EMPIRICAL PAPER

Multilingualism and Technology: A Review of Developments in Digital Communication from Monolingualism to Idiolingualism

Helen Kelly-Holmes*

University of Limerick
*Corresponding author. E-mail: Helen.kelly.holmes@ul.ie

Abstract
The focus in this article is on the evolution of language and technology in relation to multilingualism, in particular on how multilingual provision has developed in tandem with the development of the internet and the World Wide Web (WWW). In trying to understand how multilingualism has evolved, it is also necessary to understand how the technical aspects of digital technology as well as the politico-economic dimensions to that technology have changed. Four distinct periods emerge in the development: monolingualism, multilingualism, hyperlingualism, and idiolingualism. Monolingualism covers the origins of the internet and later the WWW as monolingual spaces. This was followed by a long period that charts the slow but gradual development of increased language provision and what I am terming “partial multilingualism.” Multilingualism expanded substantially, potentially limitlessly, with the development of Web 2.0. This has involved the diversification of online spaces to the point of “hyperlingualism.” I argue that we are still in this hyperlingual phase, but alongside it, a new phase is developing, that of “idiolingualism” as a result of mass linguistic customization. In this article, I discuss these phases, paying attention to both their technical and economic contexts, as well as their implications for linguistic diversity online and in wider society.

Introduction
While much research has focused on language practices online (see, e.g., Androutsopoulos, 2006a; Herring & Androutsopoulos, 2015; Herring & Paolillo, 2006; Heyd, 2016; Thurlow & Mroczek, 2011) and the multilingual content that digital technology facilitates (see, e.g., Androutsopoulos, 2006b, 2015; Barton & Lee, 2013; Danet & Herring, 2007a, 2007b; Deumert, 2014a; Deumert & Masinyana, 2008; Dyers & Davids, 2015; Lee, 2016; Leppänen, 2012; Leppänen, Pitkänen-Huhta, Piirainen-Marsh, Nikula, & Peuronen, 2009; Warschauer, Said, & Zohry, 2002), it can be argued that there has been less focus on the organization of languages online and how this organization is shaped by digital technology (although cf. Callahan & Herring, 2012; Danet & Herring, 2003, 2007a; De Bres, 2015; Ivkovic & Lotherington, 2009; Kelly-Holmes, 2006a, 2013; Wright, 2006; see also John, 2013, for a critical overview of studies of online multilingualism).

© Cambridge University Press 2019
My interest in this article is in the online organization of languages, which I understand as the ways in which languages are made available, supported, presented, and managed in digital spaces, and how this has changed as technology has developed and as the functions and economics of the web have evolved. So, questions of interest include how many languages are available for users of particular sites, applications, and technologies; how much content is available in which particular languages; what is the role of English and other dominant global languages; what hierarchies and orderings are present; and what is available for people whose main or preferred language is not provided. This is, of course, not the whole story of what is happening in terms of contemporary online language, far from it; it is just one component. However, I would argue that online language organization is highly significant in people’s everyday lives, though not everywhere in the world and not for everyone—there are, of course, still very great digital gaps and content divides between speakers and language communities (Mafu, 2004; Norris, 2001; Warschauer, 2004). But taking all of these caveats into consideration, these digital spaces and practices are influential and formative for much of the world’s population. Most of our information about language is mediated today—mainly as a result of how much of our language is mediated (see, e.g., Kelly-Holmes & Milani, 2013)—and this is increasing all the time.

The World Wide Web (WWW) was conceived of and constituted as a “social machine” (Hendler & Berners-Lee, 2010), that is a machine that could grow and learn, and this growth and learning have taken place as a result of the “feeding” that the machine gets from human usage, human input. Likewise, it follows that the web is a “sociolinguistic machine”—fueled by online language practices and choices and by widespread and commonsense ideologies and beliefs about language. My focus over the past 15 years or so has been on online language organization and management, how languages are presented and represented, and how speakers are categorized (Kelly-Holmes, 2004, 2006a, 2006b, 2013). I would argue that the ways in which this is done in online spaces give us important information about languages and speakers in our contemporary world, and that these spaces and their organizational processes are important sites for analysis. Online management of multilingualism is, of course, a type of language policy (Kelly-Holmes, 2015), and I will be making use of the theories and models of language policy in this article, in particular, Johnson (2013), Shohamy (2006), and Schiffman (1996).

