International Journal of Translation, Interpretation, and Applied Linguistics
Volume 2 • Issue 1 • January-June 2020

The Challenges of Azerbaijani Transliteration on the Multilingual Internet

Sabina Mammadzada, Institute of Information Technology, Baku, Azerbaijan

ABSTRACT

Several scripts have been adopted so far in Azerbaijan in different periods. Literary manuscripts and political documents developed in each of the adopted scripts are of great importance and have to be transliterated for the next generation. However, adoption of different scripts and their frequent changes lead to the emergence and dissemination of various transliteration versions on the web. This article touches upon the challenges of Azerbaijani-English transliteration process in real life and online. In this regard, the significance and dominating status of the English language on the Internet and throughout the globe is explored. Adoption of a unique transliteration standard for Azerbaijani language may contribute to the solution to this problem. The standards for the transliteration of the Azerbaijani language with others are analyzed, and the inevitability of a new approach to Azerbaijani-English transliteration is emphasized. Moreover, the article underlines future contribution of the proposed transliteration system for machine translation system of Azerbaijani-English language pair.

KEYWORDS

Conversion, Conversion Table, Lingua Franca, Machine Translation, Romanization, Transformation, Translation, Transliteration

INTRODUCTION

Mail, instant messaging, real-time chat protocols, asynchronous discussion forums and Web pages, which perform the textual exchanges, establish Computer-Mediated Conversation. 'Conversation', in general, refers to the spoken word, and voice communication such as Skype and other forms of Voice-over-Internet Protocol (VoIP) transmission and has recently become very famous. However, text-based Computer-Mediated communication (CMC) has been historically practiced for a long time, and it is more widespread than VoIP. The first mail was sent over networked computers in 1972 (Hafner, & Lyon, 1996). Although it was previously limited to the users in the military, government, universities, and businesses, textual CMC became well-known in the late 1980s and early 1990s, after Internet Service Providers ensured users to connect online space without leaving their homes. Text messaging on mobile phones, instant messaging, weblogs, and wikis are the most recent widespread CMC modes. On multimedia platforms, such as media sharing and social network sites, textual interactions remain important (Herring, 2004; Herring, 2009). Hence text exchange is the most typical kind of computer-mediated 'conversations.' Therefore, unlike the knowledge of any specific

DOI: 10.4018/IJTIAL.2020010105

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

language, the knowledge of at least the script of most frequently spoken languages is vital in these 'conversations.' In this case, transliteration of the scripts is used.

This article deals with the transliteration issues of the Azerbaijani scripts with the scripts of English. Transliteration in real life and online is becoming more relevant when so many new technologies emerge daily and require technological and respective communication skills (messaging, typing etc.). The goal of this article is to describe the importance of transliteration in the communication, preservation of the national identity of languages through transliteration and the need for update of available conversion (Romanization) tables.

BACKGROUND

Abovementioned 'conversations' realized in the global network are mainly performed in English. This is due to different factors, as the recently emerged notion 'English as a lingua franca' (ELF). In this communication, the key participants are the speakers whose second or third language is English. Crystal (2003) points out that only one out of every four users of English is a native speaker of this language around the world. Consequently, the largest share of ELF communications is performed among 'non-native' speakers of English. And, in many cases, this language is regarded as 'a 'contact language' among the people whose native tongue or (national) culture significantly differ, and who choose English as a foreign language of communication' (Firth, 1996).

English has been defined by several scholars so far, as 'English as an international language' (Jenkins, 2000) EIL, 'English as a global language' (Crystal, 2003; Gnutzmann, 1999), 'English as a world language' (Mair, 2003). These definitions have been used as broad terms for the usage of English within and across Kachru's 'Circles'. Inner Circle, Outer Circle, and Expanding Circle contexts (Kachru, 1992). Consequently, EIL traditionally means the use of English for international communication. Nevertheless, the term 'English as a lingua franca' is preferred more when English is used as the means of communication among non-native speakers coming from linguistic and cultural boundaries (House, 1999; Seidlhofer, 2001). In addition, English is used as a medium of intercultural communication (Meierkord, 1996) and as an international language (Jenkins, 2000), which is more explicit.

Undeniably English serves as a global lingua franca. Nonetheless, as a result of its international use, nonnative speakers are forming English at least as much as by its native speakers.

