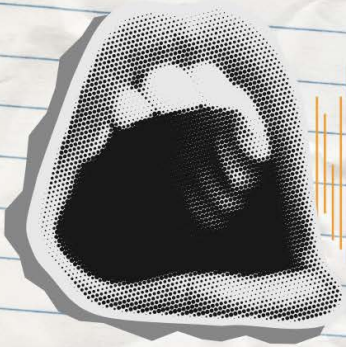

# Sounds of Identity: How Does English Pronunciation Reflect Identity, Space, and Linguistic Background?

Lida Zarearsanjani, Emad Elbana, and Ivan Aparicio Sausedo

This essay examines how linguistic identity is constructed and negotiated across spaces and how phonological variation is shaped by place, movement, and linguistic experience. It calls for deeper consideration of the less visible dimensions of language, particularly the sound system. Drawing on three research projects conducted in the fields of phonetics and phonology, this essay compares the English pronunciation of nonnative speakers from Kenya, Iran, and Argentina with that of three native English speakers from the Midwest United States. In doing so, it explores how speech production is inherently variable and influenced by a range of physical, social, and experiential factors, reflecting the uniqueness of each speaker's linguistic trajectory and identity. By reframing phonological variation as spatially produced rather than individually deficient, the essay demonstrates that English does not exist as a single, homogeneous system but rather as a constellation of situated practices shaped by where and how it is used. Through the interaction of first-language (L1) phonology, second-language (L2) acquisition and learning, and specific situational contexts, phonological variation emerges as a meaningful and systematic phenomenon. Both where one has lived and the social or situational context of speech (i.e. as formal versus informal settings) manifest spatial differences that shape pronunciation. Geographic space, linguistic exposure, and mobility all influence English pronunciation, thereby contributing to the ongoing construction of linguistic identity.

# Sounds of Identity: How Does English Pronunciation Reflect Identity, Space, and Linguistic Background?

Lida Zarearsanjani, Emad Elbana, and
Ivan Aparicio Sausedo
Ball State University

## Introduction

How can a single wave of sound both unite speakers across continents and mark profound differences between them? Spoken language is never merely a neutral vehicle for meaning. It carries histories of movement, structures of power, and traces of the lived spaces in which it is learned and practiced. As David Crystal argues, linguistic diversity preserves distinct ways of interpreting human experience, reminding us that variation is not deficiency, but cultural knowledge encoded in sound (6). In a globalized world in which English functions as a lingua franca—a prominent means of communication among speakers of other languages—pronunciation becomes one of the most visible sites where identity, geography, and social space intersect. Pronunciation is shaped not only by where a speaker has lived but also by the social and situational context in which they speak. These two manifestations of spatial difference—geographic/mobility factors and variation in register, or "variation in speech" based on social setting—interact to influence how speech is produced and perceived. What, then, can be learned from the production of English across multiple spaces by diverse speakers?

Recent scholarship in the humanities has increasingly turned to space as a critical lens of interpretation, largely through examinations of how identity is produced not only through time and narrative but also through the physical, social, and imagined spaces people inhabit.[1] Language,

1 See Bachelard, Gaston. *The Poetics of Space*. Translated by Maria Jolas, Beacon Press, 1994; Foucault, Michel. "Of Other Spaces." 1967. *Diacritics*, vol. 16, no. 1, 1986, pp. 22-27;

too, is spatial. It travels across borders and it is reshaped by the conditions of its use. Phonology offers a particularly revealing point of entry into these processes because it reflects deeply embodied habits formed through repeated interaction with specific linguistic environments. Because each individual's linguistic history is unique, speech production is never uniform. Variation is an inherent part of language use and arises from multiple factors, including physical differences and social and linguistic backgrounds (Zsiga 438 [2nd ed.]).

Variation according to place underpins the concept of dialect. As Zsiga explains, a regional dialect is a variety spoken by a group of people who live in a particular area. Since every individual is born and raised in a specific region, each person naturally speaks a dialect, reflected in accent and pronunciation (438). In this sense, differences between native and nonnative pronunciation are rarely random. Rather, they reflect systematic interactions among perception, production, articulatory habits, and cross-linguistic transfer. Research in second-language phonology further demonstrates that first-language sound systems strongly influence how new sounds and syllable structures are perceived and produced, particularly when the target language contains marked or unfamiliar structures (458).

Integrating this linguistic perspective with spatial theory allows pronunciation to be understood not merely as a technical outcome of language learning, but as an embodied record of where and how language has been lived. No dialect is inherently superior to another, nor is any variety intrinsically correct or incorrect. Dialects carry rich cultural and historical significance, reflecting the identities and lived experiences of their speakers.

