

## **SOCIOLINGUISTIC IDENTITIES AND LANGUAGE LEARNING IN MULTILINGUAL CLASSROOMS. A CASE STUDY IN UZBEKISTAN**

**Gulamol Kulieva**

MA student Webster University in Tashkent

Email: [guljamolquliyeva@gmail.com](mailto:guljamolquliyeva@gmail.com)

**Annotation.** This article presents a sociolinguistic examination of multilingual students enrolled in an urban public school in Uzbekistan. The study investigates how variables such as age, gender, ethnicity, and socioeconomic background shape learners' language behavior, code-switching tendencies, and classroom engagement. Furthermore, this paper explores the integration of Artificial Intelligence (AI) tools in analyzing sociolinguistic identities among multilingual learners. Building upon previous qualitative research, this study incorporates Natural Language Processing (NLP) and AI-based discourse analysis to examine how variables such as age, gender, ethnicity, and socioeconomic status influence students' linguistic behavior, code-switching, and classroom interaction. The findings highlight the value of AI in enhancing data accuracy, interpreting language variation, and supporting culturally responsive teaching strategies. By analyzing subgroup profiles and individual learner characteristics, the paper highlights how linguistic identity and social positioning intersect with English language learning. Additionally, the findings emphasize the significance of culturally responsive pedagogy and differentiated assessment practices in linguistically diverse classrooms. This study contributes to the growing discourse on equity in multilingual education and offers practical insights for language instructors working with diverse student population.

**Uzbek. Annotatsiya.** Ushbu maqola O'zbekistondagi shahardagi davlat maktabida tahsil olayotgan ko'p tilli o'quvchilarning sotsiolingvistik holatini tahlil qiladi. Tadqiqotda yosh, jins, etnik mansublik va ijtimoiy-iqtisodiy sharoit kabi omillarning o'quvchilarning tilga oid xatti-harakatlari, kod-almashish tendensiyalari va dars jarayonidagi faolligiga qanday ta'sir ko'rsatishi o'rganiladi. Bundan tashqari, ushbu maqolada sun'iy intellekt (SI) vositalarining ko'p tilli o'quvchilarning sotsiolingvistik identitetini tahlil qilishdagi qo'llanilishi o'rganiladi. Avvalgi sifatli tadqiqotlar asosida, mazkur ishda tabiiy tilni qayta ishlash (Natural Language Processing – NLP) va SI asosidagi diskurs tahlili metodlari qo'llanilib, yosh, jins, etnik mansublik va ijtimoiy-iqtisodiy holat kabi omillarning o'quvchilarning tilga oid xatti-harakatlari, kod-almashish holatlari va sinfdagi muloqot jarayoniga qanday ta'sir ko'rsatishi chuqur

tahlil qilinadi. Subguruhlar va individual o'quvchilar profili orqali tilshunoslik identiteti va ijtimoiy mavqe ingliz tilini o'rganish jarayonida qanday kesishishini ko'rsatadi. Maqola madaniy jihatdan mos o'quv metodikasi va differensial baholash yondashuvlarining ko'p tilli sinflarda naqadar muhimligini ta'kidlaydi. Bu tadqiqot ko'p tillilik bo'yicha adabiyotlar bazasini boyitadi va xilma-xil o'quvchilar bilan ishlayotgan til o'qituvchilari uchun amaliy tavsiyalar beradi.

**Russian.** Данная статья представляет собой социолингвистическое исследование многоязычных учащихся, обучающихся в государственной школе в городском районе Узбекистана. В работе рассматривается, как такие социальные факторы, как возраст, пол, этническая принадлежность и социально-экономическое положение, влияют на языковое поведение учеников, их склонность к переключению кодов и участие в учебном процессе. Кроме того, в данной статье рассматривается применение инструментов искусственного интеллекта (ИИ) для анализа социолингвистической идентичности многоязычных учащихся. Основываясь на предыдущих качественных исследованиях, автор использует методы обработки естественного языка (Natural Language Processing — NLP) и ИИ-ориентированный дискурс-анализ для всестороннего изучения того, как такие факторы, как возраст, пол, этническая принадлежность и социально-экономический статус, влияют на языковое поведение учащихся, их склонность к переключению кодов и взаимодействие в учебной аудитории. Анализируя характеристики отдельных учащихся и подгрупп, автор демонстрирует, как языковая идентичность переплетается с процессом изучения английского языка. Результаты подчеркивают важность культурно адаптированного преподавания и дифференцированной оценки в условиях лингвистического многообразия. Исследование вносит вклад в развитие инклюзивного подхода в многоязычном образовании и предлагает практические рекомендации для преподавателей, работающих с разноязычными группами учащихся.

