International Journal of Linguistics, Literature and Translation
ISSN: 2617-0299 (Online); ISSN: 2708-0099 (Print)
DOI: 10.32996/ijllt
Journal Homepage: www.al-kindipublisher.com/index.php/ijllt

IJLLT
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Investigating Vowel Deletion: Methodological Approaches to Production-Based Evidence

Rana Alsabhan
Assistant Professor, Department of Applied Linguistics, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
Corresponding Author: Rana Alsabhan, **E-mail**: rmalsabhan@pnu.edu.sa, **ORCID:** https://orcid.org/0000-0002-6226-3389

| ABSTRACT

Vowel deletion is a widely attested phonetic and phonological phenomenon that has attracted considerable attention in production-based speech research. Despite the growing body of empirical research, a comprehensive synthesis of the methodological approaches used to investigate vowel deletion and their contributions to our understanding of the phenomenon remains limited. Such a synthesis is essential for evaluating the strengths and limitations of existing methodologies, identifying complementary approaches, and informing future production-based research. This review synthesizes research on vowel deletion from a production-based perspective. The review first outlines the conceptual distinction between vowel reduction, devoicing, and complete deletion before examining the principal experimental approaches employed in research. It discusses articulatory techniques alongside acoustic methodologies. The review further examines how these methodologies have contributed to distinguishing vowel reduction, devoicing, complete deletion, and underlying consonant clusters, while providing empirical evidence for theoretical accounts such as gestural overlap and Articulatory Phonology. By critically synthesizing findings from studies across several languages, this review demonstrates that articulatory and acoustic methodologies offer complementary perspectives on vowel deletion and that their integration provides the most comprehensive account of vowel realization in speech production.

| KEYWORDS

Vowel Deletion, Vowel Reduction, Speech Production, Articulatory Phonology, Gestural Overlap, Acoustic Phonetics

| ARTICLE INFORMATION

ACCEPTED: 01 July 2026 **PUBLISHED:** 28 August 2026 **DOI:** 10.32996/ijllt.2026.9.9.3

1. Introduction

Vowel deletion, the apparent absence of a vowel's phonetic realization in connected speech, is one of the most extensively documented reduction phenomena across the world's languages, attested across language families from Indo-European to Mayan (Gordon, 1998; Bennett et al., 2023), yet it remains one of the most difficult to characterize with precision. Production-based research spanning several decades has used increasingly sophisticated instrumentation, from acoustic spectrography to electroglottography and automated articulatory tracking, to ask a deceptively simple question: when a vowel seems to disappear, has it truly been deleted from speech production, or does it persist in some attenuated articulatory or acoustic form (Bennett et al., 2023; Wood, 2025)? Despite this substantial and growing empirical base, no recent synthesis has drawn together the methodological approaches themselves, articulatory and acoustic, to evaluate what each can and cannot reveal about vowel realization, and how their combination has advanced the field.

This review addresses that gap. It is deliberately restricted to production-based evidence: studies that measure the physical articulation or acoustic signal of speech. Perception-based research, studies of listener judgments, categorization tasks, or intelligibility ratings, lies outside this scope, since perceptual and production evidence answer different questions and require separate methodological treatment (Reetz & Jongman, 2009); where relevant, this distinction is noted rather than developed further.

The remainder of the paper proceeds as follows. The Literature Review outlines the conceptual distinctions among reduction, devoicing, and deletion, and the theoretical debate over whether deletion is categorical or gradient. The Methodology section describes the scope and selection criteria for the studies synthesized here. The Results/Findings section surveys the principal articulatory and acoustic methodologies used to investigate vowel deletion and their specific acoustic manifestations. The Conclusion draws these findings together and identifies directions for future production-based research.

2. The Literature Review

Vowel deletion, broadly defined as the absence of phonetic realization of a vowel in speech, is a widely attested phenomenon across languages. Although frequently discussed in the literature, the terminology used to describe it varies considerably. Terms such as deletion, loss, absence, and elision are often used interchangeably to refer to vowels that are not acoustically realized, whereas reduction, shortening, and devoicing denote related but distinct phenomena (Baird et al., 2021; Dauer, 1980; Reetz & Jongman, 2009). Vowel reduction and shortening typically involve decreased duration or articulatory magnitude while retaining some acoustic evidence, often in the form of a centralized vowel. Devoicing refers to the production of a vowel without vocal fold vibration, most commonly in the context of surrounding voiceless obstruents (Delforge, 2008). By contrast, vowel deletion is characterized by the absence of acoustic evidence for the vowel (Beckman, 1982). Distinguishing these terms is essential because inconsistent terminology can obscure methodological choices and complicate comparisons across production-based studies; as recent corpus work on Chichicastenango K'iche' illustrates, the same language can show highly predictable, near-categorical deletion in one grammatical context and much more variable, gradient reduction in another, so that a single language-wide label risks masking real differences in how deletion is distributed within a grammar (Wood, 2025). Throughout this review, terms referring to vowel presence, including retention, realization, and intactness, are treated as synonymous when describing observable phonetic realization.

