

**SABINE KIRCHMEIER**

European Federation of National Institutions for Language (EFNIL)

## TRENDS IN EUROPEAN LANGUAGE POLICIES WITH A VIEW TO LANGUAGE TECHNOLOGY

### ABSTRACT

This article gives an overview of current trends in European language policies regarding the status of the official EU-languages and with a view to language technology. The article is based on data collected for the European Language Monitor (ELM) provided by EFNIL – the European Federation of National Institutions for Language – and on recent experiences with the development of the national Danish strategies for artificial intelligence and language technology. The aim is to describe the wealth of information about European language policies that is available through EFNIL and ELM, and to inspire researchers, students and policy makers to further analyse the data and use them for future policy development.

**KEYWORDS:** language policy, language planning, language technology, European languages.

### ANOTACIJA

Šiame straipsnyje apžvelgiamos dabartinės Europos kalbų politikos tendencijos, kreipiant dėmesį į oficialių Europos Sąjungos kalbų statusą ir į kalbų technologijas. Straipsnis pagrįstas Europos nacionalinių kalbos institucijų federacijos (*European Federation of National Institutions for Languages*, EFNIL) vykdomos Europos kalbų stebėsenos (*European Language Monitor*, ELM) duomenimis ir naujausia Danijos nacionalinės dirbtinio intelekto ir kalbos technologijų strategijos kūrimo patirtimi.

Straipsnio tikslas yra įvairiais aspektais išnagrinėti apibendrinamąją informaciją apie Europos kalbų politiką (remiantis ELM duomenimis, gautais iš EFNIL) ir įkvėpti tyrėjus, studentus ir politikos formuotojus toliau analizuoti šiuos duomenis bei pasinaudoti jais formuojant kalbos politiką ateityje.

**ESMINIAI ŽODŽIAI:** kalbos politika, kalbos planavimas, kalbos technologijos, Europos kalbos.

## INTRODUCTION

This article gives an overview of current trends in European language policies – a bird’s eye view on the European language landscape regarding the status of the official EU-languages and with a view to language technology. It is based on data in the European Language Monitor (ELM) that is developed and regularly maintained by the European Federation of National Institutions for Language (EFNIL), and on recent developments in Denmark regarding policies for language technology. Language technology is chosen as the focus point, since this is an area where we currently see the most interesting developments in national policies. The aim of the article is to describe the wealth of information about European language policies that is available through EFNIL and ELM, and to inspire researchers, students and policy makers to further analyse the data and use them for future policy development.

In the first sections, I will shortly describe EFNIL and its approach to language planning. Next, I will give an overview of the European Language Monitor (ELM) and give detailed examples from the sections about language legislation, about the use of English in higher education and about language technology. Finally, I will shortly describe the development regarding language technology in Denmark to illustrate important aspects of the policy development process. In the concluding chapters, I will shortly touch upon the recent developments for language technology in the European Commission and UNESCO and highlight EFNIL’s recommendations for policies on language technology.

### 1. THE EUROPEAN FEDERATION OF NATIONAL INSTITUTIONS FOR LANGUAGE (EFNIL)

EFNIL is a federation of official institutions for language in Europe both inside and outside the EU. Currently, there are 40 institutions and 29 countries represented. EFNIL is financed by an annual membership fee that is used to finance a secretariat, to arrange an annual international conference and to run data collection projects on language planning and language policy. EFNIL is pursuing the following goals:

---

- The collection and exchange of information about the officially recognized standard languages of the EU and other official European languages.
- The provision of expert advice about language policy in the EU.
- The preservation of linguistic diversity within Europe.
- The promotion of plurilingualism amongst the citizens of the EU states.

The federation also follows the development in the minority and regional languages of Europe. EFNIL works to achieve its goals through the following activities:

- Description and analysis of the current linguistic situation within the EU as well as implications for longer-term linguistic development. Currently there are 3 projects:
  - LLE: Language Legislation in Europe (a frequently updated overview of the language regime in each country).
  - ELM: European Language Monitor (a data collection about language policies and practices across Europe).
  - ELIPS: European Languages and their Intelligibility in the Public Space (a data collection about plain language strategies and practices across Europe).
- Scientifically based analysis of cross-state language problems and questions of language policy in annual conferences and publications.
- Consultation services in the field of language policy for political decision makers of the EU institutions and member states.

Propagation of the cultural and practical benefits of European linguistic diversity and plurilingualism through relevant actions and publications<sup>1</sup>.

EFNIL's approach to language planning. EFNIL's projects aim at informing decision makers and researchers about language planning activities in Europe. Language planning is generally defined as the development of policies or programs designed to direct or change language use. It typically consists of status planning, acquisition planning and corpus planning.

- **Status planning** aims at the regulation of the position of a language vs. other languages: Which languages are considered the official and non-official languages in

---

<sup>1</sup> See more on [www.efnil.org](http://www.efnil.org).

a country, in which social contexts are they spoken, and what is the legal situation of their users. Status planning also covers prestige planning.

- **Acquisition planning** aims at the regulation of which languages are taught in various educational contexts, and how these languages are taught at different educational levels.
- **Corpus planning** regulates the language as such: standardization of vocabulary, orthography, writing rules, terminology etc.

Both status and acquisition planning affect the language as such – the corpus, and so the knowledge of the language planning situation is not only relevant for decision makers but also important background knowledge for linguistic observations.

## 2. THE EUROPEAN LANGUAGE MONITOR (ELM)

ELM is a project that provides information about language policy and practice in the participating countries and is updated every 4 years. It is a scientific review of the language situation in European countries which provides exact references to the actual legislation in each country and useful background knowledge about the status of the languages of Europe.