**Digital Communication**

I am using the term “digital communication” to refer to the connectedness that has resulted from developments such as the internet and the WWW and their mainstreaming into all aspects of everyday life. Of course, language, in its basic sense as a code or system, is at the heart of digital communication: Shannon and Weaver’s (1949) encoding-decoding model is the basis for the development of digital communication. A message is encoded or translated into a bit stream (series of binaries, i.e., ones and zeros) to enable it to be transported quickly, cheaply, and widely to another point or points where it is decoded from that bit stream into the same message (Barry, Lee, & Messerschmitt, 2012). This involves transmitting modes that are inherently digital (e.g., text) as well as converting modes that are analog to digital (e.g., speech and image) before transmission. In essence, this is what facilitates email, social networking, blogging, e-commerce, online news, and everything we do in our everyday lives that involves PCs, mobile phones, smart devices, wearable technology, and the like. This
is because it enables not just one-to-many, monologic communication over a much wider (potentially limitless) geographical area, but it also enables peer-to-peer communication. Of course, as linguists, numerous questions emerge from this simple concept of digital transmission and translation of modes: Whose speech? Whose writing? How? With what consequences? Who is heard or seen? Who is not?

The main technological developments that have taken place since the switch from analog communication to digital communication have been largely concerned with making this encoding process as efficient as possible—squashing the maximum amount of information into the smallest bit stream possible—and ensuring the fidelity of the message that is sent, in other words, ensuring that the message received looks the same as the message sent (Viterbi & Omura, 2013). However, these technological developments have had massive implications for language and for multilingualism. In the next sections, I try to map the development of multilingualism to the technical and economic evolution of digital technology and, in so doing, hope to address John’s (2013) assertion that “technological blindness … is prevalent in the literature on the multilingual internet” (p. 322). As we shall see confirmed in the analysis below, language, technology, and the economy are all bound up together, and this interdependence is not something new (Kleinrock, 2010).

From my review of the development of multilingualism and digital communication, four stages or trends emerge, and I deal with each of these in detail below: monolingualism, multilingualism, hyperlingualism, and idiolingualism. Table 1 provides an overview of the eras, their key technical and economic dimensions, and their features in terms of linguistic diversity.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Technological/economic context</th>
<th>Implications for multilingualism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingualism</td>
<td>Birth of the internet and later the WWW, beginnings of e-commerce</td>
<td>Multilingualism not visible; dominance of English</td>
</tr>
<tr>
<td>Multilingualism</td>
<td>Parallel, partial multilingualism; introduction, promotion, stabilization of non-ASCII supported alphabets; mainstreaming of e-commerce</td>
<td>Increasing presence of “big” languages, growing visibility of more, but limited language; multilingual provision</td>
</tr>
<tr>
<td>Hyperlingualism</td>
<td>Web 2.0, long tail markets, gift economy</td>
<td>Multilingual display, greater visibility of “small” languages, greater levels of linguistic differentiation</td>
</tr>
<tr>
<td>Idiolingualism</td>
<td>Algorithmic mass individualization; artificial intelligence; linguistic filter bubble; market of one</td>
<td>Intensified but isolated hyperlingualism; Gradual erasure of multilingualism; linguistic isolation</td>
</tr>
</tbody>
</table>

**Monolingualism**

I am characterizing—probably crudely—the early phase of digital communication as monolingual. To understand this first era of monolingualism, it is essential to understand the technical, economic, and ideological origins of the internet and the WWW. Although histories of the internet and WWW are often largely biased toward the
early years of development (from the 1950s onward), inevitably, if perhaps unintention-
anly, casting developments since the WWW emerged in 1991 merely as refinements
(see, e.g., www.computerhistory.org), the impact of these refinements has been much
more major in terms of facilitating multilingual communication online.

What is perhaps most telling about early accounts of digital communication is the
complete absence of any consideration of language—apart from programming lan-
guages (see, e.g., Connolly, 2000; Kleinrock, 2010; www.computerhistory.org). There
would appear to be two main reasons behind this: First, as Danet and Herring
(2003) pointed out “early planners of the internet were generally American, and were
implicitly thinking only about how to facilitate communication in English; they there-
fore did not anticipate the problems that might arise when speakers of other languages
tried to communicate online” (n.p.). As an example, the original 19-node ARPANET, a
forerunner of the internet, connected only U.S.-based institutions (Kleinrock, 2010). So,
because the institutional members of this group only wished to facilitate computer-
to-computer communication between themselves, they did not think about language
problems. And, on a wider scale, because the multilingual nature of most of the
world was not part of the speech community of the original pioneers of the internet,
the ideology and culture of monolingualism became part of the culture of the internet
and later the web. This is an important point, because technological advances and
breakthroughs occur in particular ideological and cultural spaces, and the shape of
those technological advances bears the imprint of those cultural and ideological
norms. As Wei and Kolko (2005) pointed out, “the computer and network technologies
that support the internet were largely developed in a Western society … the Web was
not designed explicitly for use in isolated, peripheral cultures, nor did such cultures
participate in its inception or current design” (p. 206). A practical manifestation of
this monolingual culture is the fact that the encoding to facilitate bit streaming was
based on ASCII, which in turn is premised on the Roman alphabet and English
phonology.

