Belonging and Exclusion in the Internet Era: Estonian Case

Pille Runnel, University of Tartu, piller@jrnI.ut.ee
Pille Vengerfeldt, University of Tartu pille@jrnl.ut.ee

Abstract

When looking at the users side, language is the most important tool for structuring information and communicating in the new media environment. Mastering the language often determines access to different resources on the Internet.

For a small not-English speaking nation there is local Estonian language based Internet and international Internet where the rest of the world is. Our aim with this paper is to show that “digital divide” is not only about accessing the environment with cables and computers, but it is also about accessing resources and services available on the Internet.

Introduction

While looking at the issues of digital divide, most common is to speak of access to technical tools – of cables and computers. The aim of this article is to discuss other dimensions of digital divide – we will look at belonging and exclusion in the digital era depending on the language.

From our pilot study, covering media texts, public forums, mind maps, interviews and focus groups we gathered a number of dominant concepts of the language-based equality.

We will build our discussion by first briefly introducing the Estonian Internet and language situation. Also a brief overview of different concepts emerged from the pilot study will be given. After that we will try to look a little bit deeper into two cases: divide
on personal level while consuming news of international event and the level of society around the issues of electronic translation. We will conclude the paper by setting some hypothesis for the further research.

**Estonian case**

Estonia is a small country situated in the Northwest of Europe, close to the Scandinavian countries. With 1,4 million inhabitants, there are a little more than one million Estonian speakers in the world. Estonian territory is about 45 thousand square kilometers with 1/3 of the population gathered around capital area.

**Access to Internet**

The size of the nation also defines the character of the digital divide – with a small territory and with the small number of people, technological access is relatively easy to provide. By now there are cable connections to most of the towns in Estonia and smaller places have Internet connection using radio technology. Public Internet Access Points have been established all over the country, as a government-initiated and funded project, providing free (or very cheap) Internet access. Therefore we might conclude that cables and computers are not so much of an issue in the case of Estonia. On winter 2002, 39% of the Estonian population claimed to have used Internet within last six month (see figure 1, from EMOR, 2002).
Language

There are one million Estonian speakers. Besides being the readers of the texts in Estonian, they are the only resource possibly ever going to provide local language content for the Internet. This is obviously not enough to cover all aspects necessary. There will never be enough human power to translate all the interesting or needed information media - the news, books, movies, studies etc - available. Therefore digital divide is drawn between those who master foreign languages and those who don’t.

As Estonia belonged to the Soviet Union for almost fifty years, it had also a strong influence on the foreign language skills of the population. Learning German and English dropped radically after the Second World War, during the Soviet period learning Russian was compulsory in schools. Within last ten years, the situation is changing again – the primary foreign language taught at schools is usually English. But the majority of the population, especially the middle-aged, has very poor command of other foreign languages besides Russian (see figure 2 from Estonian Office of Statistic, 2000).
Concepts of language based inequality emerging from the pilot study

A combination of five methods was used for the pilot study: analysis of media texts, analysis of public Internet forums, focus group interviews, expert interviews and mind-maps.

At the media texts the issues, important on the society level, prevail. People discuss the question of joining EU not only in connection with preserving the language and culture, but also on the point of providing sufficient numbers of Euro-interpreters. As Estonians like to think of themselves as generally techno oriented people, the issues of language
technology – voice recognition and the complexity of Estonian language to be taught to the computer, have also been issues in Estonian media.

On the Internet-based public forum the issue of language technology has been discussed even more. Here the contradictions between levels of society and the individual level are probably most clearly visible. Discussions in those public forums are not always on the highest analytical level and most accurate, but even the fact that this kind of public debates spring from the Internet users own initiative, is important. For discussing various language related issues, a special Usenet news group named ee.keel (ee.language) has been created.

We have carried out two focus group interviews and a number of additional interviews, combined with mind mapping exercises. In the focus groups of computer experts, the English language based computing and a more general threat of disappearing Estonian speaking Internet were raised. It was claimed that the space created through the Estonian language is simply too small to have the economical efficiency essentially necessary for the computer related industry. Estonian language is linguistically so complicated both for developing translation and voice recognition software that there would be not enough devoted people (and money) to keep Estonian language capable to survive in the electronic environment.

In the focus group of university students these problems were not recognized. The students move freely around both in English and Estonian speaking electronic space or at least have good enough command of English language to master the necessities of their everyday needs.

As a result of the pilot study made with 50 informants drawing free association based mind-maps it might be said that people had “global” and “local” Internet. Among the other categories, people mentioned on their mind maps, there were also the web addresses that were their access portals to Internet. It seems that the notion of global
media space does not exist to everyone. The digital divide excludes those who due to their language skills have to limit their Internet experience only to the local web-pages.

From our data, it is possible to distinguish two dimensions describing belonging and exclusion: the national-personal dimension, showing, who is the subject of suffering the exclusion and the tool-content dimension of language, showing what causes the exclusion (see figure 3).

![Diagram](image)

Figure 3 – The dimensions of the language based inequality.

At the personal level the exclusion from news, access to international materials that are important only to that person can be seen. Usually the person was made by both media and respondents to feel responsible for the exclusion. On the other hand – in the matters that need to be understood by larger groups, the responsibility was seen to lay on the state or companies to avoid exclusion on that level.

