Linguistic Research 42(3): 685-710
DOI: 10.17250/khisli.42.3.202512.008

An ultrasound study of ambisyllabicity: The case of American English retroflex /ɹ/*

Joo-Kyeong Lee
(University of Seoul)

Lee, Joo-Kyeong. 2025. An ultrasound study of ambisyllabicity: The case of American English retroflex /ɹ/. *Linguistic Research* 42(3): 685-710. Ambisyllabicity has long been central to phonological theory, yet its articulatory basis remains unclear. This study investigates whether American English ambisyllabic retroflex /ɹ/ exhibits intermediate tongue configurations between onset and coda realizations in spatial and/or temporal dimensions. Ultrasound imaging data were collected from four native speakers producing five /ɹ/ types: three intervocalic retroflexes (an ambisyllabic retroflex preceded by a stressed lax vowel; a non-ambisyllabic retroflex preceded by a stressed tense vowel, hereafter non_A; and a non-ambisyllabic retroflex followed by a stressed vowel, hereafter non_B), as well as word-initial onset and word-final coda retroflexes. Tongue contours for each intervocalic retroflex were compared with those of onset and coda retroflexes using generalized additive mixed modeling (GAMM). Across speakers, intervocalic retroflexes followed a consistent trajectory, shifting from intermediate positions between onset and coda toward onset-like configurations. Crucially, this intermediate status was not unique to ambisyllabic retroflexes but was observed across all intervocalic contexts, suggesting that ambisyllabicity lacks a stable articulatory correlate and functions primarily as a theoretical construct. In addition, non_B retroflexes shifted toward onset-like tongue positions as early as the medial time point, whereas ambisyllabic and non_A retroflexes did so only at the final stage of articulation. This pattern indicates that intervocalic retroflexes preceded by a stressed vowel and followed by an unstressed vowel may be regarded as constituting an independent allophone, characterized by intermediate tongue contours that are distinct from onset and coda allophones. **(University of Seoul)**

Keywords ambisyllabic, ultrasound, retroflex, tongue contours

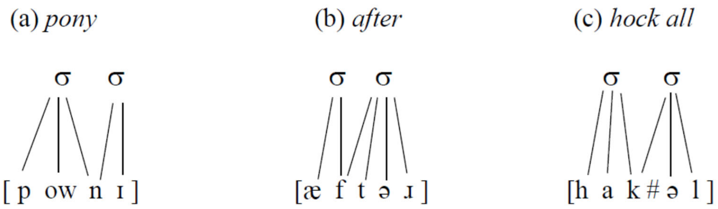
* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5A2A01070152).

1. Introduction

Ambisyllabicity refers to a syllabification pattern in which a consonant is simultaneously associated with both the coda of the preceding syllable and the onset of the following syllable. Crucially, ambisyllabicity is not determined solely by linear position—such as being intervocalic—but by its structural affiliation across syllable boundaries. That is, even if a consonant appears between vowels, it is considered ambisyllabic only if it is conditioned by prosodic and/or segmental factors (Kahn 1976; Gussenhoven 1986; Giegerich 1992; Hayes 2011). The notion of ambisyllabicity has played a significant role in phonological theory, particularly in discussions of how segments are affiliated with syllables in a prosodic structure. It has been proposed as a structural solution to account for cases where consonants appear to serve dual roles, challenging the conventional division between onset and coda positions (Kahn 1976; Giegerich 1992). At the same time, ambisyllabicity raises important questions about how such dual affiliation is phonetically realized in actual speech. While theoretical phonological research has offered detailed accounts of the conditions under which ambisyllabicity arises, much less is known about how such dual syllabic affiliation is manifested in physical articulation. In particular, the phonetic realization of ambisyllabic consonants—whether they exhibit hybrid articulatory characteristics or form either onset or coda allophonic patterns—remains an open empirical question. To address this gap, the present study examines the articulatory properties of the American English retroflex /ɹ/ in both ambisyllabic and non-ambisyllabic contexts, employing ultrasound imaging to capture tongue configuration and movement.

