

META=NET

The Future European Multilingual Information Society

Joseph Mariani

LIMSI-CNRS & IMMI, France

Joseph.Mariani@limsi.fr

META= FORUM 2011 Solutions for Multilingual Europe
Budapest, Hungary, June 27/28

eu 2011.hu



Funded by the Seventh Framework Programme and the ICT Policy Support Programme of the European Commission under contracts T4ME, CESAR, METANET4U and META-NORD (Grant Agreements 249119, 271022, 270893 and 270899)

General Statements

- ❑ Language technology (LT) is a key issue for multilingualism
- ❑ LT helps preserving culture and facilitating communication
- ❑ LT has commercial value and fosters social inclusion
- ❑ Multilingualism and LT are major political issues for the EU



There is a Language Barrier

- ❑ The single European Market is limited by the language barrier.
 - From 2006 to 2008, the share of EU e-Commerce consumers raised from 27% to 33%,
 - but EU cross-border e-Commerce remained steadily low (6% to 7%)
 - Only 33% of EU consumers say they are willing to purchase goods/ services in a different language

- ❑ The European Information Space is limited by the language barrier.
 - Only 55% of Internet users in the EU use one language other than their own
 - Only 35% write e-mail messages or post website comments in a foreign language
 - 90% prefer websites in their language

Source: European Consumer Commissioner , 2009

Source: EC DG Communication for the DG Information Society and Media , 2011



Languages are a Cultural Asset

- ❑ Native languages are an essential medium for the enjoyment of fundamental rights, such as political expression, education and participation in society.
- ❑ The Internet fosters languages
 - English is no longer the dominant language on the web.
 - Contents in Arabic and Asian languages are exploding.
- ❑ Europe has 23 official languages in the EU and 60 national, regional and minority languages.
- ❑ Europe's Digital Agenda must support Europe's languages.



LT is a big business

- ❑ The European market for Language Industries had an estimated value of EUR 8.4 billion in 2008, and is expected to grow by 10% per annum, to reach EUR 20 billion by 2015.

Source: Study on the size of the language industry in the EU, for the European Commission Directorate-General for Translation, 2009



- ❑ There is an insufficient capacity to cover the wide variety of needs:
 - languages
 - content
 - media
- ❑ It is the case for translation, interpretation, software localisation and website globalisation
- ❑ The same applies for information services, document services and media industries, as well as for communication industries.

Questions

- ❑ Should European citizens speak only one and the same language?
- ❑ Should European citizens have access to European information?
- ❑ Should European citizens communicate among themselves?
- ❑ Should the EU only rely on foreign technologies to address sovereignty issues?
- ❑ Can we accept that some European languages die?
- ❑ Are we competing actively enough in the global landscape for research and development in language technology?
- ❑ Can the European cultural background help shape the knowledge society by offering better, more secure, more precise, more innovative and more robust high-quality technology?



Our claim

- ❑ We believe in multilingualism!
- ❑ We believe in language technology!
- ❑ We believe that European language technology will significantly contribute to economic growth and social stability in Europe while establishing global leadership in technology innovation!





Social and economical background: Challenges and Opportunities

Suppress the Language Barrier

- Suppress the language barrier in commercial exchanges.
 - 59% of European retailers are prepared to carry out transactions in more than one language.
 - The first obstacle consistently mentioned by businesses, large or small, is the language barrier
 - 45% mention the additional costs arising from language differences as an obstacle
 - SMEs consider the language barrier as a main hurdle to go crossborder

Source: European Consumer Commissioner, 2009

- Suppress the language barrier in cultural exchanges.
 - 88% of EU Internet users think that websites produced in their country should be available in their country's official language
 - 81% think that websites produced in their country should also have versions in other languages.
 - 44% of European users believe they miss interesting information because websites are not available in a language that they understand.

Source: EC DG Communication for DG Information Society and Media, 2011

- **LT can help:**
 - Provide cross-lingual technologies in web browsers and e-mail applications.



- ❑ An ageing population in Europe requires assistive technologies that help the elderly master everyday situations and obtain proactive guidance.
- ❑ Ambient Assisted Living (AAL):
 - extend the time people can live in their preferred environment;
 - maintain their health and functional capabilities;
 - prevent them from social isolation.
- ❑ ***LT can help:***
 - AAL needs personalized, natural, simple human communication.
 - LT are affordable and easy to use.



Access for People with Disabilities

- ❑ Access to information should be equal to all, regardless of disabilities.
- ❑ European directive on Accessibility / e-Accessibility
- ❑ Represents a large market (approximately 50 million European citizens).

