

Endangered languages in the digital age: Supporting and studying digital language use in them¹

Abstract

Language endangerment and the need for language revitalization efforts have been at the forefront of linguistic thinking and action for over two decades now. In the same period of time, digital technology and new ways of language use mediated by it (through email, the internet and social media etc.) have become predominant. This paper discusses how language endangerment, language revitalization, and digital language use intersect, and what kind of responsibilities this leaves for the linguistic community concerned with bilingual, minority, and endangered language use in today's world. It is argued that linguists can assist endangered language communities to preserve their languages, among other ways, by focusing on language use in the digital domain (e.g. in social media): by aiding communities in creating and maintaining a digital presence and in creating internet content, as well as, more broadly, by investigating language use in the endangered language in the digital domain, since this domain has become of primary importance in recent years, so understanding how endangered languages are used in it can offer crucial insight also into how they can be further supported.

Keywords: endangered languages, minority languages, language revitalization, digital domain, digital natives

1. Introduction

Language endangerment (and the subsequent need to revitalize languages) has been in the collective consciousness of linguists for a little over twenty years now, since the early 1990s. It is now common knowledge that many of the world's languages are severely endangered, the world's linguistic diversity threatened, and that language revitalization efforts are needed to offset these trends. The same time period is also the time in which the spread of personal computers and other new digital technology have profoundly changed ways of communication and language use.

In this paper, I discuss some ways in which issues of language endangerment and digital language use intersect, and what kinds of new responsibilities these leave for the linguistic community concerned with bilingual, minority, and endangered language use in today's world.

2. Language endangerment

¹ „Research in the FinUgRevita Project reported on in this paper has been supported by the Hungarian National Research Fund, OTKA, under grant number FNN 107883.

The issue of language endangerment as such was first raised in a focused way and using this term in 1991, at the “Endangered languages symposium” of the Linguistic Society of America with the written versions of the talks published in the March 1992 issue of *Language*.

The sobering statistics presented by Krauss (1992) about the alarmingly low numbers of indigenous minority languages still taught to children in various countries and continents, the relatively low number of “safe” languages in the world, and the shocking prediction that about 90% of the world’s 6,000 languages are likely to become extinct in the 21st century –are, by now, well-known in the linguistic community although still, more than two decades later, relatively little known outside it. Krauss’s paper and the other papers in the journal issue basically outlined the agenda for linguistic work to be done in the face of this impending loss in documenting and attempting to preserve and revitalize endangered languages in their communities, and in various sorts of activism – monitoring, lobbying, and raising publicity.

The past 20 plus years have seen an overwhelming response to this call, with a proliferation of grants, scholarships and opportunities to study endangered languages, contributing to an ever increasing amount of work by a growing number of linguists and educators, a mushrooming of case studies and books on the topic and on individual languages, and, in general, a growing awareness about the issues of endangered languages and their communities.

As a very important offshoot of the development of awareness about endangered languages, a number of macro-level classifications of language endangeredness have been since proposed that measure the vitality/endangerment of languages, for the sake of cataloguing and categorizing situations and, thus, gaining a better overall view, as well as for having a checklist of characteristics and tasks to do in each particular type (or category) of situations – cf. Fishman’s Graded Intergenerational Dislocation Scale (GIDS; Fishman 1991 and 2004), later reworked by Lewis and Simons (2010) into EGIDS (an Extended GIDS), and

UNESCO’s *Methodology for assessing language vitality and endangerment*, developed a decade ago (UNESCO 2003) by a group of internationally recognized scholars and experts in the field, and its periodically updated *Atlas of languages in danger* (UNESCO 2010). (The fact that UNESCO took up the issue of meticulously assessing and publishing data on a great range of situations signals, at least on the political level, of the commitment of the UN and its organizations to dealing with the whole range of issues of language endangerment and to recognizing the need to do something about it.) As far as micro-level tools of measuring endangerment and vitality are concerned, an excellent recent example is the European Language Vitality Barometer (EuLaViBar) developed by the EL DIA project (Spiliopoulou Åkermark et al.

2013), which is a complex tool and toolkit developed and made available to researchers.

