ANŽELIKA GAIDIENĖ
Institute of the Lithuanian Language
ORCID id: orcid.org/0000-0001-8775-788X

Fields of research: semantics, lexicology, lexicography, computational linguistics.

AURELIJA TAMULIONIENĖ
Institute of the Lithuanian Language
ORCID id: orcid.org/0000-0003-0728-1856

Fields of research: language of children and youth, language preferences, computational linguistics.

DOI: doi.org/10.35321/all86-08

EUROPEAN LANGUAGE EQUALITY IN THE DIGITAL AGE: THE CASE OF LITHUANIA

Europos kalbų lygybė skaitmeniniame amžiuje: Lietuvos atvejis

ANNOTATION

The article addresses the current status of Lithuanian language technologies and presents the situation of European language equality in the digital environment. It investigates quantitative and qualitative indicators revealing language equality in the context of the European Union taking into account the number of speakers, digital language resources and technologies and the support to them, with a special focus on the case of Lithuania. The article intends to outline the work which has already been accomplished in the area of language technologies and identify the gaps and challenges which still need to be addressed in the case of Lithuanian as an official national and EU language. It provides the latest overview of the situation of Lithuanian language technologies through the analysis of digital language resources and tools/services. The results show that though a number of changes have occurred in the past ten years, there is still a shortage of language resources in education and other spheres.

KEYWORDS: European languages, language resources, language tools/services, language equality, language technologies.

European Language Equality in the Digital Age:
The Case of Lithuania

ANOTACIJA

Straipsnyje rašoma apie lietuvių kalbos technologijų būklę, supažindinama su Europos kalbų lygybės skaitmeninėje terpėje situacija. Nagrinėjami kiekybiniai ir kokybiniai kalbų lygybę atskleidžiantys rodikliai Europos Sąjungos kontekste, atsižvelgiant į kalbėtojų, skaitmeninių kalbos išteklių ir technologijų skaičių bei joms teikiamą paramą, ypatingą dėmesį skiriant Lietuvos atvejo analizei. Šiuo straipsniu siekiama pabrėžti iki šiol kalbų technologijų srityje atliktą darbą ir išryškinti spragas bei atskleisti išbandymus, su kuriais susiduria ir juos sprendžia oficiali nacionalinė ir Europos Sąjungos kalba – lietuvių kalba. Straipsnyje pateikiama naujausia lietuvių kalbos technologijų padėties apžvalga analizuojant skaitmeninius kalbos išteklius ir įrankius / paslaugas. Rezultatai rodo, kad per pastaruosius 10 metų įvyko nemažai pokyčių, bet vis dar trūksta išteklių švietimo ir kitose srityse.

ESMINIAI ŽODŽIAI: Europos kalbos, kalbos ištekliai, kalbos įrankiai / paslaugos, kalbų lygybė, kalbų technologijos.

INTRODUCTION

Language technologies are interrelated and (in)directly affect a number of daily activities: work, education, communication, etc. Language underpins most digital resources: mobile devices, social networks, virtual assistants, translation tools, spell-checkers, etc. (Pastor et al. 2017: 19). Naturally, these advances do not benefit all Lithuanian citizens on equal terms. We have to admit that the technologies adjusted to Lithuanian are still lacking, and speakers often need to revert to English. It is especially important when evaluating language viability and power, because the lack of language technologies may result in the eventual decline in language usage. The situation of Lithuanian language technologies does not yet meet all the needs.

According to the 2021 data of the Lithuanian Department of Statistics, in the 16–74 age group, almost 87% of the Lithuanian population uses the Internet (compared to almost 64.1% in 2011, and 74.4% in 2016), as many as 100% in the 16–24 age group, and 55.2% in the 65–74 age group.[1] According to the 2021 data, 81.4% of households have a personal computer, and 86.6% have Internet access.[2] The Internet is mainly used for information retrieval, communication, leisure and banking: 79% of the population aged 16–74 use

[1] Available at: https://osp.stat.gov.lt/statistiniu-rodikliu-analize?hash=b3603975-ca07-47cb-aaaf-bc3a3a403a1f#/.

[2] Available at: https://osp.stat.gov.lt/lietuvos-statistikos-metrastis/lsm-2019/mokslas-ir-technologijos/informacines-technologijos.

the Internet for communication; 74% read the news; 71% use the Internet for leisure time (watch movies or TV shows, listen to music, play or download recordings, games); 68% use online banking services. Meanwhile, 27% of the population use the Internet for learning, professional development or self-education purposes.[3] According to the 2021 data, 82.2% of 16–74 year-olds use the Internet for personal purposes, for example, 65.2% communicate on social networks, 70.5% of the population socialize in real time (e.g., via Skype, Messenger, WhatsApp, Viber, Snapchat).[4] In 2021, about 225,000 *.lt* domains were registered (compared to about 139,000 in 2012, and about 188,000 in 2018), of which more than 2,000 contain letters with Lithuanian diacritical marks (*ė*, *ž*, etc.). In addition, Lithuania remains among the leaders in fiber-optic Internet. In Lithuania, the coverage of the fiber-optic network reaches 46.8%.[5]

On 13 October 2020, the Seimas of the Republic of Lithuania approved an important document for the Lithuanian language and its future, *The Guidelines for the Development of the Lithuanian Language in the Digital Environment and the Progress of Language Technologies for 2021–2027*.[6] The Guidelines were drafted by a working group formed by the State Commission of the Lithuanian Language. These Guidelines must help to ensure the full use of the Lithuanian language in the digital environment and to establish and maintain the status of the Lithuanian language in the information society. This requires an increase in digital language resources – texts and recordings corpora, the development of language technologies and the creation of public services based on them, so that no group of society or region can feel the digital divide and foreign languages can integrate more easily into the Lithuanian society. Language technologies must help strengthen the ties between the Lithuanian society and the diaspora, and reduce the exclusion of the Lithuanian-speaking community in the global knowledge society (Jaroslavienė, Miliūnaitė 2020).

The Guidelines set out the essential tasks or challenges of Lithuanian language technologies. They stipulate what should be done in Lithuania in the near future and in which directions to work: 1. To increase the competence of specialists working in the field of language technologies and to raise the level of society's ability to use the opportunities provided by language technologies.

[3] Available at: https://osp.stat.gov.lt/skaitmenine-ekonomika-ir-visuomene-lietuvoje-2020/gyvenimas-internete.

[4] Available at: https://osp.stat.gov.lt/statistiniu-rodikliu-analize?hash=1194ea01-82ee-4bde-a222-94c5ced50f4e#.

[5] Available at: https://ivpk.lrv.lt/lt/naujienos/lietuva-islieka-tarp-pirmaujanciu-sviesolaidinio-interneto-lyderiu-2.

[6] Available at: https://www.e-tar.lt/portal/lt/legalAct/71152ab00eee11ebb74de75171d26d52.

2. To accumulate and enrich open, reliable, high-quality, reusable digital language resources and other digital language datasets. 3. To develop the language technology infrastructure, the application of language technologies in the public sector and public services, to create and improve publicly available information technology solutions and tools (Guidelines 2020).

