



Plenary sitting

A8-0228/2018

27.6.2018

REPORT

on language equality in the digital age
(2018/2028(INI))

Committee on Culture and Education

Rapporteur: Jill Evans

Rapporteur for the opinion (*):
Marisa Matias, Committee on Industry, Research and Energy

(*) Associated committee – Rule 54 of the Rules of Procedure

CONTENTS

	Page
MOTION FOR A EUROPEAN PARLIAMENT RESOLUTION	3
EXPLANATORY STATEMENT.....	13
OPINION OF THE COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY	15
INFORMATION ON ADOPTION IN COMMITTEE RESPONSIBLE	22
FINAL VOTE BY ROLL CALL IN COMMITTEE RESPONSIBLE.....	23

MOTION FOR A EUROPEAN PARLIAMENT RESOLUTION

on language equality in the digital age (2018/2028(INI))

The European Parliament,

- having regard to Articles 2 and 3(3) of the Treaty on the Functioning of the European Union (TFEU),
- having regard to Articles 21(1) and 22 of the Charter of Fundamental Rights of the European Union,
- having regard to the 2003 UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage,
- having regard to Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information¹,
- having regard to Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information²,
- having regard to Decision (EU) 2015/2240 of the European Parliament and of the Council of 25 November 2015 establishing a programme on interoperability solutions and common frameworks for European public administrations, businesses and citizens (ISA2 programme) as a means for modernising the public sector³,
- having regard to the Council resolution of 21 November 2008 on a European strategy for multilingualism (2008/C 320/01)⁴,
- having regard to the Council decision of 3 December 2013 establishing the specific programme implementing Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020) and repealing Decisions 2006/971/EC, 2006/972/EC, 2006/973/EC, 2006/974/EC and 2006/975/EC⁵,
- having regard to the UN Convention on the Rights of Persons with Disabilities (UN CRPD), ratified by the EU in 2010,
- having regard to the Commission communication of 18 September 2008 entitled ‘Multilingualism: an asset for Europe and a shared commitment’ (COM(2008)0566),
- having regard to the Commission communication of 26 August 2010 entitled ‘A Digital Agenda for Europe’ (COM(2010)0245),

¹ OJ L 345, 31.12.2003, p. 90.

² OJ L 175, 27.6.2013, p. 1.

³ OJ L 318, 4.12.2015, p. 1.

⁴ OJ C 320, 16.12.2008, p. 1.

⁵ OJ L 347, 20.12.2013, p. 965.

- having regard to the Commission communication of 11 January 2012 entitled ‘A coherent framework for building trust in the Digital Single Market for e-commerce’ (COM(2011)0942),
- having regard to the Commission communication of 6 May 2015 entitled ‘A Digital Single Market Strategy for Europe’ (COM(2015)0192),
- having regard to the opinion of the European Economic and Social Committee on the communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions ‘A Digital Agenda for Europe’ (COM(2010)0245)¹,
- having regard to the Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace adopted by the UNESCO General Conference at its 32nd session in Paris on 15 October 2003,
- having regard to the Special Eurobarometer 386 report entitled ‘Europeans and their Languages’, published in June 2012,
- having regard to the Presidency conclusions of the Barcelona European Council of 15 and 16 March 2002 (SN 100/1/02 REV 1),
- having regard to its resolution of 17 June 1988 on sign languages for the deaf²,
- having regard to its resolution of 14 January 2004 on Preserving and promoting cultural diversity: the role of the European regions and international organisations such as UNESCO and the Council of Europe³, and to its resolution of 4 September 2003 on European regional and lesser-used languages – the languages of minorities in the EU – in the context of enlargement and cultural diversity⁴,
- having regard to its resolution of 24 March 2009 on Multilingualism: an asset for Europe and a shared commitment⁵,
- having regard to its resolution of 11 September 2013 on endangered European languages and linguistic diversity in the European Union⁶,
- having regard its resolution of 7 February 2018 on protection and non-discrimination with regard to minorities in the EU Member States⁷,
- having regard to the study by the European Parliamentary Research Service (EPRS) and Scientific Foresight Unit (STOA) entitled ‘Language equality in the digital age – Towards a Human Language Project’, published in March 2017,

¹ OJ C 54, 19.2.2011, p. 58.

² OJ C 187, 18.7.1988, p. 236.

³ OJ C 92 E, 16.04.2004, p. 322.

⁴ OJ C 76 E, 25.3.2004, p. 374.

⁵ OJ C 117 E, 6.5.2010, p. 59.

⁶ OJ C 93, 9.3.2016, p. 52.

⁷ Texts adopted, P8_TA(2018)0032.