3. Language use in the digital age

The development of new digital technology and equipment in the past three decades brought to the developed world personal (and portable) computers, the internet, digital cameras, audio and video recorders, digital music players, and mobile phones, as well as various technological crossbreeds such as tablets and smartphones. With these came new ways of communication and of language use, such as email, the web, listservs, newsgroups, texting (or SMS, as it is better known in Europe), instant messaging, chat, blogging and voice-over-IP (best known to most through Skype) (Palfrey and Gasser 2008). In the early 2000s, Web 2.0 technologies brought user-generated content on the web, as well as various dynamic and collaborative ways of producing content (wikis, the best-known example of which is probably Wikipedia) (Herring 2012).

At the level of people's language use, these developments brought an increase of digitally mediated language use (perhaps in some cases at the expense of face-to-face communication, but in many respects new ways of keeping in touch with those outside of physical reach), both oral and written. While the increase of digitally mediated spoken language use (done via mobile phones or Skyping etc.) might seem all too obvious, the latter also turns out to be true (Baron 2008) once we think of all the reading and writing we do through email, texting, instant messaging, blogging, and using social media – we use these to keep in touch, to express opinions and ask questions, document personal experiences and share these with others, as well as to share practical knowledge.

Much of social media involves intensive language use, at least three of the six types of it (as identified by Kaplan and Haenlein 2010) being language based: (1) collaborative projects (such as wikis), (2) blogs and microblogs (a prime example of the latter is Twitter), and (3) social networking sites (such as the US-based Facebook, the Russia-based VKontakte, or Hungary's former iwiw), while (4) content communities (cf. YouTube for video, various sites like Picasa and Panoramio for photos), (5) virtual game worlds (such as World of Warcraft) and (6) virtual social worlds (like Second Life) tend to be primarily image based, using language only as a secondary and auxiliary mode of communication.

Probably the most striking development of the digital age is that of what has been called, following Prensky (2001a and 2001b), the emergence of digital natives, i.e. a new age- and experience-based population (rather than a generation, cf. Palfrey and Gasser 2008: 14) of young people who have only known a world that is digital and never lived in a household without computers in it. In addition to a new, digital experience in life, digital natives are characterized

by craving interactivity, multitasking, and thinking and processing information fundamentally differently than others. They are differentiated, on the one hand, in their own generation from their less privileged peers on the other side of the digital divide, for whom their access to broadband, skills and digital literacy is unavailable, and, on the other hand, from digital immigrants, i.e. the generations who saw digital technology and phenomena develop in their lifetime and who learn to adapt to their environment but “retain an accent” (Prensky 2001a). The first wave of digital natives reached college-age in the early 2000s in the US and probably somewhat later in other parts of the world.

The characteristics of language use that set digital natives apart from all others are, according to Palfry and Gasser (2008), that they use *all* the forms of digital communication (ibid. 2–4), use (only) digital technology to access information (ibid. 6) (in a way of knowledge acquisition sometimes taken to the extreme and limiting stance of “if it’s not on the web, it doesn’t exist”), thrive on collaborative ways of doing things (from creating internet content to many other things in their lives, cf. ibid. 111–129), and using digital tools and communication in their activism (ibid. 255). Examples of the latter are numerous by now, in most parts of the world, cf. the digitally based and self-organizing activists of Obama’s 2008 presidential campaign or of the 2011 Egyptian revolution, anti-government student demonstrations in Hungary in the winter of 2012/2013 (which were streamed via websites for the rest of the country to see and follow), the crowd-sourcing efforts used in an attempt to identify the Boston marathon bombers in April 2013. An outstanding example of such activism in the context of an endangered language, Udmurt (a Finno-Ugric language spoken by slightly over 300 thousand speakers, or about 60% of the ethnic Udmurts southwest of the Ural mountains in Russia; cf. 2010 Russian Census), is the internet-based contest for creating neologisms in Udmurt (to replace at least some of the loanwords from Russian, the result of a massive influx of Russian borrowings in the sphere of business- and computer-related terminology with neologisms using Udmurt stems and derivational affixes) through the Uralistica website in 2013 (Malykh et al. 2013).

