

## The neologisms in the AlCo

Besim Kabashi

Friedrich-Alexander-Universität Erlangen-Nürnberg  
Ludwig-Maximilians-Universität München  
besim.kabashi@{lmu,fau}.de

AlCo is a corpus of Albanian language for linguistic research, that contains a hundred million word tokens. It is a text corpus, but it aims to collect also speech examples that can be used for linguistic research. The corpus is a reference corpus that covers different domains of language, e. g. belletristic, economy, geography or medicine. The texts are collected according to a large set of criteria, e. g. subtype of text (columns, interviews, etc.), author gender (female or male), text length, etc.

As the work building the corpus is still in progress, and as we continue to add new texts to the corpus or to replace some of the old texts with new ones, we identify new words, some of them not listed in dictionaries respectively in lexicons. Especially the texts taken from newspapers contain neologisms. This phenomenon is very important to monitor the language change, e. g. it is crucial for the lexicography, to cover this new data. In our talk we will show some of the types of neologisms in the AlCo and speak about them.

### References

Besim Kabashi (2015): *Automatische Verarbeitung der Morphologie des Albanischen*. XVIII, 211. Erlangen, FAU University Press, 2015. ISSN: 2198-8102. ISBN [Print]: 978-3-944057-40-8. eISBN [Online]: 978-3-944057-43-9. <https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/6859>

Kabashi, Besim [2017/-18, in publication process]: “AlCo – një korpus tekstesh i gjuhës shqipe me njëqind milionë fjalë”. Në: Seminari Ndërkombëtar për Gjuhën, Letërsinë dhe Kulturën Shqiptare, XXXVI. Universiteti i Prishtinës. Prishtinë. [7 ff.]. (“AlCo – an Albanian Text Corpus with 100 million words”. In: The XXVI International Seminar for Albanian Language, Literature and Culture. University of Prishtina, Kosovo). Prishtina, 2017.

Kabashi, Besim [2017/-18, in publication process]: “Neologizmat në shqipen e sotme – njohja dhe analiza automatike”. Në: *Studimet albanistike në vendet ku flitet gjermanishtja*. Akademia e Shkencave dhe e Arteve e Kosovës & Akademia e Shkencave e Republikës së Shqipërisë. [5 ff.]. (“Neologisms in contemporary Albanian – automatic recognition and analysis” In: *Albanian studies in the German speaking countries*. Ed. by the Academy of Sciences and Arts of Kosovo & Academy of Sciences of the Republic of Albania). Prishtina, 2017.