- having regard to Rule 52 of its Rules of Procedure,
 - having regard to the report of the Committee on Culture and Education and the opinion of the Committee on Industry, Research and Energy (A8-0228/2018),
- A. whereas language technologies can make communication easier for the deaf and hard of hearing, the blind and visually impaired and those with dyslexia and whereas, for the purposes of this report, ‘language technology’ refers to technology that supports not only spoken languages, but also sign languages, recognising that sign languages are an important element of Europe’s linguistic diversity;
- B. whereas the development of language technologies (LTs) covers many research areas and disciplines, including computational linguistics, artificial intelligence, computer science and linguistics, with applications such as natural language processing, text analytics, speech technology and data mining, among others;
- C. whereas according to the Special Eurobarometer report 386 entitled ‘Europeans and their languages’, just over half of Europeans (54 %) are able to hold a conversation in at least one additional language, a quarter (25 %) are able to speak at least two additional languages and one in ten (10 %) are conversant in at least three;
- D. whereas there are 24 official languages and more than 60 national, regional and minority languages in the European Union, in addition to migrant languages and, under the UN Convention on the Rights of Persons with Disabilities (UNCRPD), the various state-recognised sign languages; whereas multilingualism presents one of the greatest assets of cultural diversity in Europe and, at the same time, one of the most significant challenges for the creation of a truly integrated EU;
- E. whereas support for local communities, such as indigenous, rural or remote communities, in overcoming geographical, social and economic obstacles to broadband access is a crucial prerequisite for an efficient, EU multilingualism policy;
- F. whereas multilingualism comes under the scope of a series of EU policy areas, including culture, education, the economy, the digital single market, lifelong learning, employment, social inclusion, competitiveness, youth, civil society, mobility, research and the media; whereas more attention needs to be paid to removing barriers to intercultural and interlinguistic dialogue, and to stimulating mutual understanding;
- G. whereas the Commission acknowledges that the Digital Single Market must be multilingual; whereas no common EU policy has been proposed to address the problem of language barriers;
- H. whereas LTs are used in practically all everyday digital products and services, since most use language to some extent, especially all internet-related products such as search engines, social networks and e-commerce services; whereas the use of LTs also has an impact on sectors of fundamental importance to the everyday well-being of European citizens, such as education, culture and health;
- I. whereas cross-border e-commerce is very low, with just 16 % of European citizens having purchased online from other EU countries in 2015; whereas language

technologies can contribute to future European cross-border and cross-language communication, boost economic growth and social stability and reduce natural barriers, thereby respecting and promoting cohesion and convergence, and strengthening the EU's competitiveness worldwide;

- J. whereas technological development is increasingly language-based and has consequences for growth and society; whereas there is an urgent need for more language-aware policies and for technological, but also genuinely multidisciplinary, research and education on digital communication and LTs and their relationship with growth and society;
- K. whereas fulfilling the Barcelona objective of enabling citizens to communicate well in their mother tongue plus two other languages would give people more opportunities to access cultural, educational and scientific content in digital form and to participate as citizens, in addition to accessing the digital single market; whereas additional means and tools, especially those provided by language technologies, are key to managing European multilingualism properly, and to promoting individual multilingualism;
- L. whereas there have been substantial breakthroughs in artificial intelligence and the pace of development in language technologies has been fast; whereas language-centric artificial intelligence offers new opportunities for digital communication, digitally enhanced communication, technology-enabled communication, and cooperation in all European languages (and beyond), giving speakers of different languages equal access to information and knowledge, and improving IT network functionalities;
- M. whereas the common European values of cooperation, solidarity, equality, recognition and respect should mean that all citizens have full and equal access to digital technologies, which would not only improve European cohesiveness and well-being but also enable a multilingual Digital Single Market;
- N. whereas the availability of technological tools such as video games or educational applications in minority and lesser-used languages is pivotal for the development of language skills, especially in children;
- O. whereas the speakers of lesser-spoken European languages need to be able to express themselves in culturally meaningful ways and to create their own cultural content in local languages;
- P. whereas the emergence of methods such as deep learning, based on increased computational power and access to vast amounts of data, are making language technologies a real solution for overcoming language barriers;
- Q. whereas language barriers have a considerable impact on the construction of the European identity and the future of the European integration process; whereas the EU's decision-making and policies should be communicated to its citizens in their mother tongue, both online and offline;
- R. whereas language makes up a very large part of the ever-increasing wealth of big data;

- S. whereas an enormous amount of data is expressed in human languages; whereas the management of LTs could enable a wide range of innovative IT products and services in industry, commerce, government, research, public services and administration, reducing natural barriers and market costs;

