Acta Globalis Humanitatis et Linguarum
ISSN: 3030-1718
Vol. 2, No. 4 (2025): Autumnus 2025

Orthographic Challenges in the Transliteration of Proper Names between the Languages with Different Spelling

[1] **Javid Babayev**
https://doi.org/10.69760/aghel.0250040018

Keywords
transliteration
proper names
duplication
orthographic challenges

Abstract

Transliteration, the process of converting words from one writing system into another, presents significant orthographic challenges, especially when applied to proper names. Unlike common nouns, proper names often carry cultural, historical, and phonetic nuances that resist straightforward conversion. This paper explores the orthographic inconsistencies and difficulties that arise in the transliteration of proper names across different language systems, highlighting the interplay between phonological representation, script limitations, and cultural context.

A central issue in transliteration is the lack of one-to-one phoneme-to-grapheme correspondence between languages. For instance, sounds present in one language may not exist in another, leading to approximations that can distort the original pronunciation or meaning of a name. This is especially problematic in cases involving names from languages with non-Latin scripts, such as Arabic, Chinese, or Russian, where multiple systems of transliteration coexist (e.g., Pinyin vs. Wade-Giles for Chinese), further complicating standardization and international recognition.

Additionally, transliteration often raises questions of identity and authenticity. A name's spelling can influence perceptions of nationality, ethnicity, or even social status, making orthographic choices politically and personally sensitive. For example, individuals may choose non-standard transliterations to retain specific cultural markers or to ease integration into a foreign linguistic environment.

The paper also discusses the role of official transliteration policies and their impact on global communication, data processing, and legal documentation. Inconsistent transliteration standards across government agencies, international bodies, and digital platforms

[1] Babayev, J. S. PhD in Philology, Senior Lecturer, Department of English and Methods, Nakhchivan State University, Azerbaijan. Email: cavidbabayev@ndu.edu.az. ORCID: https://orcid.org/0009-0009-2472-0006.

can lead to confusion, duplication, and errors in databases and legal records.
By examining case studies and cross-linguistic examples, this study underscores the need for more nuanced, context-aware transliteration frameworks. It concludes by advocating for interdisciplinary collaboration between linguists, policymakers, and technologists to create transliteration systems that balance phonetic accuracy, orthographic clarity, and cultural sensitivity.

1. Introduction

Proper names, including personal, geographical, and institutional names, are critical identifiers in communication. However, when these names cross linguistic boundaries—especially between languages with different spelling systems—transliteration becomes necessary. Transliteration aims to represent the phonetic structure of a name in the target language's script. Unlike translation, which conveys meaning, transliteration focuses on sound approximation (Newmark, 1988).

In an increasingly globalized world, the accurate and consistent representation of proper names across linguistic boundaries is more essential than ever. Proper names—such as those of people, places, and institutions—are integral to identity, legal documentation, and cultural recognition. Unlike common nouns, proper names are rarely translated for meaning; instead, they are typically transliterated, a process that attempts to reproduce the original pronunciation in a different writing system. However, this task is fraught with challenges, particularly when transliteration occurs between languages with fundamentally different orthographic systems, phonologies, or alphabets.

Transliteration differs from transcription in that it often seeks a balance between phonetic accuracy and script-based representation. While transcription focuses purely on sounds (phonemes), transliteration may also take into account the visual form (graphemes) of the source language's script. For example, transliterating the Russian name "Борис" into English as "Boris" is a relatively straightforward case due to the phonetic and alphabetic overlap. In contrast, transliterating an Arabic name like "خالد" into English could result in "Khaled," "Khalid," or "Halid," depending on phonological interpretation, regional dialects, and orthographic conventions.

The difficulties multiply when transliterating between languages with non-alphabetic scripts, such as Chinese logographs or Japanese kana, which do not correspond directly to the phonemic structure of alphabetic languages. In such cases, names are often rendered with approximations that may carry unintended meanings or associations. For instance, the English name "Clinton" is transliterated in Mandarin Chinese as 克林顿 (Kèlíndùn), a sequence chosen for its phonetic closeness rather than semantic value. These adaptations can obscure or distort the source identity, especially when retranslated into the original language.

