

# Investigating lexical input processing through vocabulary growth curves: a case study of Spanish learners of French L3

Paula Lissón<sup>1</sup>, Nicolas Ballier<sup>1</sup>

(1) Université Paris-Diderot, EA 3967 – CLILLAC-ARP, France  
nicolas.ballier@univ-paris-diderot.fr  
paula.lisson@etu.univ-paris-diderot.fr

This paper deals with the investigation of lexical input processing through the analysis of vocabulary growth curves (cf. Ballier & Gaillat, 2016). Our research is based on the preliminary results obtained from the creation of a longitudinal, written corpus of Spanish learners of French L3 (beginners), having English as L2 (intermediate level)<sup>1</sup>. We propose that Vocabulary Growth Curves (VGC) could be used to assess the amount of vocabulary retained by learners in regards to the input they have been exposed to.

From the number of tokens and hapax legomena of our corpus, frequency spectrums (see Figure 1) have been generated using the {zipfR} package (Evert & Baroni, 2006) implemented in R (R Core Team, 2016). Next, frequency spectrums were used to create VGC for the three productions collected (see Figure 2).

We present this method as a preliminary investigation of longitudinal progress in lexical acquisition while the L3 input is being scanned (the textbook used in class is taken as the baseline of learners' lexical input). Our poster will discuss potential L2 intake (cf. Murphy, 2003), mostly present in lexical and syntactic transfers (cf. Osborne, 2015). Lexical transfer analysis through VGC raises issues to be debated.

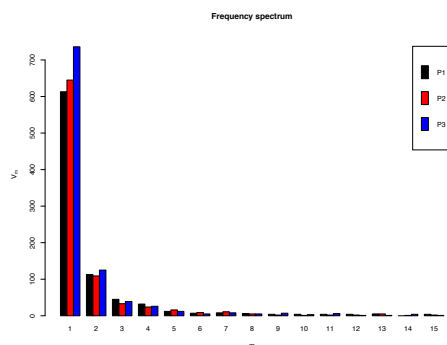


Figure 1: Frequency spectrum of hapax legomena

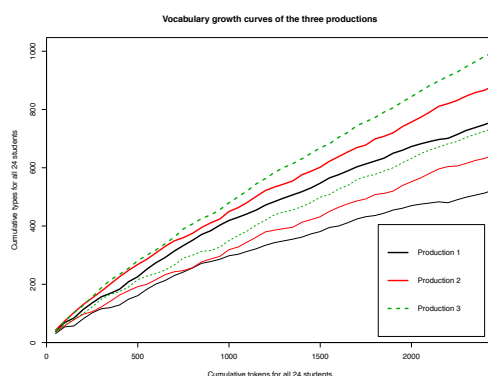


Figure 2: Vocabulary growth curves

<sup>1</sup> All the participants were University students enrolled in English studies (major) and in French studies (minor). Students are not assumed to have any prior knowledge of French. Classes of French during the first year start from the very beginning, and by the end of the first semester, students are supposed to have an A1 level of the CEFR (Verhelstel *et al.*, 2009).

## References

- Ballier, N., & Gaillat, T. (2016). Classification d'apprenants francophones de l'anglais sur la base des métriques de complexité lexicale et syntaxique (Vol. 9, pp. 1–14). Presented at the JEP-TALN-RECITAL 2016.
- Evert, S., & Baroni, M. (2006). The zipfR library: Words and other rare events in R. *Relazione Presentata all'useR*.
- Murphy, S. (2003). Second language transfer during third language acquisition. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 3(1), 1–21.
- Osborne, J. (2015). Transfer and learner corpus research. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 333–356). Cambridge University Press.
- R Core Team. (2016). R: A language and environment for statistical computing. (Version 3.3.1 (2016-06-21)). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Verhelst, N., Van Avermaet, P., Takala, S., Figueras, N., & North, B. (2009). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge University Press.