In terms of language policy, we can see this early period as an example of a de facto
or implicit (Schifman, 1996) policy of English-only monolingualism. Nothing was
written down or explicitly stated, and no policy was directly formulated. Instead, the
monolingual culture of the founders simply became the commonsense language prac-
tices and in this way a type of “covert” (Shohamy, 2006) policy. However, over and
above this implicit or de facto policy of English-only monolingualism, which was
assumed or simply occurred, though not in any official or explicit way, it can be argued
that monolingualism was also an overt policy and explicit ideology of the early internet,
because of the aim of its founders that it be a global network. As Bekele (2014) high-
lighted, “the internet we love and know today was by design built as a global network
but not a multilingual network” (n.p.). As he rightly pointed out, there is an implied
contradiction between a global internet that comprises a “single language, common
standard that facilitates connectivity in a seamless manner” (Bekele, 2014) and an inter-
net that facilitates multilingualism and so is different everywhere.

In addition, the resistance to top-down governance and authority in the manage-
ment of the internet also explains why there might be resistance to any kind of inter-
ventionist language policy to support greater linguistic diversity. The internet was
deliberately designed to be nonhierarchical and to have parity between all links and
nodes—there is no center and thus no overall authority that might impose language
policies such as promoting multilingualism. Likewise with the web, which is the content
to the physical infrastructure of the internet. In the words of Tim Berners-Lee (1998),
who invented the WWW in 1990, “The dream behind the Web is of a common information space in which we communicate by sharing information. Its universality is essential: the fact that a hypertext link can point to anything, be it personal, local or global, be it draft or highly polished” (n.p.). Berners-Lee’s use of “we” is significant here—he is talking about the small group of developers of the web; however, this “we” dictated the monolingual design of the web, with wide-reaching implications for the many.

Despite these monolingualist origins, work was starting to happen to make the internet and the web multilingual, and the impetus for this was a growing recognition of the economic potential of digital communication. The development of Unicode, happening in parallel with the development of the internet and WWW, shows that the multilingual problem was of concern from at least the 1980s. The creation of Unicode (around 1988) was designed to solve this problem by enabling “unique, universal, and uniform character encoding” across computers and networks (https://www.unicode.org/history/earlyyears.html). The developers of Unicode had a common dream that “people would not be shut off from computers because of their native language” (https://www.unicode.org/history/earlyyears.html). The adoption of the full Unicode character set by domain name and internet security group Verisign in the early 2000s can be seen to mark the start of online multilingualism.

Finally, it should be pointed out that the dominance of English language monolingualism in the early internet and web has been disputed. For example, Pimienta, Prado, and Blanco (2009) asserted that the role of English has been overstated and this overstatement has in fact contributed to enhancing the status of English as the language of the web, thus in turn further consolidating its status as global lingua franca (De Swaan, 2010). Up to the late 2000s, it was generally accepted that 80% of WWW content was in English, while growing numbers of users were not first-language English speakers (Yang, Lay, Tsao, Liou, & Lin, 2006). However, Pimienta et al. (2009) argued that, while in 1995 the presence of English was considered to be above 90%, over the following 12 years, this decreased to about 40% in 2008. Part of the reason for this overestimation of English was due to the strong bias toward English in the language recognition algorithms that were used to estimate the online presence of different languages. So, according to their account, English language dominance began to wane in the early 2000s, despite “the broad perception of a massive, pervasive and stable English dominance” (Pimienta et al., 2009, p. 28) in popular and academic accounts.

Despite disparities in estimations of linguistic diversity, the key dimension to this first phase is that English language learning and/or translation were required in order to participate in the internet and later in the early days of the web, up to the early 2000s and certainly beyond for speakers of “smaller” languages. What happened to change this is documented in the next section.