Historically, transliteration practices have been influenced not only by linguistic principles but also by colonial history, religious traditions, and political ideologies. For example, French and British colonial administrations in the Middle East employed different standards for Arabic name rendering, leading to inconsistencies that persist to this day. Similarly, Soviet transliteration standards impacted how Central Asian names were rendered in Cyrillic and then re-rendered back into Latin script following independence (Comrie, 1981). In recent years, geopolitical changes such as the Ukrainian push to de-Russify toponyms and names (e.g., "Kiev" to "Kyiv") have further emphasized the role of transliteration in national identity and cultural politics (Bilaniuk, 2005).

Transliteration challenges arise due to mismatched phoneme inventories, script differences (e.g., Latin vs. Cyrillic vs. Arabic scripts), and cultural or political sensitivities (Knight & Graehl, 1998). The lack of standardized transliteration rules can result in multiple variants of the same name (e.g., "Mohammed," "Muhammad," "Mohamad"), leading to confusion in databases, documents, and digital systems.

This paper investigates the orthographic challenges in transliterating proper names between languages with significantly different spelling conventions. The study analyzes common transliteration discrepancies and offers insights into potential standardization frameworks.

In digital communication, the stakes of transliteration inconsistency are particularly high. With the global expansion of databases, passports, academic publications, and social media platforms, a single individual might have their name represented in multiple forms across different systems, resulting in fragmented identities, searchability issues, and administrative errors (Knight & Graehl, 1998; Li et al., 2020). Despite the development of international standards (such as ISO 9 for Cyrillic or ALA-LC for Arabic), implementation is inconsistent, and deviations are often driven by localized norms, dialectal variation, or technical limitations in data encoding.

Given this complex linguistic and sociopolitical landscape, the need to better understand the orthographic challenges in transliterating proper names is both practical and theoretical. This paper investigates these challenges through a cross-linguistic lens, examining how different systems handle the transliteration of proper names and what factors influence the variation. It aims to identify patterns of inconsistency, the underlying causes of transliteration errors, and possible avenues toward more coherent, standardized approaches that respect both phonological integrity and cultural specificity (Babayev, 2024).

2. Methods

This study employed a qualitative, comparative linguistic analysis to investigate the orthographic challenges associated with the transliteration of proper names across languages with differing writing systems and spelling conventions. The research design focused on identifying patterns, inconsistencies, and linguistic constraints in transliteration practices, with attention to both systemic linguistic issues and sociopolitical influences.

2.1. Language Pair Selection

Five language pairs were selected for comparative analysis, chosen based on their:

- Distinct script systems (e.g., alphabetic vs. logographic),
- High frequency of cross-linguistic interaction (in diplomacy, media, or migration),
- Relevance in international communication and transliteration studies.

The language pairs analyzed were:

- English ↔ Russian (Latin vs. Cyrillic)
- English ↔ Arabic (Latin vs. Arabic script)
- English ↔ Chinese (Latin vs. logographic characters)
- French ↔ Arabic (due to historical colonial transliteration legacy)
- Russian ↔ Chinese (two non-Latin-based scripts)

2.2. Data Collection

Data were collected from multiple open-access and institutional sources to ensure a comprehensive and representative corpus of transliterated proper names:

- **Media sources**: Articles from international news outlets such as *BBC*, *Al Jazeera*, *Xinhua*, *RT*, and *Le Monde* were analyzed to observe real-world transliteration in journalistic contexts.
- **Government documents**: Official passports, diplomatic directories, immigration forms, and UN documentation provided transliterations used in legal and bureaucratic contexts.
- **Academic and reference databases**: Names from *Wikipedia*, *Geonames.org*, and *Linguist List* were used to trace standardized and variant forms across languages.
- **Online transliteration tools**: Resources such as Google Translate, Lexilogos, and academic transliteration engines were evaluated for machine-generated patterns.

2.3. Corpus Construction

A purposeful sampling strategy was used to select a corpus of 120 proper names, categorized as follows:

- Personal names (e.g., world leaders, historical figures, celebrities)
- Place names (e.g., capital cities, countries, regions)

- Institutional names (e.g., political parties, international organizations)

Each name was traced across at least three languages to identify transliteration variants. Preference was given to names that are frequently used in cross-linguistic contexts or have well-documented transliteration controversies.