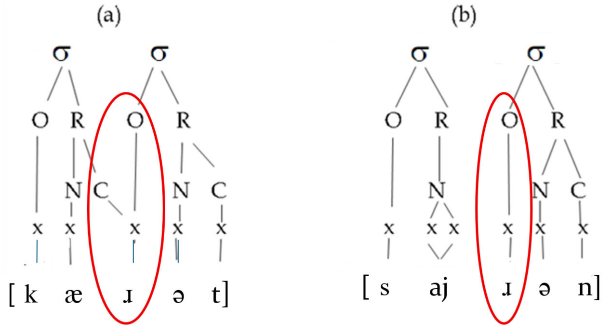
Kahn's (1976) work introduced ambisyllabicity as a structural solution to account for syllabification patterns in English, proposing that certain intervocalic consonants belong both the coda and the onset. Kahn initially posited a flap as an ambisyllabic consonant as in *city* ˈ[sɪɾi], because it was a different allophone of /t/ from the aspirated one [t^h] in onset position or the glottalized one [ʔt] in coda position. He contended that different allophones were derived from different syllable structures, allowing [ɾ] simultaneously branching to onset and coda. Expanding the scope of ambisyllabicity to all types of consonants and all prosodic hierarchies, Kahn assumed that ambisyllabicity arises under specific prosodic conditions. First, when a consonant follows a stressed syllable and precedes an unstressed one, it is often shared between syllables as in (1a). For instance, [n] is ambisyllabic in *pony* ˈ[pow.n.i], they function as the coda of the first syllable and simultaneously as the onset of the second syllable. Second, in the word *after* [ˈæ.f.təɹ], the [f] is assigned to both the coda of the stressed syllable and the onset of the following unstressed syllable as shown in (1b). Finally, ambisyllabicity can also occur across word boundaries, as in *hock all* [hɑ.k. ɔl], where the [k] bridges the final syllable of the first word and the initial syllable of the second when the latter begins with a vowel as illustrated in (1c). An important monomorphemic condition for ambisyllabicity is that the following syllable assigning the ambisyllabic consonant to be its onset should be unstressed. The sounds [n] and [f] in the examples of (1a) and (1b) are ambisyllabic, but also [ɾ] is ambisyllabic in the word of society[sə.saj.ə.ɾ.i] where the following syllable of [ɾ] is consistently unstressed. The preceding syllables are, on the other hand, stressed in *pony* and *after*, but unstressed in *society*. Stress seems to be an important condition for ambisyllabicity.

(1)



In contrast, Giegerich (1992) proposed an alternative analysis grounded in the Weight-Stress principle (Duanmu 2010), which posits that stressed syllables should be heavy. Under this view, when a stressed syllable contains a lax vowel and lacks a coda, the onset of the following syllable may be co-opted to serve as its coda, thereby satisfying the weight requirement. This process yields not as a default syllabification strategy, but as a phonological resolution to conflicting structural principles—specifically, the tension between the Onset Maximization principle, which favors parsing intervocalic consonants in to onsets, and the Weight-Stress Principle, which demands coda consonants in stressed syllables with lax vowels. In this account, a consonant is considered ambisyllabic when it occurs between a stressed lax vowel and an unstressed following syllable. For instance, in *carrot* [ˈkæɹət] as shown in (2a), the intervocalic consonant [ɹ] is first assigned to the onset of the second syllable in accordance with the Onset Maximization Principle. To satisfy the weight requirement of the first syllable, the same consonant is also affiliated with the coda, resulting in ambisyllabicity. Conversely in *siren* [ˈsajɹən], the stressed syllable contains a tense vowel, which inherently satisfies the heaviness requirement of the Weight-Stress Principle as presented in (2b). Thus, no coda is needed, and the intervocalic consonant remains solely affiliated with the onset of the following syllable.

