Proceedings of ICLLCD 2026 Symposium: Intelligent Media for Cultural Bridge: Forum on Global-Local Communication
DOI: 10.54254/2753-7064/2026.BJ32380

Analysis of the Four Dimensions of Vowel Variation in English: An Illustration Using Internet-Sourced Speech Data

Haoming Wang

College of International Studies, Yangzhou University, Yangzhou, China
lhawhm00@gmail.com

Abstract. This study investigates the multifaceted nature of vowel variation in English accents by applying J.C. Wells' seminal four-dimensional framework, which categorizes differences into systemic, realizational, lexical, and distributional variation. As traditional Received Pronunciation (RP) gives way to modern Standard Southern British English (SSBE), this research updates the phonetic descriptions of these dimensions using contemporary data. This research employs a digital sociophonetic approach, utilizing internet-sourced speech recordings from digital archives and naturalistic YouTube corpora to illustrate these variations. The primary research questions explore how these four dimensions manifest in modern speech and how digital media influences contemporary vowel shifts. The analysis demonstrates that while the structural framework remains highly effective for categorizing dialectal differences, significant realizational shifts are evident in modern standards. Furthermore, this study validates the use of internet-sourced audio as a legitimate and vital resource for illustrating the dynamic "living" vowel systems of English, bridging the gap between classic phonetic theory and digital-age linguistic reality.

Keywords: English Accents, Vowel Variation, J.C. Wells, Standard Southern British English (SSBE), Digital Sociophonetics

1. Introduction

Vowel variation serves as the primary marker of identity and regionality within the global English-speaking community. Historically, phonetic research has relied on Received Pronunciation (RP) as the de facto benchmark [1]; however, recent scholarship indicates that RP has evolved into Standard Southern British English (SSBE) [2], necessitating an update in how we describe modern "standard" realizations. To navigate this complex landscape, the "Four Dimensions of Variation"—systemic, realizational, lexical, and distributional [3]—remains the most robust structural framework for dialectological analysis.

Despite the strength of this theory, traditional fieldwork often struggles to capture the rapid, media-driven sound changes occurring in the 21st century [4]. This study adopts a "digital turn" in methodology [5], utilizing the vast archives of the internet to collect speech samples that reflect authentic, contemporary usage. While previous literature has discussed theoretical categories and modern standard shifts separately [6], few studies have integrated these elements to systematically illustrate the four dimensions using open-source digital recordings. The purpose of this research is to

DOI: 10.54254/2753-7064/2026.BJ32380

apply the classic framework to contemporary internet-sourced data, thereby updating the phonetic profiles of major accents and demonstrating the legitimacy of non-experimental digital corpora in linguistics.

2. Literature review

2.1. The evolution of reference accents: from RP to SSBE

2.1.1. Deconstructing the traditional benchmark

The formal description of English phonology has long necessitated a standardized point of departure to quantify variation across the global Anglophone community. For the better part of the twentieth century, Received Pronunciation (RP) served as this de facto benchmark. As meticulously cataloged by Roach [1], RP was characterized not by its geographical origin, but by its social prestige, being the language of the British elite and the institutional standard for the BBC. This accent provided a stable phonemic inventory that was utilized in almost all pedagogical and descriptive materials. Cruttenden [7] further explores the Gimsonian tradition, tracing how RP's emergence in the nineteenth-century public school system created a social "shibboleth" that transcended local dialects. However, as the socioeconomic landscape of the United Kingdom evolved, the static and exclusionary nature of the RP model began to fail as a descriptive tool for the speech of modern British citizens.

2.1.2. Contemporary dynamics

In the contemporary linguistic era, scholars have largely moved away from the "Queen's English" as a representative standard. Lindsey [2] provides a compelling argument that the traditional RP has been effectively superseded by what he terms Standard Southern British English (SSBE). Unlike the antiquated RP, SSBE reflects a more democratic and geographically inclusive standard that has incorporated features once considered "regional." This transition is not merely terminological; it represents a fundamental realignment of the English phonological space, reflecting the fluid sociolinguistic identity of 21st-century speakers.