A central issue in the literature concerns whether vowel deletion should be regarded as a phonological or a phonetic phenomenon. Phonological accounts typically describe deletion as a categorical process in which a vowel is removed from the underlying representation (Bürki et al., 2007). In contrast, phonetic accounts view deletion as the endpoint of a continuum of vowel realization, ranging from temporal shortening and articulatory undershoot to devoicing and, ultimately, complete acoustic absence (Bürki, Fougeron, et al., 2011; Davidson, 2006b). This perspective is consistent with Fourakis (1991), who argues that contextual and temporal constraints may compress the vowel space through articulatory undershoot rather than producing an immediate categorical loss. Consequently, what is described as vowel deletion may, in some instances, represent the extreme outcome of progressive phonetic reduction rather than a discrete phonological operation. More recent work suggests that this dichotomy may itself be incomplete: Bennett et al. (2023) argue, on the basis of electroglottographic evidence from Uspanteko, that vowel deletion can be grammatically controlled — sensitive to morphological and prosodic structure — while still being gradiently implemented at the articulatory level as gestural overlap, suggesting that the categorical/gradient distinction may reflect two levels of the same phenomenon rather than two competing analyses of it.

This theoretical distinction has important methodological implications for production research. Determining whether an apparently absent vowel reflects true categorical deletion or a gradient reduction process requires evidence that extends beyond auditory perception. Consequently, researchers have employed a range of acoustic and articulatory techniques to investigate the presence, reduction, devoicing, or apparent absence of vowels in continuous speech. The choice of methodology therefore plays a critical role in shaping interpretations of vowel deletion and forms the primary focus of the present review.

3. Methodology

This paper presents a narrative, critical review of production-based research on vowel deletion, rather than a systematic review or meta-analysis; its aim is to synthesize and critically evaluate the methodological approaches used in this literature rather than to quantify effect sizes across studies. The review's scope was defined around three criteria. First, only studies employing direct production-based measurement — articulatory instrumentation (e.g., electromagnetic articulography, ultrasound tongue imaging, video analysis, electroglottography) or acoustic analysis of the speech signal (e.g., spectrographic, durational, or statistical modelling approaches) — were included; studies relying solely on auditory transcription, listener perception, or theoretical/formal phonological argumentation without accompanying production data were excluded, consistent with the production-based focus outlined in the Introduction. Second, no restriction was placed on the language investigated or the theoretical framework adopted by the original study, in order to capture the typological breadth of the phenomenon; studies were included across Germanic, Romance, Slavic, Semitic, Japonic, and Mayan languages, among others. Third, no lower date limit was imposed, since foundational articulatory-phonological and acoustic-phonetic work from the 1980s and 1990s (e.g.,

Beckman, 1982; Browman & Goldstein, 1992; Dauer, 1980) remains theoretically central to the field; particular attention was given, however, to studies published in the last five years to ensure the review reflects the current state of the field.

Relevant studies were identified through searches of Google Scholar, JSTOR, ScienceDirect, and the Linguistics and Language Behavior Abstracts (LLBA) database, using combinations of the terms vowel deletion, vowel reduction, schwa deletion/elision, vowel devoicing, gestural overlap, and Articulatory Phonology, cross-referenced with method-specific terms including electromagnetic articulography, ultrasound tongue imaging, electroglottography, and acoustic analysis. Reference lists of identified studies were additionally hand-searched to locate further relevant work, and targeted supplementary searches were conducted in 2025 and 2026 to capture recent studies extending the methodological or typological scope of the literature.

Studies meeting these criteria were organized thematically rather than chronologically, according to the two broad methodological traditions that dominate production-based research on vowel deletion: articulatory approaches, which provide direct physical evidence of speech gestures, and acoustic approaches, which examine the consequences of those gestures in the speech signal. This thematic organization, developed in the Results/Findings section, allows the review to evaluate the specific evidentiary contribution and limitations of each methodology, and to assess how their integration has advanced understanding of vowel deletion in speech production.

4. Results/Findings

4.1 Articulatory Methods

Research on vowel deletion has undergone a substantial methodological evolution. Early investigations relied largely on impressionistic observations and auditory judgments, often treating vowel realization as a binary phenomenon in which vowels were either retained or deleted. For example, early analyses of the French schwa adopted this categorical perspective (Tranel, 1981). More recent production-based research, however, has challenged this assumption through the use of experimental phonetic methodologies. Acoustic analyses of more than 3,000 schwa tokens produced in spontaneous French speech demonstrated that schwa frequently undergoes gradual phonetic reduction rather than categorical deletion (Bürki, Fougeron, et al., 2011). This methodological shift has fundamentally changed how vowel deletion is investigated, moving the field from auditory impression toward empirical evidence obtained through articulatory measurements (Fagyal & Moisset, 1999), acoustic analyses (Topintzi & Baltazani, 2012), and X-ray imaging (Melguy, 2019). More recently still, this evolution has extended to physiological techniques such as electroglottography and to underrepresented language families, broadening the empirical base beyond the European languages that dominated earlier work (Bennett et al., 2023).