It is good to know that a significant progress has been made in adapting the Lithuanian language to the digital environment: a number of digital language resources and basic language analysis tools, complex online language services, the Lithuanian language ontology have been developed, and a number of computer programs and tools have been localized. Applied computer software relevant to the society is being Lithuanianized, standardization works of computer terms are being carried out. Lithuanian researchers actively participate in the cooperation and mobility activities of international associations, and the core of Lithuanian language specialists working in the field of Information Technology application and systematically developing innovative works in this field is formed. Lithuania also strives for all citizens to have full access to digital solutions, therefore the policy of their adaptation for the disabled is very important (Guidelines 2020).

The object of this article is the digital resources and tools/services of the Lithuanian language. The article seeks to investigate the data on Lithuanian digital resources and to show the current status of Lithuanian language development in the digital environment with due regard to the general European context. The above aim entails the following objectives: 1) to discuss the status of Lithuanian language technologies in the multilingual European context; 2) to carry out the analysis of the empirical data collected in the study; 3) to determine the major strengths, weaknesses, opportunities and threats underlying the situation of the Lithuanian language in the digital environment. The research data were collected[7] in 2021–2022 in the framework of the project "European Language Equality" funded by the European Commission[8] (hereinafter – ELE).

[7] The author expresses her gratitude to the State Commission of the Lithuanian Language, Vilnius University, Vytautas Magnus University, Tilde IT for assistance in the collection of data about language resources.

[8] Twenty-four official languages and more than 60 regional and minority languages constitute the fabric of the EU's linguistic landscape. This project answers this call and lays the foundations for a strategic agenda and roadmap for making digital language equality a reality in Europe by 2030. The primary goal of ELE is to prepare the European Language Equality Programme, in the form of a strategic research, innovation and implementation agenda and a roadmap for achieving full digital language equality in Europe by 2030. This programme was prepared jointly with the whole European Language Technology, Computational Linguistics and language-centric AI community, as well as with representatives of relevant initiatives and associations, language communities

The aim of the metadata collection activities was to discover the components that contribute to the level of technological support of the Lithuanian language.

In the first stage of the metadata collection process we focused on Lithuanian language resources and technologies, i.e., corpora a. k. a datasets (collections of raw or annotated, monolingual or bi-/multilingual, mono- or multimodal, text segments or documents, audio transcripts, scripts, audio and video recordings, etc., as well as learner corpora and sign language corpora); language descriptions comprising language models and computational grammars; lexical/conceptual resources, comprising computational lexica, terminological databases, gazetteers, ontologies, term lists, thesauri, etc.; tools/services: services offered through the web, other networks or running in the cloud, but also downloadable tools, source code, etc. (Giagkou, Piperidis 2021: 9). These include basic NLP tools for the European languages (morphological analysers, POS taggers, lemmatizers, parsers, etc.), authoring tools (e.g., spelling, grammar and style checkers), tools/services for information retrieval/extraction/mining, text and speech analytics, machine translation, natural language understanding and generation, speech technologies, etc.). In the second stage, an additional round of metadata collection was implemented. During the second stage we focused on identifying LT stakeholders in Europe and national and European LT-focused projects, public funding for LT/NLP/AI, or any other contextual factors that may emerge from the preliminary definition of the ELE (Giagkou, Piperidis 2021: 9).

The collected Lithuanian language resources were described under the single metadata form to be completed by all members of ELE, which included the following mandatory and recommended elements. Mandatory elements: resource type (a classification of resources into types: corpora, lexical/conceptual resources, grammars/(language) models, tools/services); resource name (a human-readable name or title by which the resource is known); landing page; description (a short free-text account that provides information about the resource); language(s) (the language(s) covered by a corpus or lexical/conceptual resource or the input language of a tool/service); corpus subclass; lexical/conceptual resource subclass; subclass of grammar/model; media type(s) of parts (for corpora and lexical/conceptual resources only); media type(s) of input (for tools/services only); language dependent (for tools/services only), etc. Recommended elements: resource short name; funding type; resource publication year; licence; homepage of source; organisation/provider: the organisation responsible for providing, curating, maintaining and making available the resource; etc.

and RML groups (more information about the project available at: https://european-language-equality.eu/about/).

1. EUROPEAN LANGUAGE EQUALITY: IS IT (IM)POSSIBLE?

The linguistic landscape of the European Union (hereinafter – the EU) is rich and diverse. It is composed of 24 official languages and over 60 national and regional languages. This linguistic and cultural diversity is an essential part of the heritage of every state, which is highly appreciated and promoted in the EU, with a special focus on multilingualism. Multilingualism as an EU value is guided by the motto "United in diversity", which first came into use in 2000.[9] The EU pays special attention to regional or minority languages driven by its commitment to protect them: "The European Charter for Regional or Minority Languages is the European convention for the protection and promotion of languages used by traditional minorities. Together with the Framework Convention for the Protection of National Minorities it constitutes the Council of Europe's commitment to the protection of national minorities".[10] The equality of European languages is unanimously recognized as one of the EU's greatest strengths. However, it also poses a number of challenges, as every state is in a different demographic, geopolitical, cultural, etc. situation, and inevitably faces various barriers, including linguistic; there is a substantial divide in the area of language technologies. It is also underlined in the resolution adopted by the European Parliament "Language equality in the digital age".[11]

The numbers of the speakers of the official EU languages indicate (see Fig. 1) that the following languages are the top five most spoken languages: English (approx. 30%), German (17%), French (15%), Spanish (8%), and Italian (7%).

The data showing the dominant languages (top five) do not change if all the languages of the EU are evaluated (e.g., see Fig. 2). It has been expressly declared since the EU came into existence that Europe retains its linguistic and cultural diversity; significant sums are allocated for that matter (Rehm, Uszkoreit 2013: 12–14). However, the current situation shows that the languages with fewer speakers still face various language barriers, and it is continuously discussed how to reduce this language inequality by digital means, as it is impossible to learn so many EU languages without technological support.

[9] More information available at: https://european-union.europa.eu/principles-countries-history/symbols/eu-motto_lt.

[10] More information available at: https://www.coe.int/en/web/european-charter-regional-or-minority-languages?.

[11] More information available at: https://www.europarl.europa.eu/doceo/document/TA-8-2018-0332_EN.html.

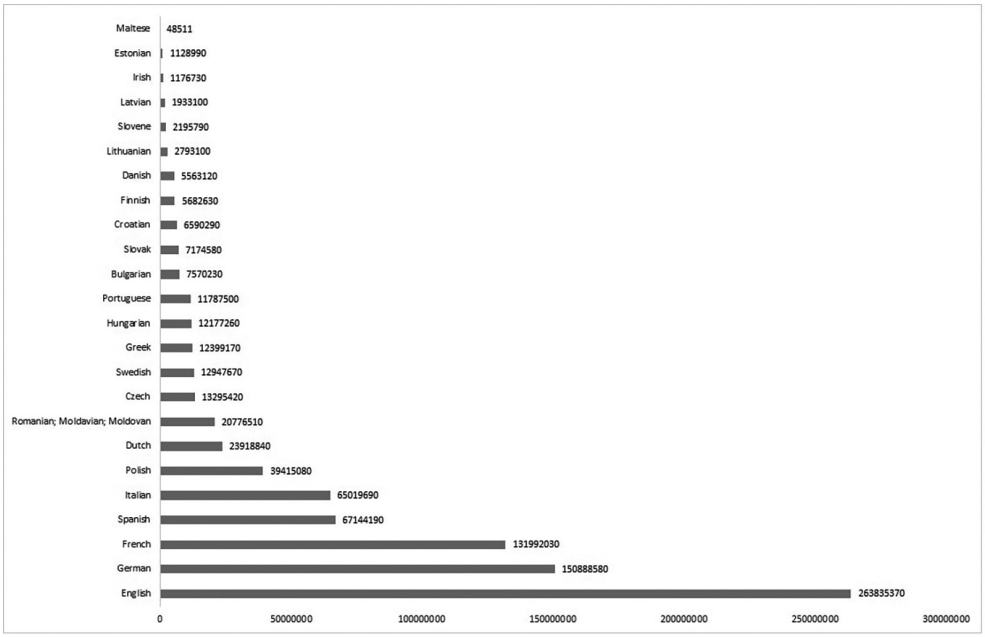

Fig. 1: Number of speakers of the official EU languages [12]

