

now largely corpus-based,¹ and has had an impact on descriptive grammars, as shown by the ground-breaking *Longman Grammar of Spoken and Written English* (Biber et al., 1999). The revolution lies precisely in the fact that corpora provide evidence not just of what is cognitively feasible, grammatically accurate, and contextually appropriate, but also of what is attested in use. This has paramount socio-cognitive implications for language learning, notably for the development of learners' *competence* and *capacity* (as described by Hymes, 1972 and Widdowson, 1983), that is, both their awareness of how the language is used and their ability to use it (see also Aston, 2001, pp. 3–5).

Based on these characteristics, corpora are a major resource in *data-driven learning* (DDL), which moves from the assumption that the task of the learner is to *discover* the foreign language, and that the task of the language teacher is to provide a context in which the learner can develop strategies for discovery – strategies through which he or she can ‘learn how to learn’ (Johns, 1991, p. 1). This approach to language learning goes hand in hand with the Observe – Hypothesize – Experiment (OHE) paradigm championed by Michael Lewis, whose pedagogic value has been acknowledged by more and more language instructors. In his work on what he called the *lexical approach*, Lewis (1993, 1997) claims that the OHE paradigm is more effective than the Present – Practice – Produce (PPP) model in second language acquisition (SLA), and calls for the inclusion of *lexical items*, that is, words or sequences of words counting as units of meaning, into the input English language teachers should present learners with. While a discussion of the merits and limitations of either the OHE or the PPP model is beyond the scope of the present paper,² we embrace the view that introducing learners to the OHE approach through carefully designed teacher-

1 The first fully corpus-based dictionary is the *Collins COBUILD English Dictionary*, which was first published in 1987 (see <https://collins.co.uk/pages/elt-cobuild-reference-the-history-of-cobuild>)

2 Suffice it to say here that Lewis' theories on language learning have been challenged among others by Ellis (1992), who posits that the PPP paradigm does not necessarily imply adopting solely a deductive approach, and by Golebiewska and Jones (2014), who point out areas of overlap between the OHE and the PPP models, and report on how, under test conditions, the two seem to yield similar results in terms of their effect on learners' receptive and productive knowledge.

