

# Multilingual challenges in forensic linguistics

 Ashurov Bobir Shakirovich

Independent researcher, Uzbekistan State World Languages University, Uzbekistan

**Received:** 20 January 2025; **Accepted:** 18 February 2025; **Published:** 17 March 2025

**Abstract:** In an increasingly globalized world, digital communication transcends linguistic boundaries, creating new challenges for forensic linguistics. Forensic linguistics, the study of language applied to legal and criminal contexts, plays a vital role in authorship attribution, deception detection, and threat analysis. However, the rise of multilingual communication poses unique obstacles in accurately identifying authorship and detecting deception. This research explores how multilingualism affects forensic linguistic methods, focusing on authorship attribution and deception detection in digital communication.

**Keywords:** Multilingualism, translanguaging, grammar interference, borrowed words, cross-linguistic analysis, multilingual corpora, code-switching.

**Introduction:** Forensic linguistics emerged as a discipline in the late 20th century, primarily used in authorship disputes, threat analysis, and legal testimony evaluation. Early studies emphasized stylistic markers, vocabulary choices, and syntactic patterns to identify authors. However, most foundational research focused on monolingual texts, often overlooking the complexities of multilingual communication.

Multilingualism in forensic linguistics refers to the use of linguistic analysis in legal contexts involving multiple languages. A forensic linguist must examine language across different linguistic systems to gather evidence in cases where individuals speak diverse languages. This is particularly important in multicultural societies where legal proceedings may take place in more than one language. In regions with significant linguistic diversity, accurately interpreting statements, confessions, or legal documents requires a deep understanding of language nuances. Misinterpretations or mistranslations can lead to serious legal consequences, making multilingual expertise essential in forensic investigations. It is crucial to equip lawyers, linguists, and language practitioners with the knowledge to recognize this important connection. This ensures that legal cases are not influenced by cultural and linguistic misunderstandings, such as instances where a specific word in an African language has no direct equivalent in English, potentially leading to misinterpretation and

unjust decisions. [Canagarajah, 2011; 401-417] Hence, Pavlenko's research on multilingual speakers highlights cognitive and linguistic variations that impact forensic analysis, focusing on how language diversity affects legal settings. She examines how multilingual individuals may experience linguistic bias during interrogations, trials, and asylum hearings. A key issue in forensic linguistics is whether a suspect or witness fully understands legal proceedings in a non-native language, as misinterpretations can lead to wrongful convictions or unjust decisions. Pavlenko's work highlights the need for forensic linguists to analyze language proficiency, code-switching, and translation accuracy to ensure fair legal outcomes. [García & Li, 2014; 46-62]

Therefore, multilingualism in authorship attribution and deception detection is also important to study, as it often occurs in these sections of forensic linguistics. Authorship attribution involves determining the author of a disputed text by analyzing linguistic patterns. Traditional methods rely on stylistic features such as word frequency, sentence length, and syntactic structures. However, in multilingual settings, authors may switch between languages, adopt borrowed words, or demonstrate varying levels of fluency, complicating the analysis. Studies like Makarova's (2023) highlight how bilingual authors exhibit distinct stylistic traits depending on the language they use. [Makarova, 2023]

In the digital world, authorship attribution plays a crucial role in maintaining content integrity by preventing deceptive social media activities, identifying account breaches, and connecting user profiles across different social platforms. Furthermore, authorship attribution techniques are essential in tackling misinformation, safeguarding intellectual property rights, and exposing fraudulent activities such as fake reviews.

Deception detection focuses on identifying false statements based on linguistic cues. Research indicates that deceptive texts often contain fewer self-references, more negative language, and increased cognitive complexity. However, these markers may vary in multilingual communication due to language proficiency, cultural norms, and translation effects. In cross-cultural studies, Vrij (2023) suggest that bilingual speakers display different deception patterns depending on the language used during interviews. [Poplack, 2001; 2062-2065] While deception detection relies on linguistic patterns, multilingual speakers introduce additional complexities. Code-switching and translanguaging, for instance, blur linguistic boundaries, making it challenging to apply standard forensic techniques.

Therefore, multilingual texts present several challenges in forensic linguistics:

1. Code-switching: Alternating between languages within the exact text.

Code-switching is the alternation between languages within the same discourse, often dictated by social, contextual, or pragmatic factors. According to Poplack, Code-switching (CS) is the practice in which bilingual or multilingual individuals alternate between two or more languages during communication, often without shifting the conversation's topic or changing the interlocutor. [Pavlenko, 2025] This phenomenon disrupts linguistic consistency in forensic analysis, making it difficult to determine authorship or detect deception. This variability complicates analysis and necessitates adaptive analytical frameworks.