Current obstacles to achieving language equality in the digital age in Europe

1. Regrets the fact that, owing to a lack of adequate policies in Europe, there is currently a widening technology gap between well-resourced languages and less-resourced languages, whether the latter are official, co-official or non-official in the EU; regrets, furthermore, the fact that more than 20 European languages are in danger of digital language extinction; notes that the EU and its institutions have a duty to enhance, promote and uphold linguistic diversity in Europe;
2. Points out that, over the last decade, digital technology has had a significant impact on language evolution, which remains difficult to evaluate; recommends that policymakers devote serious consideration to the studies showing that digital communication is eroding young adults' literacy skills, leading to grammar and literacy barriers between generations and a general depletion of language; is of the opinion that digital communication should serve to broaden, enrich and advance languages and that these ambitions should be reflected in national literacy education and literacy policies;
3. Stresses that European lesser-used languages are at a significant disadvantage on account of an acute lack of tools, resources and research funding, which is inhibiting and narrowing the scope of the work done by researchers who, even if equipped with the necessary technological skills, are unable to derive the full benefit of language technologies;
4. Notes the deepening digital divide between widely used and lesser-used languages, and the increasing digitalisation of European society, which is leading to disparities in access to information, particularly for the low-skilled, the elderly, people on low incomes and people from disadvantaged backgrounds; stresses that making content available in different languages would reduce inequality;
5. Notes that while it has a strong scientific base in language engineering and technology, and at a time when language technologies constitute an enormous opportunity for it, both economically and culturally, Europe remains far behind, on account of market fragmentation, inadequate investment in knowledge and culture, poorly coordinated research, insufficient funding and legal barriers; further notes that the market is currently dominated by non-European actors, which are not addressing the specific needs of a multilingual Europe; highlights the need to shift this paradigm and reinforce European leadership in language technologies by creating a project tailored specifically to Europe's needs and demands;
6. Notes that LTs are available in English first; is aware that large global and European manufacturers and companies often also develop LTs for the major European languages with relatively large markets: Spanish, French and German (these languages already lack resources in some sub-areas); stresses, however, that general EU-level action (policy, funding, research and education) should be taken to ensure the development of LTs for official EU languages which are less widely spoken and that special EU-level

actions (policy, funding, research and education) should be launched to include and encourage regional and minority languages in such development;

7. Insists on the need to make better use of new technological approaches, based on increased computational power and better access to sizeable amounts of data, in order to foster the development of deep-learning neural networks which make human language technologies (HLTs) a real solution to the problem of language barriers; calls, therefore, on the Commission to safeguard sufficient funding to support such technological development;
8. Notes that languages with fewer speakers need proper support from stakeholders, including type foundries for diacritical marks, keyboard manufacturers and content management systems, in order to properly store, process and display content in such languages; requests that the Commission assess how such support can be instigated and made a recommendation in the procurement process within the EU;
9. Calls on the Member States to boost the use of multiple languages in digital services such as mobile applications;
10. Notes with concern that the Digital Single Market remains fragmented by a number of barriers, including language barriers, thus hindering online commerce, communication via social networks and other communication channels, and the cross-border exchange of cultural, creative and audiovisual content, as well as the wider deployment of pan-European public services; stresses that cultural diversity and multilingualism in Europe could benefit from cross-border access to content, particularly for educational purposes; calls on the Commission to develop a strong and coordinated strategy for the multilingual Digital Single Market;
11. Notes that language technologies currently do not play a role in the European political agenda, despite the fact that respect for linguistic diversity is enshrined in the Treaties;
12. Commends the important role of previous EU-funded research networks such as FLaReNet, CLARIN, HBP and META-NET (including META-SHARE), for leading the way in the construction of a European language technology platform;

Improving the institutional framework for language technology policies at EU level

13. Calls on the Council to draft a recommendation on the protection and promotion of cultural and linguistic diversity in the Union, including in the sphere of language technologies;
14. Recommends that in order to raise the profile of language technologies in Europe, the Commission should allocate the area of ‘multilingualism and language technology’ to the portfolio of a Commissioner; considers that the Commissioner responsible should be tasked with promoting linguistic diversity and equality at EU level, given the importance of linguistic diversity for the future of Europe;
15. Suggests ensuring comprehensive EU-level legal protection for the 60 regional and minority languages, recognition of the collective rights of national and linguistic minorities in the digital world, and mother-tongue teaching for speakers of official and

non-official languages of the EU;

16. Encourages those Member States that have already developed their own successful policy strategies in the field of language technologies to share their experiences and good practices in order to help other national, regional and local authorities develop their own strategies;
17. Calls on the Member States to develop comprehensive language-related policies and to allocate resources and use appropriate tools in order to promote and facilitate linguistic diversity and multilingualism in the digital sphere; stresses the shared responsibility of the EU and the Member States, together with universities and other public institutions, in contributing to the preservation of their languages in the digital world and in developing databases and translation technologies for all EU languages, including languages that are less widely spoken; calls for coordination between research and industry with a common objective of enhancing the digital possibilities for language translation and with open access to the data required for technological advancement;
18. Calls on the Commission and the Member States to develop strategies and policy action to facilitate multilingualism in the digital market; requests, in this context, that the Commission and the Member States define the minimum language resources that all European languages should possess, such as data sets, lexicons, speech records, translation memories, annotated corpora and encyclopaedic content, in order to prevent digital extinction;
19. Recommends that the Commission consider the creation of a centre for linguistic diversity that will strengthen awareness of the importance of lesser-used, regional and minority languages, including in the sphere of language technologies;
20. Asks the Commission to review its Framework Strategy for Multilingualism and to propose a clear action plan on how to promote linguistic diversity and overcome language barriers in the digital area;
21. Calls on the Commission to make as a priority of language technology those Member States which are small in size and have their own language, in order to pay heed to the linguistic challenges that they face;
22. Emphasises that the development of language technology will facilitate the subtitling, dubbing and translation of video games and software applications into minority and lesser-used languages;
23. Stresses the need to reduce the technology gap between languages by strengthening knowledge and technology transfer;
24. Urges Member States to come up with effective ways to solidify their native languages;