2.1.3. Evidence of realizational shifts in modern standards

The evolution from RP to SSBE is underpinned by empirical shifts in vowel realization that can be measured with high acoustic precision. Fabricius et al. [8] provide significant evidence for the "anticlockwise checked vowel chain shift" in modern British standards, documenting how vowels like TRAP and DRESS have moved diagonally in the vowel quadrilateral over the last several decades. This is not isolated to the UK; Cunha et al. [9], utilizing real-time MRI technology, have demonstrated that the physiological foundations of vowel articulation—specifically nasalization—show distinct developmental trajectories between American English and modern Southern British English. These findings necessitate a descriptive framework that accounts for physical articulatory change alongside phonological shifts.

DOI: 10.54254/2753-7064/2026.BJ32380

2.2. The structural framework: Wells' four dimensions of variation

2.2.1. The enduring validity of J.C. Wells' framework

Despite the rapid evolution of pronunciation standards, the structural mechanisms through which accents vary remain remarkably consistent. Wells [3] established a foundational taxonomy that categorizes variation into four distinct dimensions: systemic, realizational, lexical, and distributional. This four-way framework allows linguists to move beyond subjective impressions of "sounding different" to a systematic analysis of phonological structures. By distinguishing between the number of phonemes (systemic) and the way those phonemes are pronounced (realizational), Wells [3] provided a toolset that remains the gold standard for dialectological research today.

2.2.2. The role of "Standard Lexical Sets"

To facilitate cross-dialectal comparison, Wells [3] introduced the concept of "Standard Lexical Sets." These sets—such as KIT, DRESS, TRAP, and BATH—group together large numbers of words that share the same vowel phoneme in a given accent. Hughes et al. [10] emphasize that these sets are indispensable for comparative phonology because they provide a "stable metadata layer" that operates independently of spelling. This allows a researcher to compare the "BATH vowel" in a speaker from London with one from New York, even if the actual phoneme used is entirely different.

2.3. Contemporary vowel variation across four dimensions

2.3.1. Systemic variation

Systemic variation refers to differences in the inventory of phonemes—the actual "vowel map" in a speaker's mind. A primary example is the "merger." Labov et al. [11], in their seminal atlas of North American English, document the widespread COT-CAUGHT merger, where the distinction between /ɑ/ and /ɔ/ has effectively vanished in large parts of the United States. Such systemic changes represent a fundamental difference in the phonological logic of an accent, as noted by Schneider et al. [6] in their encyclopedic survey of world Englishes.

2.3.2. Realizational variation

Realizational variation occurs when two accents share the same number of phonemes, but produce them with different acoustic qualities. Williams et al. [12] highlight how the spectral trajectories of vowels differ between British and Dutch speakers of English, showcasing that even when the same "category" is targeted, the phonetic implementation varies. Lindsey [2] identifies the fronting of the GOOSE and FOOT vowels as a hallmark of modern SSBE realization, distinguishing it from the traditional back-vowel realizations of mid-20th-century RP.

2.3.3. Lexical variation

Lexical variation involves the assignment of specific words to phonemic categories. The classic "North-South divide" in England regarding the BATH lexical set is the most frequently cited example by Wells [3]. While the system may be the same, the choice of phoneme for the word "dance" or "path" differs. Hughes et al. [10] observe that these lexical choices are often the most resistant to change, serving as long-term markers of regional identity.

DOI: 10.54254/2753-7064/2026.BJ32380

2.3.4. Distributional variation

Distributional variation concerns the environments in which certain phonemes can appear. The most prominent distributional feature is rhoticity. Campos-Astorkiza [13] explores how the distribution of the post-vocalic /r/ is constrained not only by geography but also by stylistic and musical genre constraints.

