Journal of Phonetics 119 (2026) 101535

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Shared structure and divergent phonetic grammars: prosodic fine-tuning of vowel nasalization in Australian and American English

Jiyoung Jang [a], Hyunjung Joo [a], Sahyang Kim [b], Anne Cutler [c], Taehong Cho [a,d,*]

[a] Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul, South Korea
[b] Department of English Education, Hongik University, Seoul, South Korea
[c] The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, NSW, Australia
[d] Department of English Language and Literature, Hanyang University, Seoul, South Korea

ARTICLE INFO

Article history:
Received 19 February 2026
Received in revised form 22 July 2026
Accepted 1 September 2026
Available online xxxx

Keywords:
Nasal coarticulation
Glottalization
American English
Australian English
Prosodic structure
Phonetic grammar
Cross-dialectal comparison

ABSTRACT

This study examines how prosodic structure regulates coarticulatory vowel nasalization in Australian English (AusE) and how this system compares with American English (AmE). Adopting a phonetic-grammar framework in which prosody serves as a structural scaffold for phonetic implementation, we manipulate focus-driven prominence and prosodic boundary strength to investigate two interrelated dimensions of nasality: nasal consonant duration and time-varying vowel nasalization. For direct comparisons, the experimental design mirrors prior work on AmE (Cho et al., 2017) and explicitly models potential phonation (glottalization) confounds. Results show that AusE and AmE share a prosodic-structural conditioning of nasality. In both dialects, prominence enhances the nasal consonant while inducing coarticulatory resistance in the adjacent vowel; and boundary effects show an asymmetry across phrase edges, with nasal duration and vowel nasalization reduced phrase-initially and enhanced phrase-finally. Crucially, however, shared prosodic organization does not entail phonetic identity. AusE shows stronger focus-induced coarticulatory resistance than AmE. Importantly, this difference is not a uniform shift in magnitude or temporal extent, but a dynamic modulation over the vowel's time course in both carryover (NVC) and anticipatory (CVN) contexts. This difference persists after including glottalization (stronger under focus in AusE) as a covariate, implying genuine differences in velopharyngeal coordination. Moreover, such modeling reveals a stronger domain-initial boundary effect on vowel nasalization (initial reduction in NVC) in AusE than in AmE. We conclude that dialectal variation operates at the level of phonetic grammar with reference to prosodic structure: while AusE and AmE share abstract phonological categories and prosodic representations, they differ systematically in the fine-tuning of control parameters at the phonetics–prosody interface, differentially balancing system-oriented efficiency and output-oriented distinctiveness.

© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

1. Introduction

This study examines how prosodic structure regulates coarticulatory vowel nasalization in Australian English (AusE), and how this system compares to American English (AmE) under tightly matched experimental conditions. Specifically, we investigate how focus-driven prominence and prosodic boundary strength shape two interrelated dimensions of nasality in AusE: nasal consonant duration and the time-varying pattern of vowel nasalization adjacent to nasal consonants. Building on prior work in AmE showing that prosodic structure contributes to systematic variation in vowel nasalization (Cho et al., 2017), we ask whether similar organizational principles characterize AusE, and whether dialects that share broad phonological and prosodic structure nonetheless diverge in their phonetic implementation. To enable direct comparison, the AusE experiment closely mirrors the design of Cho et al. (2017), employing matched materials, parallel prosodic manipulations, and comparable analytic procedures. Furthermore, because acoustic measures of vowel nasalization can be influenced by phonation-related variation—particularly glottalization—and because AusE and AmE differ in their prosodic use of glottalization (e.g., Jang et al., 2026), we explicitly account for phonation in the statistical analyses using the spectral tilt measure H1*–A3* as a covariant. This allows us to disentangle

* Corresponding author.
E-mail address: tcho@hanyang.ac.kr (T. Cho).

https://doi.org/10.1016/j.wocn.2026.101535
0095-4470/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

velopharyngeal effects on vowel nasalization, indexed by A1–P0*, from differences attributable to the source spectrum. By combining a detailed analysis of AusE with a controlled cross-dialectal comparison, the present study aims to clarify how prosody shapes the temporal organization of coarticulation and to assess the level of granularity at which phonetic grammar must be specified, both within and across dialects.

1.1. Theoretical and empirical background

Speech is not assembled from static pieces, but executed as a continuously changing coordination of articulatory movements, dynamically timed and scaled as an utterance unfolds. While this is evident at the physiological level—where gestures overlap, targets are approached and abandoned, and timing is continuously adjusted—it is equally consequential for theory. If the phonetic surface is intrinsically time-structured, the central question shifts from how categories are represented to how they are implemented: how speakers allocate temporal and articulatory resources, and how such allocation yields stable regularities amid variability. Cross-linguistic work over the past several decades has consolidated this view from multiple perspectives (e.g., Beddor, 2023; Cho & Ladefoged, 1999; Cho, 2025; Keating, 1984, 1985; Kingston & Diehl, 1994; Solé, 2007; Solé & Ohala, 2010), shifting the explanatory focus from documenting variability to explaining how such patterns are learned, controlled, and conventionalized.

A particularly transparent window onto this problem is coarticulation. Because multiple gestures must be executed within a limited time, articulatory plans overlap and the acoustic signal reflects this overlap (cf. Kühnert & Nolan, 1999). Coarticulation is theoretically informative not merely because adjacent speech gestures overlap, but because the extent and temporal organization of that overlap can reveal how those gestures are planned and coordinated in relation to other dimensions of linguistic structure. If coarticulation were solely an automatic consequence of articulatory mechanics, its realization should be governed primarily by the local segmental context. Systematic modulation by prominence and prosodic boundary, together with variation across dialects, instead indicates that speakers regulate gestural overlap as part of the phonetic implementation of higher-level linguistic structure. Examining when and how coarticulation varies therefore provides a means of probing the mapping between abstract prosodic organization and continuous speech production, as well as the extent to which that mapping is shared or differs across linguistic varieties.

This systematic variability is clearly illustrated in coarticulatory vowel nasalization adjacent to nasal consonants. Although vowel nasalization arises from the inherently overlapping velic and oral gestures, its realization is nonetheless constrained by linguistic structure. It varies with prosodic organization (e.g., syllable structure, phrasal boundaries and prominence; Cho et al., 2017; Krakow, 1999), phonological patterning and phonologization pathways (Cohn, 1990; Styler, 2017; Zellou, 2022), lexical organization (Scarborough, 2003), and segmental composition and differential resistance (Farnetani & Recasens, 1999). It also covaries with speaking style and reduction (Lindblom, 1990), cue-trading relations within a contrast system (Beddor et al., 2018), learned expectations (Zellou, 2022), and speaker-specific coordination shaped by experience (Pouplier et al., 2024). Such structured regimes differ across languages and varieties and may contribute to pathways of sound change (Beddor, 2009, 2023; Beddor et al., 2018).

Such converging evidence supports the view that coarticulation is governed by a phonetic grammar—i.e., a system of tunable, language-specific control parameters that mediate between abstract phonological representations and the continuous acoustic surface (e.g., Cho & Ladefoged, 1999; Keating, 1984, 1985; Kingston & Diehl, 1994; Pouplier et al., 2024; Solé, 2007). This view is further elaborated in the *extended model of phonetic grammar* (Cho et al., 2026; Cho, 2025) in which prosody serves as the structural scaffold for this implementation. By organizing utterances into hierarchical domains and designating positions of prominence, prosody provides a framework for the differential allocation of phonetic resources. Crucially, the realization of this structure—often discussed as prosodic strengthening—is not a unitary phenomenon. Prominence and boundary effects do not necessarily mirror one another, nor are they uniform across phonetic dimensions or languages (Cho, 2016; Fletcher, 2010). From a grammatical perspective, such patterned asymmetries are diagnostic: it indicates that the mapping from prosodic structure to phonetic detail is not mechanically fixed, but actively regulated by control settings that manage temporal coordination and gestural magnitude.

Coarticulatory vowel nasalization offers a particularly revealing testing ground for this view. Because prosodic modulation of coarticulation requires the precise regulation of intergestural timing, any systematic sensitivity to prominence or boundaries cannot be reduced to the inevitability of gestural overlap. Instead, it demonstrates that the system recruits timing and coordination as part of its phonetic organization. This leads to the specific empirical expectation guiding the present study: that coarticulatory vowel nasalization will pattern systematically with prosodic structure, but that the nature of this patterning—the “grain” of the phonetic grammar—may vary across dialects.

Recent work has made these issues especially concrete by examining vowel nasalization under experimentally controlled manipulations of prosodic structure. In American English, Cho et al. (2017) showed that both nasal consonant duration and coarticulatory vowel nasalization are systematically conditioned by focus-driven prominence and prosodic boundaries, but not in a uniform manner. Prominence enhanced the nasal consonant itself while simultaneously yielding coarticulatory resistance in the adjacent vowel, reflected in reduced vowel nasalization. This dissociation is theoretically revealing because it challenges a simple scaling view in which enhancing a source gesture automatically strengthens its coarticulatory consequences; instead, it suggests differential regulation of adjacent segments, allowing the vowel’s oral target to be selectively protected even as the nasal gesture is temporally enhanced. Parallel lines of inquiry extend this perspective beyond English. Studies in Korean (Jang et al., 2018) and Mandarin Chinese (Li et al., 2020) likewise show systematic prosodic conditioning, while differing in the temporal alignment between nasal consonants and vowels, the coupling between nasal duration and vowel nasalization, and the robustness of nasalization in boundary-adjacent contexts. Taken together,

these studies support a view in which vowel nasalization—despite its grounding in velum–oral overlap—is shaped by a learned, system-specific organization of timing and magnitude relations calibrated to a language's prosodic architecture.

A natural next question, then, is whether this prosodically structured organization is specified only at the level of language, or whether it can also vary across dialects of the same language. Such dialectal comparison is theoretically informative because it holds much of the phonological structure constant while allowing implementation settings—such as timing relations and sensitivity to prosodic modulation—to vary. Consequently, any observed differences in coarticulation provide diagnostic evidence about how the phonetics–prosody interface is implemented and about the grain at which phonetic grammar must be specified across dialects.

Indeed, dialectal and social variation in vowel nasalization has already been well documented in connection with sound change and sociophonetic conditioning. In American English, Tamminga and Zellou (2015) report structured subphonemic variation: Philadelphia exhibits greater anticipatory nasalization than Columbus, though both dialects show reduced nasalization among younger speakers, consistent with denasalization in progress. Similar dialectal differentiation in coarticulatory nasalization has been reported for Spanish (Bongiovanni, 2021) and Afrikaans (Coetzee et al., 2022). More broadly, fine-grained phonetic implementation can serve as a stable marker of dialect identity (e.g., Clopper et al., 2005; Clopper & Smiljanic, 2011, 2015; Kendall et al., 2023) without necessarily implying differences in underlying phonological representations—even while providing a potential substrate for sound change (e.g., Beddor, 2009, 2023; Zellou, 2022). This distinction mirrors findings in intonation, where dialects sharing the same phonological categories may nonetheless differ in fine-grained implementation, such as the temporal anchoring of tonal targets in German (Ladd, 2006, 2008) or the alignment and scaling of rising accents in American English (Arvaniti & Garding, 2007).

This dialect-specific variability extends beyond intonation to the coordination of segmental gestures. Of particular relevance is articulatory work by Cunha et al. (2024), which illustrates how dialectal differences in vowel nasalization can be instantiated via timing mechanisms. Comparing American English and Southern British English using real-time MRI, Cunha et al. showed that although American English exhibits greater vowel nasalization, this difference does not arise from earlier or more extensive velic lowering. Instead, the velum gesture is largely stable across varieties, and dialectal differences emerge from reduction and temporal reorganization of the oral (tongue-tip) gesture in the nasal coda. As a consequence, the internal vowel–nasal alignment is shifted, yielding greater acoustic nasalization without direct modification of the velic gesture. This finding is theoretically significant because it locates dialectal variation in coarticulation at the level of intergestural timing and coordination rather than categorical articulatory targets or phonological representations. However, because Cunha et al.'s data were elicited under a single prosodic condition, a critical question remains: how do such dialect-specific coordination regimes interact with prosodic structure—particularly prominence and prosodic boundaries, which are independently known to modulate nasal coarticulation (Cho et al., 2017)?

This brings us to the central gap the present study addresses. Despite evidence that dialects differ in temporal organization in both intonation (Ladd, 2008) and articulation (Cunha et al., 2024), relatively little is known about how dialects may differ specifically in the way prosody regulates coarticulatory processes. In particular, it remains unclear whether prosodic mechanisms that enhance distinctiveness under focus (e.g., coarticulatory resistance) operate uniformly across dialects, or whether the relative weighting of these mechanisms constitutes a dialect-specific dimension of phonetic grammar.

1.2. The present study

To empirically test this possibility, the present study focuses on Australian English (AusE) as a critical comparison case. Although AusE shares many phonological and prosodic properties with American English (AmE), varieties often differ in temporal organization and phonetic implementation. While AusE and AmE are conventionally regarded as dialects of the same language, the present study does not treat the language–dialect distinction itself as theoretically central. Rather, the comparison is motivated by their substantial similarities in the phonological and prosodic dimensions relevant to the present study, which make it possible to employ closely matched speech materials and the same experimental manipulations across varieties. Consequently, observed cross-variety differences can be interpreted more directly as differences in phonetic implementation rather than as consequences of differences in experimental design. The comparison therefore provides a controlled test case for a broader question that is not specific to the language–dialect distinction—that is, how linguistic systems with closely aligned higher-level structures may nevertheless differ in their phonetic grammars. In the present case, AusE and AmE may diverge in how nasality is allocated relative to prominence and prosodic boundaries despite their broadly comparable phonological and prosodic organization.

A central strength of the present study is its strict comparability: the AusE experiment closely mirrors the design of Cho et al. (2017) for AmE, employing matched materials and parallel prosodic manipulations. This alignment ensures that convergences reflect shared organizational principles, while divergences can be confidently attributed to differences in phonetic implementation rather than methodological artifacts. Following Cho et al. (2017), we examine two linked dimensions: nasal consonant duration and time-varying vowel nasalization. These are interdependent: the nasal consonant acts as the source of velic opening, while vowel nasalization reflects the coordination of velic and oral gestures. Measuring both allows us to determine whether prosody modulates the source and its coarticulatory consequences in parallel or dissociation, and whether these patterns differ across dialects.

A further motivation is methodological. Acoustic assessment of vowel nasalization relies on measures that are susceptible to other sources of spectral variation. While this study employs the established A1–P0* index (based on the relative magnitude of low-frequency spectral components) for vowel nasalization (e.g., Chen, 1997; Garellek et al., 2016; Styler, 2017), it explicitly addresses the confounding effects of

phonation—particularly glottalization. This precaution is critical for cross-dialectal comparison, as varieties may differ not only in nasality but also in baseline voice quality and the prosodic conditioning of phonation.

Indeed, vowel glottalization, often described as creaky voice or vocal fry (Gordon & Ladefoged, 2001; Ladefoged, 1971; Laver, 1980), is known to be systematically conditioned by prosodic structure, frequently appearing in prominent positions and at the beginnings of constituents (Dilley et al., 1996; Garellek, 2013; Mitterer et al., 2019; Pierrehumbert & Talkin, 1992; Redi & Shattuck-Hufnagel, 2001). This poses a methodological challenge because glottalization reduces the amplitude of low-frequency harmonics (H1), commonly indexed by H1*–A3*. Since H1 lies near the nasal peak P0, systematic variation in phonation can conflate source-spectrum differences with velopharyngeal coupling changes in the A1–P0* measure.

Crucially for the present study, Jang et al. (2026) have recently reported that AmE and AusE do differ in their prosodic use of glottalization, particularly highlighting that AusE speakers glottalize more in association with focus-induced prominence than AmE speakers. Thus, modeling this interaction is essential to ensure that apparent dialectal differences in nasality are not merely artifacts of differing phonation patterns. Accordingly, a key approach in the present work is comparison between baseline models and models that incorporate phonation-related variance (indexed by H1*–A3* for glottalization). This strategy is not merely a control. If dialectal differences in nasalization persist after phonation-related variance is accounted for, they are less plausibly reducible to acoustic artifacts of voice quality and more plausibly attributable to dialect-specific organization of velopharyngeal control. Conversely, if certain effects weaken or reorganize when phonation is modeled, this offers insight into how laryngeal and velopharyngeal subsystems may be coordinated together under prosodic structure, or into how surface acoustic measures reflect interactions among these subsystems.