If we look at the situation of European language equality by the number of speakers, we see that all the EU languages can be grouped as follows: the most vulnerable languages are at the very top (Fig. 2), and this list of languages is rather extensive; we could relatively draw a line up to Estonian. The middle group is represented by the languages, which are spoken by a rather low number of speakers, but nevertheless occupy higher positions by other indicators of language technologies and technological support (Fig. 3). At the bottom, we can see a list of the official EU languages taking up the strongest position by the number of speakers and by other indicators (see further). Irrespective of the fact that regionally recognised languages can function on equal terms with other EU languages, the assumption concerning the strongest status of official languages should not be rejected, and the technological divide is still very large.

[12] Data from: https://european-language-equality.eu/languages/.

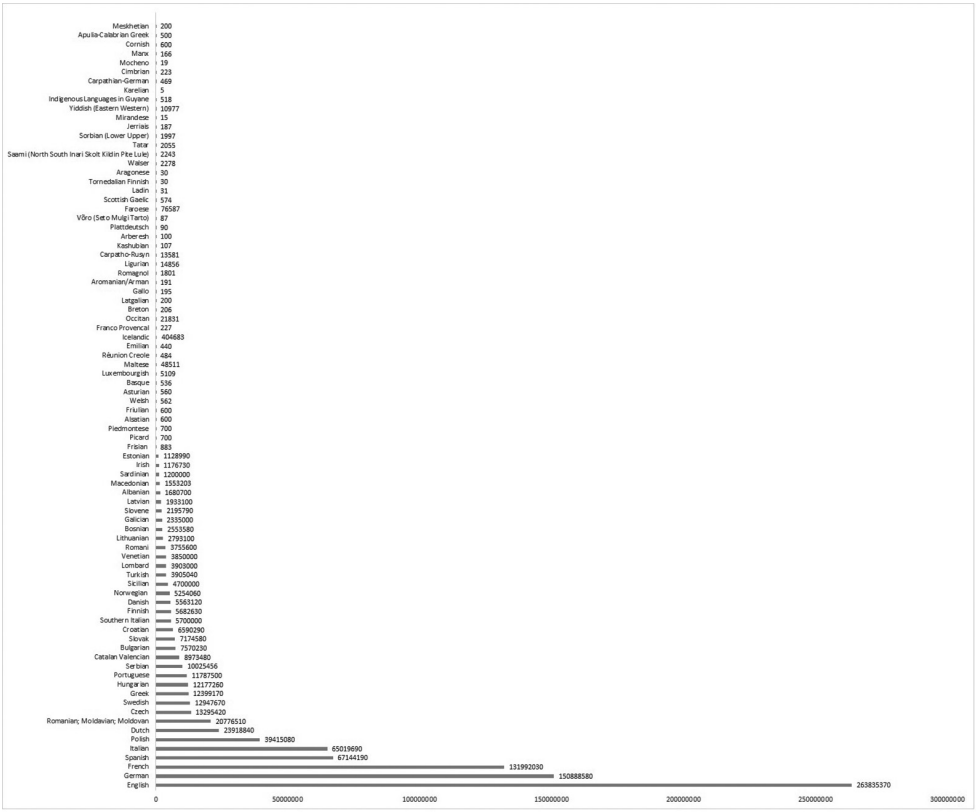

Fig. 2: Number of speakers of the EU languages

The latest 2022 data on language resources and the support allocated for their development (see Fig. 3) show that English is still the best supported language. French, German and Spanish also compete, and their results are similar to English in some dimensions. Other official EU languages are moderately supported; yet some languages have only weak or no support. National or regional languages also receive fragmentary support; other languages are borderline cases by the level of support. The analysis yields a general conclusion that no natural language is optimally supported by technologies, i.e., the required level of support has not been achieved for any language, not even English (Gaidienė, Tamulionienė 2022: 17).

Fig. 3: State of technology support, in 2022 (Figure from: Gaidienė, Tamulionienė 2022: 17). Light grey: weak/no support; grey: fragmentary support; dark grey: good support

In the framework of the ELE project, the level of progress made for each language from 2012 was compared (for more information, see Vaišnienė, Zabarskaitė 2012). We can see (see Fig. 4) that the technological level of languages changed remarkably (e.g., the widespread use of virtual assistants, chatbots, improved capacities for text analysis, etc.). The results of this analysis only show a relative status of languages, not the progress achieved for a specific language. The language technology field has made a remarkable progress in the past ten years, but the gap between best supported languages and less supported languages is still evident.

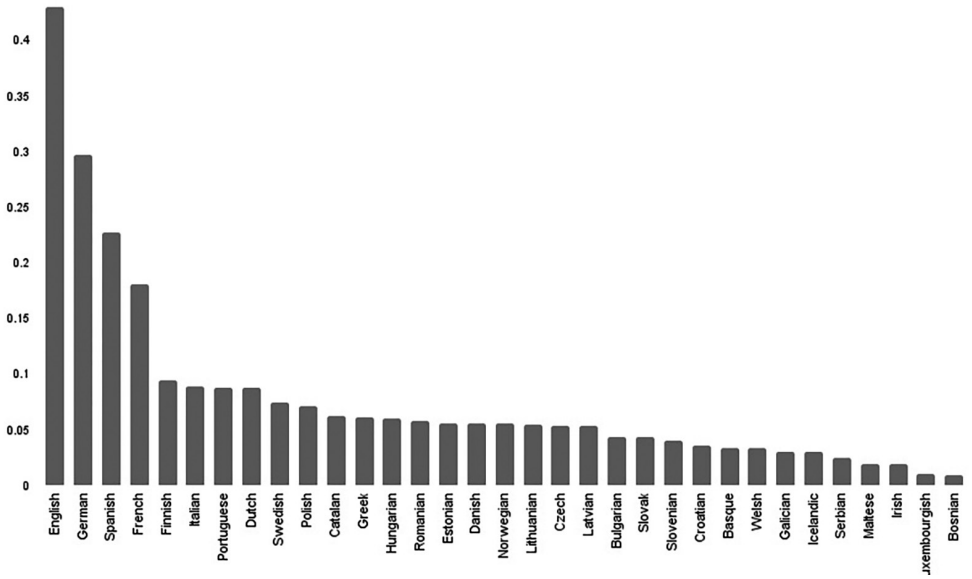

Fig. 4: Overall state of technology support for selected European languages (2022) (Figure from: Gaidienė, Tamulionienė 2022: 18)

It is difficult to say whether European language equality is possible to achieve. There is a great divide if we compare the number of speakers, language resources and technologies and the support benefited by them. It is pointed out in the report delivered within the ELE project "Report on the state of Language Technology in 2030" that two points are particularly critical to achieve by 2030: "1) Neural language models and related techniques are key to sustain progress in LTs. Therefore, being able to build neural language models for other languages with the same quality as English is key for language equality; 2) Multilingual data is the key element to train such models in a variety of languages. We should not take for granted that large amounts of publicly available corpora of good quality can be readily obtained for all European languages, rather the contrary.