Multilingualism

I am characterizing the second era as one of multilingualism, but in reality this is better described as a phase of partial and of parallel multilingualism. While studies showed that certain languages like French had a “fair representation” (Pimienta et al., 2009) in the 1990s, 2003 is seen as a turning point in the multilingual development of digital communication, with the introduction of internationalized domain names. Also, a demographic shift had taken place, whereby “already by 2003 roughly two-thirds of all internet users were nonnative speakers of English” (CyberAtlas, 2003). Thus, native
speakers of English no longer dominated, as they did for many years (Danet & Herring, 2007b, n.p.).

In fact, between 2000 and 2018, while the online presence and use of Arabic are estimated to have grown by over 8000%, in the same period the growth of English is estimated at only 650%. The other main growth languages in that time have been Russian, Malaysian and Indonesian, Chinese, Portuguese, Spanish, and French, all of which have grown more than English (www.internetworldstatistics.com). The reason, of course, for the apparent decline of English is the fact that it was already firmly established and its major growth had already happened. All of the other languages were coming from a much lower base. Internet World Statistics also gives information about internet penetration, which shows the relationship between the percentage of internet users speaking a language and the estimated total number of speakers of that language who are online. The statistics are, of course, crude, as this quote from www.internetworldstatistics.com shows: “many people are bilingual or multilingual but here we assign only one language per person in order to have all the language users add up to the total world population” (n.p.). This approach sums up the era of multilingualism and shows its main characteristics as being based on “parallel multilingualisms” (Heller, 2001), that is, bounded languages kept separate from each other with users being assigned to only one particular language largely on the basis of geographic data. Internetworldstatistics.com also acknowledged the limitation of not taking literacy rates into account in its statistics, and by so doing, also pointed to one of the failings of the early era of digital communication, that is, the extent to which participation was limited by the ability to read and later write in a particular language. Written text was the main mode of the early internet and web, with visual, aural, and multiple simultaneous modes only coming later. Apart from other issues to do with alphabets and resources, this dominance of the written mode almost certainly also impacted on minority languages and oral languages, which contributed to a “content divide” (Rogers, 2001) and may have contributed to further peripheralization of these languages (see e.g., Wei & Kolko, 2005).

I am also naming this the era of partial multilingualism, as it mainly saw growth in “big” languages. The top 10 languages that emerged in this era are all part of De Swaan’s (2013) category of “supercentral” languages. These are defined as major lingua franca languages or languages with large number of speakers and/or economic power: English, Chinese, Spanish, Arabic, Portuguese, Indonesian and Malay, French, Japanese, Russian, and German. Two of De Swaan’s supercentral languages that had not transitioned this status to the web at this point in time are Kiswahili and Hindi. Arabic took some time to catch up with the other major languages, possibly due to the lack of a key government investing in the online development of the language, whereas languages with a politically and economically strong and obvious “motherland” were resourced to develop content. Provision by governments of the infrastructure as well as the content is, of course, crucial in developing online languages. Where governments restrict this provision or access to new technology by their own citizens, this will inevitably impact on the online presence of the language. The “hypercentral” status of English (in De Swaan’s terms, it holds the global language system together, and it alone has the function of global lingua franca) in the world language system model would seem to have been confirmed by the development of the internet and the web and by early online multilingualism. Despite criticisms of the overestimation of English dominance, a crude study using the Google search engine in 2008 found that 51% of space on the internet was occupied by English, with the next nearest rival being Chinese at about
9%, German 5%, Spanish 4.1%, and French 3.5% (www.internetworldstatistics.com).

It is also not just the content in a particular language, but where it is produced. For example, the UK and Germany produced more French language content than all of the African Francophone countries in 2007, while 97% of content generated in Africa was generated in South Africa in 2007 (www.internetworldstatistics.com).

Despite the universalist founding ideology of the internet and later the web, as discussed above, which implies a de facto language policy of a global lingua franca, it was also recognized by digital designers that language, along with technical and security aspects, “is an important interface that affects people’s intention to use or access the system” (Yang et al., 2006, p. 226). Equally, while there is a lack of an official center and of a governing authority for the internet and the web, there is nonetheless implicit language governance. UNESCO, for example, has been observing and tracking multilingualism online since the early 2000s and has worked with ICANN, the internet domain name governance body, to develop multilingual international domain names. In addition, official and de facto language policies have been extended to the web in many parts of the world (see, e.g., Kelly-Holmes, 2006a). Likewise, the million book projects and other investments in alphabet stabilization across platforms by various national and regional authorities to ensure that their languages do not fall victim to the content divide can be seen as responses to de facto monolingual English language policies globally.