With these methodological controls in place, the present study examines how nasality is shaped by prosodic structure in AusE, and how this system compares to AmE under tightly matched experimental conditions. We investigate both anticipatory (CVN) and carryover (NVC) contexts, which differ in the coordination of velic and oral gestures relative to the vowel. Prosodic structure is manipulated along two dimensions—focus-driven prominence and prosodic boundary strength—allowing us to assess how prominence and phrasing independently and jointly regulate nasal consonant realization and coarticulatory vowel nasalization.

The investigation concerns three basic research questions. First, regarding *system stability*: to what extent does AusE exhibit the same prosodic-structural modulation of nasality previously reported for AmE? If prosodic conditioning is a stable component of the English phonetics–prosody interface, we expect AusE to show comparable signatures of prominence- and boundary-related modulation in both nasal duration and vowel nasalization. Second, regarding *dialect-specific implementation*: do the dialects differ in the temporal dynamics of coarticulation? Here, the emphasis extends beyond mean nasalization to ask whether dialectal differentiation manifests not merely as a static offset in magnitude, but as a distinct, time-structured modulation across the vowel—and whether focus and boundary conditions shape this temporal deployment differentially. Third, regarding *methodological control*: to what extent does phonation (glottalization) confound cross-dialectal differences in measured nasality? Given that glottalization is prosodically conditioned and varies across dialects, explicit modeling is required to determine whether observed dialect effects reflect genuine velopharyngeal control or whether they are conflated with source-spectrum differences.

From a broader perspective beyond these empirical questions surrounding coarticulatory vowel nasalization, this study bears directly on the architecture of the phonetics–prosody interface. Specifically, this investigation probes a central tenet of the *extended model of phonetic grammar* (Cho, 2025), in which prosody serves as a structural scaffold for phonetic implementation. Seen from this perspective, coarticulation serves as a diagnostic case for determining how phonetic grammar is structured across closely related systems. If prosody functions as a higher-order regulator, its effects should not be exhausted by language-general tendencies, nor should they be reducible to static enhancements of individual gestures. Rather, prosodic structure should condition—in a dialect-specific way—how multiple articulatory subsystems are coordinated over time. Just as American and Southern British English diverge in the intergestural timing of oral and velaric movements (Cunha et al., 2024), we hypothesize that such differences may also be crystallized in the way timing is regulated at the phonetics–prosody interface.

Thus, on the view adopted here, dialectal differences should emerge as systematic variations in the settings of tunable control parameters that govern the temporal implementation of shared phonological structure. While such differentiation could in principle arise along historically contingent, “arbitrary” pathways (Cho & Ladefoged, 1999), it is likely channeled by broader pressures that guide the evolution and stabilization of phonetic systems (Flemming, 2004; Keyser & Stevens, 2006; Kingston & Diehl, 1994; Lindblom, 1986, 1990; Maddieson, 1997; Ohala, 1997; see Cho, 2025, for discussion). From this perspective, dialectal variation can be conceptualized as a difference in how competing implementation pressures are balanced. Pressures promoting system-level efficiency favor gestural overlap and coordination, whereas those promoting output distinctiveness favor the recoverability of contrasts and resistance to coarticulatory intrusion. Crucially, however, this balance is not expected to be static across an utterance. Because prosodic structure specifies where contrasts are prioritized, it provides a principled context in which these pressures can be differentially reweighted: prominence and boundary structure can selectively increase the demand for distinctiveness or, conversely, relax it. We therefore ask whether dialects differ not just in their baseline settings, but in how they navigate this dynamic reweighting to shape the time course of coarticulation.

In summary, by comparing Australian and American English under tightly matched manipulations of focus-driven prominence and prosodic boundary strength, the present study therefore shifts attention away from whether nasality is simply “more” or “less” realized in one variety than another. Instead, it

asks how prosodic information is translated into time-dependent coarticulatory patterns, and whether dialects that broadly share phonological representations nevertheless differ in the fine-tuning of these prosodically conditioned implementation regimes. The results thus bear directly on the granularity at which phonetic grammar must be specified: not only at the level of languages, but also at the level of dialects, where shared prosodic structure may nonetheless support distinct, internalized patterns of vowel nasalization that have emerged through dialect-specific developmental and diachronic trajectories.

2. Methods

2.1. Participants and recording

For Australian English data, fourteen native speakers of Australian English (10 female, 4 male) participated in the current study. All participants were born and raised in Australia and were in their 20s or 30s at the time of the recording. Recordings were made in a sound-proof booth at the MARCS Institute Lab with a Tascam DR-680 multi-channel digital recorder and a Shure KSN44 condenser microphone at a sampling rate of 44.1 k Hz. American English data analyzed in Cho et al. (2017) was used for the comparison with Australian English data. Both American English and Australian English data were collected with a comparable experimental setup. (Note that regional dialect differences within the Australian English speaker groups were not systematically controlled as the present study was not designed to investigate within-variety variation.)

2.2. Speech materials

As in Cho et al. (2017), eight monosyllabic words with a nasal stop as either onset (*mob, mop, Ned, net*) or coda (*bomb, palm, den, ten*) were used.[1] The test words included only non-high vowels, such as /ɛ, ɑ/, to avoid overlap between the first formant (F1) of the vowel and the nasal peak (P0) which is often described as occurring between 250 and 450 Hz (Chen, 1997).

To situate the target words within a discourse context, the target words were embedded in carrier sentences produced in response to prerecorded questions, as shown in Table 1. The participants were asked to imagine that they had written either a word (e.g., a target word "mop") or a phrase (e.g., a target-bearing phrase "mop fast again") on a piece of paper. They were told that they would serve as Speaker B in Table 1 and correct Speaker A who would guess the word or the phrase incorrectly (see below for further explanation on the procedure). All target words were embedded in the same question-and-answer frame sentences shown in Table 1, with the target word being the only lexical item that varied across stimuli.

[1] Although item-level variation was not a focus of the present study, we note that some variability in overall nasalization was observed across lexical items in the random-effects structure (see Section 2.4. for details on model structure). In particular, *ten* exhibited relatively greater nasalization than the other target words, a pattern consistent with observations reported by Cohn (1990). However, a systematic investigation of lexical effects falls beyond the scope of the present study, and we therefore leave this issue for future research.

The target words were manipulated with respect to contrastive focus-driven prominence (Focused vs. Unfocused) and the prosodic boundary (the Intonational Phrase [IP] boundary vs. the Word [Wd] boundary) positions. For the focused condition, the target word either received phonological focus by contrasting the nasal vs. oral stop (e.g., NVC, ***m***op vs. ***b***ob; CVN, bo***mb*** vs. bo***b***) or lexical focus by contrasting a lexically relevant word (e.g., NVC, ***mop*** vs. ***wash***; CVN, ***bomb*** vs. ***war***). In the unfocused condition, the contrastive focus fell on a different word in the sentence (e.g., NVC, ***slowly*** vs. ***fast***; CVN, ***you*** vs. ***John***) so that the target word did not receive focus. In the IP boundary condition, the target word was placed at the left edge of an IP for NVC and at the right edge for CVN. In the Word boundary condition, it was produced in the middle of an IP.

2.3. Experimental procedure

In the experiment session, the mini-discourse sentences were presented on a computer screen. A pre-recorded prompt sentence of a native Australian female speaker (Speaker A) was played through a loudspeaker, and the participant answered the question as Speaker B. Speech rate of the prompt sentences was comparable in both Australian English and American English experiments, thereby helping to maintain similar speech patterns of the elicited sentences across the languages.

To yield different types of focus, speakers were asked to make contrast between words in bold in Sentences A and B, in order to induce corrective phonological contrastive focus on N (e.g., NVC, *mop* vs. *bob*; CVN, *bomb* vs. *bob*) or corrective lexical contrastive focus (e.g., NVC, *mop* vs. *wash*; CVN, *bomb* vs. *war*). To yield different boundary types, an IP boundary after the test word was obtained with a following tag question (e.g., "wasn't I", as given in Table 1) for CVN context; and an IP boundary before a test word was induced by placing the words "Not exactly" before the test word for NVC context. Finally, the Wd boundary was induced by placing the test word midway in a short phrase (e.g., "say TARGET fast again"). Test sentences were presented to the participants in a randomized order with the set of entire sentences repeated four times. A total of 2688 sentences were recorded: 2 syllable positions (NVC vs. CVN) × 4 test words × 3 focus types (PhonFOC vs. LexFOC vs. NoFOC) × 2 boundaries (IP vs. Wd) × 4 repetitions × 14 speakers. Two trained English ToBI transcribers (two of the authors) checked the prosodic renditions on the focus and boundary types. 396 tokens with unintended pitch placements and boundary markings were discarded.

2.4. Measurement

The acoustic nasal duration of CVN and NVC (henceforth, N-duration) was measured from the onset to the offset of the nasal murmur as observed on the spectrogram. As an acoustic measure of vowel nasalization (henceforth, V-nasalization), we used A1–P0* (A1 refers to the amplitude of the first formant and P0 to the amplitude of the spectral nasal peak around 250–450 Hz) as in previous studies (see Cho et al., 2017; Jang et al., 2018; Zellou, 2017). The A1–P0* values were obtained by using a Praat script by Styler (2017). Note that

Table 1
Example sentences with target word *mop* for the NVC context and *bomb* for the CVN context. Focused condition shows examples of phonological focus. Target words are underlined, and focused words are in bold.

Structure	Boundary	Focus	Example sentences
#NVC (mop)	IP-initial	Focused	A: *Did you write "**bop** fast again"?*
			B: *Not exactly. "**Mop** fast again" was what I wrote.*
		Unfocused	A: *Did you write "mop **slowly** again"?*
			B: *Not exactly. "Mop **fast** again" was what I wrote.*
	Wd-initial	Focused	A: *Did you write "say **bop** fast again"?*
			B: *Not exactly. I wrote, "say **mop** fast again."*
		Unfocused	A: *Did you write "say mop **slowly** again"?*
			B: *Not exactly. I wrote, "say mop **fast** again."*
CVN# (bomb)	IP-final	Focused	A: *Were you supposed to write "**bob**"?*
			B: *No. I was supposed to write "**bomb**", wasn't I?*
		Unfocused	A: *Were **you** supposed to write "bomb"?*
			B: *No. **John** was supposed to write "bomb", wasn't he?*
	Wd-final	Focused	A: *Did you write "say **bob** fast again"?*
			B: *No. I wrote, "say **bomb** fast again."*
		Unfocused	A: *Did you write "say bomb **slowly** again"?*
			B: *No. I wrote, "say bomb **fast** again."*

A1–P0* values are inversely related to the degree of nasalization (the higher A1–P0*, the lesser degree of nasalization). The A1–P0* values were measured at relative timepoints (i.e., 25%, 50%, 75% of the vowel duration), as shown in Fig. 1. The relative measure indicates the extent to which the coarticulatory effect spreads throughout the vowel as a process beyond a simple proximal effect. Fig. 2 shows schematized A1–P0* measures at the three relative time points in CVN. Fig. 2a shows a larger value of A1–P0* at the 75% time point (which is furthest from the nasal source), indicating a lesser degree of V-nasalization; Fig. 2c shows a lower A1–P0* value, indicating a greater degree of V-nasalization near the nasal source. It is, however, worth reiterating that although A1–P0* is a well-established acoustic correlate of vowel nasalization, it provides only an indirect measure of the underlying velic gesture and does not reveal the precise timing of velum lowering or raising. The present results should therefore be interpreted as reflecting patterns in the acoustic realization of nasalization rather than as direct evidence of velum movement or velopharyngeal aperture.

2.5. Statistical analyses

The retrieved data were analyzed using linear mixed-effects models implemented in the lme4 package (Bates et al., 2015) in R (R Core Team, 2024). Two dependent variables were examined: nasal murmur duration (N-duration, in ms) and vowel nasalization (V-nasalization), indexed acoustically by A1–P0*. The fixed effects included Dialect (Australian English [AusE], American English [AmE]), Focus (Focused, Unfocused), Boundary (Intonational Phrase [IP], Word [Wd]), and all interactions among these factors. For the purposes of the current study, PhonFOC and LexFOC conditions were collapsed and analyzed collectively as the Focused condition.

For the analysis of vowel nasalization (A1–P0*), an additional fixed factor, Timepoint (25%, 50%, 75% of the vowel), was included as a categorial factor to assess whether prosodic effects vary over the temporal course of the vowel. The reference level for Timepoint was set to the 25% point, corresponding to the vowel portion closest to the nasal source. As one reviewer also noted, a generalized additive mixed model

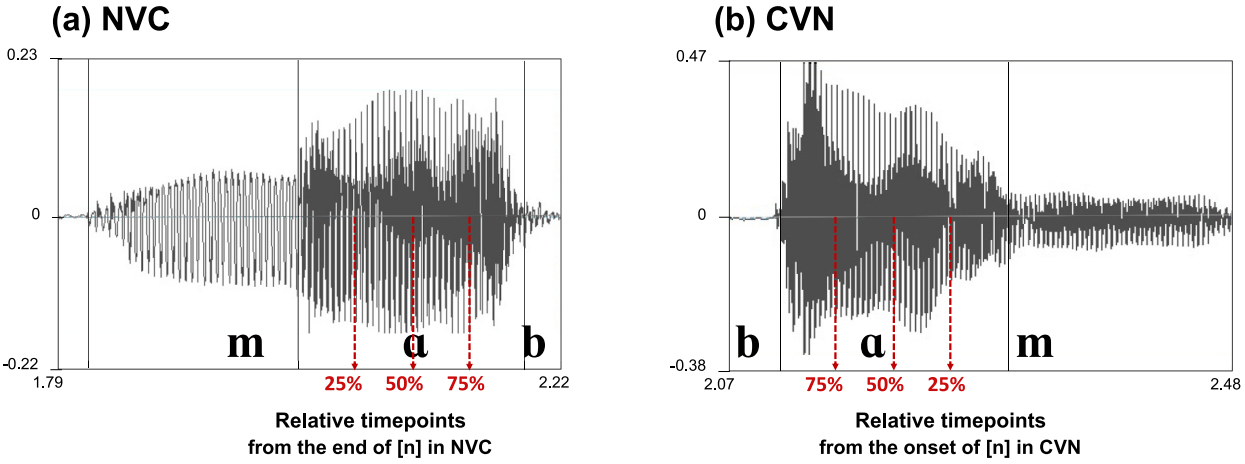

Fig. 1. Schematization of vowel nasalization measurements of A1–P0* at relative timepoints in (a) NVC and (b) CVN. Note that the order of the 25, 50, and 75% measurement points is reversed in CVN relative to NVC, such that 25% always represents the portion of the vowel closest to the nasal consonant and 75% the portion farthest from it.

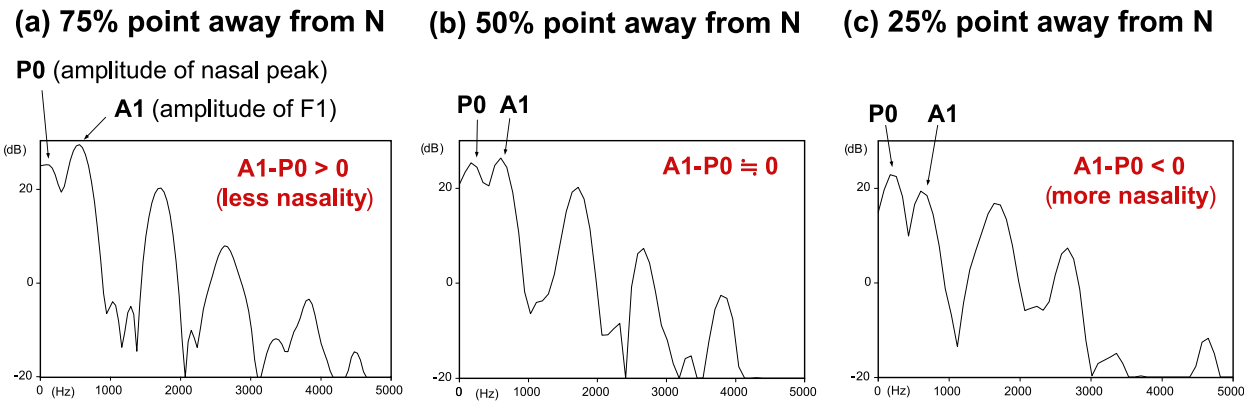

Fig. 2. Schematization of vowel nasalization measurements of A1–P0 at three different relative time points in CVN (from Jang et al., 2023). The 75% time point (furthest from the nasal source) indicates a lesser degree of V-nasalization (with a greater A1–P0) while the 25% time point (nearest from the nasal source) shows a greater degree of V-nasalization (with a smaller A1–P0).

