Historical Development of Negative Emotive Intensifiers in Hungarian

Authors : Martina Katalin Szabó, Bernadett Lipóczi, Csenge Guba, István Uveges

Abstract : In this study, an exhaustive analysis was carried out about the historical development of negative emotive intensifiers in the Hungarian language via NLP methods. Intensifiers are linguistic elements which modify or reinforce a variable character in the lexical unit they apply to. Therefore, intensifiers appear with other lexical items, such as adverbs, adjectives, verbs, infrequently with nouns. Due to the complexity of this phenomenon (set of sociolinguistic, semantic, and historical aspects), there are many lexical items which can operate as intensifiers. The group of intensifiers are admittedly one of the most rapidly changing elements in the language. From a linguistic point of view, particularly interesting are a special group of intensifiers, the so-called negative emotive intensifiers, that, on their own, without context, have semantic content that can be associated with negative emotion, but in particular cases, they may function as intensifiers (e.g.borzasztóanjó 'awfully good', which means 'excellent'). Despite their special semantic features, negative emotive intensifiers are scarcely examined in literature based on large Historical corpora via NLP methods. In order to become better acquainted with trends over time concerning the intensifiers, The exhaustively analysed a specific historical corpus, namely the Magyar TörténetiSzövegtár (Hungarian Historical Corpus). This corpus (containing 3 millions text words) is a collection of texts of various genres and styles, produced between 1772 and 2010. Since the corpus consists of raw texts and does not contain any additional information about the language features of the data (such as stemming or morphological analysis), a large amount of manual work was required to process the data. Thus, based on a lexicon of negative emotive intensifiers compiled in a previous phase of the research, every occurrence of each intensifier was queried, and the results were stored in a separate data frame. Then, basic linguistic processing (POS-tagging, lemmatization etc.) was carried out automatically with the 'magyarlanc' NLP-toolkit. Finally, the frequency and collocation features of all the negative emotive words were automatically analyzed in the corpus. Outcomes of the research revealed in detail how these words have proceeded through grammaticalization over time, i.e., they change from lexical elements to grammatical ones, and they slowly go through a delexicalization process (their negative content diminishes over time). What is more, it was also pointed out which negative emotive intensifiers are at the same stage in this process in the same time period. Giving a closer look to the different domains of the analysed corpus, it also became certain that during this process, the pragmatic role's importance increases: the newer use expresses the speaker's subjective, evaluative opinion at a certain level.

Keywords : historical corpus analysis, historical linguistics, negative emotive intensifiers, semantic changes over time

Conference Title : ICNLPMSSA 2022 : International Conference on Natural Language Processing, Morphological, Syntactic and Semantic Analysis
Conference Location : Bangkok, Thailand
Conference Dates : February 07-08, 2022