(2)



This study focuses exclusively on in intervocalic morpheme-internal American English; retroflexes that span word boundaries are excluded from consideration. While both Kahn's(1976) and Giegerich's(1992) accounts of ambisyllabicity are prosodically motivated, Giegerich's analysis imposes more restrictive conditions: the preceding syllable must be both stressed and light (i.e.,contain a lax vowel and lack a coda). In contrast, Kahn's account relies primarily on the prosodic status of the following syllable, requiring it to be unstressed, without additional constraints on the preceding syllable.

There have been a dominant number of psycholinguistic studies on ambisyllabic consonants, playing a crucial role in identifying the distributional patterns of ambisyllabicity through a variety of experimental tasks, such as pause-insertion (Derwing 1992), syllable reversal (Treiman & Danis 1988), and syllable doubling (Eddington & Elzinga 2008; Fallows 1981). These studies have consistently found that ambisyllabicity is more likely under certain phonological and orthographic conditions. For example, ambisyllabic responses were more frequent when the preceding vowel was lax rather than tense (e.g., narrow vs. mayor) and when the preceding syllable was stressed (Treiman & Zukowski 1990; Ishikawa 2002). Additionally, liquids and nasals were more readily judged as ambisyllabic than obstruents (Treiman & Danis 1988), and consonants spelled with double letters (i.e., geminates) were more likely to be perceived as ambisyllabic (Zamuner & Ohala 1999; Treiman et al. 2002).

These findings suggest that speakers possess implicit knowledge about syllable structure and consonant distribution, and that this knowledge reflects abstract syllabification tendencies. However, psycholinguistic studies are primarily perception-based and do not offer direct evidence of how ambisyllabicity is phonetically or articulatorily realized. They did not even set up phonological/phonetic definitions of ambisyllabic consonants but merely analyzed participants' responses of whether a given consonant was ambisyllabic or not. That is, they describe behavioral patterns—what speakers do when asked to manipulate syllables—but do not tell us whether such patterns correspond to physical articulatory gestures or acoustic differences.

Several acoustic studies have examined the phonetic reality of ambisyllabicity by comparing the acoustic features of intervocalic consonants in different syllabic contexts. Durvasula and Huang (2017) conducted two experiments showing that ambisyllabic consonants pattern acoustically with codas rather than onsets in their nasalization and obstruent devoicing studies. However, their method for identifying nasality relied on subjective visual inspection of spectrographic features, which raises concerns about potential experimenter bias and the reliability of their results. Nesbitt and Durvasula (2016), using data from the Buckeye Corpus, found that vowels preceding ambisyllabic consonants were shorter in duration and had lower pitch, aligning more closely with coda contexts—especially in duration. In contrast, Lee and Seo (2019) analyzed both temporal and spectral properties of ambisyllabic consonants and found no significant duration differences between ambisyllabic and non-ambisyllabic intervocalic consonants. Spectral analyses of the lateral /l/, however, revealed that both ambisyllabic and non-ambisyllabic laterals formed a distinct phonetic class, diverging from canonical onset and coda realizations. Their findings challenge phonological assumptions of ambisyllabicity as multiple linkage and suggest that certain intervocalic consonants may function as unique allophones rather than intermediate forms between onset and coda.

There have been few articulatory investigations into ambisyllabicity. Gick (2003) used electromagnetic midsagittal articulography (EMMA) to examine tongue dorsum and tip movements for English /l/ in three across-word-boundary contexts: clear syllable-initial (ha # lotter), syllable-final (hall # hotter), and hypothesized ambisyllabic positions (hall # otter). While tongue dorsum positions showed no significant differences across conditions, the ambisyllabic context exhibited an intermediate degree of tongue tip fronting, suggesting possible gestural blending. However, Gick ultimately interpreted this pattern as more indicative of resyllabified onsets rather than evidence for true ambisyllabicity, due to the absence of statistically robust distinctions.