2.4. Analytical Framework

The analysis applied a multilevel linguistic framework, including:

a. Phonological Analysis

- Identification of phonemes present in the source language but absent in the target language.
- Mapping of phoneme substitutions and approximations (e.g., /p/ → /b/ in Arabic).
- Investigation of phonotactic constraints that force changes in syllabic structure.

b. Orthographic Analysis

- Comparison of grapheme inventories between scripts.
- Identification of orthographic ambiguity in target scripts (e.g., the use of "kh" for ჯ or "zh" for ж).
- Analysis of script-specific issues, such as vowel omission in abjads (Arabic) or tone loss in logographs (Chinese).

c. Sociolinguistic Considerations

- Examination of historical, political, or ideological influences on name rendering (e.g., postcolonial renaming, national standardization efforts).
- Analysis of language-specific transliteration norms and preferences.
- Attention to diglossia and dialectal influence in the case of Arabic and Chinese transliterations.

d. Variant Mapping and Normalization

- Creation of visual maps showing multiple transliteration variants for the same name across languages.
- Categorization of variants as phonologically faithful, orthographically adapted, or politically motivated.

2.5. Limitations and Validation

To ensure analytical rigor, multiple data coders were involved in the phonological and orthographic analysis to minimize subjectivity. However, this study is limited by the absence of native speaker consultation for all language pairs, and by the variability of transliteration practices across dialects and regions. Future studies may benefit from incorporating native speaker intuitions and psycholinguistic testing of name recognition.

3. Results

The comparative analysis of transliteration across the five language pairs revealed several recurring orthographic and phonological challenges, as well as sociolinguistic factors influencing name representation. The findings are categorized below into key thematic areas.

3.1. Phoneme Gaps and Grapheme Substitution

One of the most pervasive challenges identified was the absence of certain phonemes in the target language, leading to substitutions or approximations that can significantly alter the original name's pronunciation.

- In the English → Arabic direction, the phoneme /p/ does not exist in Classical Arabic, resulting in its substitution with /b/. For instance, "Philip" (فيليب, Fīlīb) is often rendered as "Fīlīb," but the /p/ sound is effectively replaced by /b/. Similarly, "Pakistan" becomes باكستان (Bākistān), where the /p/ is substituted by /b/. This substitution can cause confusion, especially in cases where /b/ and /p/ distinctions are phonemic in the source language.

- The English → Russian transliteration revealed challenges with the English dental fricatives /θ/ and /ð/, which do not exist in Russian. Names such as "Thomas" become "Томас" (Tomas), where /θ/ is approximated as /t/, losing the fricative quality.

- In Russian → Chinese, the absence of certain consonant clusters and final consonants in Mandarin phonotactics leads to syllable restructuring. For example, "Putin" is transliterated as 普京 (Pǔjīng), where the final consonant /n/ is dropped and the consonant cluster simplified, making the name easier to pronounce in Mandarin but less faithful phonetically.

3.2. Multiple Acceptable Variants and Orthographic Inconsistencies

Across the languages studied, many proper names exhibited multiple transliteration variants, often coexisting in official and popular usage. These variations stem from different phonological interpretations, orthographic conventions, and historical legacies.

- The Arabic name "القذافي" (the Libyan leader's name) was found with multiple Latin script renderings: "Gaddafi," "Qaddafi," "Kaddafi," and "Gadhafi." The variability results from the Arabic letter ق (qāf), which lacks an exact English equivalent and is variably rendered as "q," "k," or "g" depending on dialect and transliteration norms.

- Similarly, the Chinese transliteration of foreign names often yields multiple acceptable versions due to homophonous characters. For example, the English name "Michael" is transliterated as 迈克尔 (Màikè'ěr) or 麦克尔 (Màikè'ěr), both pronounced similarly but with different semantic characters chosen by different translators or regions.

- In French → Arabic transliterations, names like "Jacques" become جاك (Jāk) or جاک (Jāk) with slight orthographic variations depending on regional Arabic dialects (Maghrebi vs. Levantine).

3.3. Script-Specific Constraints Affecting Fidelity

The inherent properties of writing systems were shown to impose constraints that affect transliteration accuracy.

- Arabic script, an abjad, omits most vowels in written form, making it difficult to convey exact vowel qualities from source languages, especially English and French. For example, the name "David" can appear as ديفيد (Dāfīd) or داويد (Dāwīd), with vowel lengths and qualities varying depending on vowel markings (harakat), which are often omitted in everyday writing.

- Chinese logographs represent meaning rather than sound, so transliteration relies on selecting characters based on phonetic approximation, often sacrificing semantic transparency. This results in names that may carry unintended meanings or obscure the original name's cultural associations. For example, the English surname "Bush" is transliterated as 布什 (Bùshí), where the characters mean "cloth" and "this," respectively, conveying no inherent meaning related to the name itself.