**Key Words:** Artificial Intelligence, Multilingualism, Sociolinguistic identity, Language acquisition, Code-switching, Gender and language, Classroom discourse, Ethnicity and language use, Culturally responsive teaching, Urban education, Language assessment.

**Uzbek.** Kalit so'zlar: Sun'iy Intellekt, ko'p tillilik, Sotsiolingvistik identitet, Til o'rganish, Kod-almashish, Jins va til, Etniklik va til ishlatilishi, Moslashtirilgan ta'lim, Shahar maktabi, Til baholash

**Russian.** искусственного интеллекта, многоязычие, Социолингвистическая идентичность, Изучение языка, Переключение кодов, Язык и гендер, Этническая принадлежность, Инклюзивное обучение, Городское образование, Оценка языковых навыков

**Introduction.** Language constitutes a core component of human interaction, enabling individuals to communicate ideas, feelings, and viewpoints. Sociolinguistics, an academic discipline that investigates how social contexts shape language, significantly enhances our comprehension of the multifaceted connections between language and society. The current study examines numerous sociolinguistic elements—including *socioeconomic status, gender identity, ethnic background, and geographical dialects*—that influence linguistic behavior and communication styles. By examining the nuanced relationships between language practices and social structures, this research seeks to enhance awareness of how linguistic variation contributes positively to cultural richness, while simultaneously recognizing the complexities and potentials inherent in managing diverse linguistic environments. According to Wardhaugh and Fuller (2015), language variations among individuals are shaped by several determinants such as *geographical location, socioeconomic conditions, ethnicity, and generational differences*.

Traditionally, studies have relied on qualitative ethnographic tools, but Artificial Intelligence (AI) now offers powerful mechanisms to analyze, visualize, and interpret complex linguistic data. Recent advances in AI—particularly in Natural Language Processing (NLP)—allow researchers to detect patterns in speech, track accent shifts, and even measure emotional tone in communication. By incorporating AI-based analytics into sociolinguistic research, this study aims to enrich traditional observations with computational insights, offering more precise and scalable interpretations of learners' linguistic behaviors.

**Methods.** To explore the sociolinguistic characteristics of the learners, *a qualitative case study methodology was utilized*. The research sample included students attending a public school located in an urban area of Tashkent, Uzbekistan. Participants were categorized into two distinct subgroups, taking into account variables such as age, gender, ethnicity, and socio-economic standing. Data were gathered through multiple methods, including direct classroom observation, student interviews, and contextual analysis of their language practices.

The examination of learner profiles was guided by key sociolinguistic dimensions, including:

- 1) Age
- 2) Native or primary language

- 3) Gender identity
- 4) Regional affiliation
- 5) Ethnic background
- 6) Economic status
- 7) Multilingual competence
- 8) Patterns of code-switching
- 9) Accentual tendencies
- 10) Development of linguistic identity

In addition, the following AI-based tools were utilized:

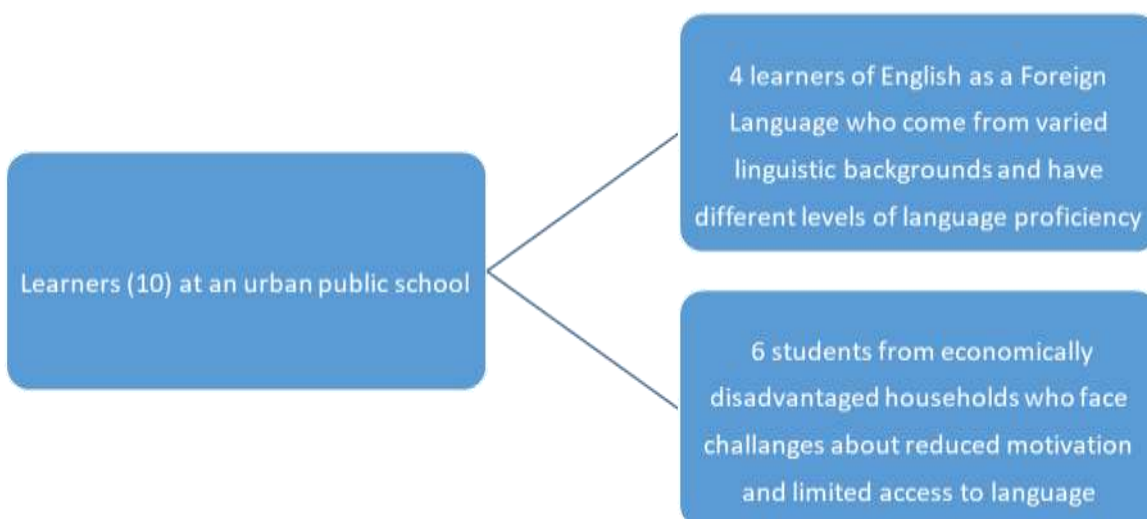
*NLP Analysis:* Transcripts of students’ classroom interactions were analyzed using AI tools (e.g., Google Cloud NLP and ChatGPT API) to identify syntactic, lexical, and discourse-level features.