The term "international language" is commonly accepted as a language used by the people from different backgrounds or nations. In this regard, many international languages have been used so far throughout the world, from Spanish and Chinese to Latin and Classical Persian, including locally constrained dialects and lingua francas.

The term "international language" can also be referred to artificial or deliberate languages such as Esperanto. These are developed unambiguously to ease international communications and understanding often being called international auxiliary languages. These languages do not have native speakers from any nation or country unlike the languages that have been spread worldwide such as English, French and Spanish.

The dominance of any language depends on several factors such as invasions, economic exploitation, military and political conquests. Language plays a role of a crucial component in the extension of authoritative governments and in the process of Europeanization of the world.

In this regard, BASIC English (BASIC = British American Scientific International Commercial), which was a constrained form of English, was supported in the period between the end of the First and the beginning of the Second World War. The creators of this project believed that minor languages would be reduced: "What the world needs is about 1000 more dead languages - and one more alive" (Ogden, 1934,). The international attitude here was that the languages other than English would be eliminated and rejected through the simplification of this language.

This linguistic strategy has been supported and continues to be supported in all continents of the world (Mühlhäusler, 1996; Phillipson, 1992). The British and Americans, in this regard, have developed a significant academic infrastructure to internationally promote English.

English is still supposed to be the unavoidable world language, the reasons for this are explained by the noticeable position of English in international affairs (Bailey, 1991: 121).

English as an "international" language is viewed as unanimously significant, even with the numerous facts that its extensive practice benefits western countries and does not provide the requirements of most population in other countries.

English is regarded as most broadly used language for international and intercultural communication by many linguists. Ammon's report on the global usage of English highlights the present dominant status of English. This dominance is based on the statistics, (1) English is spoken by nearly 1,5 billion people worldwide, (2) English is accepted as an official language of 62 nations, (3) approximately 70-80 percent of academic publications in scientific community is implemented in English, (4) English is adopted as de facto official and working language in most international organizations, and (5) English is the most taught foreign language throughout the world (Ammon, 1992)

In the dictionaries, translation is defined as the process of changing a word or text written or spoken in one language into another, while transliteration is rewriting or describe words or letters in other scripts (Wehmeier, McIntosh, Turnbull, & Ashby, 2005). Nowadays, the data found in target language, i.e., English has no sense if the non-English speakers do not understand the meaning or recognize (Birbili, 2000). It means that being generated in one language for use in another, these resources are required to be understood and compared across language borders (Birbili, 2000, Hennink, 2007).

Currently, non-English-speaking users are in search for information on the web, predominantly, researchers are very eager to explore and publish their articles in English-language journals. They mostly focus on wider audiences and scientific academic community. Sunol and Saturno (Sunol, & Saturno, 2008) mention that people are not interested in the research in the languages other than English, and therefore, those studies are not cited as much as those written in English. Moreover, the findings "published in other languages may be probably lost or, in practice, non-existent for the scientific community as a whole". Furthermore, well-known scientific data bases often exclude these studies from systematic reviews and meta-analyses owing to their language limitations.

Crystal (1991) defines translation as a process where "the meaning and expression in one language (source) is tuned with the meaning of another (target) whether the medium is spoken, written or signed". And, this practice of conversion of texts from one language to another is related to the sociocultural factors (Halai, 2007).

Unlike translation, the term transliteration is referred to the process of replacing or adding the scripts or words of one language with the scripts or words of another. The main idea here is to achieve the exact or at least the maximal correspondence of the phonology of the source language in the target language. Transliteration enables formalizing the context in which cross-cultural translation is performed. Nida and Taber (Nida, & Taber, 1969) believe that the words are often untranslatable when the word does not have its correspondence in the target language and relative correspondence is required. In this case, transliteration is used.

Many researchers consider the process of translation and transliteration very time consuming, particularly in the case of data abundance to be analyzed (Halai, 2007). Doctor of the University of Leeds, Emmel in his study (Emmel, 1998) offers to check transcripts, that is transliteration, as the only way to provide precision of translation.

Not only researchers, but also students in foreign language classes often experience complications in searching information and resources on the web in different languages. Especially when the web users are non-English speakers, they face difficulties in writing key words for information retrieval, when the relevance of the information directly depends on the accuracy of the transliteration.