- In Russian → English transliteration, letters such as щ (shch) pose challenges. For example, the Russian surname Щербаков is transliterated as "Shcherbakov," which can be cumbersome and prone to truncation or simplification in non-academic contexts.

3.4. Sociopolitical Influences on Transliteration Choices

Political and cultural factors heavily influence transliteration standards, often overriding purely linguistic considerations.

- The post-2014 Ukrainian government's initiative to promote Ukrainian transliteration norms over Russian-derived forms is a salient example. Names like "Vladimir" (Russian: Владимир) are increasingly rendered as "Volodymyr" (Ukrainian: Володимир) in international contexts, reflecting national identity and political sovereignty (Bilaniuk, 2005).

- Colonial histories explain some of the divergences between French and English transliterations of Arabic names. For instance, the Moroccan city of "Fez" is transliterated as "Fès" in French and "Fez" in English, demonstrating orthographic influence beyond phonology.
- In Chinese, political sensitivity influences the transliteration of place names and personal names, especially in Hong Kong and Taiwan, where different Romanization systems (Wade-Giles, Pinyin) coexist, leading to multiple legitimate variants (e.g., "Beijing" vs. "Peking").

3.5. Implications for Automated Systems and Data Management

The lack of consistency and the variety of transliteration variants present challenges for machine processing, database indexing, and identity verification.

- Automated transliteration tools often fail to accommodate dialectal variation or sociopolitical preferences, producing outputs that may be phonologically accurate but culturally insensitive or legally unacceptable.
- In multinational databases such as the UN or Interpol, inconsistent transliteration can cause duplicate entries or hinder record matching, impacting security and administrative processes.
- Social media and digital platforms frequently display multiple spellings for the same individual or location, complicating search engine optimization and user navigation.

In summary, the results demonstrate that orthographic challenges in transliteration are multifaceted, involving linguistic limitations, script-specific constraints, cultural dynamics, and political factors. The interplay of these elements results in a complex transliteration landscape with significant implications for cross-lingual communication and data management.

4. Discussion

The results of this study underscore the intricate complexity involved in transliterating proper names between languages with differing orthographic systems. At the core of these challenges lies a fundamental tension between phonetic fidelity and orthographic constraints, a tension that is compounded by sociopolitical and cultural factors (Sadikhova & Babayev, 2025).

4.1. Phonological and Orthographic Implications

The frequent substitution of non-native phonemes with the closest approximate sounds in the target language, as seen in the /p/ to /b/ substitution in Arabic or the /θ/ to /t/ substitution in Russian, reveals the limits of transliteration as a purely phonetic exercise. These substitutions often lead to systematic distortion of the original sound pattern, affecting name recognition and pronunciation

accuracy. This echoes observations by Knight and Graehl (1998), who noted that phoneme gaps often necessitate compromises in automated transliteration systems.

Moreover, the nature of different writing systems—alphabetic, abjad, and logographic—introduces unique orthographic constraints. The omission of vowels in Arabic script, the semantic loading of Chinese characters, and the complex grapheme-phoneme correspondences in Cyrillic to Latin transliteration all demand specialized approaches. These findings support earlier research (Comrie, 1981; Al-Jarf, 2004) highlighting script-specific barriers that hinder consistent transliteration.

4.2. Sociopolitical Dimensions

The study further highlights the sociopolitical dimensions of transliteration. The politicization of names, particularly in contexts of national identity construction or postcolonial transition, shows that transliteration is not only a linguistic act but also a cultural and ideological one. The Ukrainian shift from Russian-influenced transliteration to native Ukrainian forms exemplifies how naming practices can serve as acts of cultural assertion (Bilaniuk, 2005). This phenomenon parallels similar identity-driven transliteration reforms worldwide and underscores the need for transliteration standards to accommodate political sensitivities.

4.3. Practical Challenges in Standardization

Despite the existence of formal transliteration standards such as ISO 9 for Cyrillic or ALA-LC for Arabic, their practical implementation remains inconsistent across countries, institutions, and media outlets. The coexistence of multiple transliteration variants for the same proper name results in confusion, redundancy, and inefficiency in international communication and data systems. This inconsistency complicates legal documentation, border control, academic citation, and digital searchability.

The results suggest that standardization efforts should not aim for rigid uniformity but rather flexible frameworks that allow for phonetic, orthographic, and cultural variation while maintaining interoperability. This approach aligns with Li et al. (2020), who advocate for machine-learning-based transliteration systems capable of adapting to linguistic and cultural nuances.