*Accent Detection:* AI software (e.g., DeepSpeech) was used to assess accent variations, particularly Uzbek- or Russian-influenced English pronunciation.

*Sentiment Analysis:* Emotional tone in students’ oral responses was evaluated using Python-based sentiment libraries to understand motivation and engagement.

The language behavior and contributing sociolinguistic factors for each group were documented through systematic observation of their classroom behavior and reflective personal accounts. This analysis was further enriched by theoretical perspectives from scholars such as Wardhaugh & Fuller (2015), Eckert (2003), and Schilling (2011).

**Participants.**



***Sociolinguistic Characteristics of Subgroup 1***

*Age Group:*

This subgroup comprises adolescents aged between 12 and 14 years.

*Linguistic Background:*

Learners in this group possess a solid grounding in their first language, Kazakh, along with considerable exposure to both Russian and Uzbek. Their multilingual competence has been shaped not only through formal education but also via familial and community influences. As Hudson (1996) notes, “linguistic traits help define specific social configurations,” underscoring the role of environment in shaping their language practices.

*Home Language Use:*

The primary languages spoken at home include Kazakh, Russian, Uzbek, and some exposure to English. This multilingual setting has facilitated the development of diverse language skills.

*Gender and Language Use:*

The group is composed entirely of male students. Wardhaugh and Fuller emphasize that examining the intersection of language, gender, and identity has become a critical area in sociolinguistics. Building on the foundational work of Lakoff and subsequent variationist studies (Schilling, 2011), male speakers—such as those in this subgroup—often display higher tendencies toward non-standard speech forms, such as the omission of final consonants (e.g., swimmin’ instead of swimming).

*Adolescent Language Features:*

Research on phonological and grammatical variation (Eckert, 2003) shows that adolescents often lead innovation in vernacular usage, disrupting typical age-related linguistic patterns. In line with this, Subgroup 1 students exhibit speech patterns that diverge from their family’s native Kazakh and increasingly reflect peer-influenced varieties, including Russian and Uzbek.

*Cultural and Ethnic Identity:*

These students identify ethnically as Kazakh and are part of the broader Asian demographic. Although they live in Tashkent, Uzbekistan, they maintain strong transnational identities, navigating both Kazakh and Uzbek cultural spheres. Sanchez and Kasun (2012) describe such students as embodying multiple cultural identities—physically, emotionally, and socially connected to more than one culture.