Recommendations for EU research policies

25. Calls on the Commission to establish a large-scale, long-term coordinated funding programme for research, development and innovation in the field of language technologies, at European, national and regional levels, tailored specifically to Europe's

needs and demands; emphasises that the programme should seek to tackle deep natural language understanding and increase efficiency by sharing knowledge, infrastructures and resources, with a view to developing innovative technologies and services, in order to achieve the next scientific breakthrough in this area and help to reduce the technology gap between European languages; stresses that this should be done with the participation of research centres, academia, enterprises (particularly SMEs and start-ups) and other relevant stakeholders; further stresses that this project should be open, cloud-based and interoperable and provide highly scalable and high-performance basic tools for a number of language technology applications;

26. Believes that ICT integrators in the EU should be given economic incentives to accelerate the provision of cloud-based services, in order to enable a smooth integration of HLTs in their e-commerce applications, in particular to ensure that SMEs reap the benefits of automated translation;
27. Stresses that Europe has to secure its leadership position in the field of language-centric artificial intelligence; recalls that EU companies are the best placed to provide solutions tailored to our specific cultural, societal and economic needs;
28. Believes that specific programmes within current funding schemes such as Horizon 2020, as well as successor funding programmes, should boost long-term basic research as well as knowledge and technology transfer between countries and regions;
29. Recommends the creation of a European language technology platform, with representatives from all European languages, that enables the sharing of language technology-related resources, services and open source code packages, particularly between universities and research centres, while ensuring that any funding scheme can both work with and be accessed by the open-source community;
30. Recommends establishing or extending projects such as the Digital Language Diversity Project, among others, that carry out research into the digital needs of all European languages, including those with both very small and very large numbers of speakers, so as to address the digital divide issue and help prepare these languages for a sustainable digital future;
31. Recommends an update of the META-NET white paper series, a pan-European survey published in 2012 on the status of language technologies, on resources for all European languages, on information about language barriers and on policies related to the topic, with a view to enabling the assessment and development of language technology policies;
32. Urges the Commission to set up an HLT financing platform, drawing on the implementation of the 7th Framework Programme for Research and Technological Development, Horizon 2020 and the Connecting Europe Facility (CEF); considers, in addition, that the Commission should place emphasis on research areas needed to ensure a deep language understanding, such as computational linguistics, linguistics, artificial intelligence, LTs, computer science and cognitive science;

33. Points out that language can be a barrier to the transfer of scientific knowledge; notes that most scientific journals with high impact factors publish in English, leading to a major shift in the creation and distribution of academic knowledge; stresses the need for these knowledge production conditions to be reflected in European research and innovation policies and programmes; urges the Commission to seek solutions to ensure that scientific knowledge is made available in languages other than English and to support the development of artificial intelligence for natural language;

Education policies to improve the future of language technologies in Europe

34. Believes that owing to the current situation whereby non-European actors dominate the market in language technologies, European education policies should be aimed at retaining talent in Europe, should analyse the current educational needs related to language technology (including all fields and disciplines involved) and, based on this, provide guidelines for the implementation of cohesive joint action at European level, and should raise awareness among schoolchildren and students of the career opportunities in the language technology industry, including the language-centric artificial intelligence industry;
35. Takes the view that digital teaching materials must also be developed in minority and regional languages – which is important in terms of non-discrimination – if we wish to establish equality of opportunity and treatment;
36. Points to the need to promote the ever-greater participation of women in the field of European studies on language technologies, as a decisive factor in the development of research and innovation;
37. Proposes that the Commission and Member States promote the use of language technologies within cultural and educational exchanges between European citizens such as Erasmus+, for example Erasmus+ Online Linguistic Support (OLS), with the aim of reducing the barriers that linguistic diversity can pose to intercultural dialogue and mutual understanding, especially in written and audiovisual expression;
38. Recommends that Member States also develop digital literacy programmes in Europe's minority and regional languages and introduce language technology training and tools in the curricula of their schools, universities and vocational colleges; further stresses the fact that literacy remains a significant factor and an absolute prerequisite for progress in the digital inclusion of communities;
39. Stresses that the Member States should provide the support that educational institutions need in order to improve the digitalisation of languages in the EU;