Much of the articulatory research on vowel deletion is grounded in Articulatory Phonology, which explains phonological phenomena in terms of the temporal organization of articulatory gestures rather than discrete phonological segments (Browman & Goldstein, 1992; Zsiga, 1994). Within this framework, two complementary mechanisms have been proposed to account for vowel deletion: gestural overlap and gesture magnitude reduction. Gestural overlap refers to the temporal coordination of articulatory movements, such as those of the tongue, lips, and jaw, whose timing is governed by their coordination with one another rather than by an external rhythmic template (Zsiga, 1994). As speech rate increases or articulatory demands become more complex, greater temporal overlap between neighbouring gestures may reduce the temporal and spatial distinctiveness of individual segments, supporting the view that speech is produced as a continuous stream of coordinated gestures rather than as a sequence of discrete units (Browman & Goldstein, 1992). Gesture magnitude reduction, in contrast, involves a decrease in the spatial extent or amplitude of articulatory movements, resulting in weakened articulatory targets.

Within this theoretical framework, vowel deletion is interpreted as the outcome of increased gestural overlap, reduced gesture magnitude, or the interaction of both mechanisms. As vowel gestures become increasingly compressed, they may lose their acoustic salience despite remaining partially articulated. An alternative explanation considers apparent vowel deletion to be a consequence of high vowel devoicing, whereby glottal abduction gestures associated with adjacent voiceless consonants overlap with the glottal gestures of the vowel (Shaw & Kawahara, 2023). In such cases, the oral articulatory gestures may remain largely intact while the absence of vocal fold vibration eliminates acoustic evidence of the vowel. A parallel case for grammatically conditioned gestural overlap comes from Bennett et al. (2023), who show that unstressed vowels in Uspanteko, a Mayan language of Guatemala, delete variably within and across speakers, and that this variability is governed by foot structure and morphology rather than being purely phonetic in origin. These competing and complementary accounts underscore the importance of articulatory methodologies, which are uniquely positioned to determine whether apparently deleted vowels represent complete articulatory absence or merely reduced acoustic realization.

A variety of articulatory techniques have been employed to investigate these hypotheses, each providing complementary evidence regarding the physical realization of vowels.

Electromagnetic Articulography (EMA). Electromagnetic Articulography (EMA) enables researchers to track articulatory movements in real time with high temporal precision, making it particularly suitable for examining tongue trajectories associated with vowel production. Shaw and Kawahara (2018), for example, investigated the articulatory behaviour of devoiced /u/ in Japanese by examining tongue dorsum movements in six speakers producing minimal word pairs containing either voiced or devoiced vowels. Measurements included tongue dorsum trajectories, the duration of the first and second consonants (C1 and C2), and the inter-consonant interval (ICI). Their findings revealed two distinct articulatory patterns: some devoiced vowels maintained tongue movements comparable to those of fully voiced vowels, whereas others followed a nearly linear trajectory that lacked evidence of a vowel height target. These results suggest that /u/ deletion may occur categorically but only under specific conditions, illustrating the ability of EMA to distinguish between complete articulatory deletion and residual gestural realization.

Video-based articulatory analysis. Video recording techniques have also been employed to examine visible articulatory gestures, particularly lip rounding. In French, the schwa is frequently accompanied by lip rounding and may optionally be omitted in connected speech (Durand et al., 1987). Barnes and Kavitskaya (2002) investigated whether lip rounding disappeared entirely when schwa was deleted by analysing video recordings using NIH Image software. Their findings showed that although lip rounding was substantially reduced in deleted contexts, moderate rounding often persisted, indicating that the articulatory gesture was not completely eliminated. This study demonstrates that apparent acoustic deletion does not necessarily imply the disappearance of the corresponding articulatory gesture.

Ultrasound tongue imaging. Ultrasound imaging has become an important tool for visualizing tongue configuration during speech production. Davidson (2005) examined tongue movements using a 2.0–4.0 MHz multi-frequency convex curved linear array transducer that generated thirty mid-sagittal tongue images per second. Speakers' head positions were stabilized to ensure consistent image acquisition, and tongue movement was quantified by calculating the mean distance between tongue contours across productions. Statistical analysis using a sign test showed that tongue trajectories in /səC/ sequences frequently resembled those of consonant-only sequences, suggesting that apparent vowel realization often reflected incomplete gestural overlap rather than the presence of a fully articulated schwa. The study further demonstrated considerable inter-speaker variability, emphasizing the importance of objective articulatory measurements and appropriate statistical analyses when interpreting vowel deletion. More recent ultrasound research has begun to automate this measurement process: Strycharczuk et al. (2025) applied a deep-learning tongue-contour tracker (DeepLabCut) to midsagittal ultrasound data from forty speakers, demonstrating that multiple tongue landmarks, rather than a single highest-point measure, are needed to capture articulatory reduction accurately. This development suggests that automated contour tracking may substantially increase the scale at which articulatory evidence for vowel deletion can be gathered.

Electroglottography (EGG). Electroglottography provides a further, physiologically grounded source of articulatory evidence by measuring vocal fold contact directly, making it particularly well suited to cases in which apparent vowel deletion may instead reflect devoicing rather than articulatory absence. Bennett et al. (2023) used electroglottography alongside acoustic analysis to investigate unstressed vowel deletion in Uspanteko. Their measurements showed that even vowels classified as acoustically "deleted" frequently contributed measurable voicing to the interval between flanking consonants, indicating that these vowels remained represented in the speech production system despite their acoustic non-salience. This finding provides direct physiological support for a gestural overlap account of deletion and illustrates the value of combining articulatory and acoustic measures rather than relying on acoustic evidence alone.

