

## Exploring the development of 'verb 4-gram' sequences and grammar patterns in L2 writing: A comparison study of L1 Korean speakers and other L1 speakers

Joyce Lim  
University of Cambridge  
dol20@cam.ac.uk

Usage-based approaches (Ellis *et al.*, 2016) claim that the human's ability to construct language can be better understood by observing our cognitive sensitivity to frequency (Tyler and Ortega, 2016). This notion has made a significant contribution to the learner corpus approach to identify candidate 'constructions' and their varying levels of abstractions. The current study compared the verb sequences that shape the transition from beginner (A2) to near-native proficiency (C2) in a corpus of essays written by L1 Korean speakers (n=1,973; 734,300 words) and other L1 speakers (n=231,701; 53,819,650 words; 121 nationalities) from the Cambridge Learner Corpus. Using a bottom-up approach, I observed the frequency and distribution of 4-slot verb sequences in the English essays contributed between 1992 and 2011. In the analyses, it was found that across all proficiency levels in both the L1 Korean and other L2 speaker sub-corpora, there were 'core' sequences (e.g. VV DT NN IN, VV IN DT NN, VVN IN DT NN) that have already been acquired at the beginning level and remain stable in essays during the transition towards near-native-like proficiency. Furthermore, six 'emerging' sequences were identified, which are sequences that gradually start to surface, making it to the top 10 ranking at the C2 level.

From these 'core' and 'emerging' sequences, the grammar pattern approach (Hunston & Francis, 2000; Hunston 2019) was used to conduct case studies to regroup the identified patterns into broader categories. By drawing on the taxonomy of pattern grammar, I was able to observe the dynamic restructuring process of learners at each proficiency level and also make generalizations about how there is a transition towards more variety of complex 'constructions' and a wider group of meanings that become more prevalent in higher levels within the same verb sequence. Furthermore, it was clear that even within the CEFR level, the learner language is impacted by the L1, and that this was especially prevalent in the lower levels. However, even in the highest level, C2, there were noticeable differences in the use of verb sequences between L1 Korean speakers and other L2 speakers.

Overall, this study offers insight into the residual acquisition of form and meaning and demonstrates that the learners' implicit adjustment of their production is impacted by their mother tongue at all levels.

### References

- Ellis, N., Römer, U., & O'Donnell, M. (2016). *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar*. Oxford: Wiley.
- Hunston, S. (2019). Patterns, constructions, and applied linguistics. *International Journal of Corpus Linguistics*, 24(3), 324-353.
- Hunston S. & Francis, G. (2000). *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia, PA: John Benjamins.
- Tyler, A., & Ortega, L. (2016). Usage-based approaches to language and language learning: An introduction to the special issue. *Language and Cognition*, 8, 335-345.