(GAMM) would be well suited to modeling changes in vowel nasalization over a continuous time course. However, our primary interest was not in characterizing the continuous trajectory of nasalization itself, but rather in determining whether Timepoint interacted with the prosodic factors (Focus and Boundary) and/or with Dialect. Such interactions would indicate whether prosodic effects varied across the sampled timepoints within the vowel and whether these temporal patterns differed across dialects.

Binary predictors (Focus, Boundary, and Dialect) were contrast-coded using ±0.5. Under this coding scheme, regression coefficients represent averaged differences between factor levels rather than comparisons relative to a single reference category. The exception was the Timepoint factor in the vowel nasalization analyses, which contained three levels and was therefore treatment-coded with 25% as the reference level. As a result, coefficients involving the binary factors in those models are interpreted relative to the 25% Timepoint (i.e., the measured point closest to the nasal source). To provide a more transparent and reference-independent summary of statistical significance, fixed effects and interactions were evaluated using Type III analysis-of-variance tests on the fitted linear mixed-effects models. At the same time, regression coefficients from the model summaries are additionally reported in the text to facilitate interpretation of effect direction and magnitude.

The random-effects structure followed a maximal specification justified by the experimental design (Barr et al., 2013), including by-speaker and by-item random intercepts and slopes for the fixed effects, insofar as model convergence permitted. For models that failed to converge, random slopes with the smallest variance components were removed in a stepwise manner. Post-hoc pairwise comparisons were conducted with Tukey adjustment using the *emmeans* package (Lenth, 2020). The #NVC (domain-initial) and CVN# (domain-final) contexts were analyzed separately.

Because the acoustic measure of vowel nasalization (A1–P0*) may be influenced by source-related variation associated with glottalization (H1*–A3*), we analyzed vowel nasalization using two complementary modeling approaches. First, we fit a baseline model without H1*–A3*, corresponding to the approach commonly adopted in previous studies of coarticulatory vowel nasalization. Second, we fit a parallel model including H1*–A3* as a covariate to assess whether the observed prosodic and dialectal effects remained robust after controlling for phonation-related variation. This dual-model approach allowed both comparison with the existing literature and evaluation of the potential influence of source-related acoustic variation on the nasalization measure. The methodological motivation for this approach, including analyses of the relationship between A1–P0* and H1*–A3*, as well as between A1–P0* and F0, is presented in Appendix A.

3. Results

3.1. Nasal duration

3.1.1. Initial N-duration in NVC

Statistical results on initial N-duration in **N**VC contexts are summarized in Fig. 3 and Table 2 (see Table B2 in the Appendix for pairwise comparisons). The results show that there is no main effect of Dialect on N-duration ($F(1, 27.4) = 2.94$, $p > 0.09$) while Focus and Boundary reveal robust effects. As for the Focus effect, N-duration is longer in the focused than in the unfocused condition ($F(1, 2579.6) = 2884.72$, $p < 0.001$), showing a prominence-induced enhancement of the consonant's nasality. As shown in Fig. 3b, there is a Focus × Dialect interaction ($F(1, 2579.7) = 11.48$, $p < 0.001$). This interaction reflects a small but significant difference across dialects in the magnitude of focus-induced N-duration lengthening: the mean difference between the focused and unfocused conditions was 49.9 ms in AusE and 56.6 ms in AmE.[2] Thus, although focus significantly increased N-duration in both dialects, the magnitude of this increase was greater in AmE. The Boundary factor also shows a significant effect ($F(1, 2579.6) = 1127.79$, $p < 0.001$), but in a direction opposite to

[2] We acknowledge that the approximately 7-ms difference between these two effect sizes may not be perceptually salient in isolation. However, whether a difference of this magnitude is perceptually detectable in the present phonetic and prosodic context remains an empirical question for future perceptual research. Moreover, even if it is not independently perceptible, a difference that is reliably observed across speakers and conditions may still form part of the systematic phonetic patterning characteristic of the dialects. Importantly, the difference at issue is not a direct comparison between absolute N-duration values in the two dialects. Rather, it represents the difference between the magnitudes of focus-induced N-duration lengthening within the two dialects. It therefore cannot be straightforwardly interpreted with reference to a perceptual threshold for a direct durational contrast.

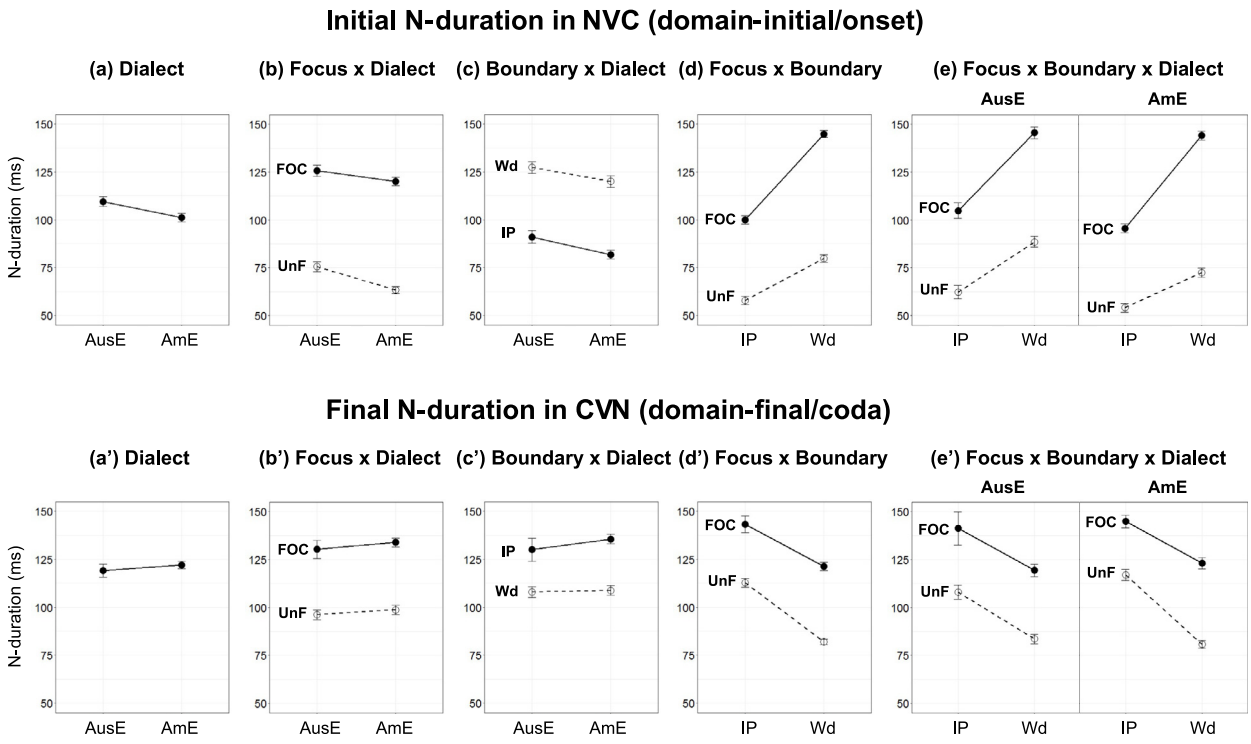

Fig. 3. N-duration in NVC at the left edge of a prosodic domain and in CVN at the right edge of a prosodic domain. Error bars refer to 95% confidence interval.

Table 2
Type III analysis-of-variance results for the linear mixed-effects model of N-duration in NVC (in domain-initial contexts) and CVN (in domain-final contexts).

	NVC (domain-initial)					CVN (domain-final)				
	NumDF	DenDF	F-value	Pr(>F)		NumDF	DenDF	F-value	Pr(>F)	
Dialect (AusE-AmE)	1	27.4	2.94	0.098		1	27.42	0.22	0.645	
Focus (FOC-UnF)	1	2579.6	2884.72	<0.001	***	1	2600.14	433.36	<0.001	***
Boundary (IP-Wd)	1	2579.6	1127.79	<0.001	***	1	2600.16	246.91	<0.001	***
Focus: Dialect	1	2579.7	11.48	<0.001	***	1	2600.14	0.01	0.916	
Boundary: Dialect	1	2579.6	0.00	0.964		1	2600.15	2.68	0.102	
Focus: Boundary	1	2579.8	132.72	<0.001	***	1	2600.13	6.47	0.011	*
Focus: Boundary: Dialect	1	2579.9	14.02	<0.001	***	1	2600.14	2.21	0.137	

the focus effect. That is, as can be seen in Fig. 3c, N-duration is shorter in the IP-initial than in the Wd-initial position, which is consistent with a general domain-initial strengthening effect that causes a shortening of the consonant's nasal murmur. There is no Boundary × Dialect interaction (F(1, 2579.6) = 0.00, p > 0.09), indicating that the boundary-related domain-initial shortening effect is comparable across the dialects (Fig. 3c).

The results further show a significant interaction between the two prosodic factors, Focus and Boundary (F(1, 2579.8) = 132.72, p < 0.001), such that, as shown in Fig. 3d, the focus-induced expansion in nasal duration is larger in the Wd-initial condition—that is, when IP-initial shortening does not constrain the focus-related expansion (FOC−UNF mean diff., IP-initial, 41.9 ms; Wd-initial, 64.7 ms). By contrast, the IP-initial shortening is larger under focus than under unfocused conditions—that is, the effect is greatest when focus-related lengthening yields longer nasal durations, leaving more room for shortening than in the unfocused condition, where nasal durations are already shorter.

Such a Focus × Boundary interaction is observed in both dialects, as shown in Fig. 3e. However, the analysis reveals a significant three-way interaction among Focus, Boundary, and Dialect (F(1, 2579.9) = 14.02, p < 0.001), which reflects a larger focus-induced difference in nasal duration in AmE than in AusE in the Wd-initial condition (FOC−UNF mean difference: AusE, 57.6 ms; AmE, 71.8 ms).

3.1.2. Final N-duration in CV*N*

Statistical results on N-duration in CV**N** are summarized in Fig. 3, and Table 2 (see Table B2 in the Appendix for pairwise comparisons). There is no significant effect of Dialect on final N-duration as shown in Fig. 3a' (F(1, 27.42) = 0.22, p > 0.09). But as was the case with initial N-duration, there is a significant effect of Focus (*F(1, 2600.14) = 433.36, p < 0.001*), such that N-duration in CVN is longer in the focused than in the unfocused condition, showing a prominence-induced enhancement of the consonant's nasality (Fig. 3b'). There is no Focus × Dialect interaction

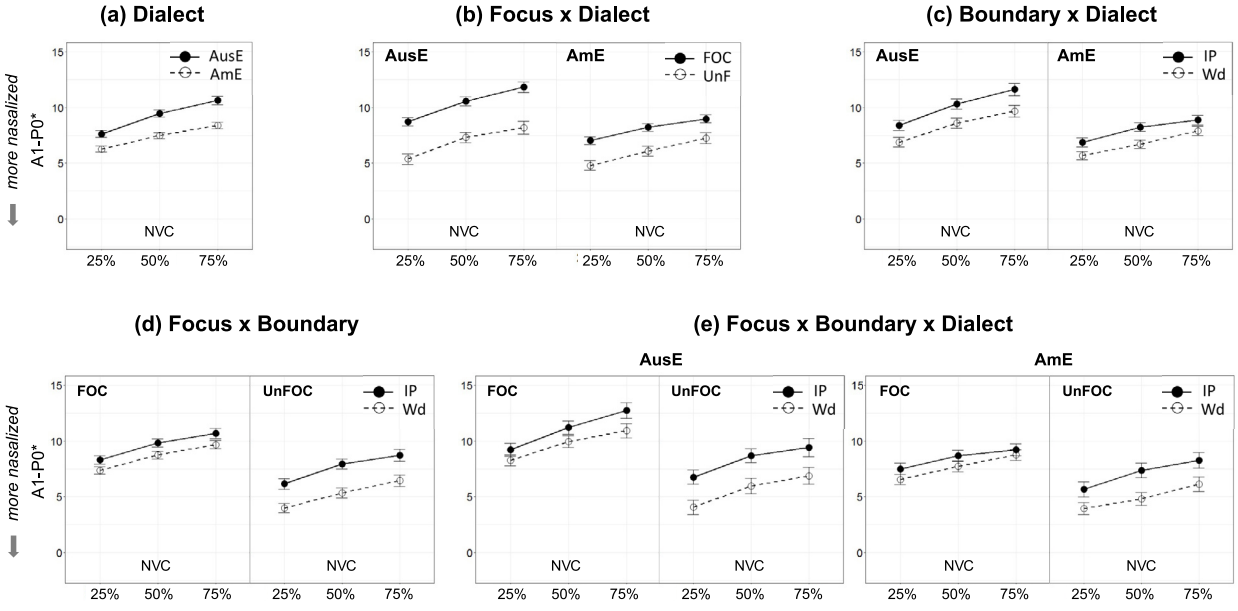

Fig. 4. Effects of Dialect, Focus and Boundary and their interactions on the carryover V-nasalization (A1–P0*) in NVC across three relative timepoints in the vowel in domain-initial contexts. Error bars refer to 95% confidence intervals.

(F(1, 2600.14) = 0.01, p > 0.09), indicating that the focus effect is comparable across the dialects. There is also a significant effect of Boundary (F(1, 2600.16) = 246.91, p < 0.001); but as shown in Fig. 3c', this boundary-related effect on final N-duration reveals a pattern opposite the case with the initial N-duration. Unlike the domain-initial shortening of N-duration at the IP boundary, the N-duration is longer in the IP-final position (at a larger boundary) than in the Wd-final position (at a smaller boundary), showing a general phrase-final (pre-boundary) lengthening effect. There is no Boundary × Dialect interaction (F(1, 2600.15) = 2.68, p > 0.09) (Fig. 3c'), indicating that the phrase-final lengthening effect is comparable across the dialects.

As was the case with the initial N-duration, the final N-duration shows a significant interaction between the two prosodic factors, Focus and Boundary (F(1, 2600.13) = 6.47, p < 0.05). As shown in Fig. 3d′, the interaction can be understood from two complementary perspectives. With respect to focus, the focus-induced expansion in final N-duration is larger in the Wd-final position than in the IP-final position (mean difference: IP-final, 30.8 ms; Wd-final, 39.4 ms). This indicates that, when IP-final lengthening is absent, there is more room for focus-induced expansion. From the boundary-related perspective, the boundary-induced difference is larger in the unfocused than in the focused condition (mean difference: Focused, 22.2 ms; Unfocused, 30.8 ms). There is no three-way interaction of Focus × Boundary × Dialect (F(1, 2600.14) = 2.21, p > 0.09).

3.2. Vowel nasalization

3.2.1. V-nasalization in NVC (carryover effect) at the left edge of a prosodic domain

Statistical results on V-nasalization (smaller A1–P0* values indicate greater nasalization) in NVC are summarized in Fig. 4 and Table 3 (see Table B3 in the Appendix for pairwise comparisons). There is no main effect of Dialect with both baseline and covariate-adjusted analysis (F(1, 27.2) = 1.94, p > 0.09; F(1, 27.7) = 0.23, p > 0.09, respectively), indicating that the two dialects do not differ uniformly in overall vowel nasalization. However, a significant Dialect × Timepoint interaction is observed in the baseline analysis (F(2, 6628.4) = 4.39, p < 0.05), but not in the covariate-adjusted analysis (F(2, 6624.4) = 1.79, p > 0.09). As shown in Fig. 4a, this interaction reflects an increasing dialectal divergence toward the later portions of the vowel (50% and 75%) relative to the earlier reference timepoint (25%): specifically, vowels in the NVC context tend to exhibit reduced nasalization in AusE compared to AmE as the vowel unfolds over time. Because the corresponding Dialect × Timepoint interaction does not reach significance in the covariate-adjusted analysis, however, this pattern cannot be interpreted as independent of variation in H1*–A3* (an index of glottalization) and should therefore be regarded as suggestive rather than conclusive. Notably, this temporal divergence becomes more transparent when considered in relation to its further interaction with Focus, as reported below.