Apparently, many web resources are Romanized to be accessible for a wider audience. In this regard, several Romanization tables are set out for different scripts. Even though, some conversion

tables are adopted by international and national establishments for the Azerbaijani language, the appropriate international standard for this language has not been accepted so far (Retrieved from https://www.bac-lac.gc.ca/eng/services/cataloguing-metadata/Documents/040006-14-e.pdf). ISO9 adopted three transliteration standards for the Azerbaijani language, the first of which was confirmed for the transliteration from Arabic scripts (ISO 233: 1984), which were used in Azerbaijan at that period (1984), into Latin scripts. The second standard, which was confirmed in 1993, was the second and simplified edition of ISO 233; however, prior to 1993, these (Arabic) scripts had already been restricted in Azerbaijan. The first two standards were very essential for the transliteration of ancient literature. The third and the principal transliteration standard ISO 9 was adopted in 1995 for the transliteration of Cyrillic scripts into Latin scripts (for Slavic and non-Slavic languages).

Most of the available transliteration systems are mainly based on graphical and phonological correspondences between the source and target languages. Nevertheless, transliteration is not systematic. It depends on the languages and varies in accordance with its field of application, i.e. conversion tables adopted in separate areas. Therefore, a user should be well aware of the conversion tables, especially if he/she is working with more than one database (Pasterczyk, 1985). A conversion table set out for one field has no wider application and is not appropriate for another. For instance, Physical Review is intended for academic usage only, or a conversion table designed for the Library of Congress is used by the general audience and adopted by most US academic libraries as a transliteration standard. It should be noted that recently, the simplified edition of LC transliteration system has been adopted by several academic publishing organizations (Edgerton, 1989).

The research activity of many students and researchers may also be affected by the transliteration problem, thus, the examination and development of transliteration methods is considered to be urgent. Numerous transliteration methods have been offered and developed for many languages by different researchers so far. A few transliteration tables have also been offered for Azerbaijani language by some organizations, however, the problem of their modification and standardization still remains. The modification of the transliteration tables may include the exclusion of diacritics and addition of some scripts. For example, the scripts, the phonology of which does not correspond to the phonology of Azerbaijani language, may be replaced. This paper efforts to address the transliteration problems for Azerbaijani-English language pair.

AZERBAIJANI-ENGLISH TRANSLITERATION CHALLENGES AND THEIR SOLUTION WAYS

Currently, the problem of transliteration is strictly interrelated to the multilingual character of the global web. Every existing and endangered language in this global web efforts to be appropriately characterized and to preserve the national identity not being constrained by other languages. Most establishments designing the conversion tables do not own transliteration systems available to handle several scripts. It is due to the unwillingness or less concern to meet the multilingual needs of community and to the fact that they do not have sufficient funds (Simsova, 1988). Moreover, many organizations are limited to the Roman scripts and are not interested in designing transliteration tables for non-Roman scripts. Therefore, users are not able to discover or access the required resources, materials and documents in their local scripts. Consequently, Romanization becomes urgent issue. The process of Romanization should be ensured with the automated cross-language transliteration system expansively representing the graphical and phonological features of a certain language.

Fortunately, some promising measures are being taken in the field of transliteration. Hence, many linguists and knowledge engineers have begun to focus on this issue analyzing and proposing significant solutions (Knight, & Graehl, 1998; Haizhou Li, Min Zhang, & Jian Su, 2004; Slaven & Hozumi, 2005; Franz Josef Och & Hermann Ney, 2003; Chen, Huang, Ding, &Tsai, 1998; Al-Onaizan, & Knight, 2002). Moreover, nowadays, International Standard ISO/IEC 10646 - a novel universal multi-script character set is available (Joint Technical Committee ISO /IEC JTC1, Information Technology-

Universal Multiple-Octet Coded Character Set, Part 1: "Architecture and Basic Multilingual Plane, ISO/IEC 10646: 1993" (Geneva: International Organization for Standardization, 1993)). The character set Unicode coincides with this standard and is applicable to almost all products of leading computer manufacturers (The Unicode Consortium, The Unicode Standard: Worldwide Character Encoding, Version 1.0 (Reading, Mass.: Addison-Wesley, 1991-92); supplemented by The Unicode Standard, Version 1.1 (prepublication ed.) (Mountain View, Calif.: Unicode Consortium, 1993). (Unicode is a trademark of Unicode, Inc.)). The advantage of this novel standard is its capability to process any script. With this feature, it will ease the inclusion of and access to any resource written in non-Roman scripts. Adoption of this standard will also contribute to cross-language machine translation systems and cross-language information retrieval.

