**The Effect of Data-driven Learning (DDL) on the Vocabulary Size of Iranian EFL Learners**

Mahrokh Sanaye

Hossein Khazaee

**Abstract**

The aim of the present study was to investigate the effect of data–driven learning (DDL) on the vocabulary size of Iranian intermediate EFL learners. The participants of this study were 25 Iranian male and female intermediate EFL learners who were studying English as a foreign language at Mehre Iran institute, Lahijan, Iran. The participants were recruited through applying a convenience sampling procedure. For this purpose, an Oxford Placement Test (OPT) was given to 45 EFL learners at Mehre Iran institute, who had been assigned as intermediate English learners by the institute’s placement procedure. Then Nation and Beglar’s (2007) vocabulary size test (version A) has been given to these 25 subjects as the pretest to assess their threshold vocabulary size level. For 8 weeks, DDL techniques through using the COCA corpus on Longman 3000 communication word list has been performed, and after this, a second vocabulary size test (version B) has been conducted. The results of the pre- and post-test have been compared using paired samples *T*-test to see whether DDL techniques had any effect on the vocabulary size of the Iranian intermediate EFL learners. The results suggested a statistically significant effect of DDL techniques on the vocabulary size of Iranian EFL learners. Due to the sample type used, further investigations are required to bring more evidence concerning the results of the current study.

**Key words**: Corpus, COCA, Data-driven learning, Vocabulary size test