Language technologies: benefits for both private companies and public bodies

40. Underlines the need to support the development of investment instruments and accelerator programmes that aim to increase the use of language technologies in the cultural and creative sector, especially targeting less-resourced communities and encouraging the development of language technology capacities in areas where the sector is weaker;

41. Urges the development of actions and appropriate funding with the aim of enabling and empowering European SMEs and startups to easily access and use LTs in order to grow their businesses online by accessing new markets and development opportunities, thereby boosting their levels of innovation and creating jobs;
42. Calls on the EU institutions to raise awareness of the benefits for companies, public bodies and citizens of the availability of online services, content and products in multiple languages, including lesser-used, regional and minority languages, with a view to overcoming language barriers and helping to preserve the cultural heritage of language communities;
43. Supports the development of multilingual public e-services in European, national and, where appropriate, regional and local administrations with innovative, inclusive and assistive LTs, which will reduce inequalities among languages and language communities, promote equal access to services, stimulate the mobility of businesses, citizens and workers in Europe and ensure the achievement of an inclusive multilingual Digital Single Market;
44. Calls on administrations at all levels to improve access to online services and information in different languages, especially for services in cross-border regions and culture-related issues, and to use existing free and open-source language technology, including machine translation, speech recognition and text-to-speech and intelligent linguistic systems, such as those performing multilingual information retrieval, summarising/abstracting and speech understanding, in order to improve the accessibility of those services;
45. Highlights the importance of text and data mining techniques for the development of language technologies; underlines the need to strengthen collaboration between industry and data owners; stresses the need to adapt the regulatory framework and ensure a more open and interoperable use and collection of language resources; notes that sensitive information should not be turned over to commercial companies and their free software, as it is unclear how they might use the knowledge gathered, such as in the case of health data;

○

○ ○

46. Instructs its President to forward this resolution to the Council and the Commission.

EXPLANATORY STATEMENT

The EU is a unique endeavour involving more than 500 million citizens sharing about 80 different languages. While multilingualism is one of the biggest assets of Europe, it also poses one of the most substantial challenges for the creation of a culturally and socially integrated EU as well as an obstacle for the fulfilment of the goals of the Digital Single Market.

Language technologies are found behind many everyday digital products, since most of them use language to some extent. Mobile communications, social media, intelligent assistants, and speech-based interfaces are transforming the way citizens, companies and public administrations interact in the digital world. Furthermore, language technologies are also helpful to develop multilingual resources and content outside of the digital sphere. Despite the fact that language technologies are critical pieces of technology to aid in this digital revolution, they are not properly represented in the agendas of the European policy-makers.

Language technologies contribute to the equality of all European citizens in their everyday experiences, regardless of the languages they speak.

Although smaller or minority languages are the ones to gain most from language technologies, tools and resources for them are often scarce —in some cases non-existent. In fact, there is a widening technology gap between large, well-resourced languages and the other official, co-official or non-official EU languages, some of which might already be facing digital extinction.

In order to bridge this technology gap, policies should focus on fostering technology development for all European languages. The preservation of a language, and thus of the culture develop around it, is critically tied to its ability to function and be useful in modern and changing environments as the digital world. So, cultural and language diversity is closely linked to the capacities and resources in the digital world.

Improvements in language technologies rely mainly on the ability to access and maintain ever larger and more finely tuned linguistic data and resources. Close collaboration between research, industry, public and private data owners becomes a necessity. Moreover, regulation of the use of such data should be made much more open and core language resources (corpora, lexicons, ontologies, etc.) should be made interoperable and shared in an open environment. The framework provided by the directive on the re-use of public sector information (Directive 2003/98/EC, known as the ‘PSI Directive’) gives the opportunity for the huge amounts of data and information generated by public administrations to be used for valuable language technology resources.

In a multilingual Europe, language barriers are expected to have profound and intertwined social and economic consequences such as: (1) fostering a language divide, (2) hampering workers’ mobility, (3) hindering the access to cross-border public services, (4) limiting citizens’ engagement and participation in the political process, (5) creating fragmented markets for cross-border trade and e-commerce, particularly for SMEs and (6) restricting access to cultural and educational resources.

Furthermore, the European language technologies community acknowledges a lack of coordination between research efforts and the market of HLT applications and services. However, recent initiatives such as META-NET, Cracking the Language Barrier federation, LT-Innovate and the Connecting Europe Facility programme have done much to bring the

fragmented community together.

However multilingualism in Europe is a complex topic involving many stakeholders in different countries with intertwined interests. No single policy will tackle the aforementioned problems. On the contrary, to truly seize the opportunities of a multilingual Europe, a joint and coordinated action is required at the European, national and regional levels, involving a variety of stakeholders.

Europe can become a world leader in the field of linguistic diversity and equality. The tools and methodologies built to facilitate widespread use of all European languages can be extended to the rest of the world. It should be noted that there are around 80 European languages and that there are 6000 languages worldwide, of which only about 2000 are written languages and only about 300 have a standardised version. The tools and methodologies built in Europe would be useful for at least other 200 other languages around the world.