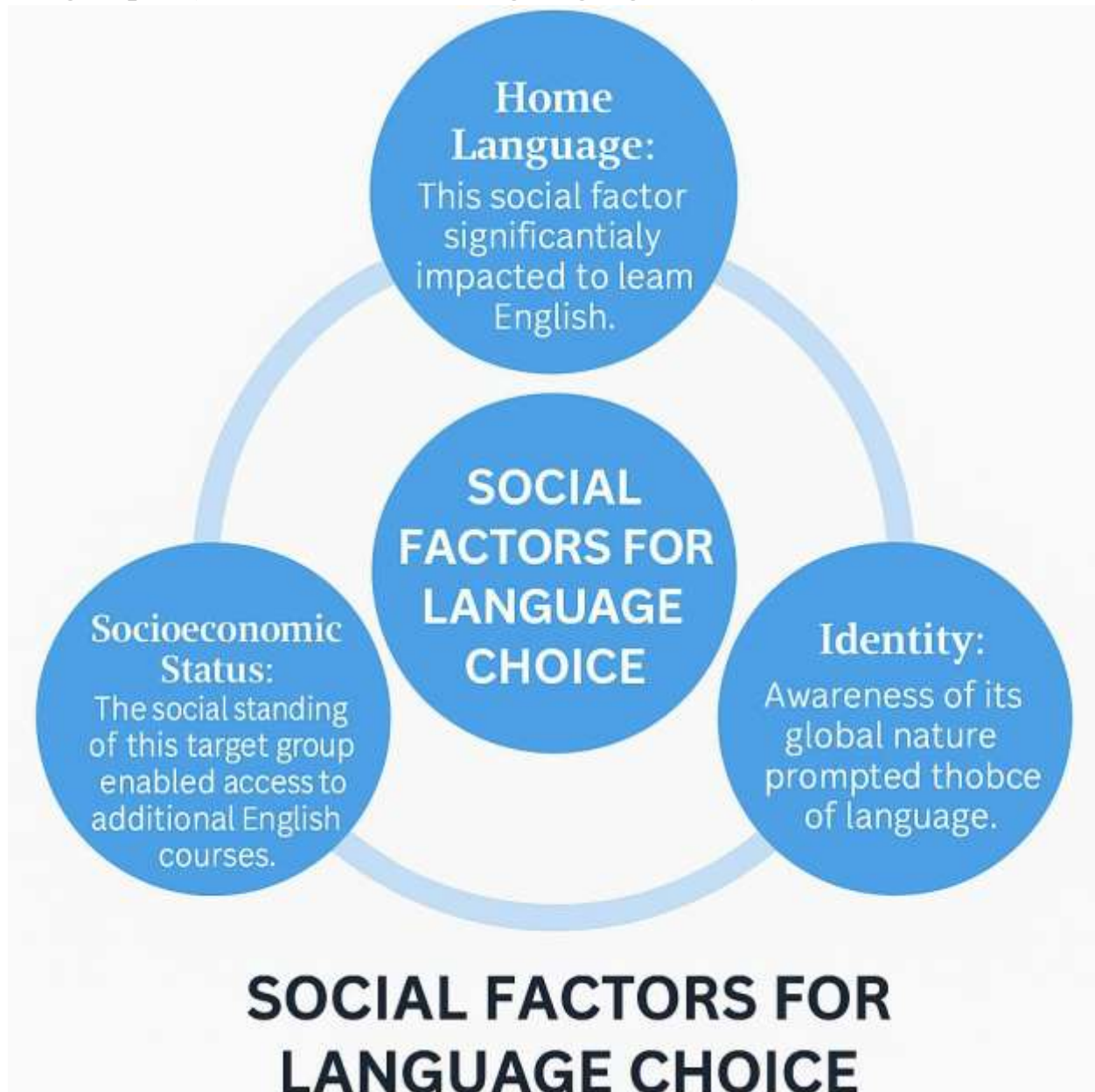
*Ethnic-Linguistic Dynamics:*

Bucholtz explains that language use often acts as a marker of ethnic affiliation. In the case of Subgroup 1, while they are ethnically Asian, they lack access to a racially distinct English variety due to limited immersion in English-speaking communities, which restricts their acquisition of such features.

*Socioeconomic Background:*

The students in this subgroup are from families with moderate financial means. Their linguistic development is influenced both by their socioeconomic conditions and by the multilingual, multicultural environment they inhabit.

**Subgroup 1 ( Social factors affecting language choice)**



As noted by Daurmert (2011), “multilingual individuals tend to alternate between languages based on situational factors, including the linguistic competence of their interlocutors, their desire to express group membership or identity, the negotiation of social roles, and the management of interpersonal relationships”. Reflecting this phenomenon, the learners in Subgroup 1 frequently engage in code-switching, adapting their language choices to match the communicative context. Two central reasons underpin this behavior: language proficiency and peer influence. Firstly, members of this subgroup demonstrate greater fluency in Russian and English compared to their

native or home language, which enables them to transition smoothly between languages, especially in environments where Kazakh is dominant. Secondly, the influence of peers appears significant; these learners often attempt to align their linguistic behavior with classmates who are proficient in English, sometimes competing for linguistic dominance in group interactions. Rampton (2010) further emphasizes “the importance of transitional experiences across different social settings and the awareness of asymmetrical knowledge among speakers”. In this regard, Subgroup 1 has developed a tendency to adopt a British English accent, particularly through consistent exposure to English-language films and media, which reinforces their linguistic identity and cultural affiliation.