Collectively, articulatory methodologies provide direct evidence of the physical realization of vowels that cannot be obtained through acoustic analysis alone. They enable researchers to determine whether vowel non-realization reflects complete articulatory deletion, residual gestural activity, or extreme gestural overlap. Nevertheless, each technique has inherent limitations. EMA offers highly precise movement trajectories but requires specialized instrumentation and technical expertise. Ultrasound imaging provides detailed visualization of tongue movement but cannot capture all articulators simultaneously, although automated tracking methods are beginning to reduce the labour cost of analysis (Strycharczuk et al., 2025). Video-based techniques are largely restricted to visible articulators such as the lips, and electroglottography, while providing direct physiological evidence of voicing, does not capture supraglottal articulation. In addition, articulatory studies often involve relatively small participant samples because of the complexity and cost of data collection, potentially limiting the generalizability of their findings. Despite these challenges, articulatory methodologies remain indispensable for advancing our understanding of the mechanisms underlying vowel deletion and provide a critical complement to acoustic analyses in production-based research.

4.2 Acoustic Methods

Acoustic analysis is one of the most widely employed methodologies for investigating vowel deletion because it examines the physical properties of the speech signal without requiring direct observation of articulatory movements. Unlike articulatory techniques, which focus on the configuration and coordination of the speech organs, acoustic methods analyse waveforms and spectrograms to identify whether vowels are fully realised, reduced, devoiced, or absent (Davidson, 2010; Delforge, 2008). Vowel presence is typically determined by spectrographic cues, including periodic waveforms, visible formant structures, and vertical voicing striations. The absence or modification of these cues provides important evidence regarding the degree of vowel realization. Owing to their relatively low cost and suitability for analysing large corpora of spontaneous speech, acoustic methods have become central to production-based investigations of vowel deletion across languages.

A common application of acoustic analysis involves measuring temporal and spectral characteristics of speech to distinguish deleted vowels from reduced or devoiced vowels. Davidson (2006b), for example, investigated schwa deletion in twenty-eight English words containing initial pretonic /CəC/ sequences (e.g., support, potato) produced at both fast and slow speech rates. The presence or absence of schwa was determined through detailed inspection of waveforms and spectrograms, accompanied by five duration measurements for each target sequence: the entire /CəC/ sequence, C1, the burst and aspiration of C1 (where applicable), the voiced vowel (when present), and C2. Deleted schwas exhibited a complete absence of voicing, aspiration, and formant structure, whereas devoiced schwas retained measurable aspiration despite lacking voicing or a visible voice bar. Fricative and aspiration durations were significantly longer when vowels were absent, suggesting that apparent deletion often results from increasing gestural overlap rather than categorical vowel loss. Davidson (2006b) argued that the acoustic evidence supports a gestural overlap account, suggesting that vowel deletion is not always categorical. This interpretation is consistent with Myers' (1995) proposal that phonological phenomena may exhibit both categorical and gradient dimensions, with vowel realization ranging from shortening and reduction to devoicing and complete acoustic absence.

Similar acoustic methodologies have been applied to other languages, although they have not always produced comparable conclusions. Torreira and Ernestus (2011), for instance, examined deletion of the French vowel /e/ in casual speech using two near-minimal pairs extracted from a spontaneous speech corpus. Their dataset comprised 450 tokens of c'était /setɛ/ ('it was'), in which /e/ between /s/ and /t/ was frequently omitted, together with 104 tokens containing underlying /st/ clusters. Vowel realization was evaluated using periodic waveforms, voicing, and formant structure. In addition to the five duration measurements adopted by Davidson (2006b), the authors also measured the duration of the /t/ closure and the spectral centre of gravity during the second half of /s/. In contrast to Davidson's findings, no significant durational differences were observed between sequences containing deleted vowels and those with underlying consonant clusters. The authors therefore concluded that deletion of French /e/ represents a categorical absence rather than a gradient process. These contrasting findings illustrate how similar acoustic methodologies may support different theoretical interpretations depending on the language under investigation.

Beyond duration-based analyses, acoustic research has also focused on detailed spectrographic classification of vowel realization. Delforge (2008), for example, analysed 16,581 unstressed vowels in conversational Andean Spanish, identifying 1,948 as reduced. Using spectrographic evidence, reduced vowels were classified into four categories: (1) weakly voiced vowels, characterised by faint voicing and poorly defined formants; (2) partially devoiced vowels, exhibiting voicing for no more than 30 ms; (3) completely devoiced vowels, displaying energy in the first and second formants but lacking glottal excitation; and (4) deleted vowels, characterised by the complete absence of both formant structure and a voice bar. This classification demonstrates that vowel realization is better viewed as a continuum than as a simple binary distinction between presence and absence.