Before the discussion of the challenges of the transliteration for Azerbaijani-English language pair, a brief review of the modified Latin script adopted for the Azerbaijani language is provided below.

The alphabet of Azerbaijani language consists of 32 Latin letters. Although it is the Latin Alfabet, it differs from the one adopted for English, French, German and other widely used languages. 7 letters (ç, ğ, ı (lower case for "I"), İ (upper case for "i"), ö, ş, ü) of the Azerbaijani Latin characters are modified and 1(ə) is adopted specifically for this language. In fact, in the process of transliteration, these letters become very challenging not only for their graphical description, but also for their pronunciations, which are unknown for English-speaking audience. This problem becomes even more complicated when these letters are represented by the diacritical signs in adopted transliteration tables, such as ġ, ž, ǰ,ˇ, ï,ˆ, ù, š, č, ĉ, which are unaccustomed even for Azerbaijani reader. Another problem is the representation of the Azerbaijani letter "ə". This unique letter is adopted in few languages, such as Abkhaz, Kazakh, Tatar, and Bashkir, being pronounced in a different way or not pronounced at all. Although, several researchers have efforted to unify all Turkic alphabets during the past two decades for the simplification of the writing system of all Turkic languages through Latin alphabet, adoption of common Turkic alphabets has failed (Altuğ, 2014).

The transliteration from English into Azerbaijani also becomes challenging due to the lack of letters in the Azerbaijani language. For example, the letter "w" does not exist in Azerbaijani alphabet and has to be replaced by letter "v" or by a letter combination "ou".

The transliteration of the scripts for Azerbaijani-English language pair is also complicated due to the significant difference in the number of the scripts in both alphabets. In this case, comprehensive transliteration is achieved through the letter combinations, thus the number of letters in the English alphabet is less than in that used for Azerbaijani language (Weinberg, 1974). However, letter combinations originate other problems. Thus, depending on the language, the letter combinations affordable for one letter with the same sound may vary. For instance, depending the conversion system, the Cyrillic letter "я" can be converted as "ia", "ja", and "ya". The point is that the characteristics of specific phonemes in Slavic languages are not written explicitly with only one Roman letter. Therefore, several equivalence for Azerbaijani letter "y" arises, which include "i", "j", and "y".

Another transliteration problem is related to the adoption of the conversion tables for the ancient alphabets, which are not used anymore. Implementation of these conversion tables is very important for the libraries, which contain rich literary sources and manuscripts in ancient alphabets. Confusion, in this regard, occurs due to the fact that the phonological and graphic features of those scripts are still being considered when designing new conversion tables, though those scripts have not been used for a long time. It obscures the adjustment, simplification, interpretation and finally, the enhancement of available transliteration standards.

Some political and social factors also affect the transliteration problems. Thus, the lands of the modern Republic of Azerbaijan have been numerously attacked by several authorities. Consequently, new culture, traditions, customs, languages and scripts has been brought with invaders. For example, the conquest of Arabic Caliphate and the spread of Islam have given impetus to the rise of Arabic script in the territory of Azerbaijan and replaced the ancient Latin script previously applied here. While in the period of Soviet Union, the Cyrillic started to be used rather than Arabic. Furthermore,

to attain the adequacy with the phonology of Azerbaijani sounds, Cyrillic alphabet used in Azerbaijan has been numerously modified in various phases. In 1991, the independent Azerbaijani government adopted modified Latin script for the Azerbaijani language. Accordingly, the main problem here is related to many literary manuscripts, political documents and materials developed in each of adopted scripts. Each of these documents and sources is of great significance and requires to be transliterated for being available to the next generation. They are transliterated as a valuable resource in terms of the conservation of the national identity. Nevertheless, adoption of several scripts and frequent changes made to them cause emergence and dissemination of various transliteration versions on the web. Adoption of a unique transliteration standard for Azerbaijani language, as well as the revision of social and psychological and moral thinking of Azerbaijani people may contribute to the solution of this problem.