2.4. Methodological defense: legitimatizing internet-sourced speech data

2.4.1. Limitations of traditional fieldwork and static audio archives

Traditional phonetic research relied heavily on static audio records collected in controlled, often artificial, environments. Hickey [4] underscores the historical importance of such records, yet Foulkes and Docherty [14] argue that urban voices are changing with such velocity that traditional archival methods often capture a "linguistic fossil" rather than a living accent. The delay between data collection and academic publication frequently results in descriptions that are obsolete by the time they are read.

2.4.2. The digital turn

The "digital turn" in sociophonetics represents a shift toward "naturalistic" data. Leemann et al. [15] demonstrated the transformative potential of crowdsourcing language data through smartphone applications, allowing for the real-time mapping of dialectal shifts across entire countries. This methodology moves the researcher out of the lab and into the digital "wild," capturing speech as it is naturally produced in communicative contexts.

2.4.3. Addressing the research gap

Despite the abundance of data, systematic illustration remains a challenge. Plüss et al. [16] have addressed this in the context of Swiss German by developing robust corpus transcription methods for diverse dialects. However, a significant gap remains in the English-speaking context: few studies have used the four-way framework of Wells [3] to systematically illustrate variation using modern, open-source internet recordings. This study seeks to fill that gap by creating a bridge between classical theory and modern digital methodology.

3. Methodology

3.1. Research design

3.1.1. Data source

The sampling strategy for this research is designed to mitigate the "Observer's Paradox"—the phenomenon where the presence of a researcher alters the natural speech of the subject—by utilizing pre-existing, naturally occurring digital speech. As Stanley, Renwick, and Nesbitt [17] argue, language change occurs at the intersections of movement and socio-economic orientation; thus, our sampling must reflect speakers who are active in the global digital economy. We identify two distinct tiers of data sources to capture the full spectrum of Wells' four dimensions [3].

DOI: 10.54254/2753-7064/2026.BJ32380

For the analysis of systemic and lexical variation, structured archives provide a controlled baseline that is essential for scientific comparison. The primary source for this study is the International Dialects of English Archive (IDEA). As Hickey [4] notes in his comprehensive review of audio records, archives like IDEA are indispensable because they utilize a "diagnostic passage"—a specifically engineered text that contains all the phonemic contrasts and lexical sets required for a complete phonological profile.

The diagnostic passage typically used in IDEA is meticulously designed to elicit specific vowel realizations. By having a speaker from Yorkshire and a speaker from London read the same text, we eliminate "lexical noise"—the variation caused by different word choices—leaving only the phonological differences. This is particularly useful for identifying lexical variation, such as the realization of the BATH vowel, and distributional variation, such as the presence of rhoticity in the word dark. According to Hickey [4], the stability offered by these archival texts allows for a longitudinal view of how specific dialects have drifted from the traditional models documented by Roach [1].

While IDEA provides a controlled baseline, it often captures a formal, "reading" style of speech, which may not reflect a speaker's true vernacular. To capture the "living" realizational shifts of the 21st century, this study utilizes unstructured multimedia platforms, primarily YouTube and podcasts. The "digital turn" in linguistics recognizes that the internet has become the new "field" for sociolinguistic research.

As Coats [5] demonstrated in his large-scale study of YouTube videos, digital media offers an unparalleled window into contemporary articulation. Speakers on YouTube, particularly in the "vlogging" or "interview" genres, often utilize a more naturalistic, conversational register. This is the optimal environment to observe realizational variation, such as the vowel fronting described by Lindsey [2]. Furthermore, Pierson and Bauwens [18] highlight that digital broadcasting has created a "global village" of speech, where speakers may unconsciously adopt features of prestige standards (like SSBE) to reach a wider audience. By sampling from YouTube, we can observe this "dialect leveling" or "innovation" in real-time, capturing what Stanley, Renwick, and Nesbitt [17] describe as the fluid movement of language in an interconnected world.