ELM has asked approximately the same questions with regular intervals for the last 10 years and thus allows researchers to track developments within the language policies in the participating countries. The questions in ELM cover the following topics:

- Country situation. Official, regional, indigenous, immigrant languages spoken within and outside the country, legal status, accordance with conventions.
- Legal situation. Language law, constitutional status, other regulations, language demands for citizenship.
- Primary and secondary education. Languages of instruction, languages offered.
- Tertiary education. Languages of instruction, languages used in publications and dissertations.
- Media. Papers, TV, film, music. Languages used and translations provided.
- Business. Regulations. Company languages, annual reports, websites.
- Dissemination of languages. Official languages taught abroad.

- Language organizations. Official, non-governmental but publicly funded, private.
- Language technology.

ELM version	ELM 1 2004	ELM 2 2012	ELM 3 2014	ELM 4 2019
Collection period	2003-2004	2008-2009	2012-2013	2017-2018
Availability of results	Not available	Articles	Online	Online + articles
Participating countries		Austria	Austria	Austria
		Belgium (FL/FR)	Belgium	Belgium (FL)
			Bulgaria	Bulgaria
		Cyprus	Cyprus	
		Czech Republic	Czech Republic	Czech Republic
		Denmark	Denmark	Denmark
		Estonia	Estonia	Estonia
		Finland	Finland	Finland
	France		France	
	Germany	Germany	Germany	Germany
		Greece	Greece	Greece
		Hungary	Hungary	Hungary
		Iceland	Iceland	Iceland
		Ireland	Ireland	
	Italy	Italy	Italy	
		Latvia	Latvia	Latvia
		Lithuania	Lithuania	Lithuania
		Luxembourg		Luxembourg
			Malta	
		Norway	Norway	Norway
		Poland	Poland	
			Portugal	Portugal
			Romania	
		Slovakia	Slovak republic	Slovak Republic
		Slovenia	Slovenia	Slovenia
	Sweden	Sweden	Sweden	Sweden
	The Netherlands	The Netherlands	The Netherlands	The Netherlands
		United Kingdom	United Kingdom	United Kingdom
Total	5	23	27	21

TABLE 1. Overview of countries participating in ELM 1-4

The collection of data for ELM 1, the first pilot project, started in 2004. Collection of data for ELM 2 started in 2008, ELM 3 2012 and ELM4 in 2017.

Since language institutions volunteer to fill in the questionnaire, which is both long and complicated, it has not always been possible to get the information from all countries. The list of participating countries is shown in Table 1. In 2008, 23 countries participated. In 2012, it was 27, and in 2017 the number of participants was 21. 17 countries have participated in all 3 rounds.

### 3. EXAMPLES OF INFORMATION IN ELM

In the following, I will present three examples to illustrate how the data in ELM can be used. I will describe the data about language legislation, the data about the use of English in higher education and the data about language technology.

#### 3.1. Language legislation

The first example deals with language legislation. The participants in each country were asked:

- Are there provisions about language in the constitution?
- Is there a language law? (a language law is a framework law which must be observed in other legislation).
- Are provisions about language made in other legislation?

The three questions may be considered a hierarchy with the constitution at the top level as the most general type of legislation, and provisions about language in other rulings as the most specific one. Figure 1 shows that in ELM 4 (2017) a little more than half of the participating countries (11 out of 21) indicated that they do have provision for language in their constitution, whereas a little less than half (9 out of 21) do not have any provisions at the constitutional level. Language provisions via language laws and other legislation is a bit more widespread: 14 out of 21 countries report that they have language provisions at these legislative levels.

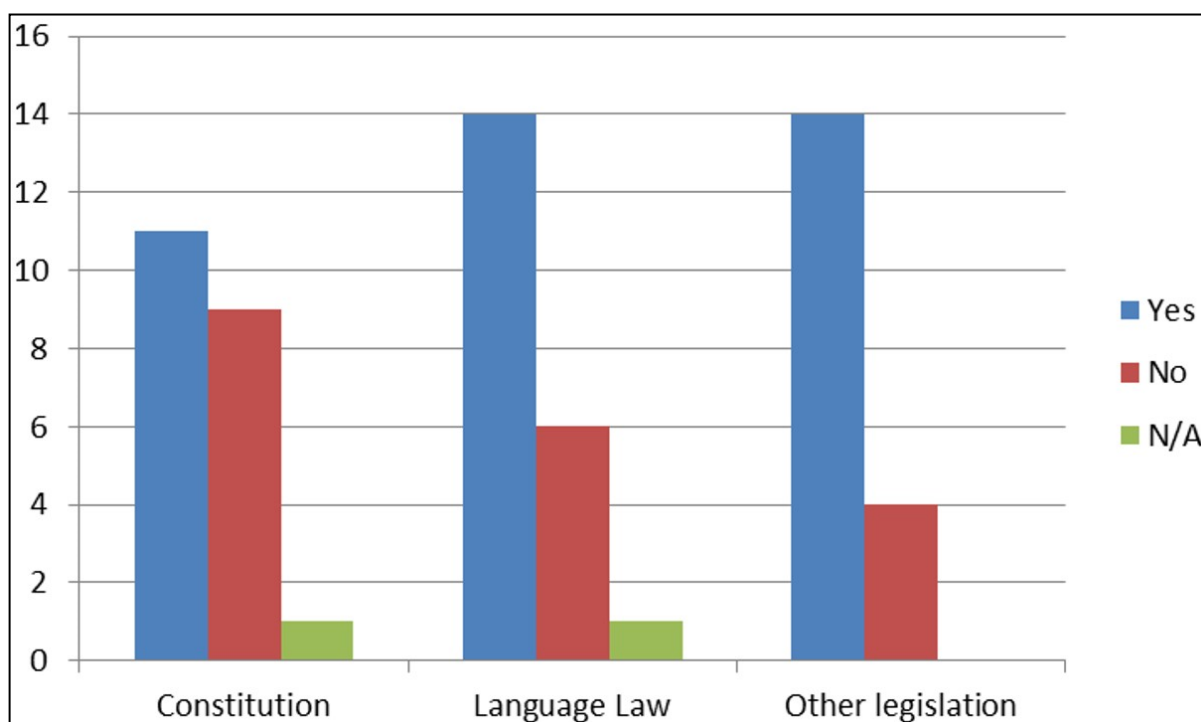


FIG. 1. Language Provisions in ELM 4 (2017)

ELM also provides detailed information about which countries have which types of regulations (See Table 2).