There is a significant main effect of Focus in both the baseline and the covariate-adjusted analyses (F(1, 6628.1) = 835.57, p < 0.001; F(1, 6368.1) = 465.72, p < 0.001, respectively), such that vowels are less nasalized (i.e., exhibit larger A1–P0* values) in the focused than in the unfocused condition. This pattern reflects increased coarticulatory resistance under focus: although the onset nasal consonant shows focus-induced enhancement in the form of longer nasal duration, its coarticulatory influence on the following vowel is attenuated when the vowel is in focus. As shown in Fig. 4b, the effect pervades the vowel across the timepoints with no substantial attenuation of the focus-induced difference as the timepoint gets more distal from the nasal source, indicating that the coarticulatory resistance effect is global rather than being localized near the source of nasality.

Table 3
Type III analysis-of-variance results for the linear mixed-effects model of V-nasalization (A1–P0*) in NVC (in domain-initial contexts).

	Without H1*-A3* as covariate					With H1*-A3* as covariate				
	NumDF	DenDF	F-value	Pr(>F)		NumDF	DenDF	F-value	Pr(>F)	
Dialect (AusE-AmE)	1	27.2	1.94	0.175		1	27.7	0.23	0.638	
Focus (FOC-UnF)	1	6628.1	835.57	<0.001	***	1	6368.1	465.72	<0.001	***
Boundary (IP-Wd)	1	6627.3	291.37	<0.001	***	1	6583.2	142.78	<0.001	***
Timepoint	2	6629.1	210.88	<0.001	***	2	6598.3	145.47	<0.001	***
H1*-A3*						1	1295.9	736.33	<0.001	***
Focus: Dialect	1	6628.1	60.00	<0.001	***	1	6626.7	15.66	<0.001	***
Boundary: Dialect	1	6627.3	4.85	0.028	*	1	6625.3	0.88	0.349	
Focus: Boundary	1	6627.4	35.27	<0.001	***	1	6625.7	25.36	<0.001	***
Dialect: Timepoint	2	6628.4	4.39	0.012	*	2	6626.4	1.79	0.166	
Boundary: Timepoint	2	6627.1	1.00	0.366		2	6624.7	1.07	0.344	
Focus: Timepoint	2	6627.2	0.02	0.978		2	6625	0.08	0.919	
Focus: Boundary: Dialect	1	6627.4	0.02	0.892		1	6625	0.01	0.924	
Boundary: Dialect: Timepoint	2	6627.1	0.59	0.554		2	6624.5	0.77	0.463	
Focus: Dialect: Timepoint	2	6627.2	1.71	0.180		2	6625.1	2.44	0.087	
Focus: Boundary: Timepoint	2	6627	0.03	0.967		2	6624.6	0.15	0.861	
Focus: Boundary: Dialect: Timepoint	2	6627	1.23	0.292		2	6624.4	1.17	0.310	

Crucially, there is also a significant Focus × Dialect interaction in both the baseline and the covariate-adjusted analyses ($F(1, 6628.1) = 60.0$, $p < 0.001$; $F(1, 6626.7) = 15.66$, $p < 0.001$). As evident in Fig. 4b, this interaction reflects a larger focus-induced reduction in vowel nasalization in AusE than in AmE (FOC − UNF mean difference, pooled across timepoints: AusE = 3.6; AmE = 2.1). Neither the baseline nor the covariate-adjusted analysis revealed a significant Focus × Dialect × Timepoint interaction (baseline: $F(2, 6627.2) = 1.71$, $p > 0.09$; covariate-adjusted: $F(2, 6625.1) = 2.44$, $p = 0.087$), indicating that the dialectal differentiation in focus-related coarticulatory resistance is consistent across the vowel.

The Boundary factor also reveals a significant effect in both the baseline and the covariate-adjusted analyses ($F(1, 6627.3) = 291.37$, $p < 0.001$; $F(1, 6583.2) = 142.78$, $p < 0.001$, respectively), such that, as shown in Fig. 4c, vowels are nasalized *less* in the IP-initial than in the Wd-initial position. Notice, however, that this boundary-induced reduction of coarticulation differs in nature from the focus-induced coarticulatory reduction. Readers are reminded that this boundary-related reduction of vowel nasalization co-occurs with phonetic reduction of the nasal consonant itself (i.e., shorter N-duration IP-initially), unlike focus-related reduction occurring with enhanced nasal source (Section 3.1.1).

With respect to the Boundary × Dialect interaction, although covariate-adjusted analysis yielded no significant interaction ($F(1, 6625.3) = 0.88$, $p > 0.09$), the baseline analysis revealed a significant effect ($F(1, 6627.3) = 4.85$, $p < 0.05$). While this divergence should be interpreted cautiously as the corresponding interaction does not reach significance in the covariate-adjusted analysis, this result indicates that coarticulatory reduction associated with prosodic boundary structure (i.e., reduced vowel nasalization IP-initially) may be more robust in AusE than in AmE. The absence of a further interaction with Timepoint (i.e., no significant Boundary × Dialect × Timepoint interaction) suggests that this effect is not confined to regions proximal to the nasal source (the onset nasal), but instead extends across the vowel as a whole.

Finally, results show a significant Focus × Boundary interaction in both the baseline and the covariate-adjusted analyses ($F(1, 6627.4) = 35.27$, $p < 0.001$; $F(1, 6625.7) = 25.36$, $p < 0.001$, respectively). As can be seen in Fig. 4d, the interaction is in part due to the boundary-induced difference being augmented in the absence of the focus effect (in the unfocused context), and in part due to the focus-induced difference being augmented in the absence of the boundary effect (in the
Wd-initial position). Under focus, vowels appear to exhibit heightened resistance to coarticulatory nasalization, such that the additional reduction otherwise associated with the Boundary factor is attenuated, showing a ceiling-like constraint on the cumulative effects of multiple prosodic factors; consequently, the boundary effect emerges more strongly in the unfocused condition. These interaction patterns are comparable across dialects, as neither the baseline nor the covariate-adjusted analysis reveals a significant three-way interaction among Focus, Boundary, and Dialect ($F(1, 6627.4) = 0.02$, $p > 0.09$; $F(1, 6625) = 0.01$, $p > 0.09$, respectively). This pattern is further illustrated in Fig. 4e, which plots the Focus × Boundary interaction separately for each dialect and shows that the boundary effect is minimized under focus (left panel for each dialect) but augmented in its absence (right panel for each dialect).

3.2.2. V-nasalization in CVN (anticipatory effect) at the right edge of a prosodic domain

Statistical results for vowel nasalization in the CVN context (with smaller A1–P0 values indicating greater nasalization) are summarized in Fig. 5 and Table 4 (see Table B3 in the Appendix for pairwise comparisons). There is no main effect of Dialect in either the baseline or the covariate-adjusted analysis ($F(1, 27.1) = 0.09$, $p > 0.09$; $F(1, 11.5) = 0.01$, $p > 0.09$, respectively), indicating the absence of a uniform cross-dialectal difference in the overall magnitude of anticipatory vowel nasalization. However, Dialect interacts significantly with Timepoint in both the baseline and the covariate-adjusted analyses ($F(2, 7275.1) = 5.74$, $p < 0.01$; $F(2, 7271) = 5.85$, $p < 0.01$, respectively). These interactions indicate that the AusE–AmE difference varies systematically over the temporal course of the vowel. Specifically, the two dialects diverge in the temporal patterning of anticipatory vowel nasalization, with AusE exhibiting reduced nasalization relative to AmE toward the earlier portions of the vowel—that is, at positions farther from the nasal source in the CVN sequence. Consistent with this trajectory-based interpretation, Fig. 5a illustrates a progressive

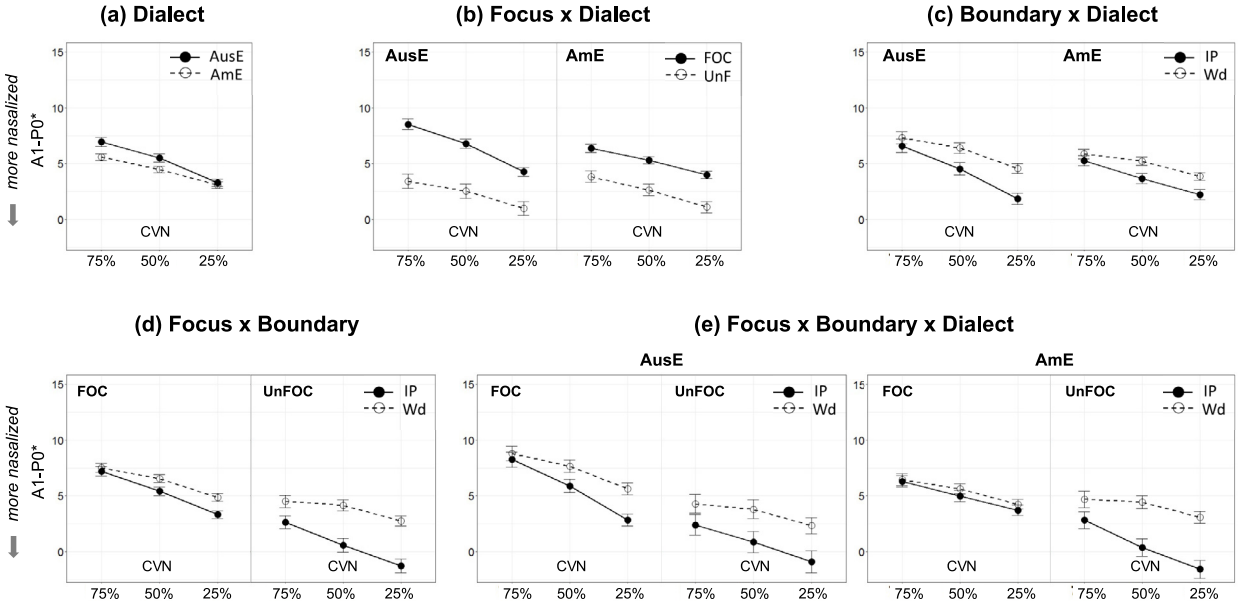

Fig. 5. Effects of Dialect, Focus and Boundary and their interactions on the anticipatory V-nasalization (A1–P0*) in CVN across three relative timepoints in the vowel in domain-final contexts. Error bars refer to 95% confidence intervals.

Table 4
Type III analysis-of-variance results for the linear mixed-effects model of V-nasalization (A1–P0*) in NVC (in domain-initial contexts).

	Without H1*-A3* as covariate					With H1*-A3* as covariate				
	NumDF	DenDF	F-value	Pr(>F)		NumDF	DenDF	F-value	Pr(>F)	
Dialect (AusE-AmE)	1	27.1	0.09	0.772		1	11.5	0.01	0.911	
Focus (FOC-UnF)	1	7275.7	1045.63	<0.001	***	1	7282.8	569.37	<0.001	***
Boundary (IP-Wd)	1	7275.6	288.33	<0.001	***	1	7272	273.94	<0.001	***
Timepoint	2	7275.1	262.82	<0.001	***	2	7271.2	316.40	<0.001	***
H1*-A3*						1	7203.1	858.64	<0.001	***
Focus: Dialect	1	7275.7	46.22	<0.001	***	1	7272.2	29.66	<0.001	***
Boundary: Dialect	1	7275.6	0.33	0.565		1	7272	1.71	0.191	
Focus: Boundary	1	7275.4	56.47	<0.001	***	1	7271.7	44.69	<0.001	***
Dialect: Timepoint	2	7275.1	5.74	0.003	**	2	7271	5.85	0.003	**
Boundary: Timepoint	2	7275	17.86	<0.001	***	2	7271.2	11.21	<0.001	***
Focus: Timepoint	2	7275	2.68	0.069	.	2	7271	2.70	0.067	.
Focus: Boundary: Dialect	1	7275.4	29.07	<0.001	***	1	7271.9	20.90	<0.001	***
Boundary: Dialect: Timepoint	2	7275	0.72	0.489		2	7271	1.15	0.317	
Focus: Dialect: Timepoint	2	7275	8.21	<0.001	***	2	7271	7.04	0.001	***
Focus: Boundary: Timepoint	2	7275	1.19	0.303		2	7271	1.51	0.221	
Focus: Boundary: Dialect: Timepoint	2	7275	4.14	0.016	*	2	7271	2.12	0.121	

increase in dialectal separation from 25% to 75%, with the effect most robust at 75%, where both the baseline and the covariate-adjusted analyses show significant Dialect × Timepoint interactions. (Note that further nuanced time-dependent dialectal differences emerge when this pattern is considered in interaction with Focus, as discussed below).

In the CVN context, as in the NVC context, there is a significant main effect of Focus in both the baseline and the covariate-adjusted analyses ($F(1, 7275.1) = 1045.63$, $p < 0.001$; $F(1, 7282.8) = 569.37$, $p < 0.001$, respectively), indicating that vowels are *less* nasalized in the focused than in the unfocused condition—i.e., focus-induced coarticulatory resistance. This focus effect differ across dialects dynamically, as indicated by the significant Focus × Dialect interaction ($F(2, 7275.7) = 46.22$, $p < 0.001$; $F(1, 7272.2) = 29.66$, $p < 0.001$, respectively), as well as Focus × Dialect × Timepoint interaction in both the baseline and the covariate-adjusted analyses ($F(2, 7275) = 8.21$, $p < 0.001$; $F(1, 7271) = 7.04$, $p < 0.001$, respectively). Inspection of the time-course patterns in Fig. 5b suggests that this interaction reflects dialectal differences in the temporal distribution of the focus effect across the vowel. In AusE (left panel), the focus-induced reduction in vowel nasalization relative to the unfocused condition (FOC–UNF difference) increases toward earlier portions of the vowel, farther from the upcoming nasal consonant. In contrast, in AmE (right panel), the magnitude of the focus effect is reduced at these earlier timepoints. Thus, while focus induces suppression of anticipatory nasalization—reflecting focus-related coarticulatory resistance in both dialects—the temporal implementation of this suppression differs across dialects. Specifically, AusE shows stronger focus-related resistance at vowel portions that are temporally more distant from the nasal source, with the effect culminating at 75% and evident in both the baseline and the covariate-adjusted analyses.

The Boundary factor also yields a significant main effect in both the baseline and the covariate-adjusted analyses ($F(1, 7275.6) = 288.33$, $p < 0.001$; $F(1, 7272) = 273.94$, $p < 0.001$, respectively). As shown in Fig. 5c, however, the direction of the boundary effect in the anticipatory CVN context is opposite to that observed in the carryover NVC context. Recall that in NVC contexts, boundary-related domain-initial coarticulatory reduction was observed, consistent with reduced nasality at the source due to initial shortening of the nasal consonant. In contrast, the boundary effect in CVN shows a coarticulatory propensity rather than reduction—i.e., vowels are nasalized *more* in CVN in the IP-final than in the Wd-final position. A significant Focus × Boundary interaction is observed both in the baseline and the covariate-adjusted analyses ($F(1, 727541) = 56.47$, $p < 0.001$; $F(1, 7271.7) = 44.69$, $p < 0.001$, respectively). The interaction is primarily driven by the boundary-induced modulation of coarticulation being more clearly expressed in the unfocused condition, in the absence of focus-induced effects, as shown in Fig. 5d. This interaction pattern parallels that observed in the carryover NVC context.

With respect to dialectal differentiation in boundary effects, a significant three-way interaction involving Focus, Boundary, and Dialect emerge in both the baseline and the covariate-adjusted analyses ($F(1, 7275.4) = 29.07$, $p < 0.001$; $F(1, 7271.9) = 20.9$, $p < 0.001$, respectively). This three-way interaction reveals important dialectal differences in how boundary-related effects are modulated by focus. As shown in Fig. 5e, this interaction stems from differences between the two dialects in the manner in which Boundary and Focus interact. In AusE (left panels of Fig. 5e), a boundary-induced reduction of vowel nasalization is evident in both focused and unfocused conditions, although this boundary effect is substantially stronger in the unfocused condition, consistent with a Focus × Boundary interaction whereby focus-related coarticulatory resistance attenuates the boundary effect. In contrast, as can be seen in the right panels of Fig. 5e, AmE shows the significant boundary effect only in the unfocused condition. This pattern indicates that in AmE the focus-induced coarticulatory resistance substantially limits the expression of boundary-induced reduction. The interaction thus suggests a stronger attenuation of the boundary effect under focus in AmE than in AusE. In AmE, boundary-related reduction is minimal under focus, whereas in AusE a reduced, but still observable, boundary effect remains present even in the focused condition.

