

# Evaluating the effectiveness of Data-driven learning in the context of Italian L2 learning and teaching

Luciana Forti

Data-driven learning (DDL) is known as the direct use of corpus data in the second language classroom (Leech, 1997), and is generally, though not exclusively, based on inductive, collaborative and pattern-hunting activities focused on concordances (Johns, 1991; Sinclair, 2003).

Over time, a variety of DDL approaches have been experimented with in terms of the corpus data used and the learning approach adopted (Boulton, 2017). A recent meta-analysis indicates that empirical studies aimed at evaluating its effectiveness show promise (Boulton & Cobb, 2017), though a number of shortcomings have been repeatedly pointed out. Two of these are:

- limited evidence concerning learners with a proficiency level other than advanced;
- limited evidence concerning learners of languages other than English.

This seminar will focus on an ongoing PhD project aimed at evaluating the effectiveness of using DDL principles and materials with pre-intermediate learners in an Italian L2 context.

More specifically, it addresses the following research questions:

1. Are there any significant differences when comparing a DDL to a traditional approach, in terms of collocational proficiency development?
2. How does transparency influence the development of collocational proficiency?
3. What is the students' general attitude towards working with concordance-based materials?

The project integrates native corpus data from PeC – Perugia Corpus (Spina, 2014) and leaner corpus data from LoCCLI – Longitudinal Corpus of Chinese Learners of Italian (Spina, 2017), and focuses on Verb-Noun collocations, widely recognised as challenging even for learners at higher levels of proficiency (Bestgen & Granger, 2014; Nesselhauf, 2005; Wang, 2016).

Preliminary results of the research project will be presented and reflections on the steps discussed, especially in relation to data analysis and interpretation.

## References

Bestgen, Y., & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing*, 26, 28–41.

Boulton, A. (2017). Research Timeline. Corpora in language teaching and learning. *Language Teaching*, 50(4), 483-506.

Johns, T. (1991). Should you be persuaded - Two examples of data driven learning materials. *Classroom Concordancing, English Language Research Journal* 4, 1–13.

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora*. Harlow: Addison Wesley Longman.

Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam-Philadelphia: Benjamins.

Spina, S. (2014). Il Perugia Corpus: una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione. In *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & the Fourth International Workshop EVALITA 2014litica* (Vol. 1, pp. 354–359). Pisa: Pisa University Press.

Spina, S. (2017). *Learner Corpus Research and the acquisition of Italian as a second language: the case of the Longitudinal Corpus of Chinese Learners of Italian (LoCCLI)*. Presented at the 4th Learner Corpus Research Conference, 5-7 October, Eurac Research, Bolzano.

Wang, Y. (2016). *The Idiom Principle and L1 Influence. A contrastive learner-corpus study of delexical verb+noun collocations*. Amsterdam; Philadelphia: John Benjamins Publishing Company.