Acoustic classification has also been informed by phonetic context. Dauer (1980) proposed a five-stage model of unstressed high-vowel reduction in which vowel realization varies systematically according to the surrounding consonantal environment. Stage 1 represents fully voiced vowels with regular waveforms and complete formant structures, typically occurring between voiced consonants. Stage 2 comprises shortened but fully voiced vowels, frequently following liquids or nasals. Stage 3 includes very short vowels characterised by irregular waveforms and only the first formant, commonly occurring between voiceless and voiced segments. Stage 4 corresponds to devoiced vowels displaying friction or weak acoustic energy after voiceless consonants, whereas Stage 5 represents complete deletion, where no acoustic evidence of vocalic energy remains, most frequently between voiceless consonants. This model illustrates how acoustic analysis can capture progressive stages of vowel reduction while demonstrating the influence of phonetic environment on vowel realization.

Acoustic evidence has also informed broader cross-linguistic typologies of vowel deletion. Based on a survey of thirty-two languages, Gordon (1998) proposed a typology situating vowel deletion along a continuum extending from fully voiced vowels through varying degrees of devoicing to complete deletion. The typology identified seven factors influencing deletion: vowel length (short > long), position (final > pre-vocalic > medial), vowel height (high > non-high), stress (unstressed > stressed),

manner of articulation (fricatives > unaspirated plosives), place of articulation (/h/ > /s/ > /f, θ, x/), and consonant voicing (voiceless > voiced). Rather than functioning as a methodological framework, Gordon's typology provides an interpretive context within which acoustic findings may be understood, complementing the production-based evidence reported by Davidson (2006b), Delforge (2008), and Dauer (1980). More recent corpus-based work has extended this typological picture to previously undocumented languages: Wood (2025) analysed acoustic spectrographic evidence for over 14,000 vowel tokens from spontaneous speech in Chichicastenango K'iche', a Mayan language, using mixed-effects logistic regression to show that deletion in content words follows predictable patterns conditioned by vowel quality, syllable structure, and stress, whereas deletion in function words is considerably more variable and sensitive to phrase-level prosodic context. This finding illustrates that the acoustic factors identified in earlier, largely Indo-European-focused research generalize only partially to other language families, underscoring the continued need for acoustic documentation of vowel deletion in typologically diverse and underrepresented languages.

Rather than representing competing approaches, articulatory and acoustic methodologies provide complementary perspectives on vowel deletion. Articulatory techniques reveal the underlying gestures involved in vowel production, enabling researchers to determine whether articulatory targets are maintained despite the absence of acoustic evidence. Acoustic analyses, in contrast, examine the observable consequences of these gestures in the speech signal by identifying patterns of voicing, formant structure, duration, and spectral energy. Together, these methodologies offer a more comprehensive account of vowel realization than either approach alone.

Acoustic methodologies have substantially advanced the investigation of vowel deletion by enabling researchers to analyse large quantities of naturally occurring speech and quantify subtle differences in vowel realization across languages, speaking styles, and phonetic environments. Their relatively low cost, non-invasive nature, and applicability to large speech corpora make them particularly well suited for examining variation in spontaneous speech while capturing both intra- and inter-speaker variability. Furthermore, advances in quantitative statistical modelling have strengthened acoustic research by allowing investigators to identify systematic patterns of vowel realization across large datasets (Torreira & Ernestus, 2011). For example, Bürki, Ernestus, et al. (2011) demonstrated the value of generalized linear mixed-effects modelling in analysing more than 4,000 tokens of French schwa from naturally produced speech. This shift toward mixed-effects and curve-based statistical models reflects a broader methodological trend in phonetics: as Tavakoli et al. (2025) document in their review of statistical practice in the field, the adoption of random-effects regression, generalized additive mixed models, and Bayesian approaches, supported by the growth of large forced-aligned corpora, is increasingly allowing researchers to model vowel reduction and deletion as continuous, speaker- and item-sensitive phenomena rather than as fixed categorical outcomes. Nevertheless, acoustic evidence alone cannot determine whether the apparent absence of a vowel reflects complete articulatory deletion or the persistence of articulatory gestures without acoustic realization. Conversely, although articulatory techniques provide direct evidence of speech movements, they are often constrained by specialised equipment, laboratory settings, and relatively small participant samples. Consequently, neither methodology independently provides a complete account of vowel deletion.

Integrating articulatory and acoustic methodologies enables researchers to relate underlying speech gestures to their acoustic manifestations, thereby distinguishing between true segmental deletion, gradient reduction, devoicing, and extreme gestural overlap. The complementary use of these methodologies therefore represents the most robust approach to investigating vowel deletion and provides a more comprehensive understanding of the mechanisms underlying vowel realization in speech production.

4.3 Acoustic Manifestations of Vowel Deletion

Vowel deletion has important acoustic consequences that extend beyond the disappearance of the vowel itself. Following deletion, the resulting surface forms often consist of adjacent consonants without any intervening vocalic material, producing sequences that may acoustically resemble underlying consonant clusters. Consequently, one of the principal challenges in production-based research is determining whether an observed consonant sequence represents a genuine underlying cluster or has arisen through vowel deletion. Addressing this question requires careful interpretation of multiple acoustic cues rather than reliance on a single diagnostic feature.

