The decision to develop a specific corpus-based language learning object was based on the need of the PhD students in the Department of Computer Science at the University of Bari (Italy) to enhance their English language abilities for writing theses and articles. This need was clearly recognized in the Department both by the discipline specialists and by the authors – the ESP language specialists – thanks to their vast experience in reading and correcting such manuscripts.

A learning object was considered the best, most innovative method for transmitting self-contained “chunks” of linguistic information to busy PhD students. In order to establish a suitable approach to the development of the learning object, the authors performed an in-depth study of the literature concerning academic, scientific and specialized discourse analysis and works regarding corpus analysis. The authors also considered it important to make comparisons between the learner (L1) 'corpora' and the native speaker (L2) 'corpora', in order to help solve problems of non-native speaker writers of academic English texts.

An accurate analysis of the students’ starting level and linguistic needs was then performed, by means of a written test. To build the ESP corpus, a selection of specialized academic writings, published in Computer Science journals, was collected, chosen by the discipline specialists. These target papers were presented to the students and each section was analysed both stylistically and grammatically, concentrating on the recurrent lexical and syntactic features of the micro-language used in Computer Science (for example, the common use of acronyms, nominal groups and specific phrasal verbs). Hence the learning object gives the students a basic model of how to build a paper/thesis in the specific field of Computer Science, through the identification of these key corpus tools. The learning object is described in detail in the core of this paper, including the guidelines, hints and useful activities provided to aid the learning process. The results of the student training are presented and some conclusions are drawn regarding the efficacy of the approach adopted and the suitability of the corpus size.