4. Language endangerment in the digital age

And while much of the language use in the digital domain (such as texting, commenting in social media and in forum discussions etc.) is ephemeral in being produced for the moment and for an immediate, short-term effect, they gain greater importance if we look at issues of language endangerment in the context of our digitalized (or just simply digital) world. With so much of the communication and language use today happening in the digital domain, it seems inevitable that those languages in which content is not available and which cannot be used in digitally mediated writing (in texting, emailing, blogging

etc.) are facing digital endangerment, i.e. non-existence in the digital domain. It is also easy to see that the languages that face such digital endangerment are most likely those languages which are endangered anyway: the languages of economically disadvantaged, socially dominated, and/or culturally repressed indigenous populations and other minority groups (of the latter, probably especially those without a supporting majority community elsewhere).

The vision of digital language death has been forcefully argued by András Kornai in a series of recent talks (e.g. Kornai 2012) and an influential paper (Kornai 2013), in which he states that language endangerment in the digital age has to be viewed, following the traditional language death aspects of loss of function, of prestige, and of competence, in terms of the loss of digital function, loss of digital prestige, and loss of digital competence: that is, whether digital functions can be performed in the language, whether digital content giving the language prestige exists (otherwise, “if it’s not on the web, it doesn’t exist”), and whether there are digital natives (and digital immigrants) who are able to use the language in the digital domain. Basing his argumentation on a complex statistical analysis of languages and their digital functions and support, Kornai convincingly claims that there are likely about 250 “digital survivor” languages (out of the 6,000 existing today). As an important and telling measure of how a language is likely to fare, he posits the existence (or lack) of Wikipedia in the given language (since the Wikipedia is always among the first active communities of language users today).

Kornai’s powerful vision has a very simple assumption at its base, namely, that language use in the digital domain (i.e. all the digitally mediated communication that we do via our phones and computers) has become an enormously important domain of language use in recent years. A good indication of this is, on the international political level, that UNESCO’s Commission on Communication and Information has made “Linguistic diversity and multilingualism on the internet” one of its main themes and focus areas of concern (for a freely available publication, see Vannini and Le Crosnier 2012).

Given the significance and prominence of digital language use, speakers of endangered languages (who are, minimally, bilingual in their minority language and the language of the dominant majority pretty much by definition), digital natives and digital immigrants alike, face a crucial choice in their lives. This choice is, of course, not whether they will use language digitally or not, but what language or languages they use digitally: will it be the minority language (which is possible if, for instance, software support – including fonts, keyboards, operating systems, and applications etc. – is available in/for it), or will it be the majority language (which, due to its position, is more likely to have all this necessary digital support)?

The consequences of the digital status quo and the importance of the digital presence of languages for their (digital) survival places several major responsibilities on linguists, language activists, and speaker communities of endangered languages and their work beyond the by now widely accepted need to document and archive endangered languages digitally.

5. Responsibilities for endangered languages in the digital age

5.1. Digital presence and creation of content

The basic realm where responsibility lies is the creation and/or strengthening of the digital presence of endangered languages, that is, the creation of digital content available in the endangered languages. This is, of course, highly dependent on the endangered language speakers and their willingness and opportunity to create and support the digital presence of their own language.

The first of the conditions, willingness, is something which only speakers of the endangered language can help themselves with: as Thomason (forthcoming) puts it, “no linguist or anthropologist, no matter how passionate s/he is about an endangered language, can or should attempt to force a community to try to save their endangered language”.

The examples of, for instance, Siberian indigenous languages with speaker communities where basic computer hardware and software is present, are encouraging in this respect. For instance, despite the fact that Mansi, a Finno-Ugric language with less than 1,000 speakers (among the 12 thousand ethnically Mansi population dominant in Russia) and spoken in western Siberia, in Russia, is a severely endangered language (UNESCO 2010), the Mansi language bi-weekly newspaper, *Luima Seripos*, is available online (<http://www.khanty-yasang.ru/luima-seripos/>) as a result of the work of a staff of a handful of enthusiastic journalists. In another, much stronger language, Sakha (also known as Yakut), an indigenous language spoken in eastern Siberia, in Russia, by about half a million speakers – which is about half of the local ethnically Sakha population – active blogging with lots of meme-creation and storytelling in Sakha has been described by Basharina (2013). The most popular social networking site in Russia, VKontakte (from *v kontakte* “in contact”; <http://vk.com/>), which has gained more than 140 million users (70 million of them active users, and 39 million of them using it daily) since its founding in 2006, is available in 70 languages, 3 of them (Russian, English and Ukrainian) official, and the rest being non-official user-generated and –created translations into languages spoken in Russia: it is available in three minority indigenous Finno-Ugric languages, Udmurt, Mari (with about 360 thousand speakers among the 547 thousand ethnic Mari) and Erzya (one of the varieties of Mordvin, which has close to 400 thousand speaker among an ethnically Mordvin population of 744 thousand) (for the demographics, see the 2010 Russian census).