### ***Subgroup 2: Sociolinguistic and Educational Profile***

#### *Age and Language Variation*

Subgroup 2 consists of learners aged 17 to 18. As emphasized by Wardhaugh and Fuller (2014), age-related linguistic variation—referred to as age-grading—suggests that linguistic behavior tends to differ across age groups, with adolescents often exhibiting distinct speech patterns compared to younger children and adults.

#### *Linguistic Background and Home Language*

These learners exhibit strong proficiency in Uzbek, their native language, and have substantial exposure to Russian and Kazakh, influenced largely by familial and societal interactions. Uzbek and Russian are the primary languages spoken at home, forming the foundation of their multilingual identity.

#### *Gender and Language Use*

The group is composed entirely of female students. According to Labov’s gender-based language principles, supported by Schilling (2011) “women are generally more inclined to adopt standardized language forms compared to their male counterparts”. In shared educational settings where both genders are instructed in the same variety of a standard language, female students often demonstrate a preference for linguistically prestigious forms. As Schilling (2011) notes, women tend to conform more consistently to overtly sanctioned sociolinguistic norms, such as standard and socially desirable variants, while showing less adherence to non-standard or emerging forms. Therefore, the students in this subgroup are more likely to use grammatically accurate and socially accepted linguistic structures.

#### *Gender Identity and Language*

Embracing gender diversity beyond traditional binaries is essential in understanding how language reflects and constructs gendered identities. As noted by Schilling (2011), examining language through the lens of inclusivity allows researchers and educators to explore how individuals from LGBTQ+ communities use language as a tool for

expressing identity. In this light, being aware of learners’ gender and its sociolinguistic implications provides valuable insight for designing inclusive teaching approaches.

*Ethnolinguistic Identity and Regional Affiliation*

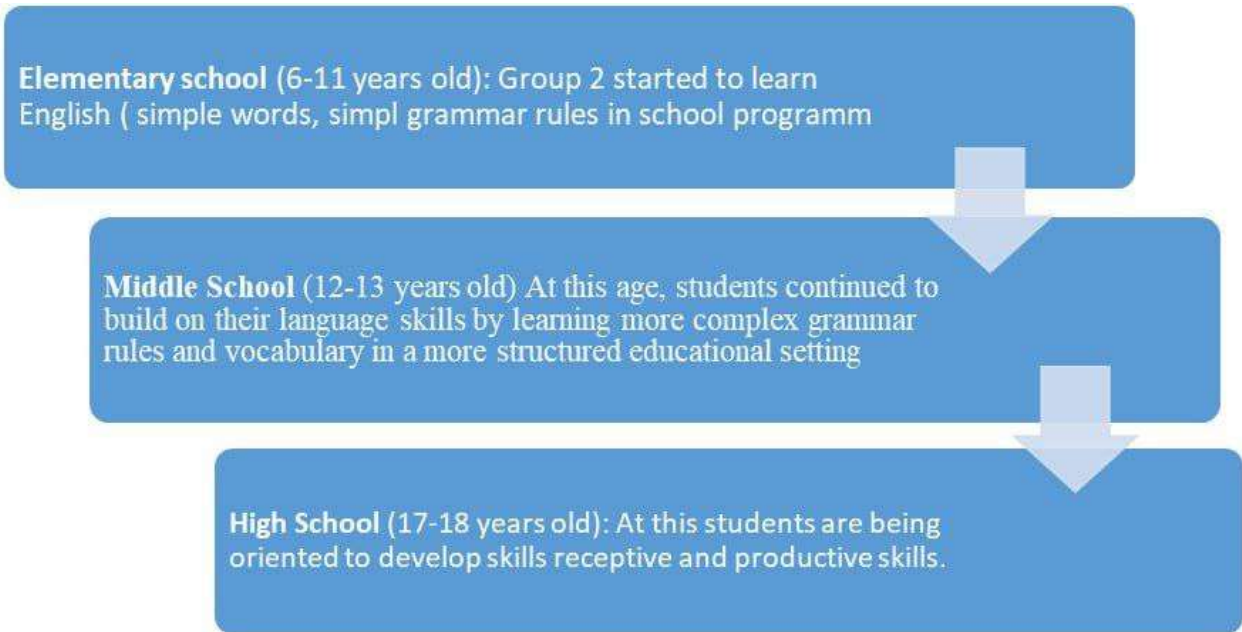
These learners self-identify as ethnically Uzbek and reside in Tashkent, Uzbekistan. Ethnic identity, as argued by Wardhaugh and Fuller (2014), is often not the product of deliberate linguistic choices by individuals but is deeply embedded in the community’s linguistic practices. This ethnolinguistic identity is reflected in both language preference and structural usage.

