	**400-600 ms**	**500-700 ms**	**300-500 ms**	**400-600 ms**	**500-700 ms**
SVM	.457	.600	.686	.600	.543	.600
LR	.486	.571	.657	.600	.514	.600
LDA	.543	.600	.629	.600	.629	.571
SE	.457	.600	.629	.600	.514	.657
EEG Conformer	.714	**.743**	.686	.629	.686	.629
DB Conformer	.571	.686	.600	.629	.629	.571

References

[1] Zhao, J., Gao, R., & Brennan, J. R. (2025). Decoding the neural dynamics of headed syntactic structure building. *Journal of Neuroscience*, *45*(17).

[2] Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, 17(2), 277–293.

[3] Ellis, N. C. (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.

[4] MacWhinney, B. (2017). Entrenchment in second-language learning. In H.-J. Schmid (Ed.), *Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge* (pp. 343–366). De Gruyter Mouton.

[5] Kim, C. E., O'Grady, W., & Deen, K. (2014). The extrinsic plural marker in Korean. *Korean Linguistics*, 16(1), 1–17.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(0670-20240045).