Therefore, this report proposes different policy options that would improve language equality in Europe through using new technologies, by (1) improving the institutional frameworks for language technology policies, (2) creating new research policies to increase the use of language technology in Europe, (3) using education policies in order to secure the future of language equality in the digital age, and (4) increasing the support for both private companies and public bodies to make better use of language technologies.

25.4.2018

OPINION OF THE COMMITTEE ON INDUSTRY, RESEARCH AND ENERGY

for the Committee on Culture and Education

on language equality in the digital age
(2018/2028(INI))

Rapporteur for opinion (*): Marisa Matias

(*) Associated committee – Rule 54 of the Rules of Procedure

SUGGESTIONS

The Committee on Industry, Research and Energy calls on the Committee on Culture and Education, as the committee responsible, to incorporate the following suggestions into its motion for a resolution:

- A. whereas the European Union has 24 official languages and more than 60 national, regional, minority, sign and migrant languages;
- B. whereas more than 20 European languages are in danger of digital language extinction; whereas language technologies (LTs) play an essential role in protecting and promoting linguistic diversity, especially in the case of lesser-used languages;
- C. whereas according to UNESCO, only 30-50 % of online content worldwide is in English; whereas high-quality machine translation and computer-aided translation will help to overcome language barriers and improve cross-lingual information access;
- D. whereas multilingualism is one of Europe's greatest assets and a founding principle of the EU, promoting increased proximity in a Europe for citizens, but is also one of its biggest challenges and should enhance the development of the gigabit society;
- E. whereas the Commission acknowledges that the Digital Single Market must be multilingual; whereas no common EU policy has been proposed to address the problem of language barriers;
- F. whereas the EU and its institutions have a duty to enhance, promote and uphold linguistic diversity in Europe;

- G. whereas the penetration of cross-border e-commerce is very low; whereas LTs can contribute to future European cross-border and cross-language communication, boost economic growth and social stability and reduce natural barriers, thereby respecting and promoting cohesion and convergence and strengthening EU competitiveness worldwide;
- H. whereas LTs will facilitate communication and cooperation in other languages, give speakers of different languages equal access to information and knowledge, and improve IT network functionalities;
- I. whereas language makes up a very large part of the ever-increasing wealth of Big Data;
- J. whereas difficulties in accessing linguistic data (including online data and research data sets) is constraining the technological development of LTs;
- K. whereas an enormous amount of data is expressed in human languages; whereas the management of LTs could enable a wide range of innovative IT products and services in industry, commerce, government, research, public services and administration, reducing natural barriers and market costs;
- L. whereas LTs in Europe are still far behind, owing to market fragmentation, inadequate investment in knowledge and culture, uncoordinated research, insufficient funding and legal barriers, but their current development is progressing at a rapid pace and there is enormous potential for further advancement;
- M. whereas technological development is increasingly language-based and has consequences for growth and society; whereas there is an urgent need for more language-aware policies and for technological, but also genuinely multidisciplinary, research and education on digital communication and LTs and their relationship to growth and society;
- N. whereas the availability of technological tools such as video games and educational applications in minority languages is pivotal for the development of language skills, especially in children;
- O. whereas European human language technologies (HLTs) constitute an enormous opportunity for the EU, both in economic terms and from a cultural perspective;
- P. whereas the top LT players are not European and do not address Europe's specific needs;
- Q. whereas the Digital Single Market strategy only makes a brief mention of multilingual services;
- R. whereas there have been substantial breakthroughs in artificial intelligence; whereas LTs are essential to ensuring that artificial intelligence is multilingual;
- 1. Underlines that the Digital Single Market cannot be implemented without technological solutions to overcome language barriers; asks the Commission to develop a strong and coordinated strategy for the multilingual Digital Single Market;