The effort to ensure that all languages have large amounts of publicly available corpora of good quality, taking into account fairness issues, should be at the center of any future efforts towards DLE[13]"[14]. Increased attention is paid to the above by the ELE project; it is planned to deliver a detailed strategic plan laying down specific actions by 2030.

2. AVAILABILITY OF LANGUAGE DATA AND TOOLS

As mentioned before, the data on the available digital Lithuanian language resources were collected as part of the ELE project. In total, 218 Lithuanian language digital resources were collected and described, the largest part of which consists of lexical/conceptual resources (139; 64%); tools/services (53; 24%) come second; corpora (24; 11%) occupy the third position, whereas grammar/language models[15] (2; 1%) comes in the fourth position (see Fig. 5).

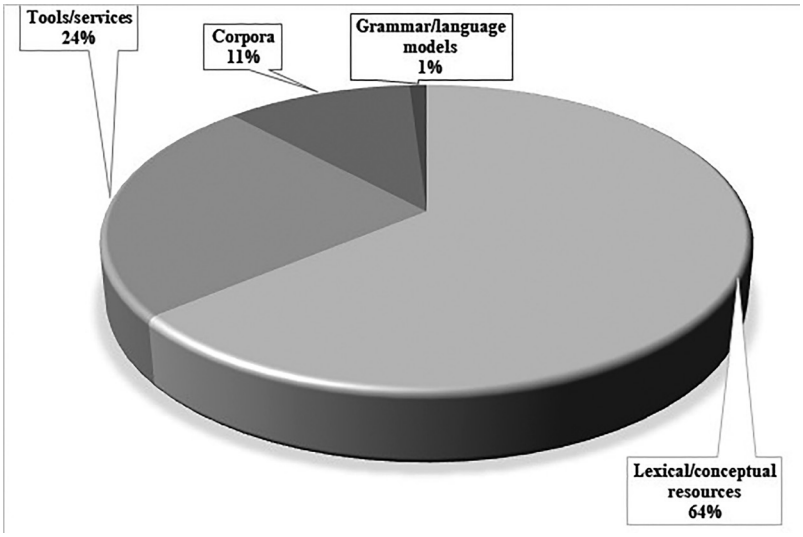

Fig. 5: Types of Lithuanian Language Digital Resources

[13] DLE: Digital Language Equality.

[14] Available at: https://european-language-equality.eu/wp-content/uploads/2022/05/ELE___Deliverable_D2_18__Report_on_State_of_LT_in_2030_.pdf.

[15] Lithuania is lagging behind by the digital resources of grammar/language model; as barely several resources were identified in this field, they are not separately discussed in the article.

Every group of digital resources will be separately discussed in the article by giving the examples of the most well-known and largest resources, tools and electronic services.

2.1. Lexical/conceptual resources

Dictionaries/lexicons account for the major part of lexical/conceptual resources (77%); they are followed by terminological resources (15%). Lexical/conceptual resources also include ontologies, wordnets (4%).

Most of the digital dictionaries are developed by the Institute of the Lithuanian Language and accessible through the information system for Lithuanian language resources E. KALBA[16].

E. KALBA provides access to the latest descriptive normative *Dictionary of the Standard Lithuanian Language*,[17] which currently contains 74,062 headwords (all dictionary entries starting with letters B, C, Č, D, E, Ę, Ė, F, G, H, J, O, R, Z, Ž and the whole nomenclature are accessible online). Thousands of new words and new meanings are added to the dictionary; headwords are illustrated with numerous authentic examples from real language usage.

Other monolingual descriptive dictionaries are also available in the information system for Lithuanian language resources E. KALBA: *The Dictionary of the Modern Lithuanian Language*[18] (8th revised and updated edition, which includes 48,342 headwords) and *The Dictionary of the Lithuanian Language*[19] (reflecting the lexis of the Lithuanian language from the 16th century to the late 20th century. The dictionary contains 310,659 headwords). E. KALBA also offers specialised dictionaries: the continuously updated *Database of Lithuanian Neologisms*[20] (featuring new words (loanwords and new coinages), word phrases and abbreviations or new meanings of words that came into Lithuanian from the late 20th century and are currently in public usage, as well as information on their origins, usage, and standardization. At present, it contains over 7,000 entries); *The Dictionary of Synonyms*,[21] *The Dictionary of Antonyms*,[22] *The*

[16] Available at: https://ekalba.lt.

[17] Available at: https://ekalba.lt/bendrines-lietuviu-kalbos-zodynas/.

[18] Available at: https://ekalba.lt/dabartines-lietuviu-kalbos-zodynas/.

[19] Available at: https://ekalba.lt/lietuviu-kalbos-zodynas/.

[20] Available at: https://ekalba.lt/naujazodziai/naujienos.

[21] Available at: https://ekalba.lt/sinonimu-zodynas/.

[22] Available at: https://ekalba.lt/antonimu-zodynas/.

Dictionary of Phraseology,[23] *The Dictionary of Comparisons*,[24] etc. We can also find various bilingual dictionaries,[25] e.g., Lithuanian-English and English-Lithuanian; Lithuanian-German and German-Lithuanian; Lithuanian-Latvian and Latvian-Lithuanian, etc.

Various card files belong to the category of lexicons,[26] e.g., the Main Card File, the Supplementary Card File and the Dialect Card File of *The Dictionary of the Lithuanian Language*, etc.

As for terminological resources, *The Term Bank of the Republic of Lithuania*,[27] the largest and most reliable source of Lithuanian terms contentwise, is most widely known. Legal acts and the glossaries of terms are the major sources of this term bank (over 255,000 term entries are available). The collections of terms from different areas are included in the bank (a total of 26 areas, such as politics, defence, finance, environment, transport, culture, health, etc., which are subdivided even further). *The Database of Terms of the Lithuanian Standards Board*[28] includes a collection of terms from standards. At present, it contains 76,000 term entries. The information system for Lithuanian language resources RAŠTIJA.LT[29] offers a search across 32 glossaries of terms. The areas of terms depend on the glossaries: there are terms from electrical engineering, computers, management, linguistic didactics, mathematics, metrology, meteorology, etc.

There are few ontologies and semantic networks in Lithuania. There is the *General Ontology of the Lithuanian Language*, the open-access ontology of Lithuanian medical terms *Snomed CT*,[30] the electronic service *E-terms*,[31] which includes ontologies in the following areas: *The Ontology of Human Anatomy Terms* (7,500 terms), *The Ontology of Economy Terms* (500 terms) and *The Ontology of Computer Hardware and Parts* (1,000 terms). We have several Lithuanian wordnets, which can be further expanded: *LitWorNet*[32] (for more,

[23] Available at: https://ekalba.lt/frazeologijos-zodynas/.

[24] Available at: https://ekalba.lt/palyginimu-zodynas/.

[25] Available at: https://ekalba.lt.