The spread of various transliteration forms, as mentioned above, contributes to confusion. This problem exists almost in all languages. For example, the last name *Хрущев* has ne following variants on the web: *Hrushev, Hrushew, Hrushyev, Hrushyew, Hrushjev, Hrushjew, Hruschev, Hruschew, Hruschyev, Hruschyew, Hruschjev, Hruschjew, Hrushchev, Hrushchew, Hrushchyev, Hrushchyew, Hrushchjev, Hrushchjew, Hruchhev, Hruchhew, Hruchhyev, Hruchhyew, Hruchhjev, Hruchhjew, Hruwev, Hruwew, Hruwyev, Hruwyew, Hruwjev, Hruwjew, Xrushev, Xrushew, Xrushyev, Xrushyew, Xrushjev, Xrushjew, Xruschev, Xruschew, Xruschyev, Xruschyew, Xruschjev, Xruschjew, Xrushchev, Xrushchew, Xrushchyev, Xrushchyew, Xrushchjev, Xrushchjew, Xruchhev, Xruchhew, Xruchhyev, Xruchhyew, Xruchhjev, Xruchhjew, Xruwev, Xruwew, Xruwyev, Xruwyew, Xruwjev, Xruwjew, Khrushev, Khrushew, Khrushyev, Khrushyew, Khrushjev, Khrushjew, Khruschev, Khruschew, Khruschyev, Khruschyew, Khruschjev, Khruschjew, Khrushchev, Khrushchew, Khrushchyev, Khrushchyew, Khrushchjev, Khrushchjew, Khruchhev, Khruchhew, Khruchhyev, Khruchhyew, Khruchhjev, Khruchhjew, Khruwev, Khruwew, Khruwyev, Khruwyew, Khruwjev, Khruwjew* (retrieved from https://www.analizfamilii.ru/Khrushchev/translit.html). Each of these variants has been adopted in accordance with the grammatical, graphical and phonological features of the languages it is converted to. Another example, is related to the Azerbaijani last name *Quliyev*, which has numerous transliterations as *Kuliyev, Quliyev, Guliyev, Kuliev, Quliev,* and *Guliev*. Due to the long-term impact of Russian language on Azerbaijani language, these variations have occurred. The phonological correspondence of the letter "Q" in Azerbaijani language is "G", however, it was wrongly transliterated into Russian as "K". Moreover, phonological correspondence of the letter combination "ye" in Russian is the letter "e", and therefore, it was transliterated into "e", neglecting the letter "y". Accordingly, numerous personal names and place names are modified losing their distinctiveness.

Thus, most of the district and city names were first adapted to the phonological features of the Russian language during the soviet period, and then to the international community. For example, *Zagatala - Zakataly, Balakan - Belokany (Belokan), Shamakhi - Shemakha (Shemaxa)* and *Nakhchivan - Nakhichevan (Nakhichivan)*. Distortion of these toponyms in the global information space results in a violation of identity. These distortions can, in some cases, be related not only to the grammatical features of transliteration languages, but also to political-ideological factors. For example, transliteration of the city of Shamakhi in the form of *Shemakha (Shemaxa)* can be identified with the name of the village located on the shore of *Shemakha* river in the Chelyabinsk region of the Russian Federation (retrieved from https://commons.wikimedia.org/wiki/Category:Sukhaya_Shemakha_River).

One of the most common difficulties translators must manage is translating proper names and technical terms. For the language pairs like Spanish/English, this presents no great challenge: a phrase like *Antonio* usually gets translated as *Antonio*. However, the situation is more complicated for language pairs that use dissimilar scripts and sound systems, such as Japanese/English and Arabic/English. Phonetic translation across these pairs is called transliteration. To write the words like *şəkərbura* and *paxlava* in Latin scripts, some compromises must be made, since English has no distinct *ə* and *x* sounds. Transliteration is clearly an information-losing operation. Human translators can often "sound

out" a phrase in Azerbaijani to guess an appropriate translation. Automating this process may have great practical importance in Azerbaijani/English machine translation.

Several explanations about back-transliteration are mentioned below to describe an idea of the difficulty of the task:

- **Back-transliteration is less forgiving than transliteration:** There are many ways to write a word in Azerbaijani like *şəkərbura* in English, all equally valid - *shakarbura*, however this flexibility does not present in the reverse direction. For example, when back-transliterating, the *h* in *shakarbura* cannot be remained. Forward-direction flexibility wreaks mess, because *sh* combination is not always transliterated as *ş* depending on the word, here, word segmentation is required;
- **Back-transliteration is harder than romanization:** A romanization scheme simply sets down a method for writing a foreign script in roman letters. For example, to Romanize q the symbol g is used. This substitution gives us *qədim – gädim*, but back-transliterated as *gədim*. Romanization schemes are usually deterministic and invertible, although small uncertainties may occur.