## Methodology

Linguists often rely on careful listening and acoustic analysis to investigate how language reflects identity and space. By attending closely to pronunciation patterns, researchers can uncover how speakers' linguistic choices index their geographic backgrounds, social experiences, and interactional contexts. Listening is therefore not a passive activity, but an analytical method that allows linguists to interpret how sound encodes lived experience and spatial history.

The research projects analyzed

and Bhabha, Homi K. *The Location of Culture*. Routledge, 1994.

here recorded speech from three L1 English speakers, and three L2 English speakers. The data was collected using open-ended questionnaires to elicit speech production in a casual manner, and text scripts that addressed possible differences in segments. Using both auditory analysis and instrumental tools such as Praat, a software for phonetic analysis (Boersma & Weenink 2025), we examined segmental and suprasegmental features in selected words and speech samples. Segmental features refer to the individual sounds of speech, such as consonants and vowels (phonemes), that occur in sequence and serve to distinguish meaning. Suprasegmental features, by contrast, extend over more than one sound and include elements such as syllable structure, stress, intonation, rhythm, and pitch, all of which shape meaning and contribute to the overall flow and pattern of spoken language. This approach allowed us to compare how speakers from different geographic and social spaces produce English sounds, and to identify distinctive features of vowel sounds and consonants using the International Phonetic Alphabet for transcription. The purpose of this alphabet is to provide a unique symbol for each distinct sound in a language, meaning every phoneme that distinguishes one word from another ("International Phonetic Alphabet").

## English as a Lingua Franca

"Lingua Franca" refers to the global status that English has acquired over time, in which it is a means of communication among speakers of other languages. The language backgrounds of the participants in the three research projects revealed how and why they learned English, as well as how these experiences shaped their views on the global purpose of the language. For the Kenyan speaker, English represented more than a language—it was a pathway to new opportunities and social mobility. She viewed English as a "ship to the world," a tool to pursue her professional ambitions and connect with broader communities beyond Kenya. She was motivated to learn English by her desire to work in English-speaking spaces, where she could teach and share knowledge as an instructor. This aspiration reflects the broader role of English as a lingua franca, functioning not just as a means of communication but as a marker of status and mobility. Her phonological choices, including careful articulation influenced

by classroom-based learning and literacy practices, demonstrate how the spaces in which she learned English—classrooms, media, and formal educational environments—shaped her speech. English in multilingual societies often develops as a language of education and economic opportunity, which is evident in how her pronunciation reflects both L1 influence and the social function of English in Kenya (Kachru 125).

The Persian speaker's experience with English similarly highlights its role as a lingua franca, though shaped by a different spatial trajectory. After moving from Iran to the United States, English became central to her ability to navigate academic and professional spaces. Rather than aiming to sound like a native speaker, her use of English is oriented toward effective communication and participation in these settings. This functional orientation is reflected in her pronunciation, which maintains features of Persian phonology—such as vowel insertion in consonant clusters, as in "Skype" pronounced [ɛskaɪp], and the clear articulation of final consonants at the ends of words—while remaining highly intelligible. These patterns suggest that, for this speaker, English operates as a practical and social resource rather than a marker of native-speaker identity. Her pronunciation illustrates how speakers using English as a lingua franca adapt the language to meet communicative needs within specific spaces, without fully suppressing the phonological patterns shaped by earlier linguistic experience.

At a very young age, the L2 English speaker from Argentina began learning English at a language school. Initially motivated by his family's encouragement and the practical benefits of learning the language, he gradually grew to enjoy learning English and developed positive attitudes toward the imagined English-speaking culture. He later pursued higher education in the field of English teaching, further consolidating his academic and professional development not only in his home country, but also in the U.K. and in the U.S. This immersive experience strengthened his language proficiency and fostered integrative motivation, a concept developed by Gardner (qtd. in Paltridge and Phakiti 404) as he developed meaningful connections with the local culture and community. Throughout his language

experience, his understanding of English expanded beyond seeing it solely as the language of native speakers. He came to view it as a global lingua franca, a tool for communication among speakers from diverse linguistic and cultural backgrounds. This perspective has shaped his professional identity and approach to language, emphasizing the role of English in intercultural communication and global engagement.

