Phonological Changes in English Pronunciation among the Digital Generation: A Corpus-Based Study

Manal Hashim Raheem

Abstract: The present study investigates phonological changes in English pronunciation as observed among the digital generation—broadly defined as individuals aged 15 to 30 who have grown up in digitally mediated communicative environments. Using a corpus-based approach, this research draws on data sourced from YouTube content, podcasts, and open-access speech datasets to identify, quantify, and interpret dominant phonological processes—including elision, assimilation, vowel reduction, and connected speech phenomena—in spontaneous digital speech. A total of 120 speakers contributed approximately 95 hours of transcribed audio, yielding a substantial body of phonological evidence. Results reveal that phonological simplification processes are highly frequent, systematic, and cross-contextual in digital speech, with elision and vowel reduction occurring at significantly elevated rates compared to benchmarks drawn from formal spoken corpora. The study argues that digital communication environments accelerate phonological reduction by reinforcing informal speech norms, providing high-frequency exposure to reduced forms, and normalizing connected speech phenomena across demographically diverse groups of speakers. Findings carry meaningful implications for corpus phonology, sociophonetics, language pedagogy, and the development of speech recognition systems.

INTRODUCTION

Language is not a static system. It evolves continuously in response to the social, technological, and communicative conditions in which it is embedded. Few forces have reshaped those conditions as dramatically or as rapidly as the rise of digital communication. In the twenty-first century, spoken language increasingly inhabits digital environments—podcasts, YouTube vlogs, TikTok commentary, live-streamed gaming, and informal video conversations—where informality, speed, and multimodal engagement redefine what it means to speak naturally. These environments are not merely contexts for the transmission of language; they are powerful agents in the formation and diffusion of new

linguistic norms. Among the most consequential and yet least systematically studied dimensions of this transformation is the phonological level: the sounds of speech, and the processes through which those sounds change.

Research on pronunciation change in digital environments has expanded rapidly over the past decade, yet the field remains conceptually and methodologically fragmented. A substantial portion of the literature is anchored in pedagogical concerns, examining how language learners can leverage digital platforms, speech analysis software, and corpus-based tools to enhance pronunciation accuracy and intelligibility (Ma, Mei, & Qian, 2024; Chen, Tian, & Chan, 2025). While these studies contribute valuable insights into second language acquisition and instructional design, they tend to frame digital technology as a neutral tool rather than as a sociophonetic force capable of reshaping speech patterns. Parallel to this, another body of research focuses on variation across regional and social dialects of English, often documenting phonological features tied to geographic or socio-economic identity (Shah & Zahra, 2025; Michael, 2025). However, such approaches rarely account for the pervasive influence of digital exposure as an independent variable, thereby limiting their explanatory scope (Ma et al., 2024; Shah & Zahra, 2025).

Despite these advances, relatively few studies have adopted a systematic corpus-based approach to investigate the phonological characteristics of naturally occurring speech among digitally immersed speakers. This represents a critical gap, as corpus linguistics offers robust tools for capturing authentic, large-scale speech data and identifying subtle phonetic and phonological patterns that may not be observable through experimental or pedagogical studies alone. Moreover, existing corpora are often not designed to capture the nuances of digitally mediated interaction, such as speech influenced by online gaming, social media discourse, or algorithmically curated audiovisual content. As a result, the phonological impact of sustained digital engagement remains underexplored, particularly in terms of how exposure to diverse accents, speech rhythms, and prosodic features may lead to emergent hybrid forms of pronunciation (Chen et al., 2025; Michael, 2025).

Addressing this gap is essential because the digital generation is not merely a passive consumer of online content but an active participant in a globalized communicative ecosystem. Continuous interaction with digital media may contribute to the diffusion, leveling, or even innovation of phonological features, creating patterns that transcend traditional geographic and social boundaries. From a theoretical perspective, this raises important questions about language change, identity construction, and the mechanisms of phonological adaptation in digitally saturated environments. Empirically, it necessitates the development of new corpus-based methodologies capable of capturing and analyzing speech as it naturally occurs within digital contexts. By reconceptualizing digital exposure as a formative factor in pronunciation change, future research can provide a more comprehensive account of contemporary phonological variation and its underlying drivers (Ma et al., 2024; Chen et al., 2025; Shah & Zahra, 2025).

The present study addresses this gap by constructing and analysing a purpose-built corpus of digital speech produced by speakers aged 15 to 30. The corpus combines data from YouTube videos, podcast recordings, and open-access spoken language datasets such as Mozilla Common Voice. It employs established phonological frameworks alongside quantitative and qualitative corpus analysis techniques to identify the predominant phonological processes at work in digital speech, examine their distribution across speakers and contexts, and interpret their significance in relation to broader theories of phonological change and variation.

The study is guided by two primary research questions. First, what phonological processes dominate the spontaneous speech of the digital generation in informal online environments? Second, are these phonological changes systematic across different speakers, digital platforms, and communicative contexts? In addressing these questions, the study aims to make a contribution both to the empirical documentation of contemporary phonological change and to the growing theoretical literature on the relationship between digital communication and language evolution.

LITERATURE REVIEW

2.1 Corpus-Based Phonology

The use of large spoken corpora to investigate phonological phenomena has a well-established and growing tradition in linguistics. Liberman (2019), in a foundational review, argued that corpus phonetics has transformed the capacity of phonologists to test theoretical claims against real-world spoken data, enabling discoveries about reduction phenomena, prosodic structure, and allophonic variation that would be inaccessible through laboratory phonetics alone. The shift from small elicited datasets to large naturalistic corpora has been particularly consequential for the study of connected speech, where phonological processes unfold in complex and probabilistic ways.

Vella and Grech (2022) demonstrate the analytical richness of corpus approaches by showing how phonetic and phonological variation—including variable deletion, lenition, and assimilation patterns—can be rigorously documented and quantified when corpora are large enough and adequately annotated. Similarly, Seifart (2021) highlights the methodological synergy between documentary linguistics and corpus phonetics, arguing that the combination of these two traditions enables advances in typological phonology that would be impossible using either approach alone. The value of combining interpretive depth with quantitative breadth is now widely recognised as a cornerstone principle of corpus-based phonological research.

Mair (2021), while focusing primarily on grammatical change, provides a methodological model directly applicable to phonological inquiry, arguing that corpus-driven analysis of ongoing linguistic change must attend carefully to

register variation, speaker demographics, and the differential rates of change across subsets of the corpus. This insight is particularly relevant to the present study, which must contend with the heterogeneous nature of digital speech data.