- ❑ ***LT can help:***
 - Screen readers
 - Dictation systems
 - Voice-activated services
 - Automatic sign language recognition and generation
 - Summarization and content simplification



- ❑ Access to information should be equal to all regardless of cultural background and language proficiency.
- ❑ Approximately 60 million migrants live in Europe.

Source: UN Department of Economic and Social Affairs

- ❑ Need to facilitate communication, provide access to information, and conduct language training in order to better integrate migrants in society.

- ❑ ***LT can help:***

- Computer Assisted Language Learning (CALL)
- Instantaneous subtitling
- Text translation
- Simultaneous speech translation
- Language simplification tools



- ❑ Personal information services are a necessary commodity.
 - Broadband access to information and services and mobile communication are now a daily routine.
 - In this 24/7 economy, we expect quick and reliable answers as well as timely online news broadcasts.
 - But information overload is common and limits exchanges.

- ❑ ***LT can help:***
 - Automatic and intelligent question answering systems
 - Personalized and trusted text and speech processing of email messages
 - Text and speech summarization
 - News on Demand

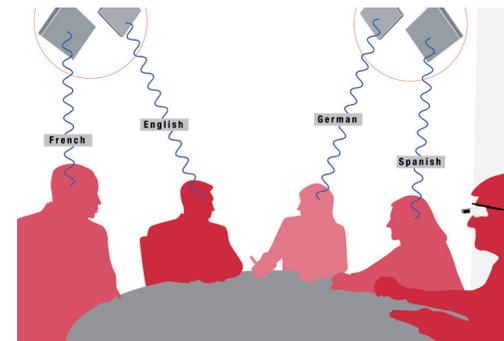


Global Cooperation and Embedded Communication

- ❑ Companies need to address new markets where multiple languages are spoken and when support teams are in different locations.
- ❑ Many jobs cannot be filled because of language barriers.
- ❑ A flexible and mobile population requires multilingual skills.

- ❑ ***LT can help:***

- Advanced video-conferencing services
- Simultaneous translation
- Automatic minute-taking
- Video indexing and search
- Advanced 3D interaction on the Internet for collaboration in the workplace, training and education, entertainment, cultural exchanges and tourism.



Social Media and Participation

- ❑ Participation in social networks is now a routine activity.
- ❑ Social media has impacted all areas of society and life.
 - Solve technical problems, research products, learn about interesting places...
 - Bring citizens together to express their political power, or discuss their position on topics like energy or foreign policy.
 - Feedback from citizens to politicians, journalists, marketing experts.
- ❑ Multilingualism is now a norm, not an exception.
 - Less than 30% of the Web in English (tbc)
 - Only 50% of Twitter messages in English (tbc)
- ❑ The digital divide is a danger.

- ❑ ***LT can help:***
 - Machine Translation of multimedia content, chats and tweets
 - Real-time media analysis and monitoring (measure of social impact)



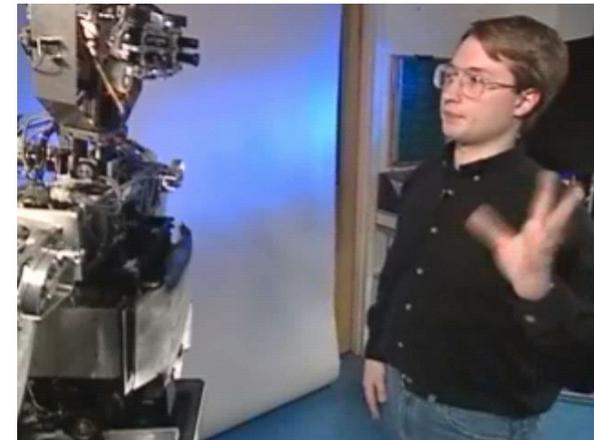
- ❑ Improved communication and information access goes together with increased cyber-criminalism (identity theft and Internet fraud).
- ❑ Need for automated tools to detect crimes and monitor offenders.



- ❑ ***LT can help:***
 - Monitor, analyze and summarize large amounts of text, audio and video data in different languages.