Acoustic analyses have shown that deleted vowels are characterized by the absence of periodic voicing, formant structure, and transitional vocalic cues between adjacent consonants (Davidson, 2006a; Kilpatrick et al., 2016). Instead, consonants are produced as tightly coordinated sequences in which the offset of the first consonant is followed directly by the articulatory and acoustic characteristics of the second consonant. Such patterns demonstrate that vowel deletion substantially alters the temporal organization of speech while preserving fluent articulatory coordination. Languages that permit complex consonant clusters, such as Czech, provide clear examples of these tightly coordinated sequences, in which both voiced and voiceless obstruent clusters occur without intervening vocalic material (Janda & Townsend, 2000). Similar patterns arise in Mayan languages such as

Uspanteko and K'iche', where variable unstressed-vowel deletion frequently produces complex word-medial consonant sequences (Bennett et al., 2023; Wood, 2025).

Distinguishing consonant clusters created through vowel deletion from underlying clusters depends on the acoustic properties of the consonants involved. Fricative–fricative (FF) sequences are generally identified through the distribution of spectral energy. Sibilant fricatives (/s, ʃ, z, ʒ/) typically display intense high-frequency noise, whereas non-sibilant fricatives (/f, v, θ, ð, h/) exhibit weaker and more diffuse spectral energy (Abu-Al-Makarem, 2005; Kent & Read, 2002; Reetz & Jongman, 2009). Among the sibilants, /s/ generally exhibits the highest concentration of high-frequency energy, while voiced fricatives are additionally distinguished by the presence of a voicing bar reflecting continuous vocal fold vibration (Davidson, 2004). These acoustic characteristics generally allow trained phoneticians to distinguish fricatives with considerable accuracy.

In contrast, vowel deletion occurring within plosive-containing sequences presents greater analytical challenges. In plosive–fricative (PF) and plosive–plosive (PP) clusters, the burst release and aspiration associated with voiceless plosives may acoustically resemble the weak energy typically associated with devoiced vowels, complicating the identification of true vowel deletion (Fujimoto, 2015; Ladefoged, 2003). To address this challenge, Davidson (2010) analysed a wide range of both voiceless (e.g., /ps/, /pf/, /tf/, /ts/, /ks/, /kf/, /pt/, /pk/, /tp/, /tk/, /kp/, /kt/) and voiced (e.g., /bd/, /bg/, /db/, /dg/, /gb/, /gd/) consonant clusters using strict acoustic criteria. A vowel was considered present only when a vocalic interval followed the plosive burst, displaying visible formant structure — particularly a distinct second formant — before a clear reduction in intensity marking the onset of the following consonant. Sequences lacking these acoustic characteristics were classified as true consonant clusters rather than consonant sequences resulting from incomplete vowel realization. This approach illustrates the importance of combining multiple acoustic parameters when evaluating apparent vowel deletion.

Beyond identifying individual instances of vowel deletion, acoustic methodologies have enabled researchers to investigate the systematic factors governing its occurrence across languages and speaking styles. Quantitative analyses of large speech corpora have demonstrated that vowel deletion is influenced by multiple interacting phonetic and prosodic variables. For example, Kohler and Rodgers (2001) compared schwa deletion in German across speaking styles and reported substantially higher deletion rates in spontaneous speech (64%) than in read speech (44%), highlighting the influence of speech style on vowel realization. Similarly, Torreira and Ernestus (2011) employed quantitative acoustic analyses to investigate French schwa deletion, examining the contribution of multiple acoustic measurements to distinguish deleted vowels from underlying consonant clusters. Extending this line of research, Bürki, Ernestus, et al. (2011) analysed more than 4,000 tokens of naturally produced French speech using generalized linear mixed-effects models and identified five significant predictors of schwa retention: word position, utterance position, speech rate, conformity to the Sonority Sequencing Principle, and consonant cluster length. These findings demonstrate the value of quantitative acoustic methodologies in identifying the linguistic and phonetic factors that systematically condition vowel deletion.

Collectively, these studies demonstrate that the acoustic manifestations of vowel deletion extend beyond the absence of vocalic energy to include systematic changes in the organization and realization of adjacent consonants. Accurate identification of deleted vowels therefore depends on the combined interpretation of spectral, temporal, and contextual acoustic cues rather than any single diagnostic property. More broadly, the findings reinforce the methodological argument advanced throughout this review: acoustic analyses provide robust evidence of how vowel deletion is realized in the speech signal, while articulatory methodologies reveal the underlying speech gestures responsible for these acoustic outcomes. Integrating these complementary approaches enables researchers to distinguish between true segmental deletion, gradient reduction, devoicing, and underlying consonant clusters, providing a more comprehensive understanding of vowel deletion in speech production.

5. Conclusion

This review has examined vowel deletion from a production-based perspective, demonstrating that the phenomenon extends well beyond the simple absence of a vowel to encompass a continuum of phonetic realizations, including reduction, devoicing, and complete acoustic deletion. The theoretical debate over whether deletion is categorical or gradient, long treated as a binary choice between competing frameworks, is better understood in light of recent evidence as a question of level of representation: deletion can be categorically controlled by grammatical structure while remaining gradiently implemented at the level of articulation (Bennett et al., 2023). Distinguishing among these realizations therefore requires not auditory impression but careful, converging interpretation of both articulatory behaviour and acoustic evidence.