The key technical development of this era, as mentioned above, was the stabilization of non-ASCII-supported alphabets. However, the changing economic model of the internet and WWW also encouraged the growth of multilingualism. The mainstreaming of e-commerce led to the necessity and the demand by consumers for greater levels of service in a variety of different languages. And it is no coincidence that as the economic potential of digital communication has expanded, so too has multilingualism—albeit in an unbalanced way and subject to the caveats outlined above.

In terms of linguistic diversity, the era of parallel and partial multilingualism entailed the appearance of a limited number of supercentral languages only at first (as discussed above), thus increasing the visibility of multilingualism online and leading in general to the provision of more languages. This involved the evolution of linguistic gateways on many sites, where the user was forced to choose from limited linguistic offerings. Such gateways comprised high levels of localization into supercentral languages and European central or national languages, with correspondingly low levels of localization for Africa, Asia, and other parts of the world (see Kelly-Holmes, 2006b, 2013). Following a period of using the Roman alphabet to represent languages not normally written in it (see Koutsogiannis & Mitsikopoulou, 2003, in relation to Greek, and also John, 2013, for an overview of the evolution of Hebrew online), different alphabets also started to appear, although it took time before these could be viewed in the same way across a range of platforms and users. As Wei and Kolko (2005) pointed out “the clash between technology and local culture may be obvious immediately when the local language is based on a non-Latin alphabet” (p. 210). It should also be highlighted that the use of the Roman alphabet to represent languages written with other alphabets is still common, particularly among youth (Elhija, 2014).

It is not just the provision of online multilingualism, but also the organization of multilingualism that has evolved over time. And this changing organization was also driven by economic, technical, and language ideological factors. Trends included the provision of flags and maps to represent language provision and prompt users to make their language choices. As the reliability of non-ASCII scripts improved and...
multi-orthographism increased in online contexts, a new trend of multilingual displays on websites emerged, including the display of multiple alphabets (see Kelly-Holmes, 2013).

The era of partial, parallel multilingualism had mixed results for minoritized languages. On the one hand, English had a “head start” (Block, 2004, p. 23), as we have seen above; on the other hand, the new technology allowed for linguistic communities dispersed over time and space to come together in new ways (Cunliffe & Herring, 2005; McLaughlin, 2014; Sheyholislami, 2010; Wright, 2006). A good example is the development of top-level domains such as .cat, which allowed for a pan-national online community of Catalan speakers to “come together” (Atkinson, 2006).

What also happened with growing multilingualism, stabilization of non-ASCII-based alphabets and growing online content, particularly with the greater enablement of user-generated, bottom-up content, was that there was much more diversity in the creators of content—so not only did more languages appear, but this in turn led to greater linguistic and cultural diversity in content (Wei & Kolko, 2005).

Growing multilingualism also coincided with the fragmentation of the internet and the WWW, and this era effectively meant the end of one global internet experienced in the same way and shared by all. Along with the fragmentation in terms of functions and domains online came a growth in boundaries and demarcated spaces on the web, and language was in fact a key tool in this. The multilingualism era brought with it a lot of boundary work (Barth, 1998), whereby countries and regions were increasingly separated from each other, with language being one of the main ways of doing this. As a result, the web, largely perceived as a global and boundaryless medium, instead became a place where the essentialized links between language, territory, and national identity in fact became reinforced—for some, not all. This proliferation and essentializing of boundaries was set to grow exponentially, and this growth is what I am referring to as the hyperlingual era.

