

# The Use of Corpus, Concordancer and Chatbot in the Teaching of Contemporary Arabic

---

Latifa Al-Sulaiti,<sup>1</sup> Andrew Roberts,<sup>2</sup>  
Bayan Abu Shawar<sup>3</sup> and Eric Atwell<sup>4</sup>

## Abstract

Theoretical and applied linguistics researchers are not the only groups who benefit from corpora, language learners and teachers can also. We have surveyed attempts at using corpora as an aid to teaching English, Spanish, and German. However, for Arabic the number of corpus-based language teaching studies is very limited: the University of Kuwait has used a parallel Arabic-English corpus for teaching, but there appear to be no other institutions adopting this approach. A factor is the limited number of freely available Arabic corpora and tools that can process the data adequately.

We report on the use of a free open-source corpus of contemporary Arabic of around one million words, covering a variety of text genres such as science, novels, religion etc.; used in conjunction with a free open-source concordance, aConCorde, which can analyse and display Arabic concordances appropriately. We demonstrate some of the difficulties learners have in learning contemporary Arabic and give samples of corpus-based concordance activities which aim to overcome these difficulties. We illustrate how the corpus of contemporary Arabic and aConCorde can be used to complement learning from a standard Arabic language course book. We also present the use of a trainable chatbot, trained with a corpus of Arabic, to provide online conversation practice for learners. We conclude that using a corpus and concordance not only exposes the learners to authentic texts but also assists them in exploring the language and finding its rules and meanings independently.

---

<sup>1</sup> University of Leeds

<sup>2</sup> University of Leeds

<sup>3</sup> University of Leeds

<sup>4</sup> University of Leeds

*e-mail:* eric@comp.leeds.ac.uk