In Table 2, countries are ordered according to the degree of language regulation in their legislation. In the top, we find the least regulated ones, and, in the bottom, those that have regulations on all 3 levels.

Only 3 countries (Denmark, Germany, and The Netherlands) have no regulations at the two highest levels: the constitution and the language law, whereas 8 countries have regulations at all three levels. The reasons differ from country to country. For instance, the constitution for Denmark is from 1849 and has only had a few minor revisions throughout the years. It was made at a time where the country was forming a nation state which was considered quite homogenous and monolingual. So, nobody thought about making special provisions in the constitution about the Danish language. In Finland, the roles of the Swedish and Finnish languages and the rights of the Swedish minority have been issues for many

years. There was a language law already in 1922, and language provisions became part of the constitution in 1999.

	<b>Rulings for official languages in:</b>		
<b>Country</b>	<b>2.1 Constitution</b>	<b>2.2. Language law</b>	<b>2.3. Other legislation</b>
Denmark	No	No	Yes
Germany	No	No	Yes
The Netherlands	No	No	Yes
Grand Duchy of Luxembourg	No	Yes	No
Greece	No	Yes	No
Czech Republic	No	Yes	Yes
Iceland	No	Yes	Yes
Norway	No	Yes	Yes
Sweden	No	Yes	Yes
UK	Not appl.	Not appl.	Yes
Portugal	Yes	No	No
Bulgaria	Yes	No	Yes
Hungary	Yes	No	Yes
Belgium	Yes	Yes	No
Austria	Yes	Yes	Yes
Estonia	Yes	Yes	Yes
Finland	Yes	Yes	Yes
Latvia	Yes	Yes	Yes
Lithuania	Yes	Yes	Yes
Slovak Republic	Yes	Yes	Yes
Slovenia	Yes	Yes	Yes

TABLE 2. Overview of legal provisions in participating countries in ELM 4 (2017)



We can see that there is no country that does not have any provision at any level. Typically, if there are no provisions at the highest level, there will be some at a lower one. The UK does not have a single constitutional code, that is why the questions about the first two types of legislation are not applicable.

### 3.2. What are language provisions about?

Many constitutions state what the official language or languages of the country are. Some cover the whole country; others restrict the legislation to cover only state institutions and schools. The provisions mainly pinpoint the rights of citizens that speak the language – such as the French laws that state that French citizens have the right to speak in their own language in every context. This is for instance the case at scientific conferences, which may not be held exclusively in English.

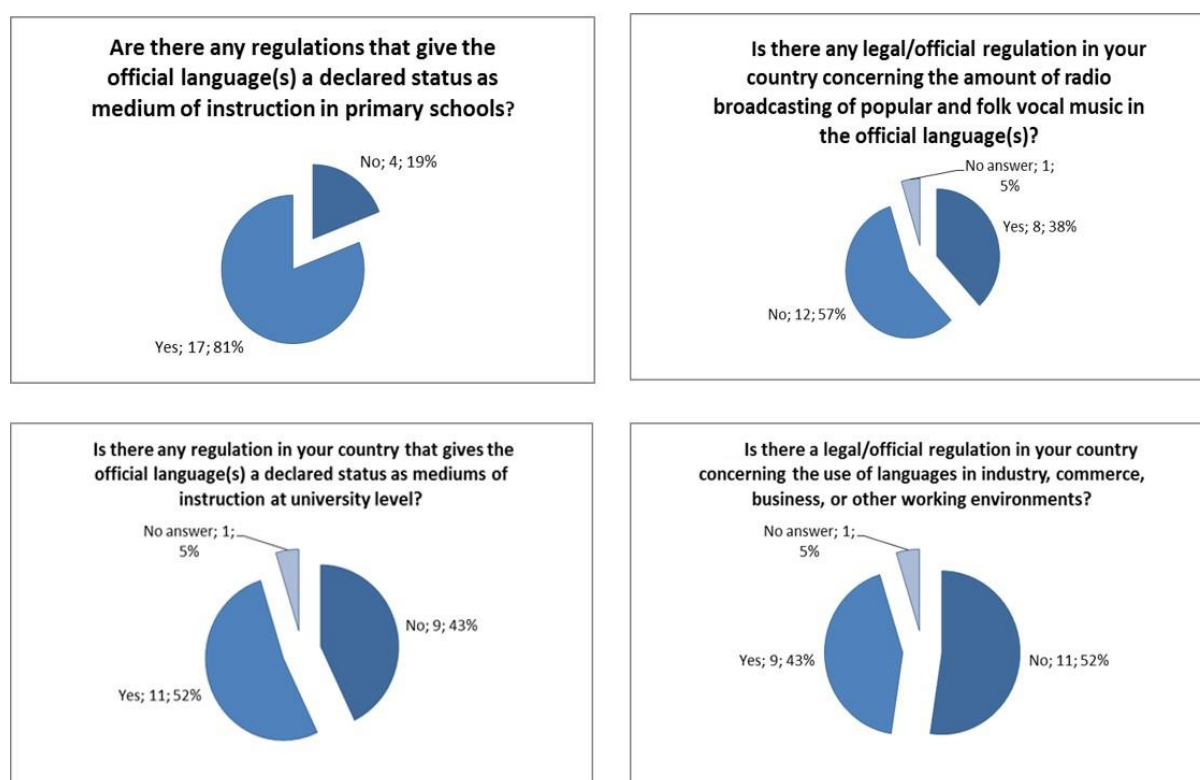


FIG. 2. Legislative measures for language use in schools, universities, mass media and business in countries participating in ELM 4 (2017)

Many provisions secure linguistic minority rights. Some establish consumer rights (e.g. they demand the translation of foreign documents such as content declarations and user manuals). Some regulate the language used in public media, and some establish the role of official institutions that monitor the language and standardize certain aspects of it, for instance orthography and terminology.