Similarly with the tool-content opposition, it was expected that the state or companies feel responsible for the exclusion and provide local language speaking computers, but the content side is not so much on anyone’s responsibility. Rather it is expected that the state takes care of producing automatic translation tools to cope with the problem. We also interviewed an expert-activist who voluntarily has been translating one freeware operating system into Estonian for several years. He pointed to the general lack of government and public support to this kind of projects.
Do you speak English? Internet and computing are essentially in English

Internet is pre-dominantly English speaking. According to the OECD report, almost 90% of secure connections are in English (German, French, Japanese and Spanish follow with tight around 1%) (See figure 4 from OECD 2000). Also the computing is essentially made in English. Just for the few last years there are different character codes for Internet pages – before that it was only 7-bit character mapping that was sufficient for English, but it lacked all kind of diaecritical signs essential for most of the other languages.

CASE 1: Access to news
A study supporting the idea that talking about language based digital divide is very important, was also a news reception study about the September 2001 crisis of terrorists attacking US. While watching the cyber community interactions it was possible to see, that people with sufficient English skills were using various sources of information,
available on the Internet, whereas others had to limit their information to what was provided by the local media. This gives us two dimensions to the divide – first the limited range of information was visible. For the second, the time factor was very important - as the information had to be translated and edited to the local language, it took time to reach the people. So those people who do not speak English were longer in informational vacuum. In normal situations half an hour would make not so much difference, but in a this kind of crisis situation it the language barrier turned into the tool of excluding people, who depended on media to be provided with updated information.

One more dimension becomes relevant in connection of news. While consuming the media a person depends on the editor, who makes the content selection. With news that need to be translated from the other language there will always be two or even more mediators selecting and editing the news content for the audience. It does not just leave out information, but it also takes time, that at some point can become a serious hindrance. An example of this could be again the September 11th crisis news, in the case of which due to the hurrying in translation, several mistakes were made in publishing them. For instance an editor from central news service made a serious mistake when translating “the flight number eleven” to “eleven planes hijacked”. Therefore the whole Estonian media spread incorrect information during the first day. So the people who spoke English, had not only more variety of news and quicker access to them, but also more accurate news.

CASE 2: Language technologies – electronic translation
The other case we will look at has sprung from Estonian media and is in its nature more public. This is the issue of the language technologies, voice recognition and automatic translation. The later has been presented in Estonian media as a part of the general topic of the enlargement of the European Union. The potentially large number of documents that will need to be translated into Estonian in a very short time has raised the question of need for machine-based translation. As Estonia is a European Union candidate country, it has to be able to take the responsibility of translation big amounts of EU-connected information, and documentation to the local language. It is obvious that there are not enough human resources to handle professionally enough this amount of materials.
Computer-based translation systems are becoming reality for the dominant European languages already, for example in the form of the automatic translation, offered by search engines. Nothing similar has been done for Estonian language. The public debate in media is mostly centered on the government’s responsibility to handle the resources for this kind of scientific development work.

The question is not so much about machine-based translating systems eliminating people from their job, as there will still be the need for human operator to check the text’s accuracy and edit it, also it is doubtful that oral translation will soon be replaced by machines.

Should Estonia ignore the need for this kind of translating system, the threat of being excluded from European Union even if we are formally included, is definitely there.

**Conclusions**

As seen from our empirical material – the language-based inequality can create important groups of belonging and exclusion for both – individuals and larger groups on more general society level. This is nothing essentially new when we think about it in general terms of experiences of societies. We can find the similar inequalities in most of the fields of economic, social and public life. But the reason we feel this being especially important for the Internet era is, that being a former soviet country, Estonia has attached many hopes and dreams to technological innovations, including the Internet. As Lauristin and Vihalemm have said, the rapid introduction of e-society in Estonia has given a chance for a small post-communist country to makes a “shortcut” to the advanced post-industrial society and to make use of the new opportunities for economic and social development (Lauristin&Vihalemm, 2001). This illustrates and summarizes well the expectations people have towards Internet. But if the language is dividing people to those who have possibility to take the “shortcut” to post-industrial society and to those who can’t, is there a reason to expect Internet in general to “save” the society? Or has a small
nation enough resources to invest into bridging the language-based exclusion to take that leap into information-based society? Where the information is plenty and always in some other language?

We have mostly been arguing the point of potential exclusion via language based divide and it may seem that we support the spreading of the English. Media scholars who look at the language issues related to media have talked about the threat of losing the language (see, for instance Findahl, 1989). The fear of English diminishing other, smaller languages at least here in Europe, is often talked about. We are not arguing against that. Rather we hope that with this paper we have managed to show, how much of the electronic world people miss when they do not speak English. Instead of English domination, it is more of a question, people voluntarily choosing to speak and write in English for not to be excluded.

A question, that might be interesting for the further research, is that is there a possibility that people are willing to adapt to English as there is already so much content for English speaking web or are people ready to be content with local language based web just for the sake of keeping the language alive? Will the English pre-dominance continue as most of the financing for computing sciences come from English speaking countries? Will any of this be changed if the fully operating electronic translation systems will be available?

References