More recent corpus-based work has extended these principles to a variety of lesser-studied phonological domains. Fleischer, Cysouw, and Speyer (2018) used a historical corpus to trace variation and reduction in German schwa across time, demonstrating that even subtle phonological phenomena leave statistically recoverable traces in well-constructed corpora. Their findings underscore the viability of corpus approaches for tracking gradual phonological change—a principle that the present study applies to the temporally compressed but socially rich context of digital communication.

Boborahimov, Akhmedjanova, and colleagues (2025), working on a Middle Turkic language corpus, further illustrate how corpus construction principles and weighted averaging methodologies can be adapted to capture phonological patterning in data-sparse or historically unusual languages, suggesting that the technical challenges of corpus construction need not preclude rigorous phonological analysis.

2.2 Phonological Variation and Change

Phonological variation is among the most extensively studied topics in sociolinguistics and linguistic anthropology. The variationist tradition, associated most prominently with Labovian paradigms, has established that pronunciation variation is never random but structured by social, demographic, and contextual factors. Elision, assimilation, and vowel reduction are among the most frequently documented and theoretically significant phonological processes in spoken English.

Elision—the deletion of segments in connected speech—has been documented across a wide range of English varieties and speaking styles. The elision of word-final /t/ and /d/ in clusters is particularly prevalent in informal registers, as is the deletion of unstressed syllables in rapid speech (e.g., probably → [ˈprɒbli]). Assimilation, whereby one segment takes on the features of an adjacent segment, operates both within and across word boundaries (e.g., that person → [ˌðæp ˈpɜːsn]) and contributes substantially to the distinctive rhythm and connected quality of natural English speech. Vowel reduction, involving the centralisation and weakening of vowels in unstressed syllables—often realised as the schwa [ə]—is a pervasive feature of natural English pronunciation that interacts intimately with both elision and assimilation.

Belando (2024) examines the social and temporal dynamics of tapped /r/ in Received Pronunciation across the twentieth century, demonstrating through corpus evidence that even prestige phonological features undergo systematic change over time in response to social pressures. This kind of sociophonetic investigation, grounded in corpus data, provides a methodological and

conceptual precedent for the present study's examination of phonological change in digitally inflected speech.

Alfaifi, Qasem, and Bokhari (2024) investigate gemination in child Egyptian Arabic through corpus data, providing evidence that phonological processes operate systematically even in developing language systems, a finding that resonates with theoretical claims about the universality and predictability of phonological reduction across languages and developmental stages.

Proctor and Cormier (2023) analyse sociolinguistic variation in British Sign Language mouthings, illustrating that the theoretical frameworks developed for spoken phonological variation are applicable—with appropriate modification—to a variety of modalities and semiotic systems. Their work reinforces the principle that phonological change is fundamentally a sociolinguistic phenomenon, driven by speaker identity, community membership, and communicative context.

Afanasev and Lyashevskaya (2025) apply corpus-based language distance measurement to diatopic variation, demonstrating that phonological distance between varieties can be quantified systematically using distributional corpus evidence. Their approach informs the comparative methodology employed in the present study when assessing whether phonological patterns observed in digital speech diverge systematically from benchmarks established for more formal registers.

2.3 Digital Communication and Pronunciation

The relationship between digital communication and phonological change is a relatively recent but rapidly expanding area of inquiry. The emergence of podcasts, YouTube, social media audio, and video calling as primary communicative genres has created new registers of spoken language that blend features of informal conversation, broadcast speech, and written-language influence in ways that are linguistically distinctive and not yet fully documented.

Jawad, Chandio, and Tahir (2026) investigate phonological patterns in digital communication by examining sound-based spellings—orthographic representations of non-standard pronunciations (e.g., "gonna," "wanna," "lemme")—arguing that these spellings serve as proxies for the pronunciation norms of digitally native speakers. Their corpus-based analysis of online texts reveals that such spellings cluster around elision and assimilation contexts, suggesting that the pronunciation patterns they reflect are both widespread and salient among digital communicators.

Shah and Zahra (2025) document phonological variations in Pakistani English as used in digital contexts, finding that digital-age speakers exhibit novel pronunciation patterns that cannot be fully explained by reference to either the traditional norms of Pakistani English or the models of standard British or American English. Their sociolinguistic analysis identifies digital exposure—particularly through social media platforms—as a significant independent variable in pronunciation variation.

Sweet and colleagues (2025) examine Bangla-English code-mixing on social media platforms in Bangladesh, revealing that digital contexts facilitate not only lexical mixing but also phonological negotiation between languages, as speakers navigate the phonological inventories of two languages simultaneously. While their focus is code-mixing rather than monolingual phonological change, their findings underscore the role of digital environments as laboratories for phonological experimentation and norm diffusion.

Bao (2024), analysing Chinese social media through corpus methods, demonstrates that digital platforms create distinctive linguistic communities with their own norms and patterns, a finding that is broadly transferable to the English-language digital speech context. Muhammad and John (2025), working on morphological productivity in Nigerian English, similarly document how contact and digital exposure interact to accelerate language change at multiple linguistic levels, suggesting that phonological change is part of a broader pattern of digitally mediated linguistic evolution.

Garcia (2023), investigating anglicisms in Urdu through phono-semantic matching, documents processes of phonological adaptation that are partly driven by digital media consumption—as borrowed English words are adapted to Urdu phonology, the patterns of adaptation reflect the pronunciations encountered in digital contexts. This finding, while focused on a borrowing context, speaks to the broader thesis that digital exposure shapes phonological norms.

2.4 AI and Corpus Tools in Phonology

The analytical toolkit available for corpus-based phonological research has expanded dramatically over the past decade, with automated speech recognition, acoustic analysis software, and machine learning-based annotation systems transforming what is computationally feasible.

Chen, Zhou, and Tian (2025) evaluate the effectiveness of integrating corpus-based and AI approaches in the context of English speaking practice, finding that AI-assisted analysis enables the detection of subtle pronunciation patterns that would be difficult to identify through manual inspection alone. Their study provides methodological support for the use of AI tools in the present investigation.

Chen, Tian, and Chan (2025) extend this framework to pronunciation production and perception, demonstrating that corpus and AI-aided pedagogy produces measurable improvements in learners' phonological awareness, with clear implications for how AI-detected pronunciation patterns can inform both research and instruction.

Yang (2020) develops a CycleGAN-based approach to pronunciation feedback generation for Computer-Assisted Pronunciation Training, demonstrating that deep learning architectures can be trained to detect and classify pronunciation variation at a fine-grained level. This work anticipates many of the AI-assisted analytical techniques now being applied to corpus phonological research.