Creation of digitally available indigenous language content can also be done by people from outside the language community. An example to this effect is a newly funded project based on the cooperation of Finnish and Hungarian linguists, for instance, aimed at creating freely accessible online lexical resources and support for the community based generation of translated material on the web (for the Wikipedia and social networking sites) for several endangered Finno-Ugric languages (Oszkó 2014).

Whether speakers of endangered languages have the opportunity to create their own digital presence is, primarily, an economic issue, and, secondarily, an issue of technological support. Given basic economic privileges and general availability of digital equipment of some sort (in the form of computers and/or smartphones), both linguistic and computer technological prerequisites need to be met, the former including a standardized orthography, and the latter including keyboards, fonts, and (ideally) computational language tools (like a spellchecker, dictionary, and phrase- and sentence-level tools).

The case of Mansi is an interesting example again: Mansi, specifically its Northern dialect, the one in which the newspaper *Luima Seripos* is written, has at least two ways of Cyrillic-based spelling (the only alphabet currently used by speakers of the language), one of them employed in scholarly publications (such as published grammars and dictionaries of the language), the other one used popularly in newspapers and on the internet. With no computational language tools available for the language at all, another recently started linguistics project, FinUgRevita (see <http://www.ieas-szeged.hu/finugrevita/>) carried out by linguists in Finland and Hungary has as one of its goals the creation of a morphological analyzer and other computational language tools for the speakers of the Mansi in order to support their written digital use of their indigenous language. The linguists involved in it had to make a decision about which orthography to use (they opted for the popular one) and yet have to overcome issues of how to encode special Mansi orthographic characters with their unique diacritic markings.

The creation of computational language tools is a labor intensive and professionally challenging process when it comes to small and endangered minority languages. In addition to the issue of existence or lack of a standardized orthography, even more basic questions that might arise is the lack of a standard variety in the language – if it exists in several varieties none of which is regarded as standard – and a lexicon heavily populated with loanwords if, as is the case with many minority languages under heavy language contact from the dominant majority language, it is undergoing massive borrowing. It is clear that, if computational language tools such as spellcheckers and analyzers etc. are to be practical for the speakers themselves, they need to reflect the real language use of the users of the language, i.e. include and be able to handle

loanwords. This, in turn, might, however, be frowned upon by people with normative attitudes towards their endangered language (usually educators), who would like to preserve it in its “purest” form, “uncontaminated” by borrowings from the majority language.

In an interesting sociolinguistic twist of things, the lack of a standard variety in the case of endangered languages can also be a blessing in disguise: it can basically mean a lack of societally “agreed-on” forms to be stigmatized – if there is no artificially elevated standard variety with forms held in high prestige, there are no stigmatized varieties and/or stigmatized forms either, as is the case in a community of Faetar speakers studied by Nagy (2009) in southern Italy.

5.2. Investigating language use in the digital domain

Language use in the digital context has become a very important domain of language use, as has been argued above. As such, it deserves the attention of descriptive linguistic and sociolinguistic work in both the monolingual and the bilingual/multilingual context. The beginnings of it in the monolingual context date back to the late 1990s, when computer-mediated communication, or CMC, was first systematically studied. In addition to general works on language and the digital domain (cf. Crystal 2006, Baron 2008), a new line of investigation has been set by one of the pioneers of the field, Susan Herring, in computer mediated discourse analysis (CMDA) (cf. Herring 2004 and 2012) and others studying discourse aspects of language use in the new media (cf. papers in Thurlow and Mroczek 2011 and Tannen and Tester 2012) as well as pragmatic aspects of language use in the digital domain (Herring et al. 2013).