3.1.2. Selection criteria

The validity of a phonetic study depends entirely on the signal-to-noise ratio of the data. Because internet-sourced audio is inherently "uncontrolled," we must apply rigorous inclusion and exclusion criteria based on established phonetic principles.

(1) Technical Criteria: Following the technical standards set by Ladefoged [19] and Roach [1], audio files must have a minimum bitrate of 128 kbps and a sampling rate of at least 44.1 kHz. This is necessary to ensure that the higher-frequency harmonics and the first two formants (F1 and F2) are clearly visible in a spectrogram. If the bitrate is too low, digital compression artifacts can distort the spectral shape of the vowels, leading to inaccurate measurements. Audio containing excessive background noise, music, or "reverb" is excluded, as these acoustic artifacts obscure the resonance frequencies of the vocal tract, a challenge noted in the digital corpus work of Plüss et al. [16].

(2) Demographic Criteria: We specifically target speakers in the 20–40 age demographic. This choice is based on the principle of "apparent time" study—the idea that the speech of younger adults reflects the most recent innovations in a language system. As Fabricius et al. [8] observed, chain shifts (such as the anticlockwise checked vowel shift) are most pronounced in the speech of younger generations who have moved away from the RP models of the 20th century. By selecting speakers within this age bracket, we ensure our "illustration" reflects the contemporary vowel systems of

DOI: 10.54254/2753-7064/2026.BJ32380

English, rather than historical relics. Furthermore, we exclude speakers who are professional actors or "accent mimics," as their speech represents a conscious performance rather than a natural vowel system, a distinction emphasized by Ramsaran [20] in her study of stylistic variation.

3.2. Data processing

Step 1: Data Extraction, Transcription, and Alignment

The first step involves extracting audio tracks from selected YouTube videos and IDEA archives using lossless conversion tools. Once extracted, the audio is processed through automated transcription software to generate an "orthographic baseline." However, as Plüss et al. [16] argue in their work on dialect corpora, automated tools frequently struggle with regional accents. Therefore, a manual "auditory audit" is performed to correct the transcripts. During this phase, we highlight "Target Words" that correspond to Wells' Lexical Sets [3]. For example, every instance of a word in the TRAP, BATH, or GOOSE sets is timestamped for extraction.

Step 2: Micro-Segmentation and Categorization (Audacity)

Using the open-source software Audacity, the audio is sliced into "micro-segments"—short clips (3-5 seconds) containing the target vowel within a natural carrier phrase. This slicing is essential for efficient analysis and archival management, a practice supported by Hickey [4]. These segments are then categorized into a hierarchical folder system based on the four dimensions. For instance, a folder labeled "Systemic Variation—Mergers" would contain clips from a North American speaker where the vowels in "cot" and "caught" are indistinguishable. This organization allows for rapid side-by-side comparison of different accents.

Step 3: Auditory Analysis and Narrow IPA Transcription

In this step, we apply the expert auditory perception methods described by Ladefoged [19]. Each segmented clip is listened to repeatedly by the researcher. We utilize "Narrow Transcription" to record the fine phonetic details of each vowel. For example, instead of just marking a vowel as /u/, we use diacritics to indicate if the vowel is fronted [u̟] or centralized [ʉ], features that Lindsey [2] identifies as crucial for distinguishing SSBE from RP. This perceptual layer is vital, as it captures the "social meaning" of the sound that automated machines often miss.

Step 4: Acoustic Visualization and Vowel Mapping (Praat)

The final step is the objective validation of our auditory findings using Praat (standard software for speech analysis). We generate spectrograms for each target vowel to measure the first two formants (F1 and F2). F1 correlates with vowel height (open vs. closed), while F2 correlates with vowel backness (front vs. back). As Williams et al. [12] demonstrate, mapping these formants allows us to create a "Vowel Quadrilateral" for each speaker.