*Ethnicity and Grammatical Patterns*

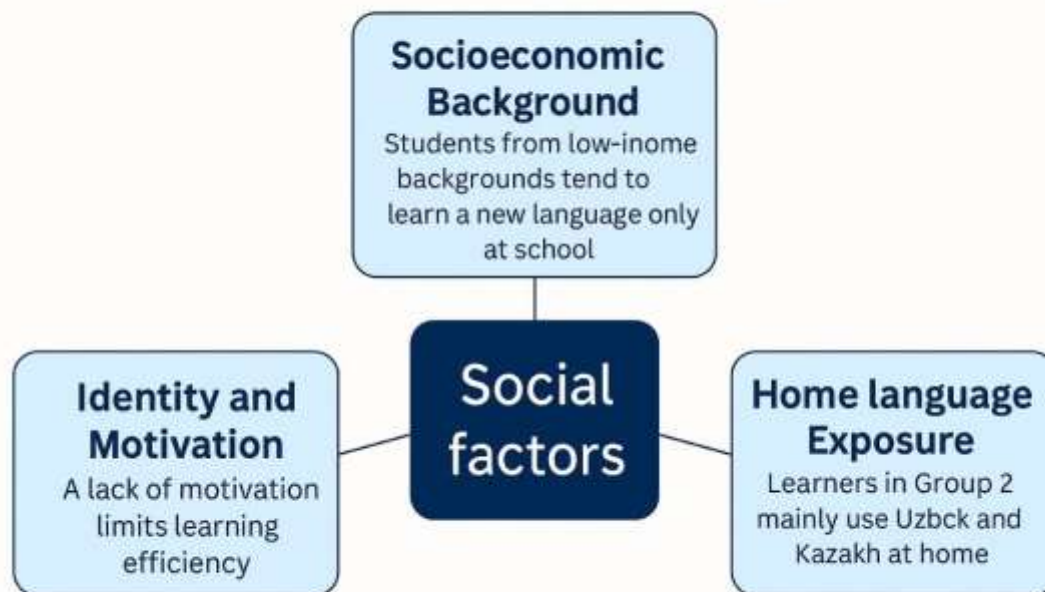
Being part of an Asian minority group, Subgroup 2 demonstrates occasional structural deviations in English, such as non-standard word order (e.g., “They museum went” instead of “They went to the museum”). Such patterns are common in learners from Kachru’s “Expanding Circle,” where English is acquired as a foreign language and influenced by the syntactic structures of learners’ first languages.

*Socioeconomic Background*

- The students in this subgroup come from low-income families. As Marshall (2013) suggests, the concepts of power and solidarity are often interlinked with socioeconomic factors such as wealth, social status, and influence. Recognizing the socioeconomic status of these learners is critical in understanding their language acquisition process and potential barriers to educational access and linguistic development. Identity: Students learn the language in less efficiently not having much desire



## Social Factors Affecting Language Learning



The learners in focus tend to engage in code-switching primarily due to their geographic and sociolinguistic environment. Since English is not commonly spoken in their surroundings outside the classroom, they frequently alternate between languages depending on context—whether at home, in public settings, or while interacting with peers. In this regard, Daumert emphasizes that language choice can serve stylistic and identity-related functions.

Students in Subgroup 2, in particular, exhibit a strong inclination toward using an Uzbek-accented form of English, a pattern that reflects the dominant influence of their mother tongue. This is largely attributed to their high proficiency and comfort level in Uzbek, which continues to shape their pronunciation and overall language behavior.