Bilingual language use in the digital domain – which brings us to close relevance for endangered languages (which, by definition, exist in a bi- or multilingual context) – in general has also been discussed by Herring (cf. Danet and Herring 2007), as well as in the context of minority language use in Wales by Daniel Cunliffe and his colleagues (Cunliffe 2004, 2009, Cunliffe et al. 2009, 2010, and Cunliffe and Honeycutt 2010), and in Udmurtia by Pischlöger (2013 and 2014). A recent book, Jones and Jongbloed (2013) studies specifically minority languages and social media.

In addition to works with a specific focus on language use in the digital domain like the above, a growing body of literature on bilingualism is discussing issues of language use in the digital domain: for instance, five chapters of *The Blackwell guide to research methods in bilingualism and multilingualism*, edited by Li and Moyer in 2008, discuss such issues, although these discussions are mostly limited to macro perspectives on language use in the media rather than focusing on the fact that digital language use is an important sociolinguistic domain of bilingual speakers’ language use. In an important handbook of bilin-

gualism, Bhatia and Ritchie (2006, and its extended and updated 2012 edition), the discussion is similarly limited to language in the media and advertising.

In stark contrast, in the most important recent handbook of the field of endangered language study, Austin and Sallabank (2011), nine of the 29 chapters, or roughly one-third of the book, engage in issues of digital technology and endangered languages, and one chapter, Holton's "The role of information technology in supporting minority and endangered languages" deals specifically with central issues.

The EuLaViBar tool developed by the ELDIA project (<http://www.eldia-project.org/>, cf. Spiliopoulou Åkermark et al. 2013) to assess the vitality or endangeredness of languages contains in its questionnaire several questions aimed at finding out whether internet content and computer software is available in the minority language in question, and also several questions that enquire about the speakers' minority language use in emailing, texting, using social media and games.

As sporadic case studies of bilingualism touching on speakers' language use in the digital domain actually already done in the 21st century demonstrate, important insight about these communities' shift or maintenance patterns can be gained by studying what role is played by which language of the community in patterns of the language use in the given community. Also, since so much of digitally mediated language use is written, investigation into the internet-based language use of minority language users can provide invaluable insight into the written language use of such communities after a long tradition of focusing on the study of spoken language use.

In an early macro-sociolinguistic study of American Hungarians' language use in Toledo, Ohio, Polgár (2001) showed that language use on the internet was an important factor of language maintenance already half a dozen years after widespread use of it started in the US. In her study of lexical interference in American and Canadian Hungarian scout homepages, Botka (2003) demonstrated a specific contact-induced effect, a proliferation of loanshift creations in this context (which did not appear to be characteristic of non-internet-based language use in the same communities), made common probably because they were less "visible" than loanwords and, therefore, less regarded as deplorable written language use in this semi-formal (back then Web 1.0) context. Investigating contact effects in Canadian Estonian in online forum discussions of the Toronto paper *Eesti Elu*, Janurik (2008) noticed heavy structural effects of English without heavy lexical borrowing – quite a contradiction in terms of Thomason and Kaufman's (2008) borrowing scale, which would predict heavy lexical borrowing in cases where heavy structural borrowing is present – and concluded that, in writing, it is probably easier for forum participants to self-monitor the use of lexical borrowing than that of structural borrowing.

Finally, in a recent macro-sociolinguistic overview of language use by Canadian Hungarians in Hamilton, Ontario, Huber (2013) found digital language use to be a highly important domain of heritage language use, and also observed that while only slightly more than half of the (typically older) first-generation speaker members of this community used Hungarian on the internet, almost all of the (typically younger) second-generation speakers did so: even though, as Canadian-born bilingual speakers they could have been predicted to use Hungarian less in this context than their Hungarian-born parents, because of their generational affiliation as digital natives, they in fact used it more – presumably helping their language maintenance while doing so. In this case (and presumably, in many other cases also) failing to look at the language use of the community members in the digital domain would have provided an incomplete, basically misleading picture of heritage language use in the community – one where heritage language use patterns conform to the pattern found in other North American Hungarian communities before, where immigrants use the language more extensively than their children in all domains. Huber's close look at heritage language use in the digital domain, however, uncovered the opposite tendency in this domain.