**Hyperlingualism**

As digital communication became increasingly participatory, an increasingly large and potentially limitless number of people could be involved in digital multilingualism (see, e.g., Androutsopoulos, 2010). In was during this time that the term “Web 2.0” was coined to describe this growing peer-to-peer and multiway communication. So, while in the era of monologic digital communication, a relatively small group of people made decisions about which languages to provide, the peer-to-peer and two-way nature of the Web 2.0 era changed the model of online multilingualism. These new ways of interfacing and contributing content have made crowdsourcing possible (see Estellés-Arolas & González-Ladrón-De-Guevara, 2012), and this in particular has had a profound effect on multilingualism, which prompts my use of the term “hyperlingualism” to describe this particular era. I am borrowing the term from Pauwels (2014), who has used hyperlingualism to refer to the proliferation of languages in global cities in the contemporary era. For her, hyperlingualism is seen as the linguistic outcome of super-diversity. For me, hyperlingualism is both the product of technological affordances that mean that it is now possible to offer an unlimited number of languages in digital contexts and the ideology underpinning the enablement of these possibilities. These technological and ideological shifts have resulted in a kind of hyper-differentiation in relation to language, whereby more and more languages are achieving their own bounded spaces and places of use on the web and in other digital contexts.
A key development here, as mentioned above, has been crowdsourcing. In the multilingualism era, the website provider decided on a limited number of necessary and/or desirable languages that should, in addition to English, be offered and then commissioned a language professional to localize the resources. This could be described as a “multilingual provision” model. With crowdsourcing in the hyperlingual era, the localization work is done by users—sometimes called “produsers” or “prosumers”—and technology makes the choice of languages largely irrelevant. A particularly well-known and early example was the Facebook translations app (see Lenihan, 2011, 2013, for an extensive study). The social network Facebook was originally available only in English; however, when the company decided to become multilingual to expand its market, instead of going down the multilingual provision route, it implemented a crowdsourcing solution, developing an app that users could download. The availability of languages on the app for localization by users was determined by user demand. The translation community consisted of users who participated in the app, and this community voted individual users’ translations up and down. While there was a final moderation step by Facebook, the process overall bypassed language professionals. There was no need to “prove” any kind of language ownership or expertise in order to participate in the translation of Facebook into that particular language.

In some ways, this was a refreshing view of language as not being bound to issues of territory, citizenship, or competence. Another interesting language ideological change promoted by the way in which the app worked was that the process was the same for all languages: Very “big” languages such as French were subject to the same crowdsourced process as very “small” languages like Irish. In reality, the French and Spanish translations happened much more quickly because of the sheer numbers of people involved; however, the process was still the same, unlike the multilingualism era, when “important” central and supercentral languages such as French and Spanish would automatically have been provided and would have been expected by consumers, whereas “small” languages would not have been. In the multilingual provision era, it would have simply been too costly to employ language professionals to provide these languages to a very small number of consumers.

Now, with users doing all the work, the number of languages that can be provided becomes potentially limitless. The “gift economy” (Cheal, 2015) is the basis of much multilingualism in online contexts as well as much contemporary content on sites such as YouTube, TripAdvisor, and the like. Users gift their work for free, and this work, like the translation of Facebook, becomes the property of the company. The market basis of much of the current online economy is the working consumer (Cova & Dalli, 2009), who creates value for products through free work (e.g., writing reviews, voting, commenting, liking, sharing, creating content, etc.). As we can see, much of this work is language work (Thurlow, 2018), albeit done by nonprofessionals. An interesting question for us is why people gift language work for free. For many, particularly speakers of minoritized languages, they are committed to reviving and revitalizing as well as documenting and developing that language (see, e.g., Budzise-Weaver, Chen, & Mitchell, 2012; McMonagle, Cunliffe, Jongbloed-Faber, & Jarvis, 2018; Nichols, Witten, Keegan, Bainbridge, & Dewsnip, 2005; Sperlich, 2005). For others, there is a reward in terms of displaying linguistic knowledge, and the many language ideological debates (Blommaert, 2010) that take place in online contexts are evidence of this (see, e.g., De Bres & Belling, 2015; Lenihan, 2011; Sherman, & Švelch, 2015; Yazan, 2015).

While national-level boundary-making and identity processes are still happening through language work, in contrast to the multilingualism era, hyperlingualism has
seen something of a decline in the importance of the language and territory link. This has given way to a focus on satisfying individual rather than collective needs, in line with contemporary market models. For example, the way in which the digital economy has developed has resulted in viable “long tail markets” (Anderson, 2006) for more and more languages. In the multilingual provision era, the cost of developing linguistically localized products, services, and information for speakers of “smaller” languages was not considered viable. Now, as a result of technology and economic models that are a consequence of the advances in digital technology (e.g., crowdsourcing and the gift economy), it is no longer just mainstream markets and large languages that are economically attractive. The “outlier” languages in economic terms have now become attractive and viable. For example, the browser Mozilla Firefox has made it possible through crowdsourcing for local, primarily oral languages to go online across Africa. Rather than waiting for language rights to be granted by the state, minoritized groups can now, as a result of technology, create online media spaces for themselves. This, of course, has positive and negative consequences (see, e.g., Deumert, 2014b, for a review of relevant issues, and Jones, Cunliffe, & Honeycutt, 2013, in relation to Welsh on Twitter). On the one hand, obviously, more languages go online and fewer languages get left behind in the increasing digitization of the world. On the other hand, if language communities bypass the state in the pursuit of language rights, opting instead for consumer rights from global media corporations, it is increasingly unlikely that the state will see a need to provide education and related rights in these languages. The development of online spaces for what have been in many cases primarily oral languages can herald a whole new era of revitalization and revival and can bring them to new and modern domains (Cunliffe & Herring, 2005; De Schryver, 2002; Mafu, 2004; Pasch, 2008). However, could it also lead to an imbalance and very little change if it is not matched by offline development of corpus, acquisition, and status planning (see Cunliffe, 2019; Jones & Uribe-Jongbloed, 2012)?