Zhao and colleagues (2025) construct a multimodal dialect corpus based on deep learning and digital twin technology, demonstrating how advanced computational methods can be brought to bear on the challenges of large-scale spoken corpus construction and analysis. Wang (2023), similarly, develops an automatic phoneme conversion algorithm using KNN classification, providing a computational foundation for the large-scale processing of phonological data.

Chohan and Yousaf (2025) conduct a meta-analysis of corpus-based approaches in phonology, identifying key trends in tool use and theoretical orientation across the field. They find that the integration of quantitative corpus analysis with AI-based detection is increasingly normative in cutting-edge phonological research, and that studies employing such integrated approaches tend to produce more robust and replicable findings. Chohan and Sawar (2025) extend this analysis specifically to AI-supported phonology teaching, cataloguing the tools and techniques that have proven most effective across a range of pedagogical and research contexts.

Kunath (2021) provides a systematic overview of tools and techniques for corpus-based analysis of non-native speech patterns, covering acoustic analysis platforms, transcription protocols, and annotation frameworks. His treatment of Praat as an acoustic analysis tool, alongside Python-based scripting environments for data processing and frequency analysis, closely mirrors the methodological approach adopted in the present study.

2.5 Empirical Studies on Digital Phonology and Pronunciation (2023–2026)

The period from 2023 to 2026 has seen a significant increase in empirical corpus-based work on pronunciation variation in digital and contemporary contexts. Al-Jarf (2026) conducts a systematic review of studies on pronunciation instruction and practice in second language contexts published between 2005 and 2025, noting a decisive shift in focus toward digital environments, online speech corpora, and computer-assisted analysis. She observes that naturalistic digital speech is increasingly recognized as a valuable data source for both research and pedagogy.

Topal (2023) proposes YouGlish, a web-sourced corpus built from YouTube content, as a resource for bolstering L2 pronunciation research and instruction, arguing that real-world digital speech data provides authentic phonological models that are more representative of contemporary norms than traditional audio resources. This perspective directly informs the corpus construction strategy adopted in the present study.

Ma, Mei, and Qian (2024) examine EFL students' pronunciation learning supported by corpus-based pedagogy, finding that exposure to naturally occurring corpus data accelerates the acquisition of connected speech features, including elision and assimilation. Their findings imply that the phonological patterns characteristic of corpus-attested digital speech are learnable and socially valued by contemporary speakers.

Michael (2025) investigates phonetic variation in English loanwords in South Korean media, using a corpus of broadcast and online content to document systematic patterns of phonological adaptation. His study illustrates the global reach of digitally transmitted English phonological norms and their influence on pronunciation standards in non-native contexts.

Xodabande, Asadi, and Valizadeh (2023) compare digital and paper-based methods for vocabulary instruction using corpus-based wordlists, a study that, while primarily pedagogical, provides evidence that digital delivery of corpus-based material is at least as effective as traditional methods—important context for any research programme that advocates the use of digital corpora in educational settings. Ormanova and Anafinova (2022), examining linguistic interference in Kazakh through corpus methods, contribute to the growing body of evidence that corpus approaches can illuminate subtle structural changes at the phonological interface, including patterns of interference and convergence that may parallel processes observed in English digital speech.

2.6 Research Gap

Despite the proliferation of corpus-based phonological research and the growing scholarly awareness of digital environments as linguistically distinctive, the field currently lacks studies that focus specifically and systematically on the phonological behaviour of naturally occurring digital speech produced by the digital generation as a defined group. Most existing studies either focus on pedagogy (what phonological models should learners be exposed to?), on regional or social variety (how do Pakistani or Nigerian English speakers pronounce English?), or on cross-linguistic comparison (how are English sounds adapted in other languages?). The intersection of corpus phonology, natural digital speech, and generational speaker identity remains underexplored.

Furthermore, existing meta-analyses (Chohan & Yousaf, 2025; Chohan & Sawar, 2025) confirm that while corpus-based phonological research has grown substantially, studies focusing on digital speech as a naturally occurring register—rather than as a pedagogical resource or as the medium through which learners interact—are rare. The integration of corpus evidence, phonological theory, and digital communication research that the present study undertakes thus addresses a genuine gap in the literature and offers findings with implications across multiple disciplinary domains.

METHODOLOGY

3.1 Research Design

The present study adopts a corpus-based, descriptive and analytical research design. The corpus-based approach was selected because it enables the systematic extraction of phonological patterns from large quantities of naturalistic spoken data, providing both the quantitative rigour and the

contextual validity that the research questions demand (Liberman, 2019; Vella & Grech, 2022). The design is descriptive insofar as it aims to document the phonological features of digital speech without experimental manipulation, and analytical insofar as it seeks to interpret those features within established phonological and sociophonetic frameworks.

The study was conducted in three phases: (1) corpus construction, involving the collection, selection, and preparation of speech data; (2) data processing, involving acoustic analysis, transcription, and annotation; and (3) data analysis, involving quantitative frequency analysis and qualitative phonological interpretation.

3.2 Corpus Construction

The Digital Generation English Speech Corpus (DGESC) was constructed specifically for this study. Data were collected from three primary sources: YouTube videos, podcast episodes, and the Mozilla Common Voice dataset (an open-access repository of spontaneous speech from volunteer speakers). Each source was selected to maximise the representativeness of digitally mediated informal speech while maintaining sufficient diversity in speaker background, topic, and communicative context.

- YouTube videos were selected from vlogs, commentary videos, and conversational content on the basis of the following criteria: (a) the speaker was identified as a native or near-native English speaker aged 15–30; (b) the content was spontaneous or semi-spontaneous rather than scripted; (c) audio quality was sufficient for acoustic analysis; and (d) the content was publicly accessible and free from copyright restrictions that would preclude research use.
- Podcast episodes were drawn from conversational and interview formats in which informal interaction between co-hosts or between a host and guest was prominent. Scripted monologue podcasts were excluded to preserve the spontaneous character of the data. Episodes were selected from platforms including Spotify and Apple Podcasts, focusing on youth-oriented content.
- Mozilla Common Voice recordings provided an additional layer of demographic diversity, including speakers from a range of English-speaking countries (United States, United Kingdom, Canada, Australia, and Ireland), ensuring that the corpus was not biased toward a single national variety.

Inclusion criteria specified that speakers must: (a) be aged between 15 and 30 at the time of recording; (b) be identified as native or near-native English speakers; and (c) have produced speech in an informal, conversational, or semi-spontaneous context. Speakers reading from a script or producing highly formal discourse were excluded.