In her detailed investigation of, specifically, the online use of Mansi, Horváth (2013) reports its use and presence on the internet in the following: a presence in news sites and webpages of businesses, and wide use in video-sharing sites and social networking. The typical creators of Mansi internet content are, unsurprisingly, young, urban speakers of the language. Horváth observes that, in fact, online texts form the majority of written material now available in Mansi, which also provide more information on things modern Mansi than printed sources do. The effect of online use of Mansi is definitely positive: it attracts the youngest generation of speakers and makes them use the language, it is raising the prestige of the language for them (“if it's on the web, it's cool!”), and it is also increasing the number of heritage language speakers of the language who are willing to learn and use the language in everyday life. The patterns of online use of Mansi are also substantially different from offline language use patterns observed before: heritage language speakers are identifying with the speaker community much more willingly, whereas native speakers happily switch the language of conversation from Russian to Mansi and freely use it in the publicity of online situations, use Mansi even with non-native speakers and in their presence (for instance, in social network discussions), and create online discourse in Mansi with other native speakers – things that they would typically not do offline. On the whole, then, Horváth (2013) concludes, the online presence of the Mansi language encourages speakers (of whatever level of proficiency) to use the language in digital space, offers “a new way for speakers to re-create

their speaker-community”, and also, happily for researchers, provides a rich source of texts in the language, nonexistent before.

In addition to actually doing research on digital language use, I believe it to be an important responsibility of older, “digital immigrant” linguists to encourage their (most likely) digital native students and advisees to do such research as well. Young digital native linguists are likely to have a better grasp on being present and active in the digital domain and, through this experience, to have insights into it that their digital immigrant professors and advisors do not have. Such insights, I believe, need to be foregrounded and, if coupled with a basic understanding of sociolinguistic, discourse and bilingualism issues, be encouraged to go in innovative directions.

6. The informality of digital language use

In addition to the technological issues (such as fonts, keyboards, etc.) and the sociolinguistic issue of availability of standardized orthography referred to above, another issue of sociolinguistic nature is highly relevant to digital language use in endangered languages: notably, the informality of such language use.

As we know from personal experience, much (or even most) of the digitally mediated writing is fairly informal: texting, tweeting, chatting, writing blogs and comments on social media and in user forums are definitely very informal ways of writing. Even though emailing can be quite formal – it can be used to apply for jobs, funding, and conferences – most of it is likely to occur also on the informal end of the spectrum. A similar tendency toward informal language use is also visible, in languages with the dichotomy of formal, *vous*-type or address vs. informal *tu*-type of address, in the predominant use of *tu* address in digitally mediated contexts such as forum discussions, auction and shopping sites (like EBay in the English language context), dating sites, and social network comments.

This is basically good news for endangered languages since informal written language use is easier to achieve and more important to pursue than formal: both Fishman (1991, 2004), in his discussion of GIDS and the stage where literacy acquisition should be introduced, and Reyhner (1999), writing about the importance of informal written language use in reversing language shift in indigenous language communities in North America, stipulate that H(igh) language use in the indigenous language should not be pursued before literacy in L(ow) language use is achieved during the process of reversing language shift.

In terms of written digitally mediated language use this means that the various informal forms of writing required of indigenous language use in the digital domain are likely to be a good “playground” for language users in their progression of developing literacy in their indigenous/minority language. If we also consider the fact that a lot of blogging and social media presence is

about sharing personal opinions and experiences as well as building and/or emphasizing group identity and relationships between members of the groups (Kietzman et al. 2011), indigenous/minority language use in such digital genres appears as the ideal first stage of written language use which also has the added benefit of strengthening indigenous/minority community life, ties and consciousness. A nice example of this is found in Basharina's 2013 insights regarding the use of the Sakha language in online storytelling of both traditional and new genres, and, through this, the local appropriation of digital technology as "evidence of cultural and linguistic adaptation, modernization, and vitality" of the Sakha community and the good service it brings to the revitalization of their language. With its emphasis on group and community ties, social media can also help maintain social relationships in contexts where face-to-face interaction is hindered by distances between speakers or groups, or by weather conditions in especially cold climates most of the year (like the Sakhas' in eastern Siberia or the Saamis' in northern Scandinavia).

7. Conclusion

In this paper I have argued that (especially written) language use in the digital domain has become a highly important domain of language use in today's world, for monolingual and bilingual speakers alike. In terms of the latter, however, this is of primary significance: the focal issue is what language bilingual speakers – and especially speakers of minority and/or endangered languages – will use in the digital domain, and if they choose to use their minority/endangered language, whether they are supported in doing so. The forms of support range from gaining a better understanding of the characteristics of bilingual digital language use to actually providing technical support for endangered languages.