**Idiolingualism**

The final era I would like to discuss, what I am terming “idiolingualism,” is happening in parallel with and can be seen to be the logical development of hyperlingualism. Idiolingualism involves intensified but isolated hyperlingualism and the increasing tailoring and personalizing of online language provision. The hyperlingualism era is in part a product of the contemporary focus on “excessive variety” in marketing, whereby “companies are diversifying their products to respond to almost every conceivable customer taste” (Gilmore & Pine, 2000).

An example here is the language learning application DuoLingo, which brings together many of the features of hyperlingualism. The aim of the app is to have as many different languages on offer as possible; the model by which the languages are learned depends on crowdsourcing and gift economics—the language work was gifted by speakers of the language, and the slogan is “learn a language for free forever.” The normal hierarchy of large and small languages is distorted by the long-tail economic model—for example, Irish is one of the largest languages being learned on the site—and the focus is on individual and customized language learning. The app offers “personalized learning.” In the company’s own words, “Everyone learns in different ways. For the first time in history, we can analyse how millions of people learn at once to create the most effective educational system possible and tailor it to each
student. Our ultimate goal is to give everyone access to a private tutor experience through technology” (www.duolingo.com). And, of course, all of this is free.

This stated goal points to the differentiating technological dimension of the current era, which can be characterized in terms of idiolingualism. The possibility to “analyse how millions of people learn” languages and to use this to tailor the experience to individual learner needs and different language learning styles is afforded through algorithmic mass customization and driven by competition in the marketplace for language learning apps (www.duolingo.com). As mentioned earlier, Tim Berners-Lee described the web as a social machine that is capable of learning—and this capability is the basis of artificial intelligence. By extension, it is also a sociolinguistic machine that learns language through the language work that is carried out by users in online contexts, as outlined above. Google Translate, for example, is based on the concept of the gift economy as well as sociolinguistic learning by the Google browser and other tools. The language that is put onto the web by ordinary people in their everyday use of the web constitutes a massive multilingual corpus that is then is manipulated by Google Translate to find appropriate translations. As with crowdsourcing, it usurps the role of the language professional and provides individualized translations to users based on their particular translation needs at any particular time and in any particular place. As the web has become increasingly multilingual, with more content in more languages, Google Translate has improved significantly. To use Tim Berners-Lee’s metaphor, we feed the web whenever we use it, and we are feeding it with language also. This language allows it to learn, to become intelligent.

The ultimate goal of contemporary marketing is often seen as the “market of one,” where there is “mass-customization, making products only in response to actual demand” (Gilmore & Pine, 2000). We can also see this at work in language provision in current digital communication. Gaming is an interesting example (see, e.g., Leppänen & Piirainen-Marsh, 2009). Minecraft is just one game that allows individuals to “make your own language.” Predictive text in messaging and other applications, which previously relied on a dictionary and prescriptivist approach by correcting the user’s language in line with standard language and grammar, instead now learns the user’s language preferences and behaviors and makes suggestions based on these. This new proliferation of tailored and individualized multilingualism or hyperlingualism frequently takes place alongside a minimizing of exposure to “real” multilingualism through mass linguistic customization. There is an illusion of increased choice, but, in fact, choice is reduced. Users no longer need to pick their language preference or put up with the languages provided. Instead, their language preferences are already known, or they are guesstimated based on their geographical location, although it is worth keeping in mind that this is still largely based on parallel monolingualisms. As a result of this, we are, it has been argued, being cocooned in linguistic “filter bubbles” (Pariser, 2011), and we are being steered through the global, multilingual web in a monolingual bubble; we see only the language it is assumed we want to see based on past linguistic behavior and choices, and we are cocooned from other languages.