Some regulations are more frequently used than others. Figure 2 shows that from the countries that participated in the ELM 4 survey, only 4 report that there are no regulations that give the official languages a declared status in primary schools. Much fewer have these kinds of regulations when it comes to universities and the language in mass media. And when it comes to the business sector, most countries do not seem to have any provisions regarding language.

### **3.3. English as medium of instruction in higher education**

The second example of the information that is accessible in ELM, is data about English as medium of instruction in higher education. All countries, except Slovakia, report that English is used as medium of instruction at the bachelor level.

But how many courses are actually taught in English? Is it up to 25 %, up to 50 %, up to 75 % or even more than 75 %? And is there a difference between different scientific fields? An overview of the information that can be found in ELM 4, is given in Table 3.

Table 3 shows the list of countries in the left column and next, the total percentage for all scientific fields, and then separate figures for the humanities, social science, and natural science. The light green colour indicates the between 0 and 25 % of the courses are taught in English. A little darker it goes up to 50 % then 75 % and the darkest colour indicates more than 75 %. So, the darker it gets, the more English is taught. The darkest colours are, of course, found in all fields in the UK, but also in social science and natural science in Luxembourg.

Country	Total percentage	Humanities	Social Science	Natural science
Austria	No answer	No answer	No answer	No answer
Bulgaria	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Czech Republic	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Estonia	Between 26-50%	Between 26-50%	Between 26-50%	Between 26-50%
Finland	Between 26-50%	Between 0-25%	Between 0-25%	Between 51-75%
Germany	Between 0-25%	Between 0-25%	Between 0-25%	Between 26-50%
Greece	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Iceland	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Latvia	Between 26-50%	Between 0-25%	Between 0-25%	Between 26-50%
Lithuania	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Luxembourg	No answer	Between 26-50%	More than 75%	More than 75%
Slovak Republic	No answer	No answer	No answer	No answer
Slovenia	Between 26-50%	Between 26-50%	Between 0-25%	Between 26-50%
UK	More than 75%	More than 75%	More than 75%	More than 75%
Sweden	Between 51-75%	Between 0-25%	Between 26-50%	Between 0-25%
Norway	No answer	No answer	No answer	No answer
Netherlands	Between 51-75%	Between 51-75%	Between 51-75%	Between 51-75%
Denmark	Between 26-50%	Between 0-25%	Between 26-50%	More than 75%
Belgium	No answer	No answer	No answer	No answer
Portugal	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Hungary	Between 0-25%	Between 0-25%	Between 0-25%	Between 26-50%

TABLE 3. English as medium of instruction at the BA level in countries participating in ELM 4 (2017)

It is easy to see, that the percentage of courses taught in English is clearly highest in natural science, whereas there is less in humanities and social science. In many countries, the figures on the actual use of English at universities are not easily available because the use of language of instruction is not registered systematically. Some of the figures, therefore, are estimates.

If we look at the use of English as medium of instruction at the master's level we see a considerable increase in the use of English, especially in natural science, but for instance in the case of the Netherlands also in the other fields. Besides Luxembourg, Denmark has more than 75 % of university courses at the master level in natural science taught in English.

Country	Total percentage	Humanities	Social Science	Natural science
Austria	No answer	No answer	No answer	No answer
Bulgaria	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Czech Republic	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Estonia	Between 26-50%	Between 26-50%	Between 26-50%	Between 26-50%
Finland	Between 26-50%	Between 0-25%	Between 0-25%	Between 51-75%
Germany	Between 0-25%	Between 0-25%	Between 0-25%	Between 26-50%
Greece	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Iceland	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Latvia	Between 26-50%	Between 0-25%	Between 0-25%	Between 26-50%
Lithuania	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Luxembourg	No answer	Between 26-50%	More than 75%	More than 75%
Slovak Republic	No answer	No answer	No answer	No answer
Slovenia	Between 26-50%	Between 26-50%	Between 0-25%	Between 26-50%
UK	More than 75%	More than 75%	More than 75%	More than 75%
Sweden	Between 51-75%	Between 0-25%	Between 26-50%	Between 0-25%
Norway	No answer	No answer	No answer	No answer
Netherlands	Between 51-75%	Between 51-75%	Between 51-75%	Between 51-75%
Denmark	Between 26-50%	Between 0-25%	Between 26-50%	More than 75%
Belgium	No answer	No answer	No answer	No answer
Portugal	Between 0-25%	Between 0-25%	Between 0-25%	Between 0-25%
Hungary	Between 0-25%	Between 0-25%	Between 0-25%	Between 26-50%

TABLE 4. English as medium of instruction as the MA level in countries participating in ELM 4 (2017)

In many countries there are worries whether this will affect the lexical inventory, the expressivity and especially the terminology of their languages in the scientific discourse. ELM provides a good point of departure for setting up comparative research projects to gain more insights in these developments.

### 3.4. Language technology

Already in 2002, David Crystal in his book *Language Death* (Crystal 2002) mentions language technology as one of the important means to preserve the vitality of a language. Since then, there have been several reports, for instance the Meta-NET White Paper Series (Rehm and Uszkoreit 2012), that are warning about languages becoming digitally extinct.