Lee (2024) examined tongue body lowering across three types of intervocalic laterals in American English: (1) those following a stressed lax vowel and preceding an unstressed vowel (e.g., *color* [ˈkʌlə]); (2) those following a stressed tense vowel and preceding an unstressed vowel (e.g., *collar* [ˈkɑlə]); and (3) those following an unstressed vowel and preceding a stressed vowel (e.g., *collapse* [kəˈlæps]). Tongue body lowering has been widely regarded as a robust articulatory correlate distinguishing the two primary allophones of /l/, with greater lowering associated with the "dark" [ɫ] in coda position and less lowering with the "clear" [l] in onset position (Sproat & Fujimura 1993; Lee-Kim et al. 2013). Using ultrasound imaging, Lee analyzed the articulation of the target intervocalic laterals and compared them to canonical word-initial onset and word-final coda laterals.

Lee's findings revealed no significant differences in tongue body gestures among the three intervocalic lateral types, and neither (1) nor (2) showed an intermediate degree of tongue body lowering. These results suggest that intervocalic laterals do not exhibit articulatory properties indicative of ambisyllabicity. Thus, the ultrasound data failed to provide empirical evidence for ambisyllabicity as an articulatorily distinct phenomenon, but it is merely a theoretical construct within phonological theory.

The present study compares ultrasound tongue images of three intervocalic retroflexes (one ambisyllabic and two non-ambisyllabic contexts) with those of word-initial onset and word-final coda retroflexes. Ambisyllabicity would receive phonetic support if the tongue contours of ambisyllabic retroflexes consistently occupy an intermediate position between onset and coda retroflexes throughout articulation, or if they shift systematically from onset-like to coda-like configurations over time. This analysis investigates whether ambisyllabicity manifests in the articulatory realization of /ɹ/, thereby providing phonetic grounding for a construct that has thus far been primarily theorized within phonological frameworks.

2. Experiments

2.1 Participants

American English speakers typically produce two primary variants of the rhotic /ɹ/: the retroflex and the bunched types. These variants differ in their articulatory configurations—retroflex /ɹ/ is characterized by tongue tip raising accompanied by tongue dorsum lowering, whereas bunched /ɹ/ involves raising of the tongue dorsum with the tongue tip lowered (Espy-Wilson et al. 2000). A preliminary pilot study was conducted to ensure that participants produced the retroflex variant of /ɹ/. Eight native speakers of American English were recruited and participated in an ultrasound recording session, during which they read an English passage containing the rhotic /ɹ/ in various phonological positions like word-initial, word-final, between vowels, etc. Of the eight, six were confirmed to consistently produce the retroflex rather than the bunched variant, and they were subsequently invited to return to the lab for the main experiment. Ultrasound data were collected from them, all of whom were English instructors at a university in Seoul. None of the participants reported any speech or hearing impairments. They received compensation for their participation. The ultrasound recordings were conducted using a probe with a relatively high frequency range (5–10 MHz), which proved unsuitable for two participants with larger body mass. In their recordings, the palate was not visible, and the tongue surface appeared too high in the midsagittal view, occasionally obscuring tongue tip raising. As a result, their data were excluded from the analysis, and consequently, tongue imaging data from four participants were included in the final dataset. The research protocol was reviewed and approved by the Institutional Review Board of the hosting university.

2.2 Stimuli

As outlined in Introduction, the present study restricts its focus to morpheme-internal retroflex segments. Accordingly, retroflexes that occur across word boundaries, such as those illustrated in (1c), are excluded from consideration (Kahn 1976). Within morphemes, Kahn (1976) characterizes a consonant as ambisyllabic when it occurs between syllables and the following syllable is unstressed, as in (1a) and (1b). In contrast, Giegerich (1992) proposes that a consonant may be ambisyllabic only when it appears intervocalically and the preceding syllable contains a stressed lax vowel, as in (2a). His account is more theoretically grounded, drawing on the Onset Maximization Principle and the Weight-Stress Principle (Duanmu 2010). Based on Giegerich's framework, ambisyllabic retroflexes are assumed to be preceded by a stressed lax vowel in this work.