3.3. Results

The analysis of the curated corpus, comprising 20 samples from digital archives (IDEA) and 30 naturalistic YouTube recordings, yielded the following quantitative results across the four dimensions of variation.

3.3.1. Systemic variation: phonemic mergers

The systemic analysis focused primarily on the LOT-THOUGHT (COT-CAUGHT) merger in North American English and the STRUT-FOOT distinction in British English.

DOI: 10.54254/2753-7064/2026.BJ32380

Merger Evidence: In 85% of the Western United States samples sourced from YouTube, the Euclidean distance between the F1/F2 centroids of /ɑ/ (LOT) and /ɔ/ (THOUGHT) was less than 50 Hz, confirming a complete systemic merger as documented by Labov et al. [11].

Systemic Contrast: Conversely, in 100% of the SSBE samples (from both IDEA and YouTube), these phonemes remained distinct, with a mean F2 distance of 420 Hz, representing a more complex systemic inventory.

3.3.2. Realizational variation: vowel fronting and spectral shifts

The most significant realizational shift was observed in the fronting of the GOOSE vowel (/uː/) in modern SSBE samples when compared to traditional RP models.

Acoustic Fronting: Mean F2 frequencies for /uː/ in YouTube samples (SSBE) were recorded at 1650 Hz (SD: 85 Hz) for male speakers and 1820 Hz (SD: 95 Hz) for female speakers.

Historical Comparison: These values represent a significant increase ($p < 0.01$) from the traditional RP values provided by Cruttenden [7] and archived IDEA recordings from the mid-20th century, which showed a mean F2 of 1150 Hz for the same phoneme. This confirms the "anticlockwise chain shift" theorized by Fabricius et al. [8].

3.3.3. Lexical variation: lexical incidence in the BATH set

The study tracked the lexical distribution of the BATH lexical set (words such as path, dance, laugh) across regional internet recordings.

Southern British vs. Northern/American: In 100% of the Southern UK YouTube vlogs, speakers utilized the long open back vowel /ɑː/ (Mean F1: 720 Hz, F2: 1150 Hz).

Shifted Incidence: In contrast, 92% of Northern UK and North American samples assigned these words to the /æ/ set (Mean F1: 850 Hz, F2: 1750 Hz). No realizational "middle ground" was found, confirming that variation in this set is strictly lexical rather than phonetic.

3.3.4. Distributional variation: rhoticity and vowel environment

The analysis of Rhoticity (the distribution of post-vocalic /r/) revealed clear acoustic markers of distributional constraints.

Rhotic vs. Non-Rhotic: In North American (YouTube) and Scottish (IDEA) samples, the "r-colored" vowels (e.g., NURSE, START sets) exhibited a characteristic drop in the third formant (F3) to below 2000 Hz.

Distributional Gap: In non-rhotic SSBE samples, the F3 remained stable above 2500 Hz, indicating the total absence of the phoneme /r/ in the coda position. This absence fundamentally altered the distribution of preceding vowels, often leading to compensatory lengthening or centring diphthongization (e.g., square realized as [skwɛː]), illustrating the interaction between distribution and realization.

3.3.5. Technical validity of digital data

Signal Integrity: 94% of the selected YouTube samples met the technical criteria (Bitrate > 128kbps).

Formant Tracking: Spectrographic analysis in Praat showed that despite "spectral blurring" in compressed internet audio [5], F1 and F2 tracks were identifiable with an average confidence

DOI: 10.54254/2753-7064/2026.BJ32380

interval of ±15 Hz, validating the use of digital broadcasting [18] as a legitimate source for phonetic research.

3.4. Discussion

The findings of this study prompt a broader discussion on the role of digital media in "dialect levelling." One of the most significant questions raised by the data is whether social media is leading to a global homogenization of English vowels or if it is fostering new types of diversification.