Many languages with a declining number of speakers are looking towards language technology hoping to preserve or even revive their language. For instance, huge efforts have

been made for creating speech recognition and speech synthesis, spell checkers, grammar checkers, machine translation software for Icelandic, Greenlandic and the Sami languages.

Artificial intelligence and robots are currently entering the market, allowing everyone to control their home environment, to access information and to interact with technology through spoken dialogue. However, speech interfaces must be available in the language of the citizen for the devices to be used properly and safely. The more advanced the dialogue becomes, the better knowledge of the culturally determined features of the local language is necessary to enable the robot to understand the citizens' requests and to make itself understood.

Communities that can put this new technology into play will have a strong advantage compared to others, as they will be better equipped to manage the increasing demands of the private and public sector in the coming years. However, if these types of technologies are available only in a few widely spoken languages, they will only benefit those that have the privilege of speaking these languages.

In 2012, the Meta-Net project (Rehm, Uszkoreit 2012) collected information on how European languages were covered by language technology. It could clearly show that large language communities, such as primarily the English-speaking countries, but also the French- and Spanish-speaking countries, were considerably more advanced than most of their smaller European neighbours. This is called the language bias.

The consequences of the language bias have been described as follows:

"Language technology has an impact on social justice, i.e. equal opportunities for individuals and groups (such as minorities) within society to access resources, get their voice heard, and be represented in society" (Hovy, Spruit 2016).

The META-NET data show, that this not only holds for individuals and minorities within societies, but also for countries and languages. The reason lies in the way artificial intelligence and language technology are produced. Most language technology products are trained on existing data sets. If data sets are not available for a given language, the product will not be available for the users of that language. If a product is not trained on relevant data set for a given language or domain, the product will not be able to reflect the linguistic and cultural features of that language or domain. Political attention is important for all lesser resourced languages if they want to overcome the language bias and keep their languages alive in the digital age.

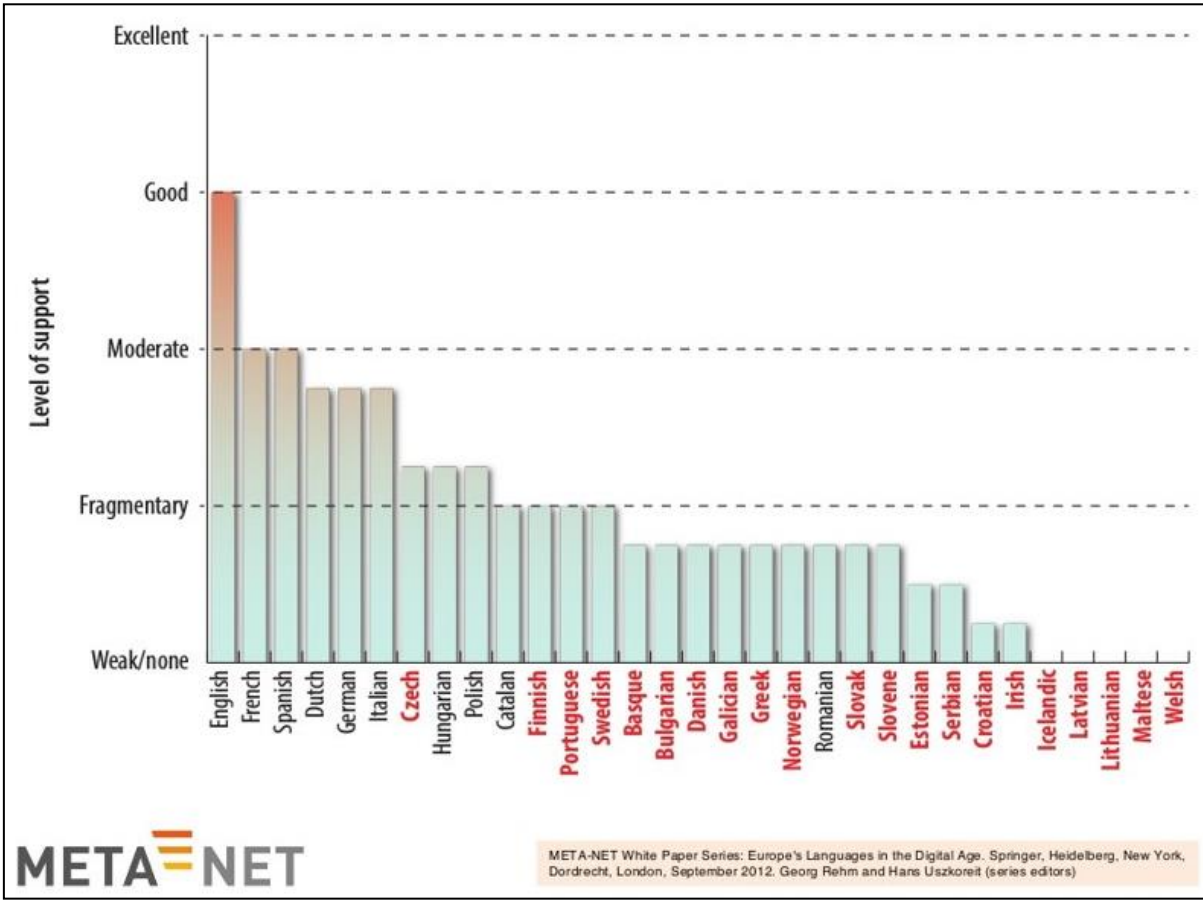


FIG. 3. Overview of language technology coverage for European Languages by META-NET (2012)

The META-NET white papers also show that political attention does make a difference. In Figure 3, we can see that the Dutch language is on the same level as German and Italian which have a considerably larger number of speakers. This is because the Dutch government already in 2004 launched a language strategy and a dedicated funding programme for language technology for Dutch called STEVIN (Spyns 2008) which systematically saw to that the relevant language resources were developed in close cooperation between industry and research institutions. Similarly, the strategic initiatives by governments, for instance in the Baltic states during the last years, show an enormous progress for language technology for the Baltic languages.