Viewed from a complementary perspective, the three-way interaction can also be understood in terms of a more robust focus effect in AusE than in AmE, particularly in the Wd context. Because Fig. 5e is plotted to highlight boundary-related differences, the focus effect itself is not immediately transparent. However, the Focus × Dialect interaction shown in Fig. 5b makes clear that the focus-induced difference is larger in AusE than in AmE. The three-way interaction further refines this pattern by indicating that the cross-dialectal difference in the focus effect (larger in AusE) is especially pronounced in the Wd-final position. As shown in Fig. 5d (compare the dotted lines contrasting focused and unfocused conditions within the Wd context across dialects), the separation between the focused and unfocused conditions is substantially greater in AusE than in AmE (FOC–UNF mean difference pooled across timepoints in Wd: AusE, 3.9; AmE, 1.3), thereby contributing to the observed three-way interaction.

Finally, the three-way interaction is further refined by its temporal evolution across the vowel, as evidenced by a significant four-way interaction among Focus, Boundary, Dialect, and Timepoint in the baseline analysis ($F(1, 7275) = 4.14$, $p < 0.05$). As can be inferred from Fig. 5e, this interaction arises in part from the fact that in AmE the boundary effect is effectively absent in the focused condition. Moreover, when a boundary-related difference does emerge in AmE in the unfocused condition, it unfolds continuously over the course of the vowel: it is most evident near the nasal source (at the 25% reference point) and progressively attenuates toward later portions of the vowel, disappearing by 75%. A comparable gradual modulation of the boundary effect over the vowel is also observed in AusE. In AusE, the boundary effect is likewise strongest near the nasal source at 25% and diminishes toward 75%, but in both focused and unfocused contexts. These dialect-specific, trajectory-based patterns of boundary effects—and their modulation by focus—collectively contribute to the observed four-way interaction.

4. General discussion

The present study demonstrates that Australian English (AusE) operates under a robust system of prosodic conditioning, in which both nasal consonant duration and coarticulatory vowel nasalization are systematically shaped by prosodic structure. Specifically, AusE speakers actively modulate nasality in response to Focus (focus-driven pitch accent) and Boundary conditions, employing these factors to enhance contrasts in ways that closely parallel the patterns previously documented for American English (AmE) in Cho et al. (2017), which used identical experimental protocols and closely matched speech materials. Crucially, however, this structural parallelism does not entail phonetic identity. While AusE shares a common prosodic system with AmE—consistent with their status as dialects of the same language—it exhibits a distinct phonetic implementation of these constraints. The divergence lies not in whether certain prosodic effects are present, but in how they are realized: specifically, AusE differs in the magnitude of coarticulatory modulation, its temporal deployment across the vowel, and the interactional dynamics between competing prosodic factors. In what follows, we discuss how Australian English—despite being rooted in a shared prosodic system with American English—exhibits both robust cross-dialectal similarity and systematic dialectal variation in the temporal and coarticulatory control of nasality.

Before turning to the shared and divergent prosodic patterns observed across dialects, it is useful to consider the present findings in relation to previous work on the temporal organization of nasal coarticulation. Beddor (2009) proposed that variation in the extent of vowel nasalization in American English may arise primarily from differences in the temporal coordination of a relatively stable velum-lowering gesture with adjacent oral gestures, rather than from changes in the gesture itself. Under this account, different degrees of coarticulatory vowel nasalization result from shifts in gestural timing that alter the extent of overlap between the vowel and the nasal gesture.

Cho et al. (2017) also considered this account in relation to prosodically conditioned nasalization. Given that prominence-induced lengthening of a coda nasal was accompanied by shorter anticipatory vowel nasalization, they suggested that this pattern was compatible with the relatively invariant velum-lowering gesture proposed by Beddor (2009), as increased nasal-consonant duration would reduce the gesture's overlap with the preceding vowel. The boundary effect, however, showed a different pattern: phrase-finally, both nasal-consonant duration and anticipatory vowel nasalization increased. Thus, although the invariant-gesture account is compatible with the prominence effect, it does not straightforwardly account for boundary-induced expansion in both consonantal and vocalic nasalization, suggesting that prominence and boundary may reorganize nasal coarticulation in different ways.

The present study, however, cannot directly determine whether the patterns observed here arise through these mechanisms because our acoustic measures quantify the degree of nasalization at predefined locations within the vowel rather than identifying the onset and offset of nasalization or estimating the duration of the underlying velum-lowering gesture. Consequently, the observed effects of focus, prosodic boundary, and dialect may reflect changes in gestural timing, changes in the duration or magnitude of the velum-lowering gesture, or some combination of these factors. The present findings should therefore not be taken as either confirming or disconfirming the relatively invariant gesture account. Rather, together with the prominence–boundary asymmetry previously reported by Cho et al. (2017), they underscore the need to determine whether different dimensions of prosodic structure reorganize velic–oral coordination through different articulatory mechanisms. Addressing this question will require future studies combining acoustic measures with direct articulatory or aerodynamic evidence.

4.1. Shared prosodic structuring of nasality in Australian English

The present results show that Australian English (AusE) exhibits a prosodic organization of nasal consonants and vowel nasalization that closely parallels American English (AmE). As in AmE, N-duration in AusE is robustly shaped by prosodic factors in both carryover (NVC) and anticipatory (CVN) contexts. Focus-induced prominence lengthens N-duration, reflecting enhancement of the nasal consonant under prominence in both contexts. Boundary effects on N-duration are parallel to those in AmE as well, with domain-initial shortening in NVC contexts and domain-final lengthening in CVN contexts. AusE also shows the same Focus × Boundary interaction in N-duration, such that the effect of one prosodic factor is strongest when the influence of the other is reduced. One exception to this complementary pattern—again paralleling AmE—occurs in domain-initial NVC sequences, where IP-initial nasal shortening is greater under focus, plausibly because focus-induced lengthening provides additional temporal room for boundary-related reduction.

A comparable cross-dialectal similarity is observed for coarticulatory vowel nasalization. As in AmE, focus in AusE induces coarticulatory resistance, yielding reduced vowel nasalization. Crucially, although the nasal consonant is enhanced under prominence, its coarticulatory influence on the vowel is not; this dissociation runs counter to accounts where coarticulation scales directly with the strength of the triggering segment (cf. Farnetani & Recasens, 1999; Fowler, 1981; Fowler & Saltzman, 1993). Following Cho et al. (2017), this pattern is best interpreted as *active* prominence-driven coarticulatory resistance: the focused vowel resists coarticulatory encroachment even as the nasal source strengthens (see also Jang et al., 2018; Zellou & Scarborough, 2012). Since focused vowels typically undergo hyperarticulation (e.g., de Jong, 1995), the observed reduction in nasalization reflects a targeted enhancement of the vowel's oral specification—an account that extends naturally from AmE to AusE.

Boundary effects on vowel nasalization in AusE closely parallel AmE, exhibiting a prosodically asymmetric pattern. At the domain's left edge (NVC), vowels show reduced nasalization, consistent with diminished input from shortened nasal duration. As discussed for AmE in Cho et al. (2017), this aligns with domain-initial strengthening, whereby enhanced oral articulatory force—plausibly involving increased velic elevation—shortens the nasal murmur and limits its coarticulatory influence. Domain-initial strengthening of nasal duration is well attested cross-linguistically (e.g., Cho, 2016; Keating et al., 2003), and parallel reductions in vowel nasalization have been reported for Korean (Jang et al., 2018, 2026) and Mandarin Chinese (Li et al., 2020). Thus, AusE conforms to a broader cross-linguistic pattern of domain-initial modulation. In contrast, at the domain's right edge (CVN), vowels exhibit increased nasalization. Here, phrase-final nasal lengthening—combined with reduced articulatory force—facilitates a more sustained, less constrained velum-lowering gesture. As suggested for AmE in Cho et al. (2017), this combined effect yields greater IP-final vowel nasalization, reflecting a boundary-related coarticulatory vulnerability rather than reduction. Again, this confirms that this phrase-final pattern observed in AmE is likewise present in AusE.

Taken together, these results demonstrate that AusE and AmE share a common prosodic system for conditioning nasality, rooted in a broadly similar phonological structure. In both dialects, the same higher-order prosodic factors—focus-driven prominence and prosodic boundary—systematically regulate nasal consonant duration and coarticulatory vowel nasalization through parallel patterns of prominence-induced enhancement, domain-initial strengthening, and phrase-final modulation. At the same time, this convergence should be interpreted within the empirical scope of the present comparison. Although it establishes that prosodic conditioning of nasality is shared by the two varieties examined here, it does not establish a uniform or stable property of English as a whole, particularly given previous evidence for variation in the realization of nasalization even within AmE (Tamminga & Zellou, 2015). Determining the extent to which this organization generalizes across varieties of English therefore requires further cross-variety investigation.

4.1.1. Phrase-final nasal coarticulation: general tendencies and language-specific control

A critical question emerging from these cross-dialectal similarities is whether the phrase-final coarticulatory patterns shared by AusE and AmE reflect a broadly general phonetic

tendency or a language-specific organizational configuration. In both dialects, phrase-final nasal lengthening co-occurs with increased vowel nasalization—a pattern plausibly attributable to general phonetic mechanisms, including the temporal expansion of velum lowering and phrase-final articulatory weakening. However, such biomechanical tendencies are not deterministic; rather, they serve as a baseline subject to extensive language-specific fine-tuning (cf. Cho, 2025). Indeed, subsequent cross-linguistic work demonstrates that the phrase-final coupling between source lengthening and increased coarticulation is far from universal. For instance, Seoul Korean exhibits increased phrase-final vowel nasalization without concomitant lengthening of the nasal consonant (Jang et al., 2018), whereas Mandarin Chinese displays the reverse pattern: the nasal consonant lengthens phrase-finally, yet this occurs without a corresponding increase in vowel nasalization (Li et al., 2020). These cross-linguistic dissociations indicate that the specific co-occurrence of nasal lengthening and enhanced vowel nasalization observed in English is not a mechanical inevitability.

These cross-linguistic disparities in the CVN context reflect differences in the temporal coordination of velic and oral gestures across languages, with such coordination shaped by language-specific phonological requirements where relevant. In English, vowel nasalization has an extensive temporal scope—often analyzed as a phonologized allophonic target under active speaker control (e.g., Cho, 2025; Cunha et al., 2024; Pouplier et al., 2024; Solé, 2007)—which allows phrase-final lengthening to be expressed in both the vowel and the nasal consonant. Australian English appears to share this broad "English pattern," though to a lesser degree in the temporal deployment of nasality (see Section 4.2 for discussion of dialect-specific fine-tuning and further implications from Cunha et al. (2024), which compare American and Southern British Englishes).

In contrast, Korean and Mandarin diverge from English through distinct intergestural timing strategies. Korean employs an "early alignment" configuration in which the velum-lowering gesture shifts leftward to overlap substantially with the preceding vowel, yielding increased phrase-final vowel nasalization without concomitant lengthening of the nasal source; here, phrase-final temporal expansion primarily affects vowel overlap rather than the consonantal portion of the gesture (Jang et al., 2018). By contrast, Mandarin as reported in Li et al. (2020), exhibits a "later alignment" strategy coupled with functional stabilization, plausibly motivated by the need to preserve perceptual recoverability in the context of ongoing nasal coda lenition. Because weakening oral closure threatens the reliability of coda nasality, speakers maintain stable vowel nasalization rather than allowing phrase-final loosening of velic control. Li et al. (2020) suggest that Mandarin speakers—like French speakers—thus actively constrain boundary-related increases in vowel nasalization, ensuring stability across prosodic conditions rather than the phrase-final enhancement observed in English.

These cross-linguistic dissociations thus establish that phrase-final nasal coarticulation is organized at the level of the phonetic grammar of individual languages, even though it may have emerged through different pathways grounded in general phrase-final phonetic mechanisms. This, in turn, motivates the question of how such timing relations may be further fine-tuned across dialects within a single language.

4.2. Beyond what is shared by Australian and American Englishes

Although AusE patterns closely with AmE in how Focus and Boundary condition nasal consonant duration and vowel nasalization, the direct comparison—made possible by using the same experimental protocols and speech materials—shows that this parallelism does not amount to phonetic identity. Crucially, these differences do not take the form of uniform magnitude or temporal shifts; rather, they emerge as gradient, time-dependent modulations over the course of the vowel, reflected primarily in the coarticulatory vowel nasalization.

For nasal duration, dialectal differentiation is limited and tightly constrained by prosodic context. The only statistical difference concerns the magnitude of the focus effect in NVC sequences. In Word-initial (phrase-medial) position—where domain-initial shortening does not apply—focus-induced nasal lengthening is smaller in AusE than in AmE, with no comparable divergence in CVN contexts. This attenuated lengthening in AusE appears to cohere with its stronger focus-induced coarticulatory resistance (i.e., reduced vowel nasalization) to be discussed further below. Apart from this specific prosodic configuration, AusE closely aligns with the AmE pattern in terms of variation in consonantal nasal duration.

The dialectal differentiation becomes particularly evident in vowel nasalization when considering how AusE deploys focus-related coarticulatory resistance under different boundary conditions. In NVC contexts, AusE shows a stronger reduction of vowel nasalization under focus than AmE, indicating a more sustained implementation of focus-related resistance. Given that nasal duration was also shorter in AusE in this context, this pattern suggests a possibly more active suppression of the velic gesture under focus, reflected in both the nasal consonant and the adjacent vowel.

In CVN contexts, AusE diverges most clearly from AmE in how focus and boundary interact. While AmE largely suppresses boundary-related modulation of vowel nasalization in the focused condition (showing a focus-related ceiling effect), AusE continues to show a boundary effect even in the focused condition, albeit in an attenuated form. This contrast is especially pronounced in the Word-final position, where the focus-induced separation in vowel nasalization between focused and unfocused vowels is substantially larger in AusE than in AmE. Moreover, boundary-related effects in AusE exhibit a robust pattern over the time course of the vowel, remaining detectable from the most proximal (25%) to the most distal (75%) portions. In AmE, however, these effects attenuate over time, becoming least evident at the distal portion.

As such, these results extend prior work on dialectal variation in English vowel nasalization—reviewed at the outset of the paper (e.g., Bongiovanni, 2021; Tamminga & Zellou, 2015)—by showing that cross-dialectal differences are not limited to overall magnitude differences. Instead, the present findings point to more fine-grained, temporally structured sources of dialectal differentiation. In line with this pattern, Cunha et al. (2024) provide a more phonetically nuanced account. As noted earlier, their real-time MRI comparisons of American and Southern British English demonstrate that greater vowel

nasalization in American English does not arise from earlier or more extensive velic lowering. Instead, dialectal differences emerge from reduction and temporal reorganization of the oral coda (tongue-tip) gesture, shifting the internal alignment between the vowel and the nasal consonant. As a consequence, acoustic nasalization increases in AmE while the velum gesture itself remains largely stable. This finding is theoretically significant as it locates dialectal variation at the level of intergestural timing and coordination, rather than in categorical articulatory targets or phonological representations, implying that such differences are governed by system-internal phonetic grammar. However, because Cunha et al.'s study examined production under a single narrow-focus prosodic condition, it captures only one specific configuration in which such timing differences arise. It therefore remains essential to investigate how these coordination patterns are further conditioned by prosodic structure.

While limited to acoustic evidence—and thus unable to directly capture the intergestural timing differences observed by Cunha et al. (2024)—the present study offers a distinct perspective. Specifically, it demonstrates that even without positing dialect-specific phonological representations, systematic differences can emerge in how prosodic structure modulates coarticulatory control. The Australian–American contrast is thus not reducible to static differences in nasalization degree; rather, it reflects differences in the temporal deployment and prosodic conditioning of coarticulation. From this perspective, nasalization appears more deeply entrenched in AmE, potentially encoded as such in the dialect's phonetic grammar. Most notably, even under prominence—where coarticulatory resistance is expected to suppress nasality—AmE exhibits weaker suppression, resulting in greater residual vowel nasalization.