2. Strongly supports the creation of a coordinated initiative with a robust, sustainable, large-scale and long-term funding scheme on LTs with the scientific goal of tackling deep natural language understanding and increasing efficiency by sharing knowledge, infrastructures and resources; emphasises that the funding scheme needs to operate at European, national and regional level, with the participation of research centres, academia, enterprises, in particular SMEs and startups, and other relevant stakeholders; stresses that this project should be open, cloud-based and interoperable and provide highly scalable and high-performance basic tools for several LT applications;
3. Notes the effort and dedication of volunteer-driven localisation, and internationalisation efforts in open-source software development; recommends ensuring that any funding scheme for HLTs can both work with and be accessed by open-source communities;
4. Stresses the need to reduce the technology gap between languages by strengthening knowledge and technology transfer;
5. Stresses that cultural diversity and multilingualism in Europe could benefit from cross-border access to content, in particular for educational purposes;
6. Urges the development of actions and appropriate funding with the aim of enabling and empowering European SMEs and startups to easily access and use LTs in order to grow their businesses online by accessing new markets and development opportunities, thereby boosting their levels of innovation and creating jobs;
7. Urges the Commission to set up an HLT financing platform, drawing on the implementation of the 7th Framework Programme for Research and Technological Development, Horizon 2020 and the Connecting Europe Facility (CEF); considers, in addition, that the Commission should place emphasis on research areas needed to ensure a deep language understanding, such as computational linguistics, linguistics, artificial intelligence, LTs, computer science and cognitive science;
8. Stresses the role of the EU, the Member States, universities and other public institutions in contributing to the preservation of their languages in the digital world and in developing databases and translation technologies for all EU languages, including languages that are less widely spoken; calls for coordination between research and industry with a common objective of enhancing the digital possibilities for language translation and with open access to the data required for technological advancement;
9. Notes the widening gap between English and other EU languages in technology, which leads to disparities in access to information between age groups, different regions and Member States, as well as between people with different levels of education; stresses that making content available in different EU languages would reduce inequality;
10. Points out that language can be a barrier to the transfer of scientific knowledge; notes that most scientific journals with high impact factors publish in English, leading to a major shift in the creation and distribution of academic knowledge; stresses the need for these knowledge production conditions to be reflected in European research and innovation policies and programmes; urges the Commission to seek solutions to ensure that scientific knowledge is made available in languages other than English and to support the development of artificial intelligence for natural language;

11. Notes that open-source and open-software systems have advantages for EU citizens, since commercial algorithms are not publicly accessible while anyone can review open-source algorithms and participate in developing the technology further, which in turn helps to keep machine languages up-to-date and functioning properly;
12. Notes the sensitive nature of some of the information entered into translation systems which, for personal data and privacy reasons, cannot be turned over to commercial companies and their free software, as it is unclear how they might use the knowledge gathered through the translation programmes, for example in the case of health data;
13. Notes that cross-sectional research and studies are needed on the consequences of language minoritisation in the digital world and its effects on equality and access to information;
14. Supports the development of multilingual public e-services in European, national and, where appropriate, regional and local administrations with innovative, inclusive and assistive LTs, which will reduce inequalities among languages and language communities, promote equal access to services, stimulate the mobility of businesses, citizens and workers in Europe and ensure the achievement of an inclusive multilingual Digital Single Market;
15. Notes that LTs are available in English first; is aware that large global and European manufacturers and companies often also develop LTs for the major European languages with relatively large markets: Spanish, French and German (these languages already lack resources in some sub-areas); stresses, however, that general EU-level action (policy, funding, research and education) should be taken to ensure the development of LTs for official EU languages which are less widely spoken and that special EU-level actions (policy, funding, research and education) should be launched to include and encourage regional and minority languages in such development;
16. Insists on the need to make better use of new technological approaches, based on increased computational power and better access to sizeable amounts of data, in order to foster the development of deep-learning neural networks which make HLTs a real solution to the problem of language barriers; calls, therefore, on the Commission to safeguard sufficient funding to support such technological development;
17. Notes that languages with fewer speakers need proper support from stakeholders, including type foundries for diacritical marks, keyboard manufacturers and content management systems, in order to properly store, process and display content in such languages; requests that the Commission assess how such support can be instigated and made a recommendation in the procurement process within the EU;
18. Underlines the need to strengthen collaboration between the industry and data owners; stresses the need to adapt the regulatory framework and ensure the more open, interoperable use and collection of language resources;
19. Believes that ICT integrators in the EU should be given economic incentives to accelerate the provision of cloud-based services, in order to enable a smooth integration of HLTs in their e-commerce applications, in particular to ensure that SMEs reap the benefits of automated translation;

20. Stresses that Europe has to secure its leadership position in the field of language-centric artificial intelligence; recalls that EU companies are the best placed to provide solutions tailored to our specific cultural, societal and economic needs;
21. Calls on the Commission to support financially the subtitling, dubbing and translation of video games and software applications into minority languages;
22. Notes that current digital tools and resources for lesser-used languages, including the digital availability of all orthographies and means of sign language translation and digitalisation, are insufficient; calls on the Commission and the Member States to support financially projects that develop, collect and promote best resources, including keyboard layouts, dictionaries, translation software and best practices for the use of languages facing the threat of digital exclusion.