Szeged

Anna Fenyvesi

References

- Austin, Peter K. – Julia Sallabank (eds). 2011. *The Cambridge handbook of endangered languages*. Cambridge: Cambridge University Press.
- Baron, Naomi S. 2008. *Always on: Language in an online and mobile world*. Oxford: Oxford University Press.
- Basharina, Olga. 2013. *The extent of the Sakha language on the internet: Digital genres and language revitalization*. Paper presented at MinorEuRus conference, Helsinki, Dec. 15–17, 2013.
- Bhatia, Tej K. – William C. Ritchie (eds.) 2006. *The handbook of bilingualism*. Oxford: Blackwell.
- Bhatia, Tej K. – William C. Ritchie (eds.) 2012. *The handbook of bilingualism and multilingualism*. Oxford: Blackwell.

- Botka, Gabriella. 2003. *Language contact online: An analysis of lexical interference in American and Canadian Hungarian scout homepages*. Szeged: University of Szeged, MA thesis.
- Crystal, David. 2006. *Language and the internet*. Cambridge: Cambridge University Press.
- Cunliffe, Daniel – C. Honeycutt. 2010. The use of the Welsh language on Facebook: An initial investigation. *Information, Communication and Society*, 13, 226–248.
- Cunliffe, Daniel – N. Pearson – S. Richards. 2010. E-commerce and minority languages: A Welsh perspective. In H. Kelly-Holmes – G. Mautner (eds.): *Language and the market*. Basingstoke: Palgrave Macmillan: 135–147
- Cunliffe, Daniel. 2004. Promoting minority language use on bilingual web sites. *Mercator media forum* 7, 42–53).
- Cunliffe, Daniel. 2009a. The Welsh language on the internet: Linguistic resistance in the age of the network society. In G. Goggin – M. McLelland (eds.): *Internationalizing internet studies: Beyond Anglophone paradigms*. New York: Routledge: 96–111.
- Cunliffe, Daniel. 2009b. Bilingual websites in jurisdictions requiring minority language use: effective implementation of policies and guidelines. In S. Pertot – T. Priestly – C. Williams (eds.): *Rights, promotion and integration issues for minority languages in Europe*. Basingstoke: Palgrave: 74–86.
- Danet, Brenda – Susan C. Herring. 2007. *The multilingual internet: Language, culture and communication online*. Oxford: Oxford University Press.
- Fishman, Joshua. 1991. *Reversing language shift: Theory and practice of assistance to threatened languages*. Clevedon, UK: Multilingual Matters.
- Fishman, Joshua. 2004. Language maintenance, language shift, and reversing language shift. In Tej K. Bhatia – William C. Ritchie (eds.): *The handbook of bilingualism*. Oxford: Blackwell: 406–436.
- Herring, Susan C. 2012. Discourse in Web 2.0: Familiar, reconfigured, and emergent. In Tannen, Deborah – Anne Marie Tester (eds.): *Georgetown University roundtable on languages and linguistics 2011: Discourse 2.0: Language and new media*. Washington, DC: Georgetown University Press: 1–27.
- Herring, Susan C. – Dieter Stein – Tuija Virtanen (eds.) 2013. *Pragmatics of computer-mediated communication*. Berlin: Mouton de Gruyter.
- Horváth, Csilla 2013. *Occupy the Web! The presence of Ob-Ugric languages on Web 2.0 domains*. Paper presented at the 14th International Conference on Minority Languages in Graz, Austria, September 11–14, 2013.
- Huber 2013. *Intergenerational transmission of Hungarian as a heritage language in Canada: The macrosociolinguistics of the Hungarian community in Hamilton, Ontario*. Szeged: University of Szeged, BA thesis.