3.3 Sample Size

The DGESC comprises contributions from 120 speakers (62 female, 56 male, 2 non-binary/unspecified), drawn from across the three data sources. The total duration of audio material in the corpus is approximately 95 hours of speech,

which, after segmentation and quality filtering, yielded approximately 74 hours of usable data for phonological analysis. On average, each speaker contributed approximately 37 minutes of analysable speech. The corpus contains approximately 680,000 words of running speech, based on orthographic transcription.

Table 1: Corpus Composition by Data Source

Data Source	No. of Speakers	Total Duration (hours)	Usable Duration (hours)	Word Count (approx.)
YouTube Videos	52	42.3	32.7	292,000
Podcast Episodes	38	31.5	24.8	221,000
Mozilla Common Voice	30	21.2	16.5	167,000
Total	120	95.0	74.0	680,000

Table 1 shows the distribution of speakers and speech data across the three source types. YouTube emerged as the largest single contributor to the corpus, reflecting both the volume of suitable content available on the platform and the centrality of YouTube to the digital communicative habits of the target age group. The Common Voice dataset provided valuable demographic balance and accent diversity.

Table 2: Speaker Demographics

Category	Subcategory	N	%
Gender	Female	62	52.7%
	Male	58	47.3%
Age Group	15–18	28	23.3%
	19–23	49	40.8%
	24–30	43	35.8%
National Variety	American English	45	37.5%
	British English	31	25.8%
	Canadian English	18	15.0%
	Australian English	16	13.3%
	Other/Mixed	10	8.3%

Table 2 provides the demographic profile of the speaker pool. The concentration of speakers in the 19–23 age bracket reflects the particularly high digital engagement and content production characteristic of late-adolescent and early-adult speakers. The distribution across national varieties ensures that findings are not attributable solely to a single dialect tradition, though American and British English together account for over 63% of the data.

3.4 Data Processing

Audio material was segmented into utterances of between 5 and 30 seconds using Audacity software. Segmentation followed natural speech boundaries (pauses, topic shifts, turn-taking) to preserve the prosodic integrity of the data.

Each segment was assigned a unique identifier encoding speaker number, data source, and segment sequence.

Transcription was carried out at two levels. First, all segments were orthographically transcribed by trained research assistants, with uncertain passages verified through a second listening. Second, a 25% stratified random sample of segments (representing all speakers, data sources, and age groups proportionally) was subjected to **broad phonetic transcription** using the International Phonetic Alphabet (IPA), enabling direct phonological analysis. This sample amounted to approximately 18.5 hours of speech and approximately 170,000 words.

Inter-annotator reliability for phonetic transcription was assessed using Cohen's kappa (κ). Two trained phoneticians independently transcribed 10% of the phonetically transcribed sample, yielding an inter-annotator agreement of κ = .81 for phoneme identification and κ = .77 for the classification of phonological processes—both values falling within the range conventionally accepted as indicating substantial to almost perfect agreement (Hernández, Molina, & Almela, 2023; Gu, Gao, Liu, & Wang, 2026).

3.5 Analytical Framework

The analytical framework is grounded in established phonological process theory (Kunath, 2021; Seifart, 2021), with a particular focus on four categories of phenomena:

1. **Elision**: The deletion or omission of one or more sounds (typically consonants or unstressed vowels) from the phonological surface form, as in the deletion of medial /t/ in exactly [ɪgˈzækli] or final cluster simplification in best friend [bes frend].

2. **Assimilation**: The modification of a phoneme to become more similar to an adjacent phoneme, including place assimilation (e.g., ten people → [ˈtem piːpəl]), voicing assimilation, and manner assimilation.

3. **Vowel Reduction**: The weakening, centralisation, or deletion of vowels in unstressed syllables, most characteristically through schwa substitution (e.g., today → [təˈdeɪ], can [kən] rather than [kæn]).

4. **Connected Speech Phenomena**: A broader category encompassing linking phenomena (r-linking, j-insertion), intrusive sounds, and rhythmically motivated reductions that operate across word boundaries.

This framework aligns with the processual approach to corpus phonology advocated by Liberman (2019) and Vella and Grech (2022), and with the connected speech focus evident in pedagogically oriented corpus research (Ma, Mei, & Qian, 2024; Al-Jarf, 2026).

3.6 Tools and Technologies

Acoustic analysis was performed using **Praat** (version 6.4), which enabled the measurement of formant values for vowel identification, waveform and spectrogram analysis for consonant classification, and duration measurements for the quantification of reduction phenomena. Praat scripts were written to automate the extraction of F1 and F2 formant values from vowels across the transcribed sample, facilitating vowel space analysis and the identification of vowel reduction patterns (Yang, 2020; Wang, 2023).

Python (version 3.11), together with the libraries NLTK, pandas, and matplotlib, was employed for orthographic corpus analysis, frequency counting, keyword-in-context (KWIC) extraction, and the visualisation of distributional patterns. The orthographically transcribed corpus was processed to extract all instances of lexemes known to be frequent sites of phonological reduction (e.g., going to, want to, have to, them, and, because), and these were then cross-referenced with their phonetic transcriptions in the stratified sample.

Additionally, AI-assisted phoneme detection was applied to a subset of the data using a pre-trained automatic speech recognition (ASR) model (Wav2Vec 2.0) to support the segmentation of phoneme sequences and the identification of reduction contexts, following methodological precedents established by Chen, Zhou, and Tian (2025) and Zhao and colleagues (2025). The ASR output was used as a supplementary rather than primary annotation source, with all AI-generated classifications subject to human verification.

3.7 Data Analysis

Analysis proceeded through two complementary strands:

- **Quantitative analysis** involved frequency counts of each of the four phonological process categories across the full transcribed sample, calculation of process rates (number of process instances per thousand running words), and examination of distributional patterns across speaker demographics (age, gender, national variety) and data source types (YouTube, podcast, Common Voice). Chi-square tests and effect size measures (Cramér's V) were used to assess the statistical significance of distributional differences.
- **Qualitative analysis** involved contextual examination of representative instances of each phonological process, including analysis of phonological environment (position in word and phrase, stress pattern, preceding and following sounds), communicative context (topic, register, interaction type), and speaker-specific patterns. The goal was to determine whether observed patterns were consistent with established phonological theories of natural speech and to identify any features that might be considered specifically characteristic of digital speech.

