

The main purpose of the thesis is to compare the existence and frequency of some words of Romani (Gypsy) origin in the actual Spanish and Hungarian language from the sociolinguistics point of view and to prove the convenience of some methods of investigation.

The thesis is divided into three parts. The first part contains a general introduction to the relevant questions concerning the Indian origin of the Roma (Gypsies), their travelling from India through Persia, Armenia and Byzantium to Europe, their final diffusion there and the basic characteristics of the European Romani dialects. Further, the history of Gypsies in Spain and Hungary follows together with a brief description of the local Romani dialects existing in both countries (in case of Spain we consider caló a hybrid between Spanish grammar and Romani lexic).

The second part of the thesis describes the methods of investigation. The methods count with the use of Internet and the investigation focuses, in fact, only at its users. The results of a recent sociological study about Internet users in Spain are exposed, according to which the main characteristics of Internet users are: lower age (80% of them are less than 45 years old), higher education and a slight prevalence of males. The innovation of the questionnaire method consists in the electronic form of the questionnaire. The existence of 57 Gypsy words in Spanish and 43 Gypsy words in Hungarian was investigated in this way. The number of the respondents, originated from all parts of the country, was higher than 1000 in Spain (184 in Hungary) and their social sample was comparable with the study mentioned above.

The second and the third method are based on a corpus investigation. Google is considered as an enormous corpus in which the words and expressions of interest are searched and, in that way, their existence is proved.