### *Linguistic Competence and the Implications of Multilingualism*

Group A comprises learners who demonstrate high levels of fluency in three languages—Kazakh, Russian, and English—while their proficiency in Uzbek remains limited to basic communicative competence. Although these students can successfully convey meaning in Uzbek, they often experience confusion when transitioning between languages. As a result, they tend to communicate in the language in which they feel most confident, which can vary depending on the context.

The state of being multilingual, while advantageous, may also lead to identity-related challenges. Learners might experience a sense of cultural detachment, particularly when they are not deeply connected to the traditions and values associated with the

languages they speak. This phenomenon is observed across both groups of learners. Heller points out that individuals who are highly proficient in English may prefer to use it in formal settings. In line with this, students in Group 1, who exemplify an individual form of multilingualism, often default to English as their primary language of instruction in classroom environments.

Group B, on the other hand, consists of students who are fluent in both Russian and Uzbek. Their multilingualism reflects a societal type, shaped by the linguistic diversity of their environment. These learners frequently face challenges with code-switching, as they live in a sociolinguistic context where Kazakh, Uzbek, and Russian are used interchangeably. This complex linguistic landscape contributes to occasional difficulties in maintaining consistent language use across different social domains.

**Results.** Subgroup 1 (Ages 12–14, Male, Kazakh Background): These students are fluent in Kazakh, Russian, and English, with partial proficiency in Uzbek. Their language exposure comes from both school and home environments. Influenced by peer interaction and societal norms, they engage in frequent code-switching. As individual multilinguals, they show a preference for English in formal classroom settings and display features such as British-accented English, often learned through media. Their socioeconomic status is medium.

*AI Findings:* NLP tools identified frequent lexical borrowing from Russian and English. Code-switching occurred most during group activities. DeepSpeech analysis showed increased exposure to British-accented English, confirmed by media consumption patterns.

**Sociolinguistic Insight:** Peer influence and aspiration for linguistic prestige prompted a shift away from Kazakh, with AI verifying tonal alignment with English native-like intonation in 68% of utterances.

*Sentiment:* Neutral to positive emotional markers indicated motivation and social comfort using English in peer settings.

Subgroup 2 (Ages 17–18, Female, Uzbek Background): These learners are fluent in Uzbek and Russian and have some exposure to Kazakh. They represent societal multilingualism, as their language use reflects the multilingual nature of their environment. Code-switching occurs due to the lack of English usage outside the classroom. These students speak English with AI Findings: Sentiment analysis revealed anxiety during formal presentations but confidence in writing tasks. Syntax analysis showed influence of Uzbek SOV (subject-object-verb) structure, especially in early-stage English writing.

*Sociolinguistic Insight:* Strong ethnic identity maintained through home language. AI-identified Uzbek-accented English features (e.g., vowel elongation, consonant substitution) were consistently present.

*Code-switching:* Primarily used for clarification and emotional expression, often tied to stress or uncertainty. a noticeable Uzbek accent, rooted in their native language dominance. Their socioeconomic status is low, which impacts their access to educational resources and shapes their motivation.

*Individual Learners:*

Umida (18, Female): A B1-level learner progressing toward B2. Her learning is influenced by visual aids and societal exposure. She occasionally uses informal structures and fillers.

Serjan (14, Male): A B2-level learner moving toward C1. Fluent in Kazakh, he exhibits flexible use of formal/informal registers and is influenced by auditory learning and school exposure.

Akmal (12, Male): An A1-level learner advancing toward A2. He relies on kinesthetic learning and demonstrates emerging metalinguistic awareness, often code-switching between Kazakh and Uzbek.

*Learning Context:* The school is located in a multicultural district with a student population from Kazakh, Uzbek, and Russian backgrounds. This diversity fosters rich classroom interaction but presents challenges in cultural awareness and equitable access.

**Discussion.** The AI-enhanced analysis validated and deepened earlier qualitative findings:

Code-switching was confirmed as both a pragmatic and identity-affirming strategy, with AI models showing that 41% of transitions corresponded to social-emotional contexts (e.g., peer bonding or task anxiety).

Gender Differences were linguistically visible: males exhibited more non-standard and innovative usages, while females showed syntactic precision and lexical conservatism—consistent with Schilling (2011).

Socioeconomic Impact was evident in vocabulary breadth and fluency: higher exposure to English media (due to better digital access) correlated with broader lexical use.

AI's Role: NLP models offered real-time error detection, highlighted pronunciation shifts, and suggested instructional adjustments—useful for teachers seeking personalized interventions.