[26] "A digital dictionary (or lexicon) is a list of entries (usually single words or multiword expressions) optionally enriched with further information" (Pastor et. al. 2017: 29).

[27] Available at: http://terminai.vlkk.lt/.

[28] Available at: https://www.lsd.lt/index.php?-452282422.

[29] Available at: https://raštija.lt.

[30] Available at: https://www.snomed.lt/snomed-ct-pritaikomumas-uzsienio-salyse-elektroninis-sveikatos-irasas-ligonines-valdymo-irankis-palaikomas-snomed-ct/.

[31] Available at: https://ekalba.lt/esavokos/.

[32] Available at: http://mackus.vdu.lt/LitWordNet/.

see Vitkutė-Adžgauskienė et al. 2015) and *WordNet*[33] (with 24,000 synsets, of which 10,000 are linked with the synsets in the *Princeton WordNet*).

2.2. Tools/services

Text and data analytics account for the major part of digital tools/services (36%); they are followed by speech recognition (15%), grammar checking, spell checking (8%); machine translation (6%), and other. As for tools/services, it should be added that they are language dependent (98%), with text (80%) as the prevalent media type of input.

First of all, machine translation systems, which have already been developed, merit a mention. "Tilde informacinės technologijos" UAB offers an opportunity to use the multilingual machine translation tool based on neural networks, *Tilde Translator*,[34] free of charge. According to its developers, "this statistic machine translation tool is the best on the market and adjusted to Lithuanian-English, English-Lithuanian, Latvian-English and English-Latvian language pairs. The rule-based machine translation system for translating from Latvian into Russian is also available for use".[35] At present, you can use *Tilde Translator* to translate from Lithuanian into German, Polish, English, Ukrainian and Russian languages.

The ALPMAVIS machine translation system is freely available to users.[36] This system covers the following language pairs: Lithuanian-English-Lithuanian, Lithuanian-French-Lithuanian, Lithuanian-German-Lithuanian, Lithuanian-Polish-Lithuanian and Lithuanian-Russian-Lithuanian. The abovementioned system is adjusted to the general, legal and IT spheres.[37]

Free translation systems that were developed outside of Lithuania should be mentioned as well, e.g., *Google Translate*, *Microsoft Bing Translator*, *eTranslation*, etc. The latter translation system is available to governmental institutions for free and is better suited to the translation of administrative and legal texts.

There are various services created for users where speech recognition technology is designed to voice-control computers, e.g., *Browser*[38] (browsing

[33] Available at: https://ekalba.lt/zodziu-prasmiu-tinklas/?p=1.

[34] Available at: https://translate.tilde.com/#/.

[35] More information at: https://www.tilde.lt/kalbines-technologijos/tilde-vertykle.

[36] Available at: https://vertimas.vu.lt/Home/About.

[37] More information at: https://vertimas.vu.lt/Home/About.

[38] Available at: https://raštija.lt/liepa/paslaugos-vartotojams/narsytuvas/.

voice control); *Controller*[39] (computer voice control); *Searcher*[40] (voice search for UNESCO heritage resources); *Recognizer*[41] (learning voice control); *Helper*[42] (voice control for the disabled); *Coaching Robot Controller*[43] (control of a humanoid robot for children); *Taxi Caller*[44]; *Caller*[45] (calling to phone contacts); *Interlingual Communicator*[46] (Lithuanian-Chinese), etc. There is also a speech recognition application available for free,[47] which converts speech into text from a pre-recorded audio file or real-time dictation.

Some services developed in Lithuania are equipped with a speech synthesis technology, e.g., *Pronouncer*[48] (an audio dictionary of Lithuanian neologisms); *The Lithuanian Speech Synthesiser for the Blind*[49] (a SAPI5-compatible Lithuanian speech synthesizer which reads out loud what is displayed on a computer monitor); *The Mobile Synthesiser for the Blind*[50]; *The Online News Reader*[51] (the collection and reading of news in a synthesised voice in Lithuanian). A speech synthesis application[52] converting text into speech is also available for free.

The major free open-source tools for the basic analysis of digital texts in Lithuanian were created in the framework of various projects in Lithuania, namely a segmentor, a lemmatiser, a morphological analyser, a part of speech tagger, a syntactic parser, a spellchecker, a text normaliser, a solution for the advanced search for Lithuanian text indices, a multi-word expression extractor, etc. (Guidelines 2020).

There are also free open-source language identification and semantic analysis solutions: the simple and aspect-based sentiment (opinion) analyzer,[53] the hate-

[39] Available at: https://raštija.lt/liepa/paslaugos-vartotojams/valdytuvas/.

[40] Available at: https://raštija.lt/liepa/paslaugos-vartotojams/ieskotuvas/.

[41] Available at: https://raštija.lt/liepa/paslaugos-vartotojams/pazintuvas/.

[42] Available at: https://raštija.lt/liepa/paslaugos-vartotojams/pagalbininkas/.

[43] Available at: https://raštija.lt/liepa-2/paslaugos-vartotojams/ugdanciojo-roboto-valdytuvas/.

[44] Available at: https://raštija.lt/liepa-2/paslaugos-vartotojams/taksi-iskviestuvas/.

[45] Available at: https://raštija.lt/liepa-2/paslaugos-vartotojams/skambintuvas/.

[46] Available at: https://raštija.lt/liepa-2/paslaugos-vartotojams/tarpkalbinis-komunikatorius/.

[47] Available at: https://www.tilde.lt/snekos-technologijos.

[48] Available at: https://liepa.rastija.lt/Tartuvas/Naujienos?_ga=2.104838787.241622064.1643131617-1654665649.1643131617.

[49] Available at: https://raštija.lt/liepa/paslaugos-vartotojams/sintezatorius-akliesiems/.

[50] Available at: https://raštija.lt/liepa-2/paslaugos-vartotojams/mobilusis-sintezatorius-akliesiems/.

[51] Available at: https://raštija.lt/liepa-2/paslaugos-vartotojams/interneto-naujienu-skaitytuvas/.

[52] Available at: https://www.tilde.lt/snekos-technologijos.

[53] Available at: https://ekalba.lt/nuomoniu-analize/.

speech recognition tool,[54] the automatic document summary service,[55] the identified entity recognition tool (Guidelines 2020).

2.3. Corpora

Lithuania proceeds to create and develop the general language data and resources required to produce language technologies and their applications. There are several corpora in Lithuania which are dominated by annotated corpora (91%) with text (96%) as the media type of parts taking the lead.

The Corpus of the Contemporary Lithuanian Language[56] is the largest corpus of the Lithuanian language. It includes 140,921,288 words (fiction – 11.6%, non-fiction – 14.2%, administrative literature – 10%, publications – 63.8%, spoken language – 0.3%). The compilation of the corpus began in 1992, and was last updated in 2011.

There are also morphologically and syntactically annotated corpora. The volume of *Morphologically Annotated Lithuanian Corpus MATAS*[57] is 1.6 million words (36% of texts from periodicals, 24% of texts from scientific literature, 19% of texts from fiction, 2.8% of administrative texts, 6.8% of verbatim reports of the Parliament of the Republic of Lithuania). The corpus was produced in a semi-automatic manner and the outcomes were reviewed by linguists. The corpus revealed the large-scale morphological polysemy of the Lithuanian language, i.e., nearly half of all the forms are morphologically polysemous. The latest version of the *Lithuanian Treebank ALKSNIS* [58] (ALKSNIS 3.0) features 3,643 syntactically annotated sentences in the PML (Prague Mark-up Language) format.