4.4. Implications for Technology and Future Research

The rise of automated transliteration tools and AI-driven language processing introduces both opportunities and risks. While technology can help manage variant forms and suggest context-sensitive transliterations, current models often lack the nuanced understanding of sociocultural factors that influence human transliteration choices. There is a pressing need for more sophisticated algorithms trained on multilingual corpora with rich metadata, including dialectal variation and political contexts.

Future research should explore user-centered transliteration models that integrate feedback from native speakers, cultural experts, and legal authorities. Additionally, psycholinguistic studies on how transliterated names are perceived and processed by speakers of target languages could illuminate how phonological distortions impact communication and social identity.

4.5. Broader Linguistic and Cultural Considerations

This study contributes to the broader discourse on language contact, script interaction, and identity construction. It reveals that transliteration is a site where language, writing systems, and power relations intersect. Proper names, as carriers of identity, demand transliteration practices that respect their linguistic and cultural significance.

Ultimately, improved transliteration practices require interdisciplinary collaboration among linguists, computer scientists, policymakers, and cultural stakeholders. Only through such cooperation can transliteration evolve from a source of confusion and error into a bridge for meaningful global communication.

5. Conclusion

This study has illuminated the multifaceted orthographic and phonological challenges inherent in the transliteration of proper names between languages with divergent writing systems. The research demonstrated that transliteration is not a straightforward process of script conversion but rather a complex negotiation between phonetic accuracy, script constraints, cultural identity, and political considerations.

Key findings reveal that phoneme mismatches frequently lead to substitutions that alter the original pronunciation, while the structural properties of writing systems—whether alphabetic, abjad, or logographic—impose significant constraints on faithful name rendering. The existence of multiple transliteration variants for a single proper name further complicates the landscape, often reflecting historical legacies, regional dialects, and ideological influences.

Importantly, the study underscored the sociopolitical power embedded in transliteration practices, as names serve not only as linguistic labels but also as markers of cultural heritage and national identity. The politicization of transliteration standards, such as seen in Ukraine's efforts to promote native Ukrainian forms, highlights the broader role of naming in cultural self-determination and international diplomacy.

From a practical standpoint, the inconsistencies identified across languages and institutions pose serious challenges for legal documentation, international communication, database management, and digital searchability. These findings point to an urgent need for flexible, context-aware transliteration frameworks that balance standardization with linguistic and cultural sensitivity.

Looking ahead, advancements in computational linguistics and AI offer promising tools for addressing transliteration challenges, but their success will depend on incorporating sociolinguistic insights and involving native speaker communities. Interdisciplinary collaboration between linguists, technologists, policymakers, and cultural representatives will be essential to develop transliteration systems that are both technically robust and culturally respectful.

In conclusion, effective transliteration of proper names requires more than phonetic transcription or script mapping—it demands an integrative approach that acknowledges the linguistic diversity, orthographic intricacies, and socio-political contexts shaping language contact. By advancing such understanding, this study contributes to improving cross-linguistic communication, fostering cultural respect, and supporting global interconnectedness in an era of increasing linguistic plurality.

References

Al-Jarf, R. (2004). *Teaching English as a Foreign Language in the Arab World*. ERIC Digest.

Babayev, J. (2024). Usage Sphere and Lexico-Stylistic Features of Anthroponomastics. *EuroGlobal Journal of Linguistics and Language Education*, *1*(2), 38-43.

Bilaniuk, L. (2005). *Contested Tongues: Language Politics and Cultural Correction in Ukraine*. Cornell University Press.

International Organization for Standardization. (1995). *ISO 9:1995 - Information and documentation — Transliteration of Cyrillic characters into Latin characters — Slavic and non-Slavic languages*. ISO.

Comrie, B. (1981). The languages of the Soviet Union. *Cambridge University Press*.

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

Li, X., Zhang, M., Wang, H., & Liu, Y. (2020). Deep learning-based name transliteration with phonological awareness. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(4), 1–23.

Newmark, P. (1988). *A Textbook of Translation*. Prentice-Hall.

Sadikhova, S., & Babayev, J. (2025). Challenges Encountered in Translation of Culture-bound and Subject-specific Terminology While Using Google Translate. EuroGlobal Journal of Linguistics and Language Education, 2(3), 119-126.

Received: 08.25.2025
Revised: 09.06.2025
Accepted: 09.18.2025
Published: 09.21.2025