## Language Identity Is Shaped by Geography and Cultures

English does more than facilitate communication; it also acts as a lens through which speakers express their identities. The way individuals pronounce English carries traces of their cultural background, educational experiences, and the regions in which they have lived. The speech patterns of the Kenyan, Persian, and Argentine participants illustrate how accent and pronunciation reflect not only language learning but also the social and geographic journeys that shape each speaker's language identity. The Kenyan speaker's identity is rooted in her African, Kenyan upbringing and cultural beliefs, which continue to influence her speech even after moving to the United States. Despite her ability to approximate native English sounds, she maintains distinct pronunciation patterns as an assertion of self and cultural belonging. For example, she pronounces words like "education" and "schedule" differently from an American speaker. The American speaker shows the expected palatalization of the /dʒ/ sequence, resulting in affrication (e.g., [ˌɛdʒuˈkeɪʃən], [ˈskɛdʒul]),[2] whereas the Kenyan speaker produces them as [ˈskɛdul] and [ˌɛduˈkeɪʃən].[3] Importantly, this does not mean that Kenyan English lacks the [dʒ] affricate because Kenyan English does include [dʒ] in many other lexical positions. The difference observed here is more plausibly explained by orthographic influence, especially in contexts where English is learned primarily as a classroom subject rather than through everyday exposure to native speakers. Moreover, this difference is not a deficit but a reflection of her identity and the linguistic space she inhabits. Her pronunciation demonstrates that language is both a personal and cultural map: it preserves her connection to Kenya while allowing her to navigate new social and professional spaces in the U.S.

2 The dʒ sounds like the letter "j".

3 The "d" in this case is pronounced like "d" in "dog".

The Persian speaker's identity similarly interacts with her spatial and social experiences. Despite achieving high intelligibility in English, she preserves certain phonological features of Farsi as a marker of cultural identity, such as full articulation of coda consonants and epenthetic vowels in onset clusters ([ɛskaɪp]).[4] Her pronunciation reflects experiences across multiple spaces—home in Tehran, classrooms in Iran, and professional spaces in the U.S. Acoustic evidence shows that even when adapting to American English norms, Persian speakers retain subtle prosodic and segmental features of their L1, highlighting the spatial and cultural dimensions of identity in pronunciation. By contrast, the Midland American English speaker's pronunciation reflects immersion in domestic, educational, and community spaces, producing context-dependent variations such as flapping[5] and vowel reduction. This comparison illustrates how both native and non-native speakers' phonological patterns are shaped by the interaction of social, cultural, and spatial environments, showing that identity is embedded in speech across contexts.

The language background of the Argentine speaker and his experience living in the U.K. supports the way he produces sounds with a clear preference for what is considered Received Pronunciation British English accent, building up a linguistic identity that reflects the process and influence of learning a target language in an EFL Classroom, since Argentina is a Spanish-speaking country. The sounds analyzed from the American English speaker align with the Midland accent of the region of his birth, serving as an example of "standard" General American English. For example, a clear distinction between R-dropping vs. R-full is noticed when pronouncing the word "car" in casual speech, the spectrogram of this speaker showed the absence of a rhotic transition, indicating deletion of the final [ɹ].[6] On the contrary, the spectrogram of the L1 English speaker produced a voiced alveolar approximant [ɹ], showing the r-full of General American English, which is pronounced at

4 A vowel sound is added before the word's initial consonant, turning it into something like "es-kaipe".

5 Flapping: the consonants /t/ and /d/ are pronounced as a quick, voiced flap sound (i.e., "butter" sounds like "budder").

6 In this case, "car" ends abruptly as "cah" without an r-colored vowel transition.

the end of a word (Zsiga 66 [2nd ed.]). This reinforces the idea that variation in sounds reveals aspects about the speaker's language identity constructed through spaces where regional accents, and language learning experience are intertwined.

## Language Pronunciation Is Shaped by Contextual Factors

Pronunciation is not static; it is shaped by the contexts in which speakers learn and use language. The Midland American English speaker acquired English natively through immersion in home, community, and educational spaces, which facilitated natural rhythm, stress, and segmental patterns such as flapping and vowel reduction. In contrast, the Kenyan speaker learned English primarily in institutional and formal educational contexts, supplemented by media exposure, which resulted in careful articulation, preserved vowel clarity, and orthography-driven pronunciation. These contextual differences explain why the Midland speaker produces "better" with [ɾ] (like a soft "d"), while the Kenyan speaker retains [t] (a crisp "t" sound), or why "sure" is pronounced [ʃɝ], as one syllable by the American speaker, but [ʃu.wa] as two, by the Kenyan speaker.

Each speaker's phonology is shaped by the spaces they inhabit: the Midland speaker's informal domestic and community spaces encourage casual, compressed forms, while the Kenyan speaker's formal educational spaces encourage careful, expanded forms. These examples illustrate that pronunciation is a spatial practice, reflecting both the physical and social environments of the speaker[7] (De Certeau 91). Instrumental phonetic research supports this view, showing that apparent consonant deletion or reduction in casual American English speech often reflects gradient phonetic processes shaped by register and reduced articulatory effort rather than categorical phonological absence. (Zsiga 243 [2nd ed.]). Considering context and register is therefore essential for interpreting pronunciation differences, particularly in L2 speech.