However, not all countries have focus on the importance of language technology for their languages. This, too, we can see in the European Language Monitor. If we look at the answers to the questions on language technology, we can see that only about half of the countries report that they have a dedicated funding programme for their languages.

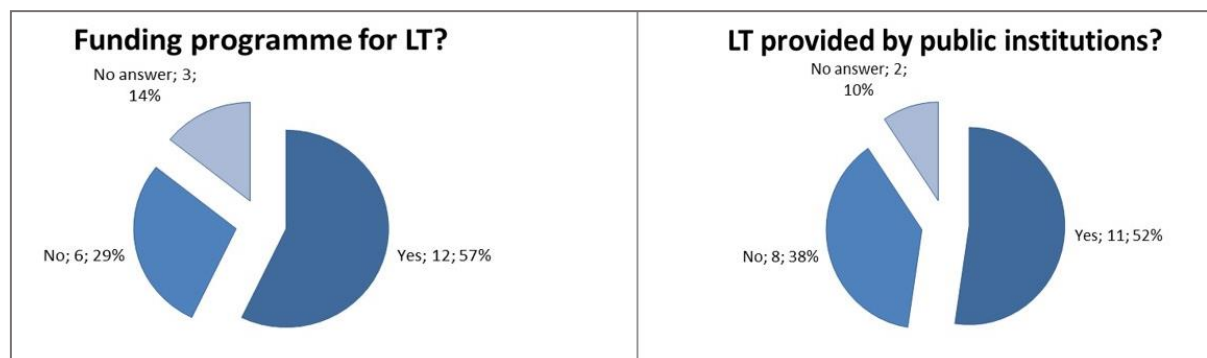


FIG. 4. Answers regarding language technology in ELM 4 (2017)

Furthermore, only about half of the countries report that language technology is provided for the citizens by public institutions. These figures show that getting political focus on language issues is far from straightforward in all European countries. However, right now, the emergence of artificial intelligence has strengthened the interest in language technology, and this may cause more countries to put this important topic on the agenda in the future. I will shortly illustrate this by describing the recent development in Denmark.

#### 4. LANGUAGE TECHNOLOGY AND AI IN DENMARK

In Denmark, there has been research and education on language technology for the Danish Language for many years – since the beginning of the 80's. Although many proposals were made by the research community, no coordinated efforts were made to strengthen language technology. By the end of the 90's, it seemed that interest in this field was even declining.

Consequently, there were only few initiatives and research projects in the beginning of the century. Official reports on the state of the Danish Language all contained recommendations for language technology, but they were practically ignored politically. Even

the META-NET series (Rehm, Uszkoreit 2012), that clearly documented that Denmark was lagging seriously behind, did not shake up the political sphere except for a few parties in the opposition.

In 2015 and 2016 the Danish Language Council published a series of articles about language technology, and this led to a proposal in the Danish parliament to establish a language technology committee (DSN LT articles). Since the proposal came from the opposition, it was decided to have a hearing in parliament first. This hearing took place about a year later, in the beginning of 2017. And finally, in the beginning of 2018, a year after the hearing, the committee could be established. In total, the process took almost 3 years.

But then, things went extremely fast. The language committee was supposed to present a report in the middle of 2019. It took care to involve all relevant actors in the field. In fact, more than 200 institutions and private enterprises participated in workshops and contributed to the final report (Kirchmeier *et al.* 2019a; 2019b).

However, only about half a year after the committee started, the focus on artificial intelligence accelerated, and already in October 2018, language technology became part of the national strategy for digital public service, which had a whole chapter dedicated to language technology for Danish. In March 2019, a month before the report of the committee was published, the ministry of finance published the national strategy for artificial intelligence and allocated 4 million Euro to Danish language resources (Denmark's National Strategy for Artificial Intelligence 2019). This was possible because there was a strong coordination and communication between the ministry and the committee on the recommendations during the whole process.

The recommendations by the language technology committee were as follows:

1. Establishment of a central organisation to plan and manage a National Danish language resource bank.
  2. The language bank should contain high quality resources of the following kind:
    - 2.1. A time annotated Danish speech technology corpus.
    - 2.2. Danish text corpora and annotated gold standards for machine learning.
    - 2.3. A comprehensive lexical database.
    - 2.4. A Danish terminology bank.
    - 2.5. A language technology tool-kit.
-



- 2.6. A language portal for the distribution of resources.
3. More education of training of experts in language technology for Danish.
4. More research on language technology for Danish.

These recommendations are currently implemented by the Danish Agency for Digitization, and more and more language resources for Danish are made available at their website [sprogteknologi.dk](http://sprogteknologi.dk).

One of the ideas put forward by the committee, was to combine all information about a given lexeme in one comprehensive lexical database.

The database should contain traditional lexical resources such as various dictionaries and databases on general language, orthographic standards, neologisms, and concepts, in combination with semantic nets, terminology databases and speech databases. It is not an easy task to combine these resources, but for an artificial intelligence program, access to all kind of linguistic information is extremely useful. Experiments to connect lexical information monolingually and multilingually involving several European languages are currently taking place in context of the ELEXIS project (ELEXIS 2020).