Accuracy is the most necessary feature of an automatic back-transliterator. Preferably, the proposed techniques should also be transferable to new language pairs like Persian/Azerbaijani with minimal effort, perhaps with the use of ready resources. Subsequently, human translators find the problem quite difficult as well, so in this paper we only try to match or possibly exceed their performance.

For comprehensive solution of transliteration problems, the following measures shall be taken:

- Available Romanization tables for Azerbaijani language shall be modified, and if necessary, a new standard transliteration table shall be set out and submitted to ISO for approval. The modifications shall be made by the experts of certain fields of sciences in accordance with the grammatical, graphical and phonological features of Azerbaijani language;
- Transliteration system of Azerbaijani language with other most commonly used languages by the Azerbaijani people shall be developed;
- Experiments and trials of abovementioned transliteration system shall be held; proposals and feedbacks from users shall be collected and processed;
- Transliteration system shall be improved to achieve maximum transliteration accuracy based on feedbacks and proposals submitted from users.

CONCLUSION

Transliteration of the Azerbaijani scripts into other languages, particularly into the Latin scripts used for English becomes urgent, with regards to the number of Azerbaijanis living throughout the world accounting for 50 million people (Report of the State Committee on Affairs with Diaspora of the Republic of Azerbaijan for year 2015). Taking into account the supremacy and the position of English as the international language, transliteration from Azerbaijani language into English and vice versa becomes an essential element of the communication in the current information society. It should be noted that according to the report provided by the Diaspora society of Azerbaijan, currently, more than 30 million of the total population of Azerbaijanis are living in Iran (retrieved from http://azerbaijan.az/portal/Society/Diaspora/diaspora_03_a.html?), where Azerbaijani language is widely spoken as the second language, although Farsi script, that is modified Arabic script, is used for this language. The key barrier, namely Farsi script between Azerbaijani populations of these two countries can be eliminated with the development of the transliteration system for Azerbaijani and Farsi language pair, which is the main topic of the next study.

This paper explored the growing dominance and status of English as an "international language" and "lingua franca". One of the key elements of communication on the web, that is translation and transliteration, their importance and challenges were highlighted. The requirements for the users to transliterate necessary resources and information written in the modified Latin scripts were explained. Since the web resources include different translations and transliterations of the same words, the factors affecting this situation were identified. First of all, these versions emerge due to several factors and cross-cultural moralities. This paper focused on and offered some strategies that could contribute to eliminate transliteration problems, preserve national identity and improve information retrieval.

REFERENCES

Al-Onaizan, Y., & Knight, K. (2002). Translating named entities using monolingual and bilingual resources. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, Philadelphia (pp. 400–408). Academic Press.

Altuğ, B. D. (2014). *The 1991 International Contemporary Turkic Alphabets Symposium and its Contributions to the Turkic Alphabet Reform*. University of Washington.

Ammon, U. (1992). *Gengo-to Sono Chii* (Y. Hieda & H. Yamashita Tokyo, Trans.). Sangensha.

Bailey, R.W. (1991). Images of English: A cultural history of the language. Cambridge: Cambridge University Press.

Birbili, M. (2000). Translating from one language to another. *Social Research Update*, *31*, 1–7.

Chen, H., Huang, S., Ding, Y., & Tsai, S. (1998) Proper name translation in cross-language information retrieval. *Proceedings of 17th COLING and 36th ACL* (pp.232–236). Academic Press.

Crystal, D. (1991). *The Cambridge encyclopedia of language*. Cambridge, UK: Cambridge University Press.

Crystal, D. (2003). *English as a Global Language* (2nd ed.). Cambridge: Cambridge University Press; . doi:10.1017/CBO9780511486999

Edgerton, W. (1989, Winter). One More Look at the Problems of Transliteration. *Slavic Review*, *48*(1), 97–99. doi:10.2307/2498689

Emmel, D. (1998). *Neighborhood, perceptions of health and the value placed on the health care delivery in the slum of Mumbai*. Leeds, UK: University of Leeds.