The studies reviewed illustrate the complementary and, increasingly, converging contributions of articulatory and acoustic methodologies to the investigation of vowel deletion. Articulatory techniques — including Electromagnetic Articulography, ultrasound tongue imaging, electroglottography, and visual analyses of speech gestures — provide direct physiological evidence of the mechanisms underlying vowel realization and deletion, evidence that acoustic analysis alone cannot supply. Acoustic

approaches, in turn, allow researchers to analyse speech signals through measurements of duration, voicing, formant structure, spectral characteristics, and, increasingly, sophisticated statistical modelling, making it possible to investigate vowel deletion across larger and more typologically diverse speech corpora than articulatory methods alone permit. Together, these methodologies have substantially improved researchers' ability to distinguish between vowel reduction, devoicing, complete deletion, and underlying consonant clusters, while providing direct empirical support for theoretical accounts such as gestural overlap and Articulatory Phonology.

Beyond synthesizing empirical findings, this review underscores that methodological integration is not merely desirable but necessary for the study of vowel deletion. Neither articulatory nor acoustic evidence alone is sufficient to explain the full complexity of vowel realization in connected speech: acoustic evidence cannot on its own establish whether an absent vowel reflects true articulatory deletion or residual, acoustically inaudible gestural activity, while articulatory evidence, constrained by cost, equipment, and sample size, cannot alone establish the generality of a finding across speakers, styles, or languages. Only the combined application of these complementary approaches yields a genuinely comprehensive account of the physical production of speech and its observable acoustic outcomes.

Three directions stand out for future production-based research on vowel deletion. First, the empirical base remains heavily weighted toward a small number of well-studied languages, chiefly French, English, and Japanese; recent work extending production-based methods to typologically distinct and previously undocumented languages, including Mayan languages such as Uspanteko and Chichicastenango K'iche' (Bennett et al., 2023; Wood, 2025), demonstrates both that this gap is tractable and that it yields theoretically consequential results, and this line of research should be substantially expanded rather than treated as exceptional. Second, methodological innovation continues to lower the cost of gathering high-quality production evidence: automated, deep-learning-based articulatory tracking (Strycharczuk et al., 2025) and the growing adoption of mixed-effects and Bayesian statistical models over large forced-aligned corpora (Tavakoli et al., 2025) together promise to make articulatory and acoustic evidence easier to collect and analyse at a scale that was previously impractical, and future research should continue to adopt and refine these tools. Third, because this review has been deliberately restricted to production-based evidence, the relationship between how vowels are produced and how they are perceived remains an open and important question; systematically connecting production-based findings of the kind reviewed here to perceptual evidence represents a natural and necessary next step for advancing a fuller account of vowel deletion in speech communication. Pursued together, these directions promise to deepen our understanding of the relationship between phonetic realization and phonological representation.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Acknowledgments: The author gratefully acknowledges Princess Nourah bint Abdulrahman University for its institutional support.
ORCID iD (if any): https://orcid.org/0000-0002-6226-3389

References

[1] Abu-Al-Makarem, A. S. (2005). The acoustics of fricative consonants in Gulf Spoken Arabic [Doctoral dissertation, Bowling Green State University]. ProQuest Dissertations Publishing. http://rave.ohiolink.edu/etdc/view?acc_num=bgsu1669646674383991
[2] Baird, A., Cristiano, A., & Nagy, N. (2021). Apocope in Heritage Italian. Languages, 6(3), 120. https://doi.org/10.3390/languages6030120
[3] Barnes, J., & Kavitskaya, D. (2002). Phonetic analogy and schwa deletion in French. Annual Meeting of the Berkeley Linguistics Society, USA.
[4] Beckman, M. (1982). Segment duration and the 'Mora' in Japanese. Phonetica, 39(2-3), 113-135. https://doi.org/10.1159/000261655
[5] Bennett, R., Henderson, R., & Harvey, M. (2023). Vowel deletion as grammatically controlled gestural overlap in Uspanteko. Language, 99(3), 399–456. https://doi.org/10.1353/lan.2023.a907008
[6] Browman, C., & Goldstein, L. (1992). Articulatory Phonology: An overview. Phonetica, 49(3-4), 155-180. https://doi.org/10.1159/000261913
[7] Bürki, A., Fougeron, C., & Gendrot, C. (2007). On the categorical nature of the process involved in schwa elision in French. Interspeech 2007, Antwerp, Belgium.
[8] Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C., & Frauenfelder, U. H. (2011). What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. Journal of the Acoustical Society of America, 130(6), 3980–3991.
[9] Bürki, A., Fougeron, C., Gendrot, C., & Frauenfelder, U. H. (2011). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. Journal of Phonetics, 39(3), 279-288. https://doi.org/10.1016/j.wocn.2010.07.003
[10] Dauer, R. M. (1980). The reduction of unstressed high vowels in modern Greek. Journal of the International Phonetic Association, 10(1-2), 17-27. https://doi.org/10.1017/S0025100300002073
[11] Davidson, L. (2004). The atoms of phonological representation: Gestures, coordination and perceptual features in consonant cluster phonotactics [Doctoral dissertation, The Johns Hopkins University]. https://go.exlibris.link/VfTfPTRs