Several parallel corpora were launched in Lithuania. *Parallel Corpus*[59] contains Czech-Lithuanian words (20.29%), English-Lithuanian words (76.6%), Lithuanian-Czech words (0.8%) and Lithuanian-English words (2.31%). *LILA parallel Corpus*[60] was created in a semi-automatic manner where texts are aligned at a paragraph and sentence level. The corpus consists of the texts published in 1991 or later. The total volume of the corpus is 8,782,050 words:

[54] Available at: http://hatespeech.vdu.lt.

[55] Available at: https://www.semantika.lt/Analysis/Summary.

[56] Available at: http://tekstynas.vdu.lt/tekstynas/.

[57] Available at: https://klc.vdu.lt/matas-morfologiskai-anotuotas-tekstynas/.

[58] Available at: https://klc.vdu.lt/alksnis-sintaksiskai-anotuotas-tekstynas/.

[59] Available at: https://klc.vdu.lt/lygiagretus-tekstynas/.

[60] Available at: https://klc.vdu.lt/lila-lygiagretusis-tekstynas/.

Lithuanian-Latvian texts occupy the major part (3,448,745 words), whereas Latvian-Lithuanian texts account for half the number (1,695,160 words). Such an asymmetric composition of data resulted from the higher number of texts translated from Lithuanian into Latvian than vice versa in recent years.

There are also other types of corpora, e.g., *The Corpus of the Spoken Lithuanian Language*,[61] *The compendium of textbook texts* KLASIUS[62]; *The Corpus of the Old Lithuanian Language*[63]; CorALit[64]: *Corpus Academicum Lithuanicum*, etc.

Open access to the data of most of the corpora is ensured.

In summary, the production and development of the existing digital language resources and tools/services evidenced a remarkable progress in Lithuania; there are even several infrastructures for Lithuanian language resources and tools/services (E. KALBA, RAŠTIJA.LT, etc.), giving a free access to different digital dictionaries, lexicons, etc. and providing various speech-controlled services, the basic analyses of the digital texts in Lithuanian, free open-code tools, etc. Various corpora and machine translation systems are produced and further developed. Nevertheless, there is a shortage of multimedia language data, parallel corpora geared to machinery translation, various speech corpora, etc. As for text semantics, there are still gaps, as qualitative wordnets, ontologies, etc. are lacking. The area of grammar is currently rather abandoned: there is a shortage of digital grammars and other technological tools that may contribute to the development of language technologies and their higher quality.

3. SITUATION OF LITHUANIAN IN THE DIGITAL ENVIRONMENT: STRENGTHS, WEAKNESSES, OPPORTUNITIES AND THREATS

The status of Lithuanian in the digital environment was evaluated taking into account the main available database of documents, as well as national infrastructures for language resources and technologies and main consortia, federations and projects. The main documents and resources are described in detail (see Guidelines 2020). This article aims to describe them in general terms in order to trace the origins of the strengths, weaknesses, opportunities, and threats.

[61] Available at: http://sakytinistekstynas.vdu.lt.

[62] Available at: https://raštija.lt/resurso-katalogas/klasius-v2/.

[63] Available at: http://coralit.lt/node/1.

[64] Available at: http://coralit.lt.

3.1. The Lithuanian government and other state institutions support several programs to promote a range of linguistic research and dissemination. The following are valid documents on language technology policies in Lithuania: 1) *The Guidelines for the Development of the Lithuanian Language in the Digital Environment and the Progress of Language Technologies for 2021–2027*[65] (for more information, see page 2); 2) *The Lithuanian Artificial Intelligence Strategy: A Vision for the Future* (2018)[66] issued by the Ministry of the Economy and Innovation of the Republic of Lithuania. The aim of the strategy is "for Lithuania to become a regional leader on the basis of the existing resources, experience and potential. It aims to increase Lithuania's competitiveness among the EU countries and to ensure its successful participation in the global AI ecosystem" (AI Strategy 2018); 3) *The strategy for Lithuania's advancement "Lietuva 2030"*.[67] This strategy outlines the "vision and development priorities of the state as well as the directions of their implementation until 2030. It is the main planning document that must be taken into account in strategic decisions and the preparation of state plans or programs" (Strategy 2012). On 2 July 2021, the European Commission gave its green light to the *Lithuania's recovery and resilience plan*. "The transformative impact of Lithuania's plan is the result of a strong combination of reforms and investments which address the specific challenges of Lithuania".[68]

3.2. Lithuania has several national technology and language data infrastructures: 1) RAŠTIJA LT[69] is the system of integrated Lithuanian language and writing resources, products and services developed by the Institute of Mathematics and Informatics of Vilnius University: knowledge base, search tools, etc.; 2) CLARIN-LT[70] – the Lithuanian National Consortium (member of CLARIN-ERIC) established in 2015. It currently consists of 5 research institutions: Vytautas Magnus University (coordinator), Kaunas University of Technology, Vilnius University, Mykolas Romeris University and the Baltic Institute of Advanced Technologies; 3) E. KALBA[71] – the Lithuanian language resources information system managed by the Institute of the Lithuanian Language. In this system one can find nine monolingual and ten bilingual dictionaries, various files and databases

[65] Available at: https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/911407f20ee911ebbedbd456d2fb030d.

[66] Available at: https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_LT(1).pdf.

[67] Available at: https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.425517.

[68] Available at: https://ec.europa.eu/info/business-economy-euro/recovery-coronavirus/recovery-and-resilience-facility/lithuanias-recovery-and-resilience-plan_en.

[69] Available at: https://raštija.lt.

[70] Available at: http://clarin-lt.lt.

[71] Available at: https://ekalba.lt.

(dialect archive; geoinformation database of Lithuanian place names, etc.), as well as electronic services (search in the Network of Meanings, E-Concepts, Opinion Analyzer, etc.); 4) SEMANTIKA[72] – the Lithuanian Syntactic and Semantic Analysis Information System (LSSAIS), which "is a unique language technology infrastructure and state information system providing speech recognition and text analysis services for the Lithuanian language. Vytautas Magnus University is the manager of the information system".[73]

3.3. Main consortia, federations and projects, which Lithuania belongs to or participates in. The representatives of *European Federation of National Institutions for Language* (EFNIL) in Lithuania are the Institute of the Lithuanian Language and the State Commission of the Lithuanian Language. The Federation pays particular attention to the languages of the EU Member States and the linguistic diversity of Europe.[74] *European Language Resource Coordination* (ELRC) "manages, maintains and coordinates the relevant language resources in all official languages of the EU and CEF associated countries. These activities will help to improve the quality, coverage and performance of automated translation solutions in the context of current and future CEF digital services".[75] *Common Language Resources and Technology Infrastructure* (CLARIN-ERIC) is "a research infrastructure that was initiated from the vision that all digital language resources and tools from all over Europe and beyond are accessible through an online environment for the support of researchers in the humanities and social sciences".[76] CLARIN-LT is a Lithuanian national consortium coordinated by Vytautas Magnus University, which has been a member of CLARIN-ERIC. *European Language Grid* (ELG) "develops and deploys a scalable cloud platform, providing, in an easy-to-integrate way, access to hundreds of commercial and non-commercial Language Technologies for all European languages, including running tools and services as well as data sets and resources".[77] Since 2019, ELG has been represented in Lithuania by the Institute of the Lithuanian Language. The primary goal of *European Language Equality* (ELE) is "to prepare the European Language Equality Programme, in the form of a strategic research, innovation and implementation agenda and a roadmap for achieving full digital

[72] Available at: https://www.semantika.lt.