INFORMATION ON ADOPTION IN COMMITTEE ASKED FOR OPINION

Date adopted	24.4.2018
Result of final vote	+: 61 -: 1 0: 2
Members present for the final vote	Zigmantas Balčytis, José Blanco López, David Borrelli, Jonathan Bullock, Cristian-Silviu Bușoi, Reinhard Bütikofer, Jerzy Buzek, Angelo Ciocca, Edward Czesak, Jakop Dalunde, Pilar del Castillo Vera, Christian Ehler, Fredrick Federley, Ashley Fox, Adam Gierek, Theresa Griffin, Rebecca Harms, Hans-Olaf Henkel, Eva Kaili, Kaja Kallas, Barbara Kappel, Krišjānis Kariņš, Seán Kelly, Jeppe Kofod, Peter Kouroumbashev, Zdzisław Krasnodębski, Miapetra Kumpula-Natri, Christelle Lechevalier, Janusz Lewandowski, Paloma López Bermejo, Edouard Martin, Angelika Mlinar, Csaba Molnár, Nadine Morano, Dan Nica, Angelika Niebler, Morten Helveg Petersen, Miroslav Poche, Julia Reda, Paul Rübig, Massimiliano Salini, Algirdas Saudargas, Neoklis Sylkiotis, Dario Tamburrano, Evžen Tošenovský, Claude Turmes, Václav Urutchev, Kathleen Van Brempt, Henna Virkkunen, Martina Werner, Lieve Wierinck, Hermann Winkler, Flavio Zanonato, Carlos Zorrinho
Substitutes present for the final vote	Pilar Ayuso, Cornelia Ernst, Francesc Gambús, Françoise Grossetête, Werner Langen, Rupert Matthews, Răzvan Popa, Dominique Riquet, Theodor Dumitru Stolojan
Substitutes under Rule 200(2) present for the final vote	Rosa D'Amato

FINAL VOTE BY ROLL CALL IN COMMITTEE ASKED FOR OPINION

61	+
ALDE	Fredrick Federley, Kaja Kallas, Angelika Mlinar, Morten Helveg Petersen, Dominique Riquet, Lieve Wierinck
ECR	Edward Czesak, Ashley Fox, Hans-Olaf Henkel, Zdzisław Krasnodębski, Rupert Matthews, Evžen Tošenovský
EFDD	Rosa D'Amato, Dario Tamburrano
ENF	Barbara Kappel
GUE/NGL	Cornelia Ernst, Paloma López Bermejo, Neoklis Sylikiotis
NI	David Borrelli
PPE	Pilar Ayuso, Cristian-Silviu Bușoi, Jerzy Buzek, Pilar del Castillo Vera, Christian Ehler, Francesc Gambús, Françoise Grossetête, Krišjānis Kariņš, Seán Kelly, Werner Langen, Janusz Lewandowski, Nadine Morano, Angelika Niebler, Paul Rübig, Massimiliano Salini, Algirdas Saudargas, Theodor Dumitru Stolojan, Vladimir Urutchev, Henna Virkkunen, Hermann Winkler
S&D	Zigmantas Balčytis, José Blanco López, Adam Gierek, Theresa Griffin, Eva Kaili, Jeppe Kofod, Peter Kouroumbashev, Miapetra Kumpula-Natri, Edouard Martin, Csaba Molnár, Dan Nica, Miroslav Poche, Răzvan Popa, Kathleen Van Brempt, Martina Werner, Flavio Zanonato, Carlos Zorrinho
VERTS/ALE	Reinhard Bütikofer, Jakop Dalunde, Rebecca Harms, Julia Reda, Claude Turmes

1	-
EFDD	Jonathan Bullock

2	0
ENF	Angelo Ciocca, Christelle Lechevalier

Key to symbols:

- + : in favour
- : against
- 0 : abstention

INFORMATION ON ADOPTION IN COMMITTEE RESPONSIBLE

Date adopted	19.6.2018
Result of final vote	+ : 21 - : 0 0 : 4
Members present for the final vote	Dominique Bilde, Andrea Bocskor, Silvia Costa, Angel Dzhambazki, Jill Evans, María Teresa Giménez Barbat, Petra Kammerer, Svetoslav Hristov Malinov, Curzio Maltese, Rupert Matthews, Stefano Maullu, Luigi Morgano, Momchil Nekov, Michaela Šojdrová, Yana Toom, Julie Ward, Bogdan Brunon Wenta, Bogdan Andrzej Zdrojewski, Milan Zver, Krystyna Łybacka
Substitutes present for the final vote	Marlene Mizzi, Liliana Rodrigues, Algirdas Saudargas, Remo Sernagiotto, Francis Zammit Dimech

FINAL VOTE BY ROLL CALL IN COMMITTEE RESPONSIBLE

21	+
ALDE	María Teresa Giménez Barbat, Yana Toom
GUE/NGL	Curzio Maltese
PPE	Andrea Bocskor, Svetoslav Hristov Malinov, Stefano Maullu, Algirdas Saudargas, Michaela Šojdrová, Bogdan Brunon Wenta, Francis Zammit Dimech, Bogdan Andrzej Zdrojewski, Milan Zver
S&D	Silvia Costa, Petra Kammerevert, Krystyna Łybacka, Marlene Mizzi, Luigi Morgano, Momchil Nekov, Liliana Rodrigues, Julie Ward
VERTS/ALE	Jill Evans

0	-

4	0
ECR	Angel Dzhambazki, Rupert Matthews, Remo Sernagiotto
ENF	Dominique Bilde

Key to symbols:

- + : in favour
- : against
- 0 : abstention