One possible alternative interpretation of the present dialectal differences, however, is that coarticulatory vowel nasalization may occupy somewhat different representational statuses in the two dialects. In particular, the greater resistance of coarticulatory nasalization to focus-induced modulation, and the generally weaker prosodic modulation observed in AmE, may reflect a greater integration of anticipatory nasalization into the representation and phonetic target of the vowel. This pattern may indicate a more advanced stage along a pathway of phonologization or transphonologization, in which nasality cues are progressively redistributed from the coda nasal to the preceding vowel (cf. Beddor, 2023; Pouplier et al., 2024; Solé, 1995, 2007). Under this account, nasalization would persist more consistently across prosodic conditions because its temporal realization is less dependent on contextually induced changes in overlap between velic and oral gestures. As Beddor (2023) reviews, such progressive cue transfer may render originally gradient coarticulatory nasalization increasingly systematic and targeted, though without necessarily making it contrastive.

Further evidence that American English may be relatively advanced along this trajectory comes from Pouplier et al. (2024), who compared anticipatory nasal coarticulation across American English, French, and German. They found that the temporal extent of anticipatory nasalization was greater in American English than in either of the other languages (see also Solé, 1995, 2007, for related findings). If AmE is likewise further along this trajectory than AusE, its comparatively weaker prosodic modulation may reflect a greater stabilization of anticipatory nasalization as a targeted phonetic property of the vowel.

Nonetheless, a phonologization account cannot fully explain the present findings. The study also examined carryover nasalization in NVC contexts, where broadly comparable dialectal patterns emerged. Because cue transfer and phonologization have primarily been proposed for anticipatory nasalization in VN sequences, the parallel findings across the two contexts suggest a broader dialectal difference in how prosodic structure shapes nasal coarticulation and temporal coordination. Thus, the weaker prosodic modulation observed in AmE may reflect not only greater integration of anticipatory nasalization into the vowel target, but also a generally more temporally stable and less prosodically malleable organization of nasal coarticulation. One possibility, although admittedly speculative, is that the reduced prosodic sensitivity associated with the phonologization of anticipatory nasalization in VN sequences may not be limited to that context, but may also extend to carryover nasalization in NV sequences, reflecting a broader dialect-specific organization of vowel–nasal coarticulation.

4.2.1. Shared phonology and divergent phonetic grammars

The present findings, as discussed thus far—showing fine phonetic detail that varies systematically with prosodic structure—are most naturally interpreted as reflecting differences at the level of phonetic implementation rather than differences in phonological representation. Specifically, they point to variation in how shared phonological categories are realized through prosodically structured articulatory control. These differences can be framed within the *extended model of phonetic grammar* (Cho, 2025; Cho et al., 2026), in which prosody functions as a central organizing scaffold for phonetic realization. In this view, phonological categories do not map onto fixed outputs; instead, they are instantiated through gradient, temporally organized phonetic operations whose magnitude, timing, and coordination are regulated by prosodic structure and remain under active speaker control.

Within this framework, the observed dialectal divergence can be understood as reflecting differences in how competing implementation pressures are weighted. In particular, system-oriented pressures favoring articulatory continuity (coarticulatory overlap) compete with output-oriented pressures favoring the perceptual distinctiveness of the oral vowel (cf. Cho, 2025; Flemming, 2004; Lindblom, 1986, 1990; Maddieson, 1997). The results suggest that AmE assigns a relatively higher weight to coarticulatory continuity (i.e., overlap): even under focus, where prominence-driven enhancement typically suppresses coarticulation, suppression remains partial, allowing greater overlap between vocalic and velic gestures. AusE, by contrast, appears to weight prominence more strongly toward output-oriented distinctiveness, resulting in tighter regulation of nasality—effectively reallocating prominence-related enhancement toward the vowel's oral specification.

Importantly, these differences are neither categorical nor confined to a single temporal point; rather, they emerge as gradient modulations in the coordination of oral and velic gestures, reflected in systematic variation in spectral tilt over the vowel's time course. This pattern points to a dialect-specific

reweighting of timing relations and interactional priorities within a shared prosodic–phonetic system, without necessitating a change in underlying phonological contrast. In other words, while Australian and American English plausibly generate similar abstract prosodic frames during speech planning—defining prominence and phrasing—this structure is interpreted through distinct phonetic grammars at the phonetics–prosody interface (Cho, 2025). It is at this level that fine-grained tuning occurs: the AusE grammar treats prominence as a stronger directive for oral enhancement, whereas the AmE grammar permits a greater degree of coarticulatory overlap. These findings therefore confirm that the detailed temporal organization of coarticulation is prosodically regulated and constitutes part of speakers' grammatical knowledge, specified in the system-internal phonetic grammar in reference to prosodic structure.

This interpretation should not, however, be taken to imply a rigid separation between phonological and phonetic representations. Rather, the extended model of phonetic grammar is invoked here to characterize how abstract prosodic structure is mapped onto dialect-specific patterns of gradient phonetic coordination. The function attributed here to phonetic grammar overlaps substantially with what some approaches to phonetically based phonology assign to phonological grammar, particularly in frameworks that incorporate scalar phonetic representations and gradient patterns into phonological analysis (Flemming, 2001; see also Hayes et al., 2004). Whether such patterns are located within phonetic or phonological grammar therefore depends partly on the theoretical framework adopted and on where the boundary between the two domains is drawn (cf. Cohn, 2006). Under a broader conception of phonology that encompasses systematically controlled, language- and dialect-specific gradient patterns, the organization observed here could equally be analyzed as part of phonological grammar. Our use of the term phonetic grammar is thus intended to highlight the systematic organization of gradient implementation patterns subject to speaker control, rather than to presuppose that phonological and phonetic levels are strictly autonomous.

4.3. When vowel nasalization interacts with vowel glottalization

To validate these theoretical interpretations regarding dialect-specific phonetic strategies, it is essential to disentangle the potential acoustic interaction between the nasal and laryngeal sources. A critical methodological step is the systematic comparison between baseline and covariate-adjusted models. As noted earlier, vowel glottalization (indexed by H1–A3) is itself prosodically structured—patterning with prominence and boundary conditions (e.g., Dilley et al., 1996; Garellek, 2013; Mitterer et al., 2019; Pierrehumbert & Talkin, 1992; Redi & Shattuck-Hufnagel, 2001)—and differs between Australian and American English (Jang et al., 2026). Given that phonation-driven modulation can confound A1–P0 independently of velopharyngeal coupling (e.g., Chen, 1997; Garellek et al., 2016; Hanson et al., 2001; Styler, 2017), comparing baseline and covariate-adjusted analyses is necessary to isolate true nasalization from source characteristics. This comparison reveals two complementary outcomes regarding dialectal differentiation: the temporal reorganization of shared variance and the unmasking of suppressed differentiation.

The first outcome is the *temporal reorganization* of effects that were previously conflated with source characteristics. This is observed where effects that are significant in the baseline analysis become attenuated or lose statistical support once the covariate is included. A clear example is the Focus × Dialect interaction in the NVC (carryover) context. While significant in the baseline analysis, this interaction is not supported in the covariate-adjusted model, indicating that the apparent "larger" focus-induced coarticulatory resistance in AusE is partly attributable to that dialect's greater tendency toward focus-induced glottalization (Jang et al., 2026). Crucially, however, this adjustment does not eliminate the dialectal contrast; rather, it reveals that the difference is temporally structured. The disappearance of the global interaction is accompanied by the emergence of a significant Focus × Dialect × Timepoint interaction, showing that once phonation is controlled, the dialectal divergence manifests not as a static difference, but as a dynamic, gradual unfolding of resistance over the course of the vowel.

The second complementary outcome is the *unmasking* of dialectal differentiation that was obscured in the baseline model. This pattern is most evident in the Boundary × Dialect interaction in the NVC (carryover) context. In the baseline analysis, this interaction was absent, leading to the impression that AusE and AmE do not differ in how boundaries condition carryover nasalization. However, in the covariate-adjusted analysis, this interaction emerges as significant. This "unmasking" indicates that concomitant variation in phonation served to suppress the acoustic visibility of the dialectal effect: reduced H1 amplitude associated with glottalization likely counteracted the spectral consequences of velopharyngeal closure in the baseline measures. Once this source-related variance is statistically controlled, it becomes clear that the boundary-induced reduction of vowel nasalization (i.e., domain-initial strengthening) is, in fact, significantly more robust in AusE than in AmE.

Finally, distinct from these two patterns of change, the results in the anticipatory (CVN) context demonstrate *stability.* The higher-order dialectal interaction structure (e.g., Focus × Boundary × Dialect) remains supported across both baseline and covariate-adjusted analyses, with effects showing only minor scalar shifts rather than reorganization or unmasking. This confirms that, unlike the carryover context, dialectal differences in anticipatory coarticulatory organization are broadly robust and not straightforwardly attributable to source-related covariation.

These findings underscore the multidimensional nature of prosodic realization and the critical importance of methodologically disentangling source–filter interactions in cross-dialectal research. The fact that distinct dialectal patterns persist—either through temporal reorganization or unmasking—after accounting for phonation confirms that the observed divergence between Australian and AmE is not merely an acoustic artifact of differing voice qualities. Rather, it reflects a sophisticated level of speaker control where the phonetic grammar actively coordinates distinct laryngeal and velopharyngeal gestures. Thus, while the two dialects may differ in their concomitant deployment of glottalization, they nonetheless maintain separable, dialect-specific strategies for the prosodic modulation of velopharyngeal control. This reinforces the view that

phonetic implementation involves the precise, parallel fine-tuning of multiple articulatory subsystems, governed by a linguistically structured, dialect-specific phonetic grammar regulated with reference to prosodic structure.

5. Conclusion

This study examined how nasality is regulated by prosodic structure in Australian English (AusE) and how this system compares with American English (AmE) under tightly matched experimental conditions. By analyzing both nasal consonant duration and time-varying coarticulatory vowel nasalization across systematically manipulated Focus and Boundary conditions, we provide an integrated account of prosodically structured nasality and its dialectal variability.

At a general level, AusE participates in the same prosodic organization previously documented for AmE. In both dialects, prominence and phrasing exert coordinated but asymmetric effects on the nasal source and its coarticulatory consequences: focus enhances the nasal consonant while inducing coarticulatory resistance in the adjacent vowel, and boundary effects pattern differentially at phrase edges. This cross-dialectal convergence supports the view that the prosodic conditioning of nasality is a stable property of the English phonetics–prosody interface, compatible with cross-linguistic evidence that prosody provides a structured basis for organizing fine-grained phonetic detail.

Crucially, however, shared prosodic organization does not entail phonetic identity. The Australian–American contrast lies not in whether prosodic effects occur, but in how they are implemented over time. Relative to AmE, AusE shows tighter prominence-related regulation of nasality, expressed as stronger focus-induced coarticulatory resistance and a more sustained suppression of vowel nasalization across the vowel time course, alongside more limited temporal enhancement of the nasal source in the relevant prosodic configurations. These differences are gradient and temporally structured rather than categorical, pointing to dialect-specific fine-tuning of prosodically conditioned implementation within a shared phonological and prosodic system.

Methodologically, the study underscores the importance of disentangling interacting phonetic dimensions in cross-dialectal work. Because phonation-related variation can systematically influence acoustic measures of nasalization, we modeled glottalization alongside vowel nasalization to evaluate which dialect effects reflect source-related covariance and which remain once such variance is controlled. The persistence—and in some cases unmasking—of dialectal differences after this control supports the interpretation that the observed divergence is not reducible to phonation-driven artifacts, indicating a high degree of independent speaker control over the coordination of laryngeal and velopharyngeal subsystems.

Taken together, the findings support the *extended model of phonetic grammar*, in which prosody organizes the allocation and coordination of phonetic operations while allowing dialect-specific tuning of their magnitude, timing, and interaction. On this view, dialectal differentiation in vowel nasalization is best characterized not as a difference in underlying phonological representations, but as variation in the prosodically regulated implementation of shared categories—visible most clearly in the time-dependent structure of coarticulation as it is dynamically shaped by prominence and boundaries.

CRediT authorship contribution statement

Jiyoung Jang: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hyunjung Joo:** Methodology, Formal analysis, Data curation. **Sahyang Kim:** Writing – review & editing, Validation, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Anne Cutler:** Resources, Investigation, Conceptualization. **Taehong Cho:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank all the Australian English speakers for their participation in the experiment. We are also grateful to the editor, Dr. Sam Kirkham, as well as Dr. Jonathan Harrington and two anonymous reviewers for their constructive comments and suggestions, which greatly strengthened the manuscript. A preliminary descriptive portion of this research, which did not include glottalization as a covariate or provide an in-depth theoretical discussion, was presented at ICPhS 2019 (Joo et al., 2019). This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5C2A02086884) awarded to T. Cho (corresponding author). This work was also supported by the research fund of Hanyang University (HY-202600000002605).

Appendix A

A.1. Preliminary covariation analysis of vowel nasalization (A1–P0) and glottalization (H1*-A3*)*

Given the prosodic conditioning of glottalization and its variability across English dialects discussed in the Introduction, a critical methodological prerequisite for the present study is to operationalize the acoustic interaction between the laryngeal source and the velopharyngeal output. The necessity for this step lies in the specific spectral configuration of the measures employed: the nasal peak (P0), typically observed in the 250–450 Hz range, lies in close spectral proximity to the first harmonic (H1). Because glottalization is acoustically indexed by a reduction in low-frequency amplitude (particularly H1), phonation-driven variation can mechanically alter the apparent magnitude of P0 independently of velopharyngeal coupling (cf. Garellek et al., 2016; Hanson et al., 2001).

This acoustic interaction poses a specific threat to the validity of A1–P0* as an independent measure of nasality. Since A1–P0* is a relative metric, source-related attenuation of H1 (captured by H1*–A3*) can bleed into the quantification of the nasal peak. Specifically, when H1 is attenuated due to creaky phonation, it may lower the amplitude of the

acoustically adjacent nasal peak, creating a spurious correlation where increased glottalization mathematically mimics reduced nasalization. This overlap is distinct from the general coupling of harmonics by vocal tract resonances; while the transfer function of a resonance augments neighboring frequencies (Stevens, 1998), the issue here is that the source characteristic itself (H1 amplitude) is a primary driver of the energy available to be radiated at the nasal pole. Consequently, the extent to which source-driven variance confounds the nasalization measure cannot be determined *a priori*—it must be empirically assessed based on the degree of spectral alignment in the specific speech materials used.

To validate the decision to model these dimensions as separate but interacting factors, preliminary linear mixed-effects analyses were conducted between A1–P0* (nasalization) and H1*–A3* (glottalization) across dialects and prosodic conditions, with random intercepts for speaker and item. As shown in Fig. A1, the analysis revealed a significant inverse relationship, with correlation coefficients ranging from −0.45 to −0.56. Crucially, however, the marginal coefficients of determination (R^2m = 0.12–0.31) indicate that only approximately 10–30% of the variance in vowel nasalization is shared with vowel glottalization independent of by-speaker and by-word variability. This correlational evidence indicates that, although the measures are acoustically associated—possibly because both are influenced by low-frequency spectral structure—the relationship is only partial. That is, much of the variance in A1–P0* remains unexplained by variation in the phonation-related measure H1*–A3*, so we can treat glottalization as a covariate that helps account for part of the variation in A1–P0* rather than as a factor that determines it.

which A1–P0* was analyzed without adjustment for H1*–A3*. Second, parallel models were fitted that included H1*–A3* as a covariate. This dual modeling strategy allowed us to evaluate whether prosodic and dialectal effects on vowel nasalization persist once variance associated with glottal configuration is statistically controlled for, and to assess the degree to which observed effects on A1–P0* are independent of concomitant variation in phonation. In addition, the baseline model (without H1*–A3*) corresponds to the analytical approach adopted in much of the previous literature on coarticulatory vowel nasalization (e.g., Cho et al., 2017; Jang et al., 2018; Zellou & Scarborough, 2012; Zellou, 2017), whereas the covariate model provides a more conservative estimate by accounting for variation associated with phonation. Reporting both models allows direct comparison with earlier studies while also evaluating the extent to which controlling for source-related variation influences the observed patterns.