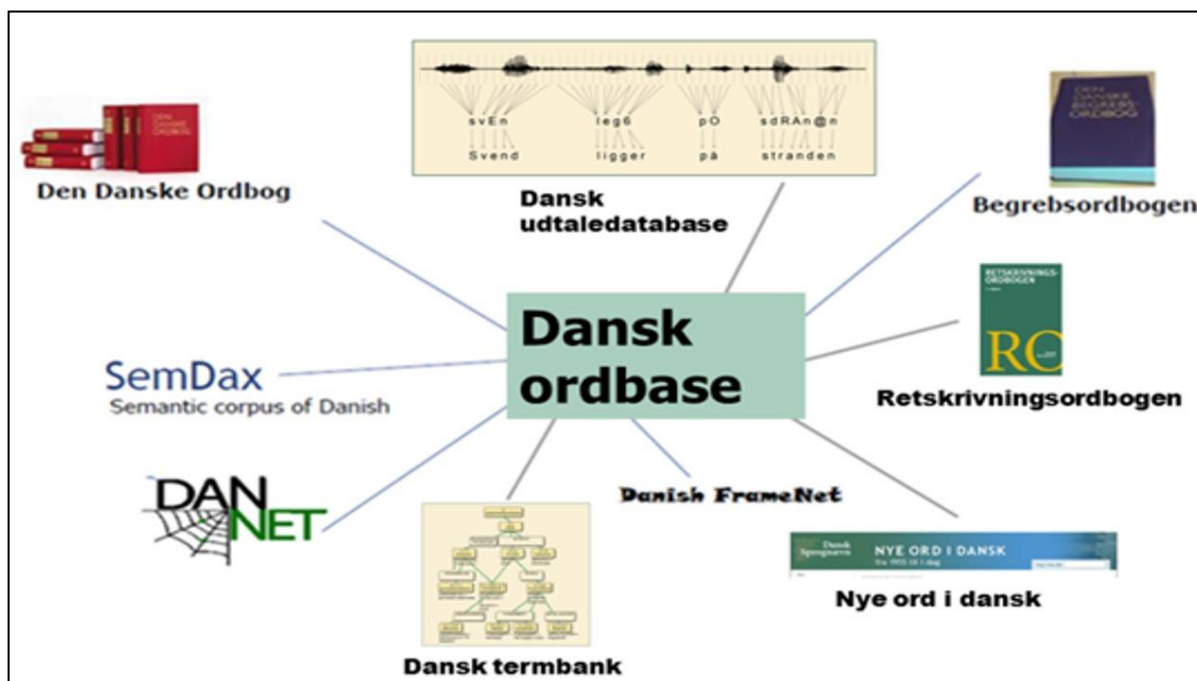


FIG. 5. Database combining the information about Danish lexemes from different lexicographical and terminological resources

Keeping track of language data especially in the public sector and making them available to developers of artificial intelligence systems sounds tedious and bureaucratic. But it proves to be of enormous importance. Therefore, the following supplementary recommendations by the language technology committee were made:

- A change in the law on copyright deposit with a view to making deposited digital and printed material available for language technology

- Guidelines for collecting texts in public institutions and in publicly financed projects

Mapping of translation tasks in public institutions and revision of translation contracts for public institutions to facilitate sharing of translation data.

## 5. CONCLUDING REMARKS

Multilingualism has always been a central issue in the EU, and for many years, the EU has made use of language technology to build its own language services and terminology databases. Now the EU is making its data and services, such as the eTranslation system, available to all public institutions in Europe to bring down the language barriers. It is a part of the EU strategy for the digital single market and its Connecting Europe Facility (CEF), which will make citizens all over Europe able to trade and to find information in other countries in their own language (ELRC 2019).

In the ELRC project, the EU encourages all countries to contribute actively to this by making language resources such as terminologies, lexical databases and translated texts freely available. The work is done in close cooperation with European projects and institutions such as ELRA, CLARIN and META-NET.

This trend is not only European, and it does not only cover the official languages. Lately, the EU and UNESCO have joined forces launching the conference LT4all (LT4all 2019) to explore what can be done for all languages in the world, including minority languages. A roadmap for developing a language technology strategy for all European languages is expected to be ready by the end of 2022 (European Call LANGEQ-2020(PPPA-AG)).

The linguistic research that takes place in many countries is of great importance, not only for the citizens and for the languages, but also for language technology. The fact is that

we only know a trifle of what will be needed for the development of artificial intelligence in the future.

Facts about language semantics, language variation, language change and dialects are still only poorly described for many languages, and research into the nature of human dialogues, in particular, are desperately needed in order to improve communication interfaces to robots, chatbots etc. For these purposes, we need more transcribed dialogues, translated texts and terminological and lexical descriptions. Very often, data collected in linguistic research projects cannot be shared for other purposes because there is a lack of formal consent from the informants and no clarity about copyright restrictions. New projects should take care to get permission to use the data not only for one research group and one project, but for general use, or at least for language technology development.

This is also the content of EFNIL's recommendations for language technology.

- All countries should have a language technology strategy and take on the responsibility for collecting and curating public language data and for ensuring the availability of the linguistic and technological expertise that is necessary to develop high quality language technology for their languages.

- The EU should have strategic long-term focus on research and development of language technology for all European languages based on close cooperation between the EU and all European states.

EFNIL will continue projects that monitor the development of language policies and strategies. The next ELM survey is scheduled to be carried out in 2021 and will be published in 2022. Perhaps this time, it will be possible to see more developments in language technology and AI. EFNIL is actively pursuing the interest of official and minority languages in Europe by partnering in the EU-project ELE (European Language Equality) aiming at developing a strategic research, innovation and implementation agenda and a roadmap for achieving full digital language equality in Europe by 2030 (European Call LANGEQ-2020(PPPA-AG)).