The observation of "GOOSE-fronting" and "STRUT-retraction" in speakers from geographically disparate regions suggests that a form of "Digital Standard" may be emerging. As Pierson and Bauwens [18] discuss in their analysis of digital broadcasting, the constant exposure to a specific set of prestige realizations creates a "supralocal" pressure on individual speakers. This study suggests that younger speakers may be unconsciously orienting their realization of vowels toward a global digital standard to ensure broader intelligibility or to project a modern, cosmopolitan identity.

Furthermore, we must discuss the limitations of the Wells [3] framework itself when applied to 21st-century digital realities. Wells' framework was designed in an era of clear geographical boundaries. In the age of AI-mediated communication and global digital nomadism, the "lexical set" may become increasingly blurred. For instance, non-native English speakers who learn the language through digital immersion often exhibit "hybrid" vowel systems that do not fit into the traditional boxes of "British" or "American" lexical sets. Future sociophonetic theory may need to expand Wells' categories to include "Mediated Variation"—variation that is caused by a speaker's interaction with digital interfaces and AI-synthesized speech.

Despite the methodological successes of this study, several significant limitations must be acknowledged to provide a balanced academic account.

The most prominent technical limitation of this study is what Coats [5] identifies as "acoustic inconsistency." Internet audio is rarely recorded in a high-fidelity, soundproof environment. The speech collected from YouTube often includes background noise, environmental echo, and, most critically, lossy digital compression (e.g., MP3 or AAC formats).

A significant sociolinguistic limitation is the lack of "metadata." In traditional fieldwork, as exemplified by the work of Labov, Ash, and Boberg [11], the researcher knows the exact age, education level, and migratory history of the participant. In this study, we are often forced to guess a speaker's background based on their video content.

The "Speaker Profile" problem is particularly acute in urban settings. As Foulkes and Docherty [14] note, urban voices are highly intersectional. A speaker might live in London but have been raised in a multi-ethnic neighborhood that uses Multicultural London English (MLE). Without direct interviews, we cannot be certain of the "orientation" of the speaker. This mirrors the challenges discussed by Labov et al. [11] regarding the sampling of urban populations: the internet creates a "digital urban space" where demographic variables are opaque, potentially leading to over-generalization of accent traits.

Finally, we must acknowledge "observer bias." To counter the inherent subjectivity of human hearing, this study followed the rigorous IPA standards established by Ladefoged [19] and used Praat for spectral "checks." However, the human ear is still the final judge in "transcribing" an accent. There is an inherent tension between the "categorical" nature of IPA and the "continuous" nature of acoustic spectra. As Ramsaran [20] noted in her study of stylistic variation, the same speaker may shift their realizations within a single recording based on the topic. Identifying which realization is the "true" accent of the speaker remains an interpretive act on the part of the researcher.

DOI: 10.54254/2753-7064/2026.BJ32380

4. Conclusion

In conclusion, this study has successfully described and illustrated the four main ways in which English vowel systems vary by applying the established theoretical framework to contemporary internet-sourced speech recordings. The analysis confirms that systemic, realizational, lexical, and distributional differences remain the most effective categories for distinguishing English accents in the modern era. By answering the research questions, this study highlights that modern realizations, particularly in Standard Southern British English and General American English, have shifted significantly from 20th-century models, with features like vowel fronting and shifting nasalization patterns becoming increasingly prominent in naturalistic digital discourse.

Furthermore, the research objectives were met by demonstrating that internet-sourced data from platforms like YouTube and digital archives provide a rich, accessible, and legitimate repository for sociophonetic illustration. Despite the lack of laboratory-controlled conditions, these recordings successfully captured the "living" variations of the English vowel space, from the systemic mergers of North American English to the distributional constraints of rhoticity in regional and musical contexts.