3.8 Reliability and Validity

Corpus representativeness was maximised through the tripartite source structure (YouTube, podcasts, Common Voice), the inclusion of speakers from

multiple national varieties and age sub-groups, and the exclusion of scripted or highly formal speech (Topal, 2022; Staples, 2019). The breadth of the corpus ensures that findings are not an artefact of any single platform, speaker community, or demographic group.

Internal validity was supported by the consistent application of the phonological process classification framework across all annotators, backed by a detailed coding manual with worked examples drawn from pilot data. External validity is qualified by the acknowledged limitations of the corpus (discussed in Section 5) but is strengthened by the broad demographic and dialectal coverage of the speaker sample (Proctor & Cormier, 2023; Afanasev & Lyashevskaya, 2025).

RESULTS AND DISCUSSION

4.1 Results

4.1.1 Overall Frequency of Phonological Processes

Analysis of the phonetically transcribed sample (approximately 18.5 hours; approximately 170,000 words) yielded a total of **47,832 identifiable instances** of the four phonological processes under investigation. The distribution across process categories is shown in Table 3.

Table 3: Overall Frequency and Rate of Phonological Processes in the DGESC

Phonological Process	Raw Count	Rate per 1,000 Words	% of Total Processes
Vowel Reduction	18,944	111.4	39.6%
Elision	14,217	83.6	29.7%
Assimilation	8,829	51.9	18.5%
Connected Speech Phenomena	5,842	34.4	12.2%
Total	47,832	281.4	100.0%

Vowel reduction emerges as the most frequent phonological process in the corpus, accounting for nearly 40% of all identified process instances. This finding is consistent with theoretical expectations, given that English has a particularly complex system of vowel reduction tied to its stress-timed rhythmic structure. The high rate of elision (83.6 per 1,000 words) is particularly notable and suggests that digital speech environments may be amplifying or accelerating the elision tendencies already documented in informal spoken English.

4.1.2 Distribution of Elision

Elision was documented across three primary environments: (1) word-final consonant cluster simplification (e.g., asked → [ɑːst], and → [ən]); (2) medial syllable deletion (e.g., interesting → [ˈɪntrɪstɪŋ], comfortable → [ˈkʌmftəbəl]); and (3) function word reduction (e.g., of → [ə], have → [əv] or [ə]). Table 4 presents the distribution of elision across these environments.

Table 4: Distribution of Elision by Phonological Environment

Elision Type	**Raw Count**	**% of Total Elision**	**Examples (IPA)**
Final Cluster Simplification	5,834	41.0%	best [bes], left [lef], and [ən]
Medial Syllable Deletion	4,129	29.0%	interesting [ˈɪntrɪstɪŋ], probably [ˈprɒbli]
Function Word Reduction	4,254	29.9%	of [ə], to [tə], for [fə]
Total	14,217	100.0%	

Final cluster simplification is the most frequent form of elision, a pattern documented across multiple varieties of English and consistent with well-established phonological principles governing cluster reduction. The high rate of medial syllable deletion (29.0%) is particularly striking and may reflect the rapid speech rates characteristic of informal digital communication, where prosodic pressure and articulatory economy converge to produce maximally reduced forms.

4.1.3 Distribution of Assimilation

Assimilation was analysed according to place, voicing, and manner dimensions. Place assimilation—whereby a word-final alveolar consonant acquires the place features of a following bilabial or velar onset—was the most frequent type, accounting for 61.3% of all assimilation instances. Voicing assimilation (e.g., have to → [hæftə]) accounted for 22.8%, and manner assimilation for the remaining 15.9%.

Table 5: Assimilation Types in the DGESC

Assimilation Type	**Count**	**%**	**Illustrative Examples**
Place Assimilation	5,412	61.3%	ten boys [tembɔɪz], good girl [gʊggɜːl]
Voicing Assimilation	2,013	22.8%	have to [hæftə], used to [juːstə]
Manner Assimilation	1,404	15.9%	this year [ðɪʃɪə], in here [ɪŋɪə]
Total	8,829	100.0%	

The preponderance of place assimilation is consonant with results from corpus phonological studies of informal spoken English more broadly. The particular frequency of regressive place assimilation in this corpus—occurring more often than in formal speech corpora—suggests that the communicative conditions of digital speech (notably the pressure for rapidity and fluency) create conditions especially favourable to articulatory economy through place-based convergence.

4.1.4 Vowel Reduction Patterns

Vowel reduction analysis focused on three principal environments: (1) grammatical function words (determiners, prepositions, auxiliary verbs); (2) unstressed syllables in polysyllabic content words; and (3) clause-final position. In all three environments, schwa substitution was the dominant realisation of reduction.

Table 6: Vowel Reduction by Phonological Context

Context	Count	Rate per 1,000 Words	% of Vowel Reduction
Grammatical Function Words	9,812	57.7	51.8%
Unstressed Syllables (Content Words)	7,108	41.8	37.5%
Clause-Final Position	2,024	11.9	10.7%
Total	18,944	111.4	100.0%

Grammatical function words account for over half of all vowel reduction instances, a finding that is phonologically expected given their characteristically low stress weight and high frequency in running speech. What is notable, however, is the rate of reduction in content word unstressed syllables (41.8 per 1,000 words), which in the DGESC exceeds rates reported in formal register corpora by approximately 34%, suggesting a register-driven amplification of reduction processes in informal digital speech.

4.1.5 Connected Speech Phenomena

Beyond the three major process categories, 5,842 instances of broader connected speech phenomena were documented, including r-linking (e.g., the idea of → [ðɪ aɪˈdɪərəv]), j-insertion (e.g., they are → [ðeɪjɑː]), and intrusive sounds (e.g., draw it → [ˌdrɔːrɪt]). Additionally, distinctive digital speech phenomena were noted, including the high frequency of lengthened vowels for emphasis (so [soːːː]) and glottal stop substitution for /t/ in intervocalic position (better [ˈbeʔə]), both of which appeared with notable regularity across speakers.

4.1.6 Distributional Patterns Across Speakers and Contexts

Chi-square analysis revealed statistically significant differences in process rates across age groups ($x^2 = 47.3$, df = 6, $p < .001$), data source types ($x^2 = 31.9$, df = 6, $p < .001$), and national variety ($x^2 = 29.4$, df = 12, $p = .003$). Effect sizes were modest (Cramér's V = .18 for age, .14 for source, .12 for variety), suggesting that while these variables do influence process rates, the overall pattern of high phonological reduction is robust across subgroups.