[73] More information available at: https://www.semantika.lt/Help/About.

[74] More information available at: http://www.efnil.org.

[75] More information available at: https://www.lr-coordination.eu.

[76] More information available at: https://www.clarin.eu.

[77] More information available at: https://www.european-language-grid.eu.

language equality in Europe by 2030".[78] Since 2021, ELE has been represented in Lithuania by the Institute of the Lithuanian Language.

During the period of 2014–2020, the funding for the implementation of Lithuanian language solutions in the digital space was received from the Operational Programme for the EU Funds Investments, priority axis 2 (*Promoting the Information Society*). Five projects were launched in 2018: "Development of Lithuanian Speech-Controlled Services" (LIEPA-2), "Development of the Public Services of the Information System of Syntactic-Semantic Analysis of Lithuanian Texts" (SEMANTIKA-2), "Enhancement and Development of Machine Translation Systems and Localisation Services", "Development of the Information System of Integrated Lithuanian Language and Written Resources (RAŠTIJA 2), and "Development of the Information System of Lithuanian Language Resources" (E. KALBA). A total of 21 public e-services were created by the end of 2020. Research and educational establishments and business enterprises were the most active participants in the programme (Guidelines 2020). In addition to the above projects, other language technology projects funded by the EU structural funds, the European Commission, national project funds, etc. were implemented as well.

3.4. A SWOT (strengths, weaknesses, opportunities and threats) analysis is provided in the Programme of the Lithuanian Language in Information Society for 2007–2010 approved by the Seimas of the Republic of Lithuania.[79] It is an excellent starting point for the analysis of the current situation. It is gratifying that a remarkable progress has been made, and most of the weaknesses or threats of that period have long been eliminated and forgotten. Significant progress has been made in adapting the Lithuanian language to the digital environment: a number of digital language resources and basic language analysis tools have been developed, complex online language services have been created, an ontology of the Lithuanian language has been developed and localization of many computer programs and tools has been achieved. Computer applications relevant to the society have been localised, computer terms standardised. Lithuanian researchers actively participate and cooperate in the mobility activities of international associations. Many Lithuanian language specialists focus on the field of information technology and systematically develop innovative work in this area. Lithuania is very interested to have a full access to digital solutions for all citizens, which makes the adaptation of access to resources for persons with disability very important (Gaidienė, Tamulionienė 2022: 17–18).

[78] More information available at: https://european-language-equality.eu.

[79] Available at: https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.294883?jfwid=32wf9h0y.

The general analysis of the current status of European languages, the innovations coming up on a daily basis, the growing needs of society, etc. are posing new challenges and open new opportunities as well as new threats.

Lithuania still lacks language resources in the electronic environment for faster integration of the Lithuanian language and Information Technologies, as well as standards for managing the Lithuanian language with Information Technologies. There is a wide range of relevant software that has not yet been Lithuanianized and adapted to the needs of society; it is necessary to further ensure the uniform use of computer terms, vocabulary and phrases in software, and to continue to take care of the use of the spoken Lithuanian language (speech) in the electronic, computer, and computerized device environment.

At a national level, a clear legal framework is essential to ensure an equal treatment of research (non-commercial and commercial) innovation based on the automatic extraction and analysis of data from electronic unstructured information sources (texts). Only after a sufficient amount of relevant resources reflecting the phenomena of the current Lithuanian language has been gathered, will it be possible to develop effective tools for the application of the Lithuanian language in Information Technologies. Particular attention must be paid to adapting to the opportunities and needs of all consumers, so as not to program social exclusion, which will have its implications on the society as a whole. It is important to create Lithuanian interfaces that would directly reduce the social linguistic segregation, promote the legal use of software, and reduce the gap with the old EU member states in the use of Information Technology. It is also important to create conditions for the use of the Lithuanian language on computers, computer-controlled devices, computers controlled in the spoken Lithuanian language (speech), to improve the means of computer voice control so that disabled people and other persons should have an unrestricted access to electronic services. Computational linguistics and language technologies, as a separate subject, are not yet established in the Lithuanian tertiary education system. No university offers language technology studies at all levels. This needs to change (Gaidienė, Tamulionienė 2022).

In Lithuania, there is a need to increase the competence of specialists working in the field of language technologies and to raise the level of the society's ability to use the opportunities that language technologies have to offer. It is also important to train specialists who know the specifics of language and information technologies, to finance fundamental and applied research, to support scientific and technical infrastructures. Moreover, it is crucial to accumulate and increase the availability of open, reliable, high-quality, sustainable digital language resources and other digital language datasets. There is a need to develop the language technology infrastructure, the application of language technologies

in the public sector and public services, teaching and learning institutions, and to develop and improve publicly available IT solutions and tools. Lithuania needs to become even more actively involved in the European Community and other international language technology programs. It is important to update infrastructures with the necessary digital resources, to upgrade and maintain infrastructure hardware, to integrate infrastructures into larger national, European and international language resource systems, and to ensure the openness of technologies and data stored therein (Guidelines 2020).

To sum up, the development of language technologies is essential for Lithuanian as well as any other language. To achieve prosperity in this area, Lithuania should work on the developments that were and were not covered in this article. To achieve the goals, it is necessary to unite the efforts of the state, science and business.

CONCLUSIONS

After a review of the situation of Lithuanian language technologies in the multilingual European context and a systematic analysis of the Lithuanian language resources and tools/services, we may formulate several major conclusions concerning the status of the Lithuanian language in the digital environment.

1. The results show that a significant progress has been made in Lithuania since 2012 in developing various digital language resources and tools/services. Though Lithuanian is grouped as the language with a low number of speakers, it is progressing rapidly in the area of language technologies. As for digital resources and tools/services, there are still areas requiring further advances.

Though a number of Lithuanian language digital resources are already available, considering the demands of language technologies and of the public, new monolingual dictionaries (dictionaries of synonyms, antonyms, phraseology, etc.) and bilingual dictionaries as well as various lexicons still have to be developed or updated. Ontologies, wordnets, corpora have to be enlarged and expanded; multilingual parallel corpora required for machine translation need to be developed, etc. Concerning terminology, additional and updated compendia or terms are needed; the structure and technological solutions of the databases of terms vary, making it more difficult to utilize data for other technological solutions; there is also a shortage of open terminological data. Lithuanian is in need of digital grammars, annotated speech databases and other resources that would accelerate the progress of language technologies.

2. A remarkable progress has been achieved in language technologies in Lithuania, but the innovations showing up every minute, the growing needs, etc. not only open new opportunities and challenges but also show weaknesses and threats.

2.1. The infrastructure of open-access language resources is still under development in Lithuania; the culture of data sharing is still stalling; the questions of licencing are still pending: the regulations of intellectual property rights and the GDPR need to be more flexible and permitting a wider use of the data, which are subject to intellectual property rights, for the development of language technologies and resources so that authors' interests would not be infringed.