## REFERENCES

- Crystal David 2002: *Language Death*, Cambridge University Press.
- Danish Digitization Agency 2019: *Denmark's National Strategy for Artificial Intelligence*: <https://en.digst.dk/policy-and-strategy/denmark-s-national-strategy-for-artificial-intelligence/>.
- DSN LT articles: <https://dsn.dk/vi-arbejder-ogsaa-med/sprogteknologi-og-fagsprog/hvordan-virker-sprogteknologi>.
- Diderichsen Philip 2015-1: Stavekontrol. – *Nyt fra Sprognævnet* (1), 5–8.
- Diderichsen Philip, Kirkedal Anders S. 2015: Talesyntese og talegenkendelse. – *Nyt fra Sprognævnet* (2), 6–11.
- Diderichsen Philip 2015-2: Maskinoversættelse. – *Nyt fra Sprognævnet* (3), 8–15.
- Diderichsen Philip 2016: Hverdagens sprogteknologi nu og i fremtiden. – *Nyt fra Sprognævnet* (1), 8–13.
- Diderichsen Philip, Henrichsen Peter J., Kirchmeier Sabine, Pedersen Bolette S. 2016, September 26th: *Sproget er digitaliseringens sorte guld*. Politiken.
- ELEXIS 2020: *European Lexicographic Infrastructure*: <https://elex.is/>.
- ELM 4, 2019: *European Language Monitor 4*. Final version: <https://juniper.nytud.hu/elm4>.
- ELM 3, 2014: *European Language Monitor*. Articles and general information: <http://efnil.org/projects/elm>.
- ELRC 2019: *European Language Resource Coordination*: <http://lr-coordination.eu/>.
- Hovy Dirk, Spruit Shannon L. 2016: The Social Impact of Natural Language Processing. – *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Berlin, Germany, August 7-12, 2016. Marseille 2020, 591–598. <https://www.aclweb.org/anthology/2020.lrec-1.403>.
- Kirchmeier-Andersen *et al.* 2012: Kirchmeier-Andersen Sabine, Robustelli Cecilia, Spetz Jennie, Stickel Gerhard, Teigland Nina. European Language Monitor (ELM). – Stickel, Gerhard (ed.): *National, Regional and Minority Languages in Europe. Contributions to the Annual Conference 2009 of EFNIL in Dublin*. Duisburger Arbeiten zur Sprach- und

Kulturwissenschaft / Duisburg Papers on Research in Language and Culture, Berlin. Peter Lang, 181-190.

- Kirchmeier *et al.* 2019a: Kirchmeier Sabine, Henrichsen Peter Juel, Diderichsen Philip, Hansen Nanna B. *Dansk Sprogteknologi i verdensklasse. Rapport fra sprogteknologiudvalget under Dansk Sprognævn nedsat af Kulturministeriet*. Dansk Sprognævn. <https://dsn.dk/nyt/nyheder/2019/dansk-sprogteknologi-verdensklasse-rapport-frasprogteknologiudvalget/?searchterm=dansk%20sprogteknologi>.
- Kirchmeier *et al.* 2019b: Kirchmeier Sabine, Pedersen Bolette, Nimb Sanni, Diderichsen Philip, Henrichsen Peter Juel. World Class Language Technology – Developing a Language Technology Strategy for Danish. – *Proceedings of the 12th Language Resources and Evaluation Conference*. <http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.403.pdf>.
- LLE (Language Legislation in Europe): <http://efnil.org/projects/lle>.
- LT4all 2019: Language Technologies for All: <https://en.unesco.org/LT4All>.
- Rehm Georg, Uszkoreit Hans (eds.) 2012: *META-NET White Paper Series: Europe's Languages in the Digital Age*. Springer, Heidelberg, New York, Dordrecht, London. 31 volumes on 30 European languages. <http://www.meta-net.eu/whitepapers>.
- Spyns Peter, D'Halleweyn Elisabeth, Cucchiarini Catia 2008: The Dutch-Flemish comprehensive approach to HLT stimulation and innovation: STEVIN, HLT Agency and beyond. – *Proceedings of the 6th International Conference on Language Resources and Evaluation, Marrakesh, Morocco*, 1511–1517.

Received 30 11 2020

Accepted 21 12 2020

## TRENDS IN EUROPEAN LANGUAGE POLICIES WITH A VIEW TO LANGUAGE TECHNOLOGY

### *Summary*

This article gives an overview of current trends in European language policies – a bird’s eye view on the European language landscape regarding the status of the official EU-languages and with a view to language technology. It is based on data in the European Language Monitor (ELM) that is developed and regularly maintained by the European Federation of National Institutions for Language (EFNIL), and on recent developments in Denmark regarding policies for language technology. Language technology is chosen as the focus point, since this is an area where we currently see the most interesting developments in national and EU policies. The aim of the article is to describe the wealth of information about European language policies that is available through EFNIL and ELM, and to inspire researchers, students and policy makers to further analyse the data and use them for future policy development.

The first section describes EFNIL and its approach to language planning. EFNIL’s projects aim at informing decision makers and researchers about language planning activities in Europe. Language planning is generally defined as the development of policies or programs designed to direct or change language use. It typically consists of status planning, acquisition planning and corpus planning.

All three aspects of language planning are addressed in the European Language Monitor (ELM) described in section 2 and 3, which also contains samples of the data in ELM about language legislation, about the use of English in higher education and about language technology. The samples disclose interesting facts about language policies in Europe, for instance: only about half of the participating countries have provisions about language in their constitution; Denmark, Finland, Luxembourg and the Netherlands are the countries where the use of English as medium of instruction in natural science at the MA level is most widely

spread; and only half of the participating countries report to have an official strategy for language technology.

Section 4 describes the development regarding the recently adopted strategy for language technology in Denmark and illustrates important aspects of the policy development process – especially the interaction between political actors and experts in the field, and the recommendations that were finally adopted.

In the concluding chapter, the recent plans for language technology in the European Commission and the UNESCO are highlighted, and it is shown that these are very much in tune with EFNIL's recommendations for policies on language technology.

SABINE KIRCHMEIER

Vice president of EFNIL

Kirchmeier.dk

Kålungsvej 2B

DK-3520 Farum

[sabine@kirchmeier.dk](mailto:sabine@kirchmeier.dk)