Table 7: Mean Process Rate (per 1,000 Words) by Age Group and Data Source

	Age 15–18	Age 19–23	Age 24–30
YouTube	301.4	294.2	271.8
Podcast	278.9	283.6	269.4
Common Voice	262.7	271.3	264.1
Overall Mean	281.0	283.0	268.4

The youngest speakers (15–18) and the YouTube data source together show the highest process rates, a convergence that is intuitively plausible: younger speakers are more deeply embedded in informal digital communicative norms, and YouTube vlogs and commentary represent the least formal of the three data source types. The relative stability of process rates across the 19–23 and 24–30 groups suggests that the phonological patterns established in adolescence are largely maintained into early adulthood.

DISCUSSION

4.2.1 Interpretation Within Phonological Theory

The patterns documented in the DGESC are, in their broad outlines, consistent with established phonological theory. Elision, assimilation, and vowel reduction are well-documented features of casual informal spoken English, and their presence in digital speech is not surprising. What is theoretically significant is their rate and systematic consistency across a demographically diverse speaker pool. The overall process rate of 281.4 per 1,000 words is substantially higher than rates reported for formal or semi-formal spoken corpora, suggesting that digital speech does not simply replicate casual spoken English but amplifies its reduction tendencies.

This amplification is consistent with a phonological economy model: speakers in digital contexts are exposed to and produce large quantities of fast, informal speech in which articulatory effort is minimised and communicative efficiency is maximised. Over time, and particularly for speakers who have grown up immersed in such environments, these reduced forms become the default—not marked departures from a more careful norm, but the norm itself (Mair, 2021; Fleischer, Cysouw, & Speyer, 2018).

The particular salience of vowel reduction in the DGESC data resonates with findings from tone and prosody research. Lu, Chuang, and Baayen (2026), working on Taiwan Mandarin, demonstrate that prosodic patterns in spontaneous speech diverge substantially from citation forms, with reduction and contextual modification operating across a range of tonal and vowel contrasts. While their language and phonological system differ considerably from English, the underlying dynamic—spontaneous speech as a site of systematic phonological deviation from canonical forms—is directly analogous.

Similarly, Belando's (2024) documentation of changing /r/ realisations in RP across the twentieth century provides a diachronic parallel to the present synchronic findings: phonological change in prestige varieties is gradual, socially driven, and detectable through corpus evidence. The digital generation's speech may represent an acceleration of long-term tendencies toward connected speech reduction, with digital environments functioning as a social context that reinforces and diffuses those tendencies at scale.

4.2.2 Evidence of Simplification and Reduction

The evidence from the DGESC strongly supports the hypothesis that the phonological profile of digital speech is characterised by systematic simplification. Simplification here refers not to a deficit or degradation of phonological competence, but to a principled reorientation toward forms that minimise articulatory effort, maximise rhythm regularity, and align with the informal communicative norms of digital communities.

The dominance of final cluster simplification in the elision data echoes findings from studies of other varieties undergoing change. Alfaifi, Qasem, and Bokhari (2024), in their corpus-based study of gemination in Egyptian Arabic, note that phonological simplification in child language and informal adult speech shares a common basis in articulatory economy principles—an observation that, while drawn from a different language, supports the cross-linguistic generalisability of the simplification hypothesis.

The reduction of function words is particularly consequential from a pedagogical standpoint. If the primary phonological model encountered by new learners of English is digital speech—as is increasingly likely given the penetration of platforms such as YouTube into language learning contexts—then learners who are not explicitly taught the reduced forms of to, of, for, and them, and related function words will encounter significant comprehension difficulties in real-world interaction (Al-Jarf, 2026; Topal, 2023; Staples, 2019).

4.2.3 Comparison with Recent Corpus-Based Studies

Comparison with recent corpus-based studies confirms and contextualises the present findings. Jawad, Chandio, and Tahir (2026), in their analysis of sound-based spellings in digital communication, identify the same clusters of lexical targets for reduction as are most frequently implicated in the DGESC data—particularly the words going to, want to, because, probably, and definitely. The convergence between their text-based evidence and the present audio-based evidence strengthens the claim that these forms represent genuinely established pronunciation norms rather than sporadic deviations.

Shah and Zahra (2025) document digital-age phonological variation in Pakistani English that shares several features with the digital speech patterns documented in the DGESC, including elevated rates of vowel reduction and final consonant weakening. While Pakistani English is a distinct variety, the similarity of digital-context phonological behaviour across nationally distinct varieties suggests that digital exposure may be functioning as a cross-variety convergence force, promoting shared reduced forms among speakers who share a digital communicative environment even if they do not share a regional speech community.

This cross-variety convergence resonates with findings from Sweet and colleagues (2025), who demonstrate that digital platforms facilitate the mixing and blending of norms across language and variety boundaries. At the phonological level, this may manifest as the diffusion of reduction patterns that are salient in globally prominent varieties of digital English (notably American English, given its dominance on platforms like YouTube) into the speech of digitally exposed speakers from other English-speaking backgrounds.

The multilingual dimension of digital phonological influence is further illustrated by Garcia (2023), whose analysis of English loanwords in Urdu reveals phonological adaptation patterns that reflect the pronunciations of English as

encountered in digital media rather than in traditional contact situations. Michael (2025), in his study of English loanwords in South Korean broadcast media, similarly documents phonological patterns shaped by exposure to digital English models. Together, these studies constitute converging evidence that digital English is emerging as a phonological reference norm with reach well beyond native English-speaking communities.

4.2.4 Role of Digital Communication in Accelerating Change

The results of the present study are consistent with the hypothesis that digital communication environments accelerate phonological change through multiple mechanisms. First, the sheer volume of exposure to informal reduced speech that digital media provides means that reduced forms have a greater probability of being encountered, internalised, and produced (Ma, Mei, & Qian, 2024; Chen, Tian, & Chan, 2025). Second, the interactive and participatory nature of digital platforms—where users not only consume but produce content and receive feedback—creates conditions for rapid norm diffusion and social reinforcement of reduced forms (Bao, 2024; Ormanova & Anafinova, 2022).

Third, the multimodal character of digital communication—in which speech, text, image, and gesture co-occur and interact—may create particular conditions for the spread of phonological norms through mechanisms not fully captured by traditional speech community models. Pavesi (2018), examining corpus-based audiovisual translation research, notes that the relationship between speech and other semiotic modes in digital media creates distinctive conditions for the reception and normalisation of spoken language patterns. Pischedda (2022), analysing the translation of ideophones in picture books, similarly highlights how phonological expressivity is shaped by multimodal contexts—a perspective that extends productively to the analysis of phonological patterns in multimodal digital communication.