Firth, A. (1996). The discursive accomplishment of normality. On "lingua franca" English and conversation analysis. *Journal of Pragmatics*, *26*(2), 237–259. doi:10.1016/0378-2166(96)00014-8

Gnutzmann, C. (Ed.). (1999). *Teaching and Learning English as a Global Language*. Tübingen: Stauffenburg.

Hafner, K., & Lyon, M. (1996). *Where wizard stay up late: The origins of the Internet*. New York: Simon & Schuster.

Halai, N. (2007). Making use of bilingual interview data: Some expressions from the field. *Qualitative Research Report*, *12*(3), 344–355.

Hennink, M. (2007). Language and communication in cross-cultural qualitative research. In P. Liamputtong (Ed.), *Doing cross-cultural research: Ethical and methodological perspectives* (pp. 21–48). London: Springer Science.

Herring, S. C. (2004). Slouching toward the ordinary: Current trends in computer-mediated communication. *New Media & Society*, *6*(1), 26–36. doi:10.1177/1461444804039906

Herring, S. C. (2009, October). Convergent media computer-mediated communication: Introduction and theory. *Paper presented at the panel on Convergent Media Computer-Mediated Communication*, Milwaukee. Academic Press.

House, J. (1999). Misunderstanding in intercultural communication: Interactions in English as a lingua franca and the myth of mutual intelligibility. *Teaching and learning English as a global language, 7389*.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University press.

Kachru, B. (Ed.). (1992). *The Other Tongue* (2nd ed.). Urbana, Chicago: University of Illinois Press.

Knight, K., & Graehl, J. (1998). Machine transliteration. *Computational Linguistics*, *24*(4), 599–612.

Li, H., Zhang, M., & Su, J. (2004). *A joint source channel model for machine transliteration*. Association for Computational Linguistics.

Meierkord, C. (1996). *Englisch als Medium der interkulturellen Kommunikation. Untersuchungen zum non-native-/non-native speaker—Diskurs*. Frankfurt/Main: Lang.

Mühlhäusler, P. (1996). *Linguistic ecology. Language change and linguistic imperialism in the Pacific region*. London: Routledge.

Nida, E., & Taber, C. (1969). *The theory and practice of translation*. Leiden: EJ Brill.

Och, F. J., & Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, *29*(1), 19–51. doi:10.1162/089120103321337421

Pasterczyk, C. (1985, February). Russian Transliteration Variations for Searchers. *Database*, *8*, 68–75.

Phillipson R. (1992). Linguistic imperialism. Oxford: Oxford University Press.

Report of the State Committee on Affairs with Diaspora of the Republic of Azerbaijan for year 2015. (n.d.). Retrieved from http://azerbaijan.az/portal/Society/Diaspora/diaspora_03_a.html?

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics*, *11*(2), 133–158. doi:10.1111/1473-4192.00011

Simsova, S. (1988). Coping with Foreign and Nonstandard Character in Libraries. In J. Eyre (Ed.), *Small Computers in Libraries* (pp. 13–15). London: Meckler.

Slaven, B., & Hozumi, T. (2005). Direct combination of spelling and pronunciation information for robust back-transliteration. *Proceedings of the Conferences on Computational Linguistics and Intelligent Text Processing* (pp. 413–424). Academic Press.

Sunol, R., & Saturno, P. (2008). Challenge to overcome language barriers in scientific journals: Announcing a new initiative from the ISQua journal. *International Journal for Quality in Health Care*, *20*(1), 1–2. doi:10.1093/intqhc/mzm072

Wehmeier, S., McIntosh, C., Turnbull, J., & Ashby, M. (2005). *Oxford Advanced Learner's Dictionary of current English*. Oxford: Oxford University Press.

Weinberg, B. (1974, March). Transliteration in Documentation. *The Journal of Documentation*, *30*(1), 18–31. doi:10.1108/eb026567

Sabina Mammadzada received her MS in Translation Study and Linguistics from the Azerbaijan University of Languages in 2010. In 2007 she was hired as an interpreter by the Institute of Information Technology of Azerbaijan National Academy of Sciences (ANAS) and is currently a senior researcher at the Institute of Information Technology and Ph.D. student of the Nasimi Institute of Linguistics of ANAS. She writes and presents widely on issues of transliteration and internet linguistics.