For the Persian speaker, contextual factors also play a significant role in shaping pronunciation, though their effects interact closely with first-language phonological constraints. While her English consistently reflects Persian phonotactic patterns, the

7 See de Certeau

degree of articulatory care varies across contexts. In more formal or academic situations, her speech is characterized by slower tempo, clear syllable boundaries, and deliberate consonant articulation, reflecting heightened attention to intelligibility and professional expectations.
In less formal contexts, speech becomes more fluid, yet reductions common in American English—such as extensive vowel reduction or consonant deletion—remain limited due to L1 constraints. Acoustic analysis of words such as "through," "interview," "button," and "Skype" demonstrates that even when speaking casually, the Persian speaker maintains systematic pronunciation strategies shaped by Persian phonology. These patterns suggest that contextual space influences how speech is produced—tempo, clarity, and effort—while first-language phonology constrains which phonological alternations are available to the speaker. From a spatial perspective, this illustrates how speakers navigate social expectations of formality while carrying embodied linguistic habits across settings.

In the case of the Argentine speaker, his language learning background informed that all English classes were taught by EFL teachers from Argentina (British and American accents target preferences). Most of the EFL course books were designed by British Publishers like Oxford University Press and Cambridge University Press and integrated the four macro skills (reading, writing, listening, and speaking). His language instruction aimed to develop communicative competence as well as specific aspects of language as he advanced his language education. In the segmental and suprasegmental differences, he clearly shows features of Received Pronunciation British English resulting from (1) his learning experience in academic settings where English is taught as a Foreign Language, (2) his preference towards British accent target, and also (3) his exposure to real-world sounds, people and places during his time in the U.K. When reading the word "mountain," the spectrogram of this speaker showed that the Argentine participant pronounced the consonant sound [t], in [maʊntɪn] with no reduction intended (using a sharp "t" sound). On the contrary, the spectrogram of the American speaker showed that he used the glottal stop [ʔ], where the consonant

sound /t/ was reduced to the glottal stop (replacing "t" with a quick throat catch) before the nasal [n], a characteristic of many American English speakers (Zsiga 23 [2nd ed.]).

## Variation: Phonological Alternations and Register

Phonological alternations refer to systematic variations in the way sounds are pronounced when influenced by their surrounding linguistic environment, revealing the underlying rules that shape a language's sound system (Zsiga 244 [2nd ed.]). One type of alternation is lenition, where sounds become weaker or more open, and in American English, this usually happens with intervocalic /t/ or /d/ ("t" or "d" sound between two vowels), which becomes a tap [ɾ] (making words like "butter" and "ladder" sound the same). This process makes speech faster and more fluid without changing the meaning of the word. For example, when pronouncing the word "city" in casual speech the American speaker produced the voiced, alveolar, tap [ɾ], [sɪɾi] (sounds like "siddy"), whereas the Argentine speaker produced [t], following British English patterns of consonant sounds, [sɪti] (sounds like "sit-ee").

Another important type of alternation introduced by Zsiga is deletion, which occurs when a segment that is present in the environment of a word is not pronounced in actual speech. Deletion can also result from the lenition of vowels: in unstressed syllables they are not pronounced with full distinctive quality resulting in a reduction to schwa, and eventually in a deletion of the vowel sound. This alternation usually leads to a reduction in the number of syllables that a word contains. When participants pronounced the word "vegetable" in casual speech, the L2 English speaker from Argentina produced a four-syllable word, syllabified as [vɛ.ʤɪ.tə.bəl] while the L1 English speaker from the U.S. produced a three-syllable word by deleting the vowel sound after the consonant sound [ʤ], therefore syllabified as [vɛʤ.tə.bəl] (or "vej-tuh-buhl"). This suprasegmental difference portrayed that the Argentine speaker pronounced each segment following general patterns of the learned target language of Received Pronunciation British English, while the American speaker applied lenition of the vowel sound, which resulted in the deletion of a segment, following

general patterns of his regional accent.