Finally, it is worth noting that digital platforms have created new mechanisms for the spread of specific phonological features through viral content. A distinctive pronunciation feature adopted by a popular content creator may be heard by millions of followers and, through social reinforcement, come to be perceived as a norm—a process that has no clear analogue in pre-digital media landscapes. The concentrated exposure of young digital speakers to a relatively small number of highly influential content creators may thus be contributing to a degree of phonological homogenisation that accelerates change in ways not predicted by traditional diffusion of innovation models (Chohan & Yousaf, 2025; Chohan & Sawar, 2025).

The methodological implications of this acceleration are significant. Corpus phonological research must develop more responsive and dynamic tools to track change in near-real-time across digital platforms. Afanasev and Lyashevskaya (2025), working on language distance measurement, suggest that corpus-based computational approaches can provide exactly this kind of dynamic tracking—a

direction that future research on digital phonological change should actively pursue.

CONCLUSION

5.1 Summary of Key Findings

The present study set out to investigate phonological changes in English pronunciation among the digital generation through the systematic analysis of a purpose-built corpus of naturalistic digital speech. The key findings can be summarised as follows. First, the dominant phonological processes in digital speech are vowel reduction, elision, assimilation, and connected speech phenomena, occurring at a combined rate of approximately 281 instances per 1,000 words. Second, these processes are systematic across speakers, contexts, and national varieties—not idiosyncratic features of individual speakers or isolated platforms, but consistent phonological tendencies characteristic of the register as a whole. Third, the rates of phonological reduction in the DGESC substantially exceed those documented in formal spoken English corpora, confirming that digital speech represents a phonologically distinctive register with its own characteristic patterns. Fourth, younger speakers (15–18) and YouTube content show the highest process rates, while the differences across age groups and data sources, though statistically significant, are effect sizes that underscore the overall consistency of the pattern. Fifth, digital communication appears to accelerate phonological change through exposure volume, norm diffusion, and social reinforcement mechanisms operating across national variety boundaries.

5.2 Contributions to Corpus Phonology and Sociophonetics

This study makes several contributions to its field. In terms of corpus phonology, it demonstrates the feasibility and productivity of constructing purpose-built digital speech corpora and subjecting them to multi-process phonological analysis. The DGESC represents a new type of corpus—one specifically designed to capture the phonological character of digitally mediated informal speech—and its methodological design offers a template for future corpus construction in this area (Liberman, 2019; Kunath, 2021; Seifart, 2021).

In terms of sociophonetics, the study provides empirical support for the hypothesis that digital environments function as active agents of phonological change, not merely passive channels for the transmission of existing variety-based norms. The cross-variety consistency of reduction patterns, in particular, points toward an emerging digital speech register with its own phonological identity—a finding with significant implications for theories of dialect levelling, norm diffusion, and language change in the digital age (Belando, 2024; Proctor & Cormier, 2023; Muhammad & John, 2025).

5.3 Practical Implications

For language teaching, the findings suggest that pronunciation instruction must devote far greater attention to reduced forms, connected speech phenomena, and the phonological norms of informal digital speech. Learners who aspire to interact successfully in digitally mediated English-language contexts—podcasts, video content, online professional communication—will be inadequately prepared by instruction focused exclusively on careful citation-form pronunciation (Al-Jarf, 2026; Topal, 2023; Ma, Mei, & Qian, 2024). Resources such as YouGlish, which Topal (2023) proposes as a digital corpus for pronunciation instruction, are particularly well-suited to closing this gap. The integration of AI-assisted pronunciation feedback tools (Chen, Zhou, & Tian, 2025; Yang, 2020) into digital speech training programmes is a further practical implication of the present findings.

For speech recognition systems, the high rates of phonological reduction documented in digital speech have direct implications for the acoustic models underlying ASR technology. Systems trained primarily on careful or formal speech will systematically underperform when applied to digitally informal speech, producing transcription errors at precisely the phonological loci—function word reduction, cluster simplification, assimilation environments—where digital speakers diverge most markedly from canonical forms. The DGESC and the phonological analysis conducted here can serve as a training and evaluation resource for ASR developers aiming to improve performance on informal digital speech. The computational approaches developed by Wang (2023) and Zhao and colleagues (2025) for phoneme conversion and dialect corpus analysis are particularly relevant tools for this development work.

5.4 Limitations

Several limitations must be acknowledged. First, the corpus, at 74 usable hours and 120 speakers, is substantial but not exhaustive. Future research should aim for larger corpora spanning a wider range of digital genres, including live-streamed gaming commentary, social media voice notes, and video call speech—genres that were underrepresented or absent in the DGESC. Second, the speaker pool, while nationally diverse, is weighted toward speakers of standard or semi-standard varieties; regional and non-standard varieties are underrepresented, and future corpora should aim for greater regional diversity. Third, the cross-sectional design of the study captures variation at a single point in time and cannot directly document change over time; longitudinal corpus designs would be better suited to tracking the diachronic trajectory of digital phonological change. Fourth, the AI-assisted annotation component, while methodologically valuable, introduces a layer of uncertainty where ASR error rates may interact with genuine phonological variation in ways that are difficult to fully disentangle.

5.5 Future Research

The present study opens several avenues for productive future research. First, cross-linguistic comparison of phonological reduction patterns in digital speech across multiple languages would illuminate whether the patterns

documented here are specific to English or reflect a broader digital register phenomenon. Studies such as those by Alfaifi and colleagues (2024), Han (2025), and Boborahimov and colleagues (2025), which apply corpus methods to phonological phenomena in Arabic, Chinese, and Turkic languages respectively, suggest that the comparative corpus infrastructure for such work is increasingly available. Second, AI-based phonological modelling using the DGESC as a training corpus could produce powerful computational tools for the detection, classification, and tracking of digital phonological change—a direction directly suggested by the AI phonetics research of Chen and colleagues (2025), Yang (2020), and Chohan and Sawar (2025). Third, longitudinal corpus studies tracking individual speakers' phonological development across their digital exposure histories could shed light on the causal mechanisms through which digital environments shape phonological norms—a question that the present cross-sectional design is unable to fully address.

In conclusion, the evidence assembled in this study makes a compelling case that the digital generation speaks in a phonologically distinctive way, and that this distinctiveness is neither incidental nor idiosyncratic, but systematic, cross-speaker, and theoretically interpretable. Digital communication is not merely changing how we write; it is changing how we speak. Corpus phonology is uniquely positioned to document and theorise this transformation.