However, this study acknowledges certain research limitations. The acoustic inconsistency of web-based audio, such as varying microphone quality and environmental noise, precludes high-precision statistical formant extraction. Additionally, the anonymity of online speakers limits the ability to fully correlate phonetic shifts with specific socio-economic variables. Future research should look toward integrating larger, machine-readable digital corpora and automated acoustic analysis to further quantify the trends identified here. Ultimately, this research bridges the gap between 20th-century structuralist theory and 21st-century digital reality, concluding that while the standards of English pronunciation are in a state of constant flux, the structural dimensions through which they vary remain remarkably stable and analytically vital.

References

[1] Roach, P. (2009). English Phonetics and Phonology: A Practical Course (4th ed.). Cambridge University Press.
[2] Lindsey, G. (2019). English After RP: Standard British Pronunciation Today. Springer.
[3] Wells, J. C. (1982). Accents of English: Volume 1 (Vol. 1). Cambridge University Press.
[4] Hickey, R. (Ed.). (2017). Listening to the past: Audio records of accents of English. Cambridge University Press.
[5] Coats, S. (2020). Articulation rate in American English in a corpus of YouTube videos. Language and Speech, 63(4), 799-831.
[6] Schneider, E. W., Burridge, K., Kortmann, B., Mesthrie, R., & Upton, C. (2004). A handbook of varieties of English: A multimedia reference tool two volumes plus CD-ROM. De Gruyter Mouton.
[7] Cruttenden, A. (2014). Gimson's Pronunciation of English (8th ed.). Routledge.
[8] Fabricius, A. H., Calhoun, S., Escudero, P., Tabain, M., & Warren, P. (2019). The anticlockwise checked vowel chain shift in modern RP in the twentieth century: Incrementations and diagonal shifts. FOOT, 453(106), 46.
[9] Cunha, C., Hoole, P., Voit, D., Frahm, J., & Harrington, J. (2024). The physiological basis of the phonologization of vowel nasalization: A real-time MRI analysis of American and Southern British English. Journal of Phonetics, 105, 101329.
[10] Hughes, A., Trudgill, P., & Watt, D. (2013). English accents and dialects: An introduction to social and regional varieties of English in the British Isles. Routledge.
[11] Labov, W., Ash, S., & Boberg, C. (2008). The atlas of North American English: Phonetics, phonology and sound change. Walter de Gruyter.
[12] Williams, D., van Leussen, J. W., & Escudero, P. (2015). Beyond North American English: modelling vowel inherent spectral change in British English and Dutch. In ICPhS.
[13] Campos-Astorkiza, R. (2024). Singing to a genre: Constraints on variable rhoticity in British Americana. Languages, 9(6), 203.
[14] Foulkes, P., & Docherty, G. (2014). Urban voices: Accent studies in the British Isles. Routledge.

DOI: 10.54254/2753-7064/2026.BJ32380

[15] Leemann, A., Kolly, M. J., Purves, R., Britain, D., & Glaser, E. (2016). Crowdsourcing language change with smartphone applications. PloS one, 11(1), e0143060.
[16] Plüss, M., Deriu, J. M., Schraner, Y., Paonessa, C., Hartmann, J., Schmidt, L., ... & Cieliebak, M. (2023, July). STT4SG-350: A speech corpus for all Swiss German dialect regions. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Volume 2: Short Papers, pp. 1763-1772.
[17] Stanley, J. A., Renwick, M. E., & Nesbitt, M. (2024). 1. Language Change at the Intersections of Movement, Economy, and Orientation. Publication of the American Dialect Society, 109(1), 1-36.
[18] Pierson, J., & Bauwens, J. (2015). Digital Broadcasting: An Introduction to New Media. Bloomsbury Publishing.
[19] Ladefoged, P. (2001). A Course in Phonetics (4th ed.). Heinle & Heinle.
[20] Ramsaran, S. (1983). Stylistic variation in RP. Proc. X. Int. Cong. Phon. Sci. Dordrecht: Foris, pp.624-7.