Un corpus DIY pour l'étude du roumain en diachronie. Stratégies de constitution et stratégies de recherche
A DIY corpus for investigating historical Romanian. Its creation and research strategies

Ana Zisman

Abstract: The present paper aims to provide an overview of some of the advantages of creating and working with a DIY corpus, *i.e.* a corpus compiled by the linguist, as groundwork for a PhD thesis. Collected in order to investigate the grammatical and pragmatical behavior in historical Romanian of some so-called parenthetical verbs: *a zice/ a spune* 'to say', *a crede* 'to think', *a ști* 'to know', within 5 types of texts from the 16th/17th to the 20th centuries, this DIY corpus represents a necessary alternative as a database of Romania texts. Although its creation demanded some additional steps (*e.g.* the selection of the texts, which is determined by various diachronical factors), such a corpus proves to be relevant for investigating parenthetical verbs in literary, historical and law texts, as well as in formal and informal letters. In order to do so, the paradigm of the afore-mentioned verbs has to be systematized in relation to a precise word frequency per text type.

Key words: DIY corpus, grammar, pragmatics, parenthetical verbs, Romanian.