REFERENCES

1. Afanasev, I., & Lyashevskaya, O. (2025). The application of corpus-based language distance measurement to diatopic variation study. In Proceedings of the Workshop on Language Resources.
2. Al-Jarf, R. (2026). A systematic review of studies on pronunciation instruction and practice in L2 (2005–2025). Journal of English Language Teaching and Applied Linguistics.
3. Alfaifi, A., Qasem, F., & Bokhari, H. (2024). Gemination in child Egyptian Arabic: A corpus-based study. Languages.
4. Bao, K. (2024). Comparative analysis of representations of feminism across Chinese social media: A corpus-based study of Weibo and Zhihu. Social Media + Society.
5. Belando, D. (2024). Tapped /r/ in RP: A corpus-based sociophonetic study across the twentieth century. Linguistics Vanguard.
6. Boborahimov, B. I., Akhmedjanova, D. A., & others. (2025). A corpus-based model of the Middle Turkic language using weighted averaging. Problems of Computational Linguistics.
7. Chen, H. C., Tian, J. X., & Chan, C. H. J. (2025). Enhancing English pronunciation production and perception through corpus and AI-aided pedagogy. Computer Assisted Language Learning.
8. Chen, H. C., Zhou, X., & Tian, J. X. (2025). Examining the effectiveness of integrating corpus-based and AI approaches for English speaking practice. ReCALL.
9. Chohan, M. N., & Sawar, M. (2025). Trends and tools in AI-supported phonology teaching: A corpus-based meta-analysis. International Bulletin of Linguistics and Literature.

10. Chohan, M. N., & Yousaf, M. (2025). Meta-analysis of corpus-based approaches in phonology: Trends, tools, and theoretical implications. Journal of Advanced Corpus-Oriented Research.
11. Fleischer, J., Cysouw, M., & Speyer, A. (2018). Variation and its determinants: A corpus-based study of German schwa in the letters of Goethe. Zeitschrift für Sprachwissenschaft.
12. Garcia, M. I. M. (2023). A corpus-based study of anglicism: Neologism in Urdu through phono-semantic matching. Journal of Policy Research.
13. Gu, J., Gao, J., Liu, L., & Wang, W. (2026). A corpus-based comparative study of errors in Russian–Chinese bidirectional translation by native Chinese speakers. SAGE Open.
14. Han, Y. (2025). Normalization or creation? A corpus-based study of normalization in the Chinese translation of English children's literature. Humanities and Social Sciences Communications.
15. Hernández, J. L., Molina, F. M., & Almela, A. (2023). Analysis of context-dependent errors in the medical domain in Spanish: A corpus-based study. SAGE Open.
16. Jawad, M., Chandio, P. J., & Tahir, Z. (2026). Phonological patterns in digital communication: An analysis of sound-based spellings. Liberal Journal of Language and Literature.
17. Kunath, S. A. (2021). Tools and techniques for corpus-based analysis of non-native speech patterns. Georgetown University Press.
18. Liberman, M. Y. (2019). Corpus phonetics. Annual Review of Linguistics.
19. López-Hernández, J., & Molina-Molina, F. (2022). Analysis of context-dependent errors in the medical domain in Spanish: A corpus-based study.
20. Lu, Y., Chuang, Y. Y., & Baayen, R. H. (2026). The realization of tones in spontaneous spoken Taiwan Mandarin: A corpus-based survey and theory-driven computational modeling. Corpus Linguistics and Linguistic Theory.
21. Ma, Q., Mei, F., & Qian, B. (2024). Exploring EFL students' pronunciation learning supported by corpus-based language pedagogy. Computer Assisted Language Learning.
22. Mair, C. (2021). Recent advances in the corpus-based study of ongoing grammatical change in English. Text & Talk.
23. Michael, J. C. (2025). Phonetic variation in the pronunciation of English loanwords in South Korean media: A corpus-based study. Journal of Linguistics, Mass Communication, and Media Studies.
24. Muhammad, K. B., & John, K. (2025). Morphological productivity and borrowing in Nigerian English: A corpus-based study. Zamfara International Journal of Linguistics.
25. Ormanova, A. B., & Anafinova, M. L. (2022). Linguistic interference in information space terms: A corpus-based study in Kazakh. Theory and Practice in Language Studies.
26. Pavesi, M. (2018). Corpus-based audiovisual translation studies: Ample room for development. In The Routledge handbook of audiovisual translation. Routledge.
27. Pischedda, P. S. (2022). A corpus-based study on the translation of English ideophones in Italian picture books: The case of Diary of a Wimpy Kid. Languages.
28. Proctor, H., & Cormier, K. (2023). Sociolinguistic variation in mouthings in British Sign Language: A corpus-based study. Language and Speech.

29. Seifart, F. (2021). Combining documentary linguistics and corpus phonetics to advance corpus-based typology. In Doing corpus-based typology with spoken language.
30. Shah, S. F., & Zahra, M. F. T. (2025). Digital age phonological variations in Pakistani English: A sociolinguistic analysis of modern pronunciation patterns. Journal of Applied Linguistics and TESOL.
31. Staples, S. (2019). Using corpus-based discourse analysis for curriculum development: Creating and evaluating a pronunciation course for internationally educated nurses. English for Specific Purposes.
32. Sweet, S. A., Emon, M. N. K., Mim, S. K., & Mahim, F. I. (2025). Banglalish on social media: A corpus-based study of Bangla–English code-mixing across four platforms in Bangladesh. Digital Applied Linguistics.
33. Topal, I. H. (2022). A small-scale corpus-based study for the varietal differences in American and British English: Implications for language education. Language Related Research.
34. Topal, I. H. (2023). YouGlish: A web-sourced corpus for bolstering L2 pronunciation in language education. Journal of Digital Educational Technology.
35. Vella, A., & Grech, S. (2022). What can a corpus tell us about phonetic and phonological variation? In The Routledge handbook of corpus linguistics. Routledge.
36. Wang, Y. (2023). Research on automatic generation algorithm of phoneme conversion learning corpus based on KNN algorithm. In Proceedings of the International Conference on Image Processing and Artificial Intelligence.
37. Xodabande, I., Asadi, V., & Valizadeh, M. (2023). Teaching vocabulary items in corpus-based wordlists to university students: Comparing the effectiveness of digital and paper-based flashcards. Journal of China Computer-Assisted Language Learning.
38. Yang, S. (2020). Pronunciation variation analysis and CycleGAN-based feedback generation for CAPT (Doctoral dissertation). Seoul National University.
39. Zhao, Y., Ding, Y., & Min, X. (2025). Construction of a multimodal dialect corpus based on deep learning and digital twin technology: A case study on the Hangzhou dialect. Methods in Sciences and Engineering.