[12] Davidson, L. (2005). Addressing phonological questions with ultrasound. Clinical Linguistics & Phonetics, 19(6-7), 619-633. https://doi.org/10.1080/02699200500114077
[13] Davidson, L. (2006a). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. Journal of Phonetics, 34(1), 104-137. https://doi.org/10.1016/j.wocn.2005.03.004
[14] Davidson, L. (2006b). Schwa elision in fast speech: Segmental deletion or gestural overlap? Phonetica, 63(2-3), 79-112. https://doi.org/10.1159/000095304
[15] Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. Journal of Phonetics, 38(2), 272-288. https://doi.org/10.1016/j.wocn.2010.01.001
[16] Delforge, A. M. (2008). Unstressed vowel reduction in Andean Spanish. Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology, Victoria College at the University of Toronto.
[17] Durand, J., Slater, C., & Wise, H. (1987). Observations on schwa in southern French. Linguistics, 25(5), 983-1004. https://doi.org/10.1515/ling.1987.25.5.983
[18] Fagyal, Z., & Moisset, C. (1999). Sound change and articulatory release: Where and why are high vowels devoiced in Parisian French. 14th International Congress of Phonetic Sciences (ICPhS-14), San Francisco, CA, USA.
[19] Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. The Journal of the Acoustical Society of America, 90(4), 1816–1827. https://doi.org/10.1121/1.401662
[20] Fujimoto, M. (2015). Vowel devoicing. In H. Kubozono (Ed.), Handbook of Japanese phonetics and phonology (Vol. 2, pp. 167-214). De Gruyter. https://doi.org/10.1515/9781614511984.167
[21] Gordon, M. (1998). The phonetics and phonology of non-modal vowels: A cross-linguistic perspective. Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals, USA.
[22] Janda, L. A., & Townsend, C. E. (2000). Czech. Languages of the World, 125. http://eridanus.cz/id32402/jazyk/mnohojazykove(1_fondy/The%20Slavic%20and%20East%20European%20Language%20Resource%20Center/CS/stand_alone_czech.pdf
[23] Kent, R. D., & Read, C. (2002). The acoustic analysis of speech (2nd ed.). Singular/Thomson Learning.
[24] Kilpatrick, A., Bundgaard-Nielsen, R. L., & Baker, B. J. (2016). Japanese vowel deletion occurs in words in citation form. Proceedings of the 16th Australasian International Conference on Speech Science and Technology, Parramatta, Australia.
[25] Kohler, K. J., & Rodgers, J. (2001). Schwa deletion in German read and spontaneous speech. In Spontaneous German speech: Symbolic structures and gestural dynamics (AIPUK 35, pp. 97–123). Institute of Phonetics and Digital Speech Processing, University of Kiel.
[26] Ladefoged, P. (2003). Phonetic data analysis: An introduction to fieldwork and instrumental techniques. Blackwell Publishing.
[27] Melguy, Y. V. (2019). Phonetic strengthening or phonological substitution? An X-ray microbeam investigation of acoustic and articulatory variability in the production of American English dental fricatives. Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia.
[28] Myers, J. (1995). The categorical and gradient phonology of variable t-deletion in English. Paper presented at The International Workshop on Language Variation and Linguistic Theory, University of Nijmegen.
[29] Reetz, H., & Jongman, A. (2009). Phonetics: Transcription, production, acoustics, and perception. Wiley-Blackwell.
[30] Shaw, J. A., & Kawahara, S. (2018). The lingual articulation of devoiced /u/ in Tokyo Japanese. Journal of Phonetics, 66, 100-119. https://doi.org/10.1016/j.wocn.2017.09.007
[31] Shaw, J., & Kawahara, S. (2023). Limits on gestural reorganization following vowel deletion: The case of Tokyo Japanese. Laboratory Phonology. https://doi.org/10.16995/labphon.8543
[32] Strycharczuk, P., Kirkham, S., Gorman, E., & Nagamine, T. (2025). Dimensionality reduction in lingual articulation of vowels: Evidence from lax vowels in Northern Anglo-English. Language and Speech, 68(3). https://doi.org/10.1177/00238309251320581
[33] Tavakoli, S., Matteo, B., Pigoli, D., Chodroff, E., Coleman, J., Gubian, M., Renwick, M. E. L., & Sonderegger, M. (2025). Statistics in phonetics. Annual Review of Statistics and Its Application, 12, 133–156. https://doi.org/10.1146/annurev-statistics-112723-034642
[34] Topintzi, N., & Baltazani, M. (2012). The acoustics of high-vowel loss in a Northern Greek dialect and typological implications. In P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, & B. Kühnert (Eds.), Consonant clusters and structural complexity (Vol. 26, pp. 373–402). De Gruyter. https://doi.org/10.1515/9781614510772
[35] Torreira, F., & Ernestus, M. (2011). Vowel elision in casual French: The case of vowel /e/ in the word c'était. Journal of Phonetics, 39(1), 50-58. https://doi.org/10.1016/j.wocn.2010.11.003
[36] Tranel, B. (1981). Concreteness in generative phonology: Evidence from French. University of California Press.
[37] Wood, E. A. (2025). Predictable vowel deletion in content words and variable vowel deletion in function words in Chichicastenango K'iche'. Journal of the International Phonetic Association, 55(3), 465–502. https://doi.org/10.1017/S0025100325100765
[38] Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. Journal of Phonetics, 22(2), 121-140.