So, it could be posited that we have come to an era where we can live in relative linguistic isolation—communicating with and through our own linguistic preferences and habits. Linguistic isolation is analogous to intellectual isolation, which, it has been argued, is the result of algorithmic filtering and customization (Pariser, 2011). Wearable language devices are the next logical step in this process. While seamless, real-time translation that erases multilingualism has long been a standard of science fiction dramas (e.g., the “babel fish” of Douglas Adams’s Hitchhiker’s Guide to the Galaxy),
Australian company Lingmo International launched the Translate One2One in 2017 with a promise “to understand both the words and context being spoken, to deliver a near-instant translation” (Nelson, 2017, n.p.). This wearable device has an earpiece and a microphone and is reported by testers to work very well “across English, Japanese, French, Italian, Spanish, Brazilian Portuguese, German and Chinese.” It is also claimed to be able to cope with multiple accents and dialects. In addition, while even more languages are available with internet connection, it can also be used offline, which offers particularly good potential across the digital divide, according to its founder: “As the first device on the market for language translation using AI that does not rely on connectivity to operate, it offers significant potential for its unique application across airlines, foreign government relations and even not-for-profits working in remote areas” (Nelson, 2017, n.p.). The futuristic and fantastical notions of science fiction seem set to come true in the form of real-time, simultaneous, and personalized translation due to advances in artificial intelligence, machine learning, wearable technology, and real-time data analysis.

Conclusion

*In the future humans will facilitate machine-to-machine communication.*

—Anonymous

*Human v machine is no more.*

—Deloitte advertising billboard, seen in Dublin Airport, Ireland, August 2018

Although I have presented a simplistic chronological progression of these four eras (see Table 1 for summary), they are better understood as overlapping and contiguous to a certain extent. Multilingual provision is still taking place (see, e.g., Berezkina, 2018) and will continue to take place alongside hyperlingualism and idiolingualism. And, while there is no reason—technically—for multilingual provision not to happen, monolingual ideologies do still pervade much technological development, although this is changing as the players in these industries become somewhat more linguistically diversified. And still, for many in the world, accessing digital technology, where that access is even physically or economically possible, involves language shift to a “bigger” language. Of course, in a brief review such as this, it is not possible to cover all aspects of the topic, and hopefully the article provides inspiration for future studies that could examine, for example, the influence of Python (as suggested by one reviewer), which is an important programming language and based on English grammar, in terms of exclusion of certain agents and shaping of content. Another suggestion was to review the intersection of online multilingualism with the growth in the digital accessibility movement and associated policies.

What became apparent to me in writing this review is that the eras have speeded up; the times between the eras and speed of change within the eras have all increased. What is also clear in bringing together these developments is that from the original starting point of a global internet, there has, within a relatively short period of time (not quite 30 years), been an increasing and increasingly rapid level of fragmentation of that internet and WWW first to nationally based spaces, then on gradually to individually styled spaces. And language—together with changing technology and economic models—has been a key agent of this.

Finally, what are the implications of these changes for language professionals, for language learning and applied linguistics in general? What we have seen in this brief
overview of the development of multilingualism in digital communication is that the role of the professional linguist has significantly changed as digital communication has become increasingly multilingual. The current era of globalization is characterized by declining trust in professionals and a do-it-yourself approach to everything from medical problems to make-up and hair styles, and language is no exception here. This decline is facilitated by the technology and changing economic models, which are in turn enabled by this changing technology. Although it is unlikely that we will ever dispense with human language teachers, language learners no longer need real, live teachers, and instead personalized learning is mediated by technology, as the DuoLingo example shows. Likewise with translation, there has been a shift from trained professional translators to do-it-yourself real-time translation. In relation to our own role as linguists, it may be time to rethink our own language. While we do have a much increased focus on commodification and on mediated language—and I think we would all agree that online language constitutes mainstream applied and sociolinguistics—our nomenclature has yet to adapt. We still add qualifiers like online, computer-mediated, and commodified; but in some ways, and again this is not everywhere or all of the time, it is the offline, non-computer-mediated, face-to-face language that takes place outside of noncommodified spaces that is almost the outlier now in many contexts and situations. And yet, we still think of the former as extraordinary. Perhaps our qualifiers and norms need to change, even as these new developments raise issues about equity, access, and privilege. And perhaps how we teach applied linguistics and sociolinguistics needs even more rethinking as artificial intelligence goes mainstream.

Author ORCIDs. Helen Kelly-Holmes, 0000-0002-5110-6409

References


Annual Review of Applied Linguistics 37