Future Trends

- ❑ Cloud computing, social media, mobile apps, Web services...
- ❑ Advanced environments require advanced interaction
 - 3D, virtual and augmented environments
immersive applications
 - Autonomous and domestic robots and agents
- ❑ ***LT must help:***
 - Spoken and multimodal dialog
 - Emotions analysis and generation



A Difficult Market

- ❑ Market awareness
 - LT market difficult to estimate.
 - LT is often hidden in products.
- ❑ Customer Acceptance
 - High customer acceptance.
 - 60% user acceptance for voice commands in Ford vehicles.
 - 32% of Ford car buyers attracted by voice commands.
- ❑ Need to address a global multilingual market.
 - Support of App Store to facilitate LT diffusion.
 - Cf Google Translate
 - More than 3,000 language pairs



A major political issue for the EU

- ❑ Major political issue for the EU
 - Single Market, Many Languages
- ❑ The 23 official EU languages must be taken into account by EU Institutions, in their relations with the EU citizens and within themselves.



Source: Lisbon Treaty

- ❑ Strong interest of regions for regional languages
- ❑ However the language dimension still less recognized in EU than in:
 - India: TDIL program (22 “official” languages)
 - South Africa: NHN program (11 “national” languages)
 - Develop LT to allow for multilingualism for text and speech (mobile devices)



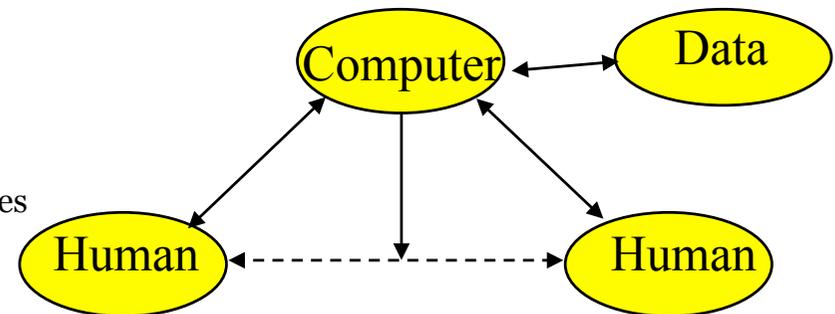
Three Visions for a Multilingual Europe

- ❑ The web is now multilingual and multimedia
- ❑ Most of the content remains hidden
 - Language barrier
 - Media barrier
- ❑ Conduct a stronger integration of social applications, media use and information access
- ❑ Envision a language-transparent and media-transparent Web for the user, regardless of age, education, profession, culture, language proficiency, technical skills
- ❑ **All information produced in the EU should be made accessible to EU citizens in their language**
- ❑ Truly multilingual online communication, commerce, education
 - Cross-lingual information access to web and to media in all languages
 - Support 200 to 1,000 languages
 - Crosslingual queries, Q&A, ICA, MT in social media (chats, tweets, etc)
 - Multimedia Multi-language subtitling
 - Personalized to language capacities (minorities, migrants)
 - Making document understandable
 - Adapt to language proficiency, technical skills



A Natural and Inclusive Interaction

- ❑ Digital communication doesn't have any border
- ❑ The global technical infrastructure exists, but it lacks natural interaction
- ❑ LT can help overcome the invisible borders in human-to-human / human-to-machine interaction, while addressing the social and cultural diversity in Europe
 - Natural interaction with agents and robots
 - Autonomous, Self-learning, Context aware, Personalized agents and robots
 - Able to achieve low-level tasks (emails, voice messaging, telephone calls)
 - Assistive applications – technologies that help
 - Reduced motor control, e-accessibility for the disabled (incl. Sign Languages)
 - Support education, rehabilitation and training for the disabled
 - Cross-lingual e-learning
 - Remote cooperation of students all over the world
 - CALL
 - Cross-lingual meeting assistants
 - With mimicking speech-to-speech translation abilities
 - Headphone and microphone-free
 - Including MT of slide presentations



- ❑ Information is growing without limits
 - Growth of high quality, free information
- ❑ However almost impossible to manage information: information value is therefore in danger
- ❑ LT can help delivering personalized Information/Knowledge access and management
- ❑ Make users able to aggregate, evaluate and share information
 - Federated multilingual audiovisual search
 - Including OCR, ASR
 - Personalized information assistant
 - Search similar documents over the web
 - Aid to the decision
 - Handle information overload
 - Life logging
 - Capture conversation into a gazeteer of concepts



Summary

- ❑ Language technology is a key enabling technology for a multilingual Europe.
- ❑ Language technology can contribute to the European economic growth and stability.
- ❑ Efforts could be devoted among 3 axes:
 - Language-transparent Web and Media
 - Natural and Inclusive Interaction
 - Efficient Information Management
- ❑ We now have to transform those visions into a Strategic Research Agenda (SRA), in the Horizon of 2020



Questions and Answers



Thank you very much.

office@meta-net.eu

<http://www.meta-net.eu>

<http://www.facebook.com/META.Alliance>