At the same time, we note that including H1*–A3* as a covariate may, in some cases, remove not only artifactual variance arising from shared acoustic dependencies in the low-frequency spectrum, but also phonetic variation that is legitimately part of a coordinated prosodic adjustment. In particular, if prosodic focus induces a temporally coordinated laryngeal and velopharyngeal configuration, variation in glottalization and vowel nasalization may reflect different acoustic consequences of the same underlying prosodic gesture rather than separable phonetic processes. Under such circumstances, controlling for H1*–A3* may attenuate effects on A1–P0* that are prosodically meaningful but acoustically inseparable from phonation. Importantly, therefore, as we discussed earlier, the inclusion of H1*–A3* as a covariate does not assume a

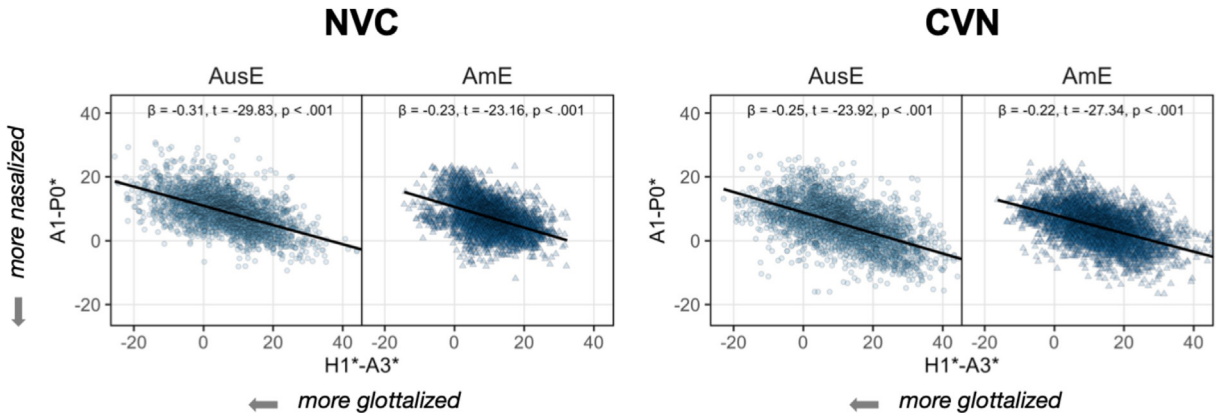

Fig. A1. Effects of the correlation between A1–P0* (V-nasalization) and H1*–A3* (V-glottalization) in both NVC and CVN.

However, this partial but non-negligible covariation nonetheless reflects some degree of measurement-level interdependence, and we therefore adopted a conservative analytic strategy to ensure methodological robustness. Specifically, analyses of vowel nasalization were carried out in two complementary ways. First, linear mixed-effects models were fitted in

causal relationship between glottalization and nasalization, nor does it presuppose that glottalization-related variance constitutes a nuisance factor. Rather, comparing models with and without H1*–A3* serves to clarify whether observed prosodic and dialectal effects on A1–P0* are robust to controlling for phonation-related variation, or whether they reflect a system-

atic coupling between nasalization and glottalization that may itself form part of the phonetic realization of prosodic structure within each dialect.

A.2. Preliminary covariation analysis of vowel nasalization (A1–P0) and F0*

A second potential source–filter interaction concerns fundamental frequency (F0). Because A1–P0* relies on the amplitude of the nasal peak (P0) within a low-frequency region (typically 250–450 Hz), its estimation can be influenced by harmonic sampling of the vocal tract transfer function. As F0 increases—most notably under prosodic focus—the wider spacing between harmonics increases the likelihood that a harmonic aligns more closely with the nasal pole, potentially boosting the measured amplitude of P0 and thereby lowering A1–P0* (cf. Styler, 2017). In principle, if dialects differ systematically in the magnitude or temporal deployment of F0 under focus, such pitch-related sampling effects could inflate apparent differences in vowel nasalization, confounding harmonic alignment with genuine differences in velopharyngeal coupling.

To evaluate this possibility, we examined the relationship between F0 and A1–P0*. Specifically, we tested a potential confound whereby increases in F0 might be associated with lower A1–P0* values, which would in turn suggest greater apparent nasalization. Contrary to this expectation, the results revealed only a weak association in the opposite direction (β = 1.06–2.36, *R2m* = 0.04–0.09), indicating that higher F0 was in fact associated with slightly less apparent nasalization (Fig. A2). These findings suggest that F0 accounts for only a modest proportion of the variance in A1–P0* in the present dataset, and that its observed effect—if anything—runs counter to an interpretation in which variation in A1–P0* is driven by variation in F0 in the manner previously identified as a potential confound. To further corroborate this conclusion, we fit mixed-effects models including F0 as a covariate (see Table B1 in the Appendix). These models yielded effect patterns that were largely comparable to those of the baseline models, particularly with no notable changes to the prosodic or dialectal effects of interest.[3] Given the limited variance in A1–P0* attributable to F0, F0 was not included as a covariate in the final analyses.

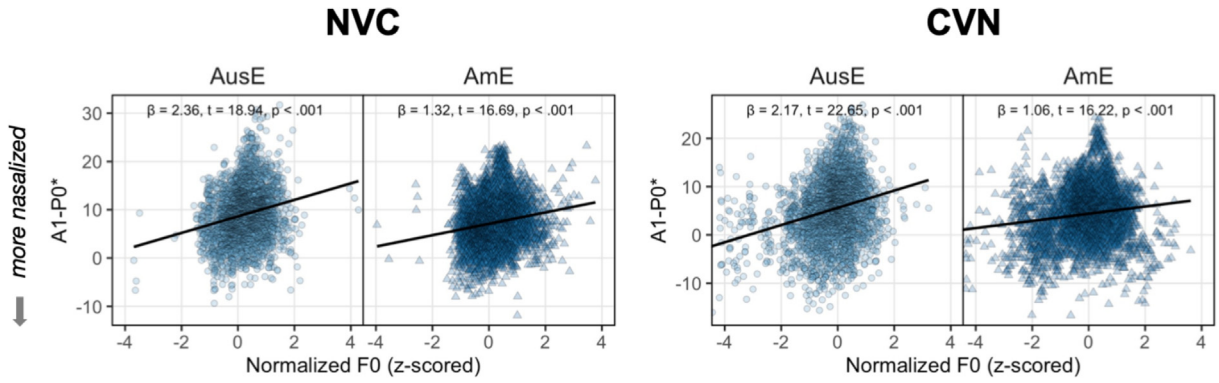

Fig. A2. Effects of the correlation between A1–P0* (V-nasalization) and F0 in both NVC and CVN.

[3] Methodologically, unlike glottalization—which constitutes a secondary dimension of voice quality that can obscure velopharyngeal effects by introducing spectral changes independent of nasal coupling—F0 is an intrinsic component of the prominence manipulation itself. As such, treating F0 as a covariate risks partialling out variance that is central to the expression of focus, rather than isolating an extraneous source of acoustic contamination. Thus, independently of the fact that F0 exhibited no substantial covariation with A1–P0* that would affect the interpretation of the results, omitting F0 preserves the interpretability of focus-related effects while maintaining the validity of A1–P0* as a measure of vowel nasalization.

Appendix B

Table B1
Type III analysis-of-variance results for the linear mixed-effects model of V-nasalization (A1–P0*) in NVC (in domain-initial contexts) and CVN (in domain-final contexts) with F0 as covariate.

	NVC (domain-initial)					CVN (domain-final)				
	NumDF	DenDF	F-value	Pr(>F)		NumDF	DenDF	F-value	Pr(>F)	
Dialect (AusE-AmE)	1	27.4	1.081	0.308		1	12.3	0.000	0.988	
Focus (FOC-UnF)	1	6472.7	443.891	<0.001	***	1	6925.3	539.073	<0.001	***
Boundary (IP-Wd)	1	6604.2	180.859	<0.001	***	1	6996.1	212.567	<0.001	***
Timepoint	2	6587.5	192.621	<0.001	***	2	6975.5	251.882	<0.001	***
F0	1	4294.8	62.840	<0.001	***	1	6092.3	26.805	<0.001	***
Focus: Dialect	1	6586.9	53.080	<0.001	***	1	6975	49.762	<0.001	***
Boundary: Dialect	1	6581.4	4.599	0.032	*	1	6973.2	1.233	0.267	
Focus: Boundary	1	6583.1	40.500	<0.001	***	1	6994.1	37.576	<0.001	***
Dialect: Timepoint	2	6583	5.203	0.006	**	2	6971.4	6.604	<0.001	**
Boundary: Timepoint	2	6581.7	0.859	0.424		2	6974.4	16.681	<0.001	***
Focus: Timepoint	2	6585	0.400	0.671		2	6971.6	4.142	0.016	*
Focus: Boundary: Dialect	1	6582.1	0.071	0.789		1	6971.9	23.786	<0.001	***
Boundary: Dialect: Timepoint	2	6581.1	0.431	0.650		2	6971.2	1.176	0.309	
Focus: Dialect: Timepoint	2	6582.2	3.263	0.038	*	2	6971.2	8.045	<0.001	***
Focus: Boundary: Timepoint	2	6581.1	0.046	0.955		2	6972.6	0.787	0.455	
Focus: Boundary: Dialect: Timepoint	2	6581.2	0.969	0.379		2	6971.1	5.116	0.006	**

Table B2
Post-hoc pairwise comparisons of N-duration across Dialect, Focus, Boundary, and their interactions. Pairwise dialect contrasts within levels of Focus, Boundary, and their interactions were also examined but yielded no significant differences and are omitted here for brevity.

	NVC (domain-initial)	CVN (domain-final)
Dialect	AusE – AmE: β = 9.8, *SE* = 5.7, $p > 0.09$	AusE – AmE: β = 3.1, *SE* = 6.6, $p > 0.09$
Focus	FOC – UnF: β = 53.3, *SE* = 0.9, $p < 0.001$	FOC – UnF: β = 35.1, *SE* = 1.6, $p < 0.001$
Boundary	IP – Wd: $\beta = -33.3$, *SE* = 0.9, $p < 0.001$	IP – Wd: β = 26.5, *SE* = 1.6, $p < 0.001$
Focus:Dialect	AusE, FOC – UnF: β = 49.9, *SE* = 1.4, $p < 0.001$	AusE, FOC – UnF: β = 34.9, *SE* = 2.4, $p < 0.001$
	AmE, FOC – UnF: β = 56.6, *SE* = 1.3, $p < 0.001$	AmE, FOC – UnF: β = 35.3, *SE* = 2.2, $p < 0.001$
Boundary:Dialect	AusE, IP – Wd: $\beta = -33.3$, *SE* = 1.4, $p < 0.001$	AusE, IP – Wd: β = 23.7, *SE* = 2.4, $p < 0.001$
	AmE, IP – Wd: $\beta = -33.4$, *SE* = 1.3, $p < 0.001$	AmE, IP – Wd: β = 29.2, *SE* = 2.2, $p < 0.001$
Focus:Boundary	IP, FOC – UnF: β = 41.9, *SE* = 1.4, $p < 0.001$	IP, FOC – UnF: β = 30.8, *SE* = 2.3, $p < 0.001$
	Wd, FOC – UnF: β = 64.7, *SE* = 1.3, $p < 0.001$	Wd, FOC – UnF: β = 39.4, *SE* = 2.4, $p < 0.001$
Focus:Boundary: Dialect	AusE, IP, FOC – UnF: β = 42.2, *SE* = 1.9, $p < 0.001$	AusE, IP, FOC – UnF: β = 33.1, *SE* = 3.4, $p < 0.001$
	AusE, Wd, FOC – UnF: β = 57.6, *SE* = 2.0, $p < 0.001$	AusE, Wd, FOC – UnF: β = 36.7, *SE* = 3.5, $p < 0.001$
	AmE, IP, FOC – UnF: β = 41.5, *SE* = 1.9, $p < 0.001$	AmE, IP, FOC – UnF: β = 28.5, *SE* = 3.2, $p < 0.001$
	AmE, Wd, FOC – UnF: β = 71.8, *SE* = 1.9, $p < 0.001$	AmE, Wd, FOC – UnF: β = 42.0, *SE* = 3.2, $p < 0.001$

Table B3
Post-hoc pairwise comparisons of V-nasalization across Dialect, Focus, Boundary, Timepoint, and their interactions. Pairwise dialect contrasts within levels of Focus, Boundary, and their interactions were also examined but yielded no significant differences and are omitted here for brevity.