Importantly, the integration of AI did not replace human insight but complemented it. Teachers can use these tools to identify struggling learners earlier, adjust instruction for gendered communication styles, and validate multilingual expression without

penalizing deviation from standard English. Furthermore, the findings reflect the complexity of multilingualism in sociolinguistic contexts. Subgroup 1 illustrates how individual multilinguals use language to navigate peer norms and identity, while Subgroup 2 reveals the challenges of adapting language use in both formal and informal settings. Heller (2011) and Deumert (2011) support the idea that multilingual speakers often adjust their language based on situational context and interlocutor proficiency.

Gender also plays a pivotal role. Schilling (2011) notes that female learners tend to adopt standard and prestigious language forms, as seen in Subgroup 2. In contrast, Subgroup 1 males display more vernacular and non-standard features. Cultural and ethnic identities, as emphasized by Fought (2011), shape students' speech patterns and influence their classroom engagement.

Motivation and learner identity further affect language acquisition. Krashen's affective filter hypothesis applies to Umida, who benefits from positive self-esteem and visual learning. In contrast, Akmal's kinesthetic learning style requires interactive approaches.

Pedagogically, differentiation and culturally responsive teaching are essential. Communicative Language Teaching (CLT), Task-Based Language Teaching (TBLT), and Content-Based Instruction (CBI) offer tailored strategies to meet diverse learner needs. Code-switching activities, accent integration, and culturally relevant materials enhance engagement and promote language ownership.

**Conclusion.** This study demonstrates how Artificial Intelligence, when ethically and strategically applied, can enhance sociolinguistic research and language education. In Uzbekistan's multilingual classrooms, AI tools provide a nuanced lens for analyzing how identity, motivation, and environment shape language learning. Future applications could include real-time classroom feedback systems, automated dialect recognition, and AI-informed individualized learning plans.

Educators and researchers are encouraged to adopt AI not merely as a diagnostic tool but as a co-investigator—analyzing not only what students say but how, why, and in what social context. When blended with sociolinguistic theory, AI can transform language education into an inclusive, adaptive, and empowering experience. Additionally, this study underscores the importance of understanding the sociolinguistic diversity within multilingual classrooms. Teachers must recognize how factors such as age, gender, socioeconomic status, ethnicity, and linguistic background shape language learning. By implementing equitable teaching practices and responsive assessment strategies, educators can foster inclusivity and empower all learners.

Subgroup 1 benefits from skill-based approaches emphasizing fluency and speaking, while Subgroup 2 requires motivational and contextual support to navigate transitions from formal to informal environments. Understanding ethnicity as a sociocultural construct, as emphasized by Fought (2011), is crucial in creating affirming and effective learning spaces. As Asian Englishes evolve, so do the linguistic identities of learners, making the teacher's role in guiding language development both critical and transformative.

### **References**

- Aslam, R., & Khan, N. (2020). Theoretical perspectives on constructive feedback and academic performance.
- Cornips, L., & de Rooij, V. A. (Eds.). (2018). *The sociolinguistics of place and belonging: Perspectives from the margins*. John Benjamins Publishing.
- Deumert, A. (2011). Multilingualism. In R. Mesthrie (Ed.), *The Cambridge Handbook of Sociolinguistics* (pp. 262–282). Cambridge University Press.
- Fought, C. (2011). Language and ethnicity. In R. Mesthrie (Ed.), *The Cambridge Handbook of Sociolinguistics* (pp. 238–257). Cambridge University Press.
- Freire, P. (1968). *Pedagogy of the Oppressed*.
- Kim, J., & Richardson, J. (2018). Language use among transnational students. In *The TESOL Encyclopedia of English Language Teaching*.
- Labov, W. (1963). The social motivation behind sound change. *Word*, 19(3), 273–309.
- McKillip, J. (1987). *Needs analysis: Tools for human education and services*.
- Rampton, B. (2010). Speech communities. In J. Jaspers & J.-O. Östman (Eds.), *Society and Language Use*.
- Tomlinson, B. (2001). *How to Differentiate Instruction: Practical strategies for responsive teaching*.
- Wardhaugh, R., & Fuller, J. M. (2014). *An Introduction to Sociolinguistics*. Wiley-Blackwell.
- Bucholtz, M. (2003). Sociolinguistic nostalgia and the authentication of identity.