Three distinct types of intervocalic /ɹ/ tokens were selected for analysis: one ambisyllabic and two non-ambisyllabic. (1) Ambisyllabic retroflexes were defined as those occurring after stressed lax vowels, as in /ˈnæ.ɹ.ow/ 'narrow.' (2) The non-ambisyllabic categories (hereafter, non_A) are included when intervocalic retroflexes are preceded by stressed tense vowels, as in /ˈaj.ɹɪʃ/ 'Irish,' and (3) those that are preceded by unstressed vowels (hereafter, non_B), as in /də.ˈɹajv/ 'derive.' Including both types of non-ambisyllabic retroflexes enables a more precise assessment of the role of stress. If retroflexes in conditions (1) and (2) exhibit similar articulatory behavior, the effect is more likely due to stress rather than ambisyllabicity. The stimuli words used in the experiment were listed in Table 1.

Table 1. English word stimuli used for the five retroflex types

Types of retroflex	Words
Ambisyllabic	*arrow, mirror, narrow, terror, barrel*
Non-ambisyllabic A (non_A)	*sorry, sorrow, siren, Irish, taurine*
Non-ambisyllabic B (non_B)	*arrive, surround derive, marine, direct*
Word-initial onset	*rise, read, rest, rust, rat*
Word-final coda	*car, bar, beer, gear, soar*

Word-initial onset and word-final coda retroflexes were included in the analysis to examine articulatory differences in tongue gestures between onset and coda allophones of /ɹ/. While previous empirical studies on American English rhotics have primarily focused on the articulatory distinction between retroflex and bunched variants (Delattre & Freeman 1968; Espy-Wilson & Boyce 1994; Hagiwara 1995; Ong & Stone 1998; Westbury et al. 1998; Guenther et al. 1999; Mielke et al. 2016), these studies highlight that the two variants involve opposing configurations of the tongue tip and dorsum, with both being acoustically characterized by a lowered third formant (F3). In contrast, allophonic variation between onset and coda /ɹ/ has been discussed primarily in relation to lip rounding. It is widely recognized that English /ɹ/ tends to be produced with notable lip rounding in syllable-initial (onset) and stressed position, but not in syllable-final (coda) or unstressed positions (Espy-Wilson et al. (2000). However, articulatory differences—particularly in tongue gestures—between onset and coda retroflexes remain underexplored. The present study addresses this gap by analyzing ultrasound images of the tongue during the production of word-initial and word-final retroflexes.

Five types of retroflexes were used for stimuli; three intervocalic (one ambisyllabic and two non-ambisyllabic) retroflexes and word-initial onset and word-final coda retroflexes. Five real English words for each type were embedded in a carrier sentence 'Please say ______ for me,' and randomized for recordings. The native speakers repeated a list of 25 sentences three times. In total, 900 ultrasound images were successfully splined (5 retroflex types × 5 words × 3 repetitions × 3 time points × 4 speakers)

2.3 Procedure

During the ultrasound recordings, participants were seated and maintained a comfortable posture while reading the sentences. A customized headset (Articulate Instruments Ltd. 2008) was used to secure the ultrasound transducer (or probe) beneath the chin. Ultrasound data were collected using the Articulate Assistant Advanced (AAA) software (Articulate Instruments Ltd. 2012), which also recorded audio signals simultaneously and synchronized them automatically with the ultrasound data.

The target /ɹ/ sounds were annotated based on aligned waveform and spectrogram data. Smoothing splines were manually applied at three temporal points—initial, midpoint, and final—in order to examine whether tongue gestures change over time from coda to onset, or whether a gestural blend of onset and coda variants emerges at a certain time point of the production of ambisyllabic retroflexes. Tongue surface contours were extracted from ultrasound images captured