	NVC (domain-initial)	CVN (domain-final)
Dialect	AusE – AmE: β = 1.6, *SE* = 1.2, $p > 0.09$	AusE – AmE: β = 0.3, *SE* = 1.6, $p > 0.09$
	25%, AusE – AmE: β = 1.2, *SE* = 1.2, $p > 0.09$	25%, AusE – AmE: $\beta = -0.1$, *SE* = 1.6, $p > 0.09$
	50%, AusE – AmE: β = 1.9, *SE* = 1.2, $p > 0.09$	50%, AusE – AmE: β = 0.5, *SE* = 1.6, $p > 0.09$
	75%, AusE – AmE: β = 1.8, *SE* = 1.2, $p < 0.08$	75%, AusE – AmE: β = 0.7, *SE* = 1.6, $p > 0.09$
Focus	FOC – UnF: β = 2.8, *SE* = 0.1, $p < 0.001$	FOC – UnF: β = 3.4, *SE* = 0.1, $p < 0.001$
	25%, FOC – UnF: β = 2.8, *SE* = 0.1, $p < 0.001$	25%, FOC – UnF: β = 3.1, *SE* = 0.1, $p < 0.001$
	50%, FOC – UnF: β = 2.8, *SE* = 0.1, $p < 0.001$	50%, FOC – UnF: β = 3.5, *SE* = 0.1, $p < 0.001$
	75%, FOC – UnF: β = 2.9, *SE* = 0.1, $p < 0.001$	75%, FOC – UnF: β = 3.8, *SE* = 0.1, $p < 0.001$
Boundary	IP – Wd: β = 1.7, *SE* = 0.3, $p < 0.001$	IP – Wd: $\beta = -1.8$, *SE* = 0.4, $p < 0.001$
	25%, IP – Wd: β = 1.5, *SE* = 0.1, $p < 0.001$	25%, IP – Wd: $\beta = -2.4$, *SE* = 0.4, $p < 0.001$
	50%, IP – Wd: β = 1.8, *SE* = 0.1, $p < 0.001$	50%, IP – Wd: $\beta = -2.1$, *SE* = 0.4, $p < 0.001$
	75%, IP – Wd: β = 1.7, *SE* = 0.1, $p < 0.001$	75%, IP – Wd: $\beta = -0.9$, *SE* = 0.4, $p < 0.001$
Focus:Dialect	AusE, FOC – UnF: β = 3.6, *SE* = 0.1, $p < 0.001$	AusE, FOC – UnF: β = 4.1, *SE* = 0.1, $p < 0.001$
	AmE, FOC – UnF: β = 2.1, *SE* = 0.1, $p < 0.001$	AmE, FOC – UnF: β = 2.7, *SE* = 0.1, $p < 0.001$
	AusE, 25%, FOC – UnF: β = 3.5, *SE* = 0.2, $p < 0.001$	AusE, 25%, FOC – UnF: β = 3.3, *SE* = 0.2, $p < 0.001$
	AusE, 50%, FOC – UnF: β = 3.5, *SE* = 0.2, $p < 0.001$	AusE, 50%, FOC – UnF: β = 4.2, *SE* = 0.2, $p < 0.001$
	AusE, 75%, FOC – UnF: β = 3.8, *SE* = 0.2, $p < 0.001$	AusE, 75%, FOC – UnF: β = 5.0, *SE* = 0.2, $p < 0.001$
	AmE, 25%, FOC – UnF: β = 2.3, *SE* = 0.2, $p < 0.001$	AmE, 25%, FOC – UnF: β = 2.9, *SE* = 0.2, $p < 0.001$
	AmE, 50%, FOC – UnF: β = 2.1, *SE* = 0.2, $p < 0.001$	AmE, 50%, FOC – UnF: β = 2.7, *SE* = 0.2, $p < 0.001$
	AmE, 75%, FOC – UnF: β = 1.8, *SE* = 0.2, $p < 0.001$	AmE, 75%, FOC – UnF: β = 2.6, *SE* = 0.2, $p < 0.001$
Boundary:Dialect	AusE, IP – Wd: β = 1.9, *SE* = 0.4, $p < 0.001$	AusE, IP – Wd: $\beta = -1.9$, *SE* = 0.4, $p < 0.001$
	AmE, IP – Wd: β = 1.4, *SE* = 0.4, $p < 0.001$	AmE, IP – Wd: $\beta = -1.7$, *SE* = 0.4, $p < 0.001$
	AusE, 25%, IP – Wd: β = 1.6, *SE* = 0.4, $p < 0.001$	AusE, 25%, IP – Wd: $\beta = -2.7$, *SE* = 0.5, $p < 0.001$
	AusE, 50%, IP – Wd: β = 1.9, *SE* = 0.4, $p < 0.001$	AusE, 50%, IP – Wd: $\beta = -2.0$, *SE* = 0.5, $p < 0.001$
	AusE, 75%, IP – Wd: β = 2.1, *SE* = 0.5, $p < 0.001$	AusE, 75%, IP – Wd: $\beta = -0.9$, *SE* = 0.5, $p < 0.0$
	AmE, 25%, IP – Wd: β = 1.3, *SE* = 0.4, $p < 0.01$	AmE, 25%, IP – Wd: $\beta = -2.1$, *SE* = 0.5, $p < 0.001$
	AmE, 50%, IP – Wd: β = 1.6, *SE* = 0.4, $p < 0.001$	AmE, 50%, IP – Wd: $\beta = -2.2$, *SE* = 0.5, $p < 0.001$
	AmE, 75%, IP – Wd: β = 1.3, *SE* = 0.4, $p < 0.01$	AmE, 75%, IP – Wd: $\beta = -0.9$, *SE* = 0.5, $p < 0.08$
Focus:Boundary	FOC, IP – Wd: β = 1.0, *SE* = 0.3, $p < 0.01$	FOC, IP – Wd: $\beta = -1.0$, *SE* = 0.4, $p < 0.05$
	UnF, IP – Wd: β = 2.2, *SE* = 0.3, $p < 0.001$	UnF, IP – Wd: $\beta = -2.6$, *SE* = 0.4, $p < 0.001$
	FOC, 25%, IP – Wd: β = 0.9, *SE* = 0.3, $p < 0.05$	FOC, 25%, IP – Wd: $\beta = -1.5$, *SE* = 0.4, $p < 0.001$
	FOC, 50%, IP – Wd: β = 1.1, *SE* = 0.3, $p < 0.01$	FOC, 50%, IP – Wd: $\beta = -1.1$, *SE* = 0.4, $p < 0.01$
	FOC, 75%, IP – Wd: β = 1.0, *SE* = 0.3, $p < 0.01$	FOC, 75%, IP – Wd: $\beta = -0.3$, *SE* = 0.4, $p > 0.09$
	UnF, 25%, IP – Wd: β = 2.0, *SE* = 0.3, $p < 0.001$	UnF, 25%, IP – Wd: $\beta = -3.3$, *SE* = 0.5, $p < 0.001$
	UnF, 50%, IP – Wd: β = 2.4, *SE* = 0.3, $p < 0.001$	UnF, 50%, IP – Wd: $\beta = -3.0$, *SE* = 0.5, $p < 0.001$
	UnF, 75%, IP – Wd: β = 2.3, *SE* = 0.4, $p < 0.001$	UnF, 75%, IP – Wd: $\beta = -1.5$, *SE* = 0.4, $p < 0.01$
Focus:Boundary: Dialect	AusE, FOC, IP – Wd: β = 1.3, *SE* = 0.4, $p < 0.01$	AusE, FOC, IP – Wd: $\beta = -1.6$, *SE* = 0.4, $p < 0.001$
	AusE, UnF, IP – Wd: β = 2.5, *SE* = 0.4, $p < 0.001$	AusE, UnF, IP – Wd: $\beta = -2.1$, *SE* = 0.5, $p < 0.001$
	AmE, FOC, IP – Wd: β = 0.8, *SE* = 0.4, $p < 0.08$	AmE, FOC, IP – Wd: $\beta = -0.3$, *SE* = 0.4, $p > 0.09$
	AmE, UnF, IP – Wd: β = 2.0, *SE* = 0.4, $p < 0.001$	AmE, UnF, IP – Wd: $\beta = -3.1$, *SE* = 0.5, $p < 0.001$
	AusE, FOC, 25%, IP – Wd: β = 0.9, *SE* = 0.4, $p < 0.08$	AusE, FOC, 25%, IP – Wd: $\beta = -2.7$, *SE* = 0.5, $p < 0.001$
	AusE, FOC, 50%, IP – Wd: β = 0.9, *SE* = 0.4, $p < 0.01$	AusE, FOC, 50%, IP – Wd: $\beta = -1.7$, *SE* = 0.5, $p < 0.01$
	AusE, FOC, 75%, IP – Wd: β = 0.5, *SE* = 0.4, $p < 0.01$	AusE, FOC, 75%, IP – Wd: $\beta = -0.5$, *SE* = 0.5, $p > 0.09$
	AusE, UnF, 25%, IP – Wd: β = 1.7, *SE* = 0.5, $p < 0.001$	AusE, UnF, 25%, IP – Wd: $\beta = -2.6$, *SE* = 0.6, $p < 0.001$
	AusE, UnF, 50%, IP – Wd: β = 2.3, *SE* = 0.5, $p < 0.001$	AusE, UnF, 50%, IP – Wd: $\beta = -2.3$, *SE* = 0.6, $p < 0.001$
	AusE, UnF, 75%, IP – Wd: β = 2.2, *SE* = 0.5, $p < 0.001$	AusE, UnF, 75%, IP – Wd: $\beta = -1.3$, *SE* = 0.6, $p < 0.05$
	AmE, FOC, 25%, IP – Wd: β = 0.9, *SE* = 0.4, $p < 0.08$	AmE, FOC, 25%, IP – Wd: $\beta = -0.3$, *SE* = 0.5, $p > 0.09$
	AmE, FOC, 50%, IP – Wd: β = 1.3, *SE* = 0.5, $p < 0.05$	AmE, FOC, 50%, IP – Wd: $\beta = -0.6$, *SE* = 0.5, $p > 0.09$
	AmE, FOC, 75%, IP – Wd: β = 1.6, *SE* = 0.5, $p > 0.09$	AmE, FOC, 75%, IP – Wd: $\beta = -0.1$, *SE* = 0.5, $p > 0.09$
	AmE, UnF, 25%, IP – Wd: β = 2.4, *SE* = 0.5, $p < 0.01$	AmE, UnF, 25%, IP – Wd: $\beta = -3.9$, *SE* = 0.6, $p < 0.001$
	AmE, UnF, 50%, IP – Wd: β = 2.5, *SE* = 0.5, $p < 0.001$	AmE, UnF, 50%, IP – Wd: $\beta = -3.8$, *SE* = 0.6, $p < 0.001$
	AmE, UnF, 75%, IP – Wd: β = 2.6, *SE* = 0.6, $p < 0.001$	AmE, UnF, 75%, IP – Wd: $\beta = -1.7$, *SE* = 0.5, $p < 0.01$

References

Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. In J. Cole & J. I. Hualde (Eds.), *Papers in laboratory phonology 9: Change in phonology* (pp. 547–576). Mouton de Gruyter.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., & Bolker, M. B. (2015). Package 'lme4'. *Convergence, 12*(1), 2.

Beddor, P. S. (2009). A coarticulatory path to sound change. *Language, 85*(4), 785–821.

Beddor, P. S. (2023). Advancements of phonetics in the 21st century: Theoretical and empirical issues in the phonetics of sound change. *Journal of Phonetics, 98*. https://doi.org/10.1016/j.wocn.2023.101228 101228.

Beddor, P. S., Coetzee, A. W., Styler, W., McGowan, K. B., & Boland, J. E. (2018). The time course of individuals' perception of coarticulatory information is linked to their production: Implications for sound change. *Language, 94*(4), 931–968.

Bongiovanni, S. (2021). Acoustic investigation of anticipatory vowel nasalization in a Caribbean and a non-Caribbean dialect of Spanish. *Linguistics Vanguard, 7*(1) 20200008.

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America, 102*(4), 2360–2370.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass, 10*(3), 120–141.

Cho, T. (2025). Advancements of phonetics in the 21st century: Phonetic universals, language variation, and phonetic grammar. *Journal of Phonetics, 111* 101426.

Cho, T., Kim, D., & Kim, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics, 64*, 71–89.

Cho, T., Kim, S., & Mitterer, H. (2026). Linguistic and cognitive functions of fine phonetic detail underlying sound systems and sound change. *Journal of Phonetics, 114* 101470.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*(2), 207–229.

Clopper, C. G., Pisoni, D. B., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America, 118*(3), 1661–1676. https://doi.org/10.1121/1.2000774.

Clopper, C. G., & Smiljanic, R. (2011). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics, 39*(2), 237–245. https://doi.org/10.1016/j.wocn.2011.02.003.

Clopper, C. G., & Smiljanic, R. (2015). Regional variation in temporal organization in American English. *Journal of Phonetics, 49*, 1–15. https://doi.org/10.1016/j.wocn.2014.10.002.

Coetzee, A. W., Beddor, P. S., Styler, W., Tobin, S., Bekker, I., & Wissing, D. (2022). Producing and perceiving socially structured coarticulation: Coarticulatory nasalization in Afrikaans. *Laboratory Phonology, 13*(1), 13. https://doi.org/10.16995/labphon.6450. 1-43.

Cohn, A. C. (1990). *Phonetic and phonological rules of nasalization* (PhD dissertation). UCLA.

Cohn, A. C. (2006). Is there gradient phonology? In G. Fanselow, C. Féry, R. Vogel, & M. Schlesewsky (Eds.), *Gradience in grammar: Generative perspectives* (pp. 25–44). Oxford University Press.

Cunha, C., Hoole, P., Voit, D., Frahm, J., & Harrington, J. (2024). The physiological basis of the phonologization of vowel nasalization: A real-time MRI analysis of American and Southern British English. *Journal of Phonetics, 105* 101329.

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America, 97*(1), 491–504.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24*(4), 423–444.

Farnetani, E., & Recasens, D. (1999). Coarticulation models in recent speech production theories. In W. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 31–65). Cambridge: Cambridge University Press.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology, 18*(1), 7–44. https://doi.org/10.1017/S0952675701004006.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 232–276). Cambridge University Press.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 523–602). Oxford: Wiley-Blackwell.

Fowler, C. A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech, Language, and Hearing Research, 24*(1), 127–139.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36*(2–3), 171–195.

Garellek, M. (2013). *Production and perception of glottal stops* (PhD dissertation). Los Angeles, CA: University of California Los Angeles.

Garellek, M., Ritchart, A., & Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics, 59*, 110–121.

Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics, 29*(4), 383–406.

Hanson, H. M., Stevens, K. N., Kuo, H. K. J., Chen, M. Y., & Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics, 29*(4), 451–480.

Hayes, B., Kirchner, R., & Steriade, D. (Eds.). (2004). *Phonetically based phonology.* Cambridge University Press.

Jang, J., Kim, S., & Cho, T. (2018). Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America, 144*(1), EL33–EL39. https://doi.org/10.1121/1.5044641.

Jang, J., Kim, S., & Cho, T. (2023). Prosodic structural effects on non-contrastive coarticulatory vowel nasalization in L2 English by Korean learners. *Language and Speech, 66*(2), 381–411.

Jang, J., Kim, S., Cutler, A., & Cho, T. (2026). When vowels glottalize without strictly aligning with prosodic boundaries: Marking prosodic structure in American and Australian English. *Linguistics Vanguard*. https://doi.org/10.1515/lingvan-2025-0005.

Jang, J., Lee, J., Lee, J., Kim, S., & Cho, T. (2022). Unveiling denasalization as an ongoing sound change: The role of prosody and gender in Seoul Korean. *Language and Speech*. https://doi.org/10.1177/00238309251390431.

Joo, H., Jang, J., Kim, S., Cho, T., & Cutler, A. (2019). Prosodic structural effects on coarticulatory vowel nasalization in Australian English in comparison to American English. In *19th International Congress of Phonetic Sciences (ICPhS 2019)* (pp. 835–839). Australasian Speech Science and Technology Association Inc.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60*, 286–319.

Keating, P. A. (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of peter ladefoged* (pp. 115–132). Orlando FL: Academic Press.

Keating, P. A., Cho, T., Fougeron, C., & Hsu, C. S. (2003). Domain-initial strengthening in four languages. In *Laboratory phonology VI: Phonetic interpretation* (pp. 145–163). Cambridge University Press.

Kendall, T., Pharao, N., Stuart-Smith, J., & Vaughn, C. (2023). Advancements of phonetics in the 21st century: Theoretical issues in sociophonetics. *Journal of Phonetics, 98*. https://doi.org/10.1016/j.wocn.2023.101226 101226.

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language, 82*(1), 33–63.

Kingston, J., & Diehl, R. L. (1994). Phonetic knowledge. *Language, 70*(3), 419–454.

Krakow, R. A. (1999). Physiological organization of syllables: A review. *Journal of Phonetics, 27*(1), 23–54.

Kühnert, B., & Nolan, F. (1999). The origin of coarticulation. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 7–30). Cambridge University Press.

Ladd, D. R. (2006). Segmental anchoring of pitch movements: Autosegmental association or gestural coordination? *Italian Journal of Linguistics, 18*(1), 19–38.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.

Ladefoged, P. (1971). Phonological features and their phonetic correlates. *UCLA Working Papers in Phonetics, 21*, 3–12.

Laver, J. (1980). *The phonetic description of voice quality.* Cambridge: CUP.

Lenth, R. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.8. https://CRAN.R-project.org/package=emmeans.

Li, H., Kim, S., & Cho, T. (2020). Prosodic structurally conditioned variation of coarticulatory vowel nasalization in Mandarin Chinese: Its language specificity and cross-linguistic generalizability. *The Journal of the Acoustical Society of America, 148*(3), EL240–EL246. https://doi.org/10.1121/10.0001743.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology* (pp. 13–44). Academic Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling, Vol. 55*, 403–439.

Maddieson, I. (1997). Phonetic universals. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 619–639). Oxford: Blackwell.

Mitterer, H., Kim, S., & Cho, T. (2019). The glottal stop between segmental and suprasegmental processing: The case of Maltese. *Journal of Memory and Language, 108*(104034), 1–19. https://doi.org/10.1016/j.jml.2019.104034.

Ohala, J. J. (1997). Aerodynamics of phonology. *Proceedings of the Seoul international conference on linguistics* (Vol. 92, pp. 97). Seoul: Linguistic Society of Korea.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In Gerard J. Docherty & D. Robert Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 90–117). Cambridge, New York & Victoria: Cambridge University Press.

Pouplier, M., Rodriquez, F., Lo, J. J., Alderton, R., Evans, B. G., Reinisch, E., & Carignan, C. (2024). Language-specific and individual variation in anticipatory nasal coarticulation: A comparative study of American English, French, and German. *Journal of Phonetics, 107* 101365.

R Core Team (2024). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics, 29*, 407–429.

Scarborough, R. A. (2003). Degree of coarticulation and lexical confusability. In P. M. Nowak, C. Yoquelet, & D. Mortensen (Eds.), *Proceedings of the 29th meeting of the Berkeley linguistics society* (pp. 367–378).

Solé, M.-J. (1995). Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization. *Language and Speech, 38*(1), 1–23.

Solé, M.-J. (2007). Controlled and mechanical properties in speech: A review of the literature. In M.-J. Solé, P. S. Beddor, & M. Ohala (Eds.), *Experimental approaches to phonology* (pp. 302–321). Oxford University Press.

Solé, M.-J., & Ohala, J. J. (2010). What is and what is not under the control of the speaker: Intrinsic vowel duration. *Papers in Laboratory Phonology, 10*, 607–655.
Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge: MIT press.
Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America, 142*(4), 2469–2482. https://doi.org/10.1121/1.5008853.
Tamminga, M., & Zellou, G. (2015). Cross-dialectal differences in nasal coarticulation in American English. In The Scottish Consortium for ICPhS 2015 (Eds.), *Proceedings of the 18th international congress of phonetic sciences* (Paper No. 745). University of Glasgow.
Zellou, G. (2017). Individual differences in the production of nasal coarticulation and perceptual compensation. *Journal of Phonetics, 61*, 13–29.
Zellou, G. (2022). *Coarticulation in phonology*. Cambridge University Press.
Zellou, G., & Scarborough, R. (2012). Nasal coarticulation and contrastive stress. In *Interspeech 2012* (pp. 2686–2689). ISCA. https://doi.org/10.21437/Interspeech.2012-667.