- Janurik, Boglárka. 2008. *Contact induced features of Canadian Estonian: An analysis of forum discussions of the online newspaper Eesti Elu*. Szeged: University of Szeged, MA thesis.
- Kaplan, Andreas M. – Michael Haenlein. 2010. Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53, 59–68.
- Kietzmann, J.H. – K. Hermkens, I. P. McCarthy – B.S. Silvestre. 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54, 241–251.
- Kornai, András. 2012. *Digital language death*. Talk presented at the 8th Conference of PhD students in Computer Science, Szeged, Hungary, June 28–30, 2012.
- Kornai, András. 2013. Digital language death. *PLoS ONE*, 8, e77056. doi:10.1371/journal.pone.0077056.
- Krauss, Michael. 1992. The world's languages in crisis. *Language*, 68, 4–10.
- Lewis, M. Paul – Gary F. Simons. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique*, 2, 103–120.
- Malykh, Artyom – Aleksey Shklyayev – Olga Urasinova. 2013. Веме берыктон как социальная технология. Paper presented at the MinorEuRus conference in Helsinki, Finland, December 16–18, 2013.
- Nagy, Naomi. 2009. The challenges of less commonly studied languages: Writing a sociogrammar of Faetar. In Stanford, James N. – Dennis R. Preston (eds.): *Variation in indigenous minority languages*. Amsterdam: Benjamins: 397–417.
- Oszkó, Beatrix. 2014. *Introducing the "Finno-Ugric digital natives: Linguistic support for Finno-Ugric digital communities in generating online content" project*. Talk presented at the Annual Finno-Ugric Seminar, Budapest, January 3, 2014.
- Palfrey, John – Gasser, Urs. 2008. *Born digital: Understanding the first generation of digital natives*. New York: Basic Books.
- Pischlöger, Christian. 2013. *Udmurt and Welsh in Web 2.0: Language revitalization and social media*. Paper presented at MinorEuRus conference, Helsinki, Dec. 15–17, 2013.
- Pischlöger, Christian. 2014. *Udmurt on the web: Urban Udmurts resisting language shift*. Paper presented at the 7th Budapest Uralic Workshop, February 3–5, 2014.
- Polgár, Etelka. 2001. *Language maintenance and language shift: A sociolinguistic analysis of a Hungarian-American community*. Szeged: University of Szeged, MA thesis.
- Prensky, Marc. 2001a. Digital natives, digital immigrants. *On the Horizon*, 9(5), 1–6.

- Prensky, Marc. 2001b. Digital natives, digital immigrants, Part 2: Do they really think differently? *On the Horizon*, 9(6): 1–6.
- Reyhner, Jon. 1999. Introduction: Some basics of indigenous language revitalization. In Reyhner, Jon – Gina Cantoni – Robert N. St. Clair – Evangeline Parsons Yazzie (eds.): *Revitalizing indigenous languages*. Flagstaff: Northern Arizona University: v-xx.
- Russian census 2010. <http://www.perepis-2010.ru/results_of_the_census>. [22 April 2014].
- Spiliopoulou Åkermark, Sia – Johanna Laakso – Anneli Sarhimaa – Reetta Toivanen – Eva Kühhirt – Kari Djerf. 2013. ELDIA EuLaViBar. <<http://www.eldia-project.org/index.php/eulavibar>> [20 January 2014].
- Tannen, Deborah – Anne Marie Tester, eds. 2012. *Georgetown University Round Table on Languages and Linguistics 2011: Discourse 2.0: Language and new media*. Washington, DC: Georgetown University Press.
- Thomason, Sarah Grey – Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thomason, Sarah Grey. *Forthcoming*. *Endangered languages*. Cambridge: Cambridge University Press.
- Thurlow, Crispin – Kristine Mroczek (eds.) 2011. *Digital discourse: Language in the new media*. New York: Oxford University Press.
- UNESCO. 2003. A methodology for assessing language vitality and endangerment. <<http://www.unesco.org/new/en/culture/themes/endangered-languages/language-vitality/>> [20 January 2014].
- UNESCO. 2010. Atlas of the world's languages in danger. <<http://www.unesco.org/new/en/culture/themes/endangered-languages/atlas-of-languages-in-danger/>> [19 April 2014].
- Vannini, Laurent – Hervé Le Crosnier (eds.) 2012. NET.LANG: Towards a multilingual cyberspace. Caen: C&F editions, available <http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/netlang_EN_pdfedition.pdf> [22 April 2014].
- Wei, Li – Melissa Moyer. 2008. *The Blackwell guide to research methods in bilingualism and multilingualism*. Oxford: Blackwell.