2.2. The required human resources are lacking in Lithuania: there is a shortage of IT specialists and researchers working in the field of language technologies; there are no specialized study programs.

2.3. The development of Lithuanian language technologies requires a strong national and international support, including relevant long-term language technology programs supporting research and business on equal terms. It is important to synchronize national and international activities, with due regard to research infrastructure and research priorities.

2.4. What would happen in Lithuania if the development of language technologies stopped altogether? Why is it repeatedly highlighted by the EU that language equality can be achieved through promoting the technological development of all EU languages? There is a broad consensus on the strength of language diversity and on the impact of native language technologies on the protection of linguistic and cultural diversity. Therefore, if no actions are taken in this area, the technological, linguistic divide would increase and the prestige of the Lithuanian language would reduce, Lithuania would no longer be competitive in the areas of artificial intelligence and other cutting-edge technologies; other foreign languages would gradually occupy the place of Lithuanian in the digital environment.

REFERENCES

AI strategy 2018: *The Lithuanian Artificial Intelligence Strategy: A vision for the future*. Available at: https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_LT(1).pdf.

Gaidienė Anželika, Tamulionienė Aurelija 2022: *Report on the Lithuanian Language*. Available at: https://european-language-equality.eu/wp-content/uploads/2022/03/ELE___Deliverable_D1_23__Language_Report_Lithuanian_.pdf.

Giagkou Maria, Piperidis Stelios 2021: *European Language Equality. Guidelines for T1.3 contributors*. Internal to ELE network.

Guidelines 2020: *The guidelines for the development of the Lithuanian language in the digital environment and the progress of language technologies for 2021–2027*. Available at: https://www.e-tar.lt/portal/lt/legalAct/71152ab00eee11ebb74de75171d26d52.

Jaroslavienė Jurgita, Miliūnaitė Rita 2020: Beribis lietuvių kalbos pasaulis skaitmeninių išteklių sistemoje „E. kalba". – *Pasaulio lietuvis*. Available at: https://pasauliolietuvis.lt/beribis-lietuviu-kalbos-pasaulis-skaitmeniniu-istekliu-sistemoje-e-kalba/.

Pastor Rafael Rivera 2017: European Parliament, Directorate-General for Parliamentary Research Services, *Language equality in the digital age: towards a human language project*, European Parliament. Available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2017/598621/EPRS_STU(2017)598621_LT.pdf.

Rehm Georg, Uszkoreit Hans, editors, 2013: *The META-NET Strategic Research Agenda for Multilingual Europe 2020*, Springer, Heidelberg, New York, Dordrecht, London. Available at: https://lirias.kuleuven.be/retrieve/501271.

Strategy 2012: *The Strategy for Lithuania's Advancement "Lietuva 2030"*. Available at: https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.412512.

Vaišnienė Daiva, Zabarskaitė Jolanta 2012: *The Lithuanian language in the digital age*, vol. 385, Springer. Available at: http://www.meta-net.eu/whitepapers/volumes/e-book/lithuanian.pdf.

Vitkutė-Adžgauskienė Daiva, Dainauskas Justinas Juozas, Amilevičius Darius, Utka Andrius 2015: Lietuvių kalbos žodžių tinklas – *LitWordNet*. – *Darbai ir dienos* 64, 101–114. Available at: https://cris6.vdu.lt/cris/bitstream/20.500.12259/27751/1/ISSN2335-8769_2015_N_64.PG_101-114.pdf.

Europos kalbų lygybė skaitmeniniame amžiuje: Lietuvos atvejis

SANTRAUKA

Straipsnyje rašoma apie lietuvių kalbos technologijų būklę, supažindinama su Europos kalbų lygybės skaitmeninėje terpėje situacija. Nagrinėjami kiekybiniai ir kokybiniai kalbų lygybę atskleidžiantys rodikliai Europos Sąjungos kontekste, atsižvelgiant į kalbėtojų,

skaitmeninių kalbos išteklių ir technologijų skaičių bei joms teikiamą paramą, ypatingą dėmesį skiriant Lietuvos atvejo analizei. Šiuo straipsniu siekiama pabrėžti iki šiol kalbų technologijų srityje atliktą darbą ir išryškinti spragas bei atskleisti išbandymus, su kuriais susiduria ir juos sprendžia oficiali nacionalinė ir Europos Sąjungos kalba – lietuvių kalba. Straipsnyje pateikiama naujausia lietuvių kalbos technologijų padėties apžvalga analizuojant skaitmeninius kalbos išteklius ir įrankius / paslaugas.

Gauti rezultatai rodo, kad nuo 2012 m. Lietuvoje padaryta didžiulė pažanga plėtojant įvairius skaitmeninius kalbos išteklius ir įrankius / paslaugas. Nors lietuvių kalba yra priskiriama prie mažai kalbėtojų turinčių kalbų, ji gana sparčiai tobulėja kalbos technologijų srityje. Nepaisant to, kad esama sukurta nemažai lietuvių kalbos skaitmeninių išteklių, atsižvelgiant į kalbos technologijų ir visuomenės poreikius, būtina kurti ir atnaujinti vienakalbius (sinonimų, antonimų, frazeologijos ir pan.) ir dvikalbius žodynus, įvairius leksikonus. Būtina gausinti ir plėsti įvairiais kalbos duomenimis ontologijas, žodžių tinklus, tekstynus, kurti mašininiam vertimui reikalingus keliakalbius lygiagrečiuosius tekstynus ir kt. Kalbant apie terminiją, trūksta daugiau ir naujesnių terminų rinkinių, terminų bazių struktūra ir technologiniai sprendimai skiriasi, o tai apsunkina galimybes panaudoti duomenis kitiems technologiniams sprendiniams, taip pat trūksta atviրųjų terminologinių duomenų. Lietuvių kalbai labai trūksta skaitmeninių gramatikų, anotuotų garsynų ir kitų išteklių, kurie prisidėtų prie spartesnės kalbos technologijų pažangos.

Lietuvoje vis dar tvarkoma atviրųjų prieigų kalbos išteklių infrastruktūra, ne visai išspręsti licencijavimo klausimai – intelektinės nuosavybės teisių ir BDAR reglamentai, kurie turi būti lankstesni ir leidžiantys plačiau naudoti intelektinės nuosavybės teisėmis apsaugotus duomenis kalbos technologijų plėtrai ir išteklius taip, kad nebūtų pažeisti autorių interesai. Lietuvoje trūksta būtinųjų žmogiškųjų išteklių: trūksta kalbų technologijų IT specialistų, taip pat šios srities mokslininkų, nėra specializuotų studijų programų. Lietuvių kalbos technologijų plėtrai reikalinga stipri nacionalinė ir tarptautinė parama, įskaitant tam skirtas ilgalaikes kalbų technologijų programas, kurios vienodai remia tiek mokslinių tyrimų, tiek verslo veiklą. Svarbu sinchronizuoti nacionalinę ir tarptautinę veiklą, ypač mokslinių tyrimų infrastruktūros ir mokslinių tyrimų prioritetų atžvilgiu.

Įteikta 2022 m. gegužės 18 d.

ANŽELIKA GAIDIENĖ
AURELIJA TAMULIONIENĖ
Lietuvių kalbos institutas
Petro Vileišio g. 5, LT-10308 Vilnius, Lietuva
anzelika.gaidiene@lki.lt
aurelija.tamulioniene@lki.lt