Register—variation in speech based on social context—further demonstrates how spatial and relational factors shape pronunciation. During the recordings, the Midland American English speaker, recorded in a familiar domestic space and speaking with someone she knew intimately, produced speech that was relaxed, fluid, and informal. In contrast, the Kenyan English speaker, recorded in a more formal academic setting with a socially distant researcher, spoke carefully and deliberately. These differences in register led to observable phonological variation: the American speaker reduced vowels and produced flaps and compressed syllables, while the Kenyan speaker maintained full consonants, inserted glides, and emphasized syllables. This pattern highlights how emotional and social distance between speakers creates distinct linguistic spaces, influencing how identity is expressed through sound. As Bhabha and De Certeau argue, identity is performed in "in-between" spaces and emerges through practice; here, pronunciation reflects not only geography and background but also the social and interpersonal environment of the interaction. By shaping speech based on closeness or formality, speakers adapt to the immediate spatial context, showing that language is an embodied reflection of both physical and social space. Findings from sociophonetic research similarly show that speakers adjust pronunciation to manage formality, align with interlocutors, and project competence or authority (Zsiga). In L2 speech, these social adjustments interact with L1 transfer producing systematic and predictable variation rather than random deviation.

For the Persian speaker, words such as "through," "interview," "button," and "Skype" revealed systematic L1-influenced phonological patterns across formal and casual speech. Intervocalic /t/ often appeared as a flap ([bʌɾɛn] for button),[8] while vowel insertion in clusters, as in [ɛskaɪp] for "Skype," reflected Persian phonotactic constraints. In more formal contexts, articulation was deliberate, whereas casual speech allowed slightly faster, more fluid production, though still constrained by L1 habits. These patterns show that the Persian speaker negotiates identity and intelligibility

8 Like saying "budden" instead of a sharp "t."

simultaneously, balancing Persian phonological influence with English norms. Compared with the American speaker, whose reductions were highly context-dependent, the Persian speaker's pronunciation illustrates how L1 background and spatial experience shape consistent yet adaptable speech patterns.

## Conclusion

This study demonstrates that English pronunciation is far more than a technical skill; it is a spatially and socially situated practice shaped by geography, mobility, and lived experience. Comparing the Kenyan, Persian, Argentine, and American speakers shows that phonological variation arises from the interaction of first-language systems, second-language acquisition, learning environments, and social contexts, rather than from individual deficits. The Persian speaker's pronunciation illustrates how L1 phonology shapes English production, with systematic adaptations such as vowel insertion and flapping reflecting both intelligibility and cultural identity across formal and informal settings. The Kenyan speaker's dialectal variation reflects phonological and spatial awareness, suggesting that linguistic identity is shaped through contextual demands and individual agency. The Argentine speaker's language identity is shaped by the target language preference, and both imagined and lived places of the target culture. We also gained insight on how American speakers showed distinctive characteristics of the American English accent depending upon how they acquired and learned their first language, and the region they were born and lived in the U.S.

By analyzing pronunciation as a form of linguistic space-making, this research reveals that English is not a single, uniform system but a constellation of practices influenced by where, how, and with whom it is spoken. Spoken language encodes speakers' histories, identities, and experiences, highlighting the meaningful role of sound in social and cultural negotiation. Recognizing these spatial and social dimensions provides deeper insight into language learning, teaching, and intercultural communication, emphasizing that linguistic identity is continuously constructed and expressed through the sounds we produce.

# Works Cited


Bhaba, Homi K. *The Location of Culture*, Routledge, 1994.

Boersma, Paul, and David Weenink. *Praat: Doing Phonetics by Computer*. Version 6.4.52, 2025, https://www.fon.hum.uva.nl/praat/.

British Council. "Pronunciation Features of Kenyan English." *British Council*, 2022, www.britishcouncil.org. Accessed 01 December, 2025.

Crystal, David. *Language Death*. Cambridge Up, 2000.

De Certeau, Michel. *The Practice of Everyday Life*. Trans. Steven Rendall. Berkeley: University of California Press. 1984

"International Phonetic Alphabet." Encyclopaedia Britannica, *Encyclopaedia Britannica,* www.britannica.com/topic/International-Phonetic-Alphabet. Accessed 13 February, 2026.

Kachru, Braj B. *The Alchemy of English: The Spread, Functions, and Models of Non-Native Englishes*. University of Illinois Press, 1992.

Paltridge, Brian, and Aek Phakiti. *Research Methods in Applied Linguistics: A Practical Resource.* 2nd ed., Bloomsbury Academic, 2015.

Seidlhofer, Barbara. *Understanding English as a Lingua Franca*. Oxford UP, 2011.

Zsiga, Elizabeth. *The Sounds of Language: An Introduction to Phonetics and Phonology*. Wiley-Blackwell, 2013.

--. *The Sounds of Language: An Introduction to Phonetics and Phonology*. 2nd ed., John Wiley & Sons, 2024.