2. Translanguaging: Mixing linguistic resources from multiple languages.

Translanguaging, in contrast, involves the fluid integration of linguistic resources from multiple languages, reflecting a speaker's dynamic communicative competence rather than mere alternation. Canagarajah (2011) defines translanguaging as the ability of multilingual speakers to move fluidly between languages, viewing their diverse linguistic repertoire as an integrated system [6]. In contrast, García and Li (2014) challenge this perspective, arguing that translanguaging is not simply

a transition between languages but rather the construction and use of complex, interrelated discursive practices. [Vrij, 2023] These practices, according to them, cannot be neatly classified under traditional language definitions but instead form a speaker's complete linguistic repertoire. Unlike code-switching, translanguaging challenges forensic analysis by defying traditional linguistic boundaries, requiring more nuanced and adaptable analytical techniques.

3. Borrowed Words: Incorporating words from one language into another.

Borrowed words are terms or phrases directly taken from another source and incorporated into a suspect text. In forensic linguistics, these words can serve as key indicators of authorship attribution. By examining the presence of such borrowed language, a forensic linguist can potentially connect an anonymous text to a particular individual by analyzing distinctive vocabulary patterns and borrowed words that consistently appear in the known writings of that author.

4. Grammar Interference: Influence of one language's grammar on another.

Grammar interference, where native language structures influence another language, is a crucial forensic marker. This linguistic footprint can reveal a suspect's linguistic background or identity, even when they attempt to disguise their language use. By examining syntactic anomalies such as incorrect verb conjugation or unusual sentence structures investigators can establish links between disputed texts and potential authors. Forensic linguists analyze such interference to identify patterns that could serve as linguistic evidence.

Examples of grammatical interference include incorrect verb conjugation, which occurs when a speaker applies their native language's verb rules incorrectly in another language (e.g., "He go to school yesterday" instead of "He went to school yesterday"). Unusual sentence structure can indicate interference when word order or phrasing does not follow the typical structure of the target language (e.g., Interesting this case is instead of This case is interesting). Incorrect word order happens when a speaker unintentionally arranges words in a way that reflects their native grammar rather than the target language (e.g., "I very like this book" instead of "I really like this book"). Misuse of prepositions or articles can also signal interference (e.g., She is in home instead of She is at home).

These factors make it difficult to determine authorship or detect deception as they introduce variability that traditional monolingual forensic techniques may not account for.

With the rapid expansion of digital text and the growing need to identify the authors of online content, the demand for accurate authorship attribution methods has significantly increased in recent years. Various computational techniques have been developed for different languages and text types, including Twitter posts, Facebook statuses, SMS messages, and chat conversations. A significant portion of authorship attribution research has focused on widely spoken languages such as English, Chinese, Spanish, and Arabic. However, literature suggests that researchers primarily direct their efforts toward English, with relatively few studies conducted on other languages. This is mainly due to the challenges associated with analyzing multilingual documents, as techniques designed for English may not be directly applicable to different languages. These differences impact the feature extraction process, making it challenging to apply English-based authorship analysis methods to other linguistic contexts. This gap necessitates the development of multilingual forensic databases and adaptable linguistic models that can handle diverse writing systems and syntactic structures.

## CONCLUSION

The impact of multilingualism on forensic linguistics is profound, as linguistic diversity introduces complexities that traditional monolingual forensic techniques often fail to address. Authorship attribution becomes more challenging when individuals switch between languages, use borrowed words, or exhibit grammatical interference. Similarly, deception detection methods must consider cross-linguistic differences in communication patterns to avoid misinterpretations. As digital communication expands, forensic linguistics must adapt by incorporating multilingual analysis methods and computational approaches that cater to diverse linguistic contexts. Future research should focus on developing multilingual corpora, improving computational models, and training forensic linguists in cross-linguistic analysis. Enhancing linguistic databases and forensic techniques for underrepresented languages will improve accuracy in multilingual investigations.

## REFERENCES

- Canagarajah, S. (2011). Codemeshing in academic writing: identifying teachable strategies of translanguaging. *Mod. Lang. J.* 95, 401–417. doi: 10.1111/j.1540-4781.2011.01207.x
- García, O., and Li, W. (2014). “Language, bilingualism and education,” in *Translanguaging: Language, Bilingualism and Education*. London: Palgrave Pivot, 46–62.
- Makarova, A. (2023). Authorship Attribution in

Bilingual Contexts. *Journal of Forensic Linguistics*.

Poplack, S. (2001). Code-switching: Linguistic. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 2062-2065). Elsevier. <https://doi.org/10.1016/B0-08-043076-7/03031-X>

University of Oslo. Aneta Pavlenko: Multilingualism and forensic linguistics. University of Oslo. Retrieved March 10, 2025, from <https://www.hf.uio.no/iln/english/research/multiling/news-and-events/news/2017/aneta-pavlenko%3A-multilingualism-and-forensic-linguistics.html>

Vrij, A. et al. (2023). Cross-Cultural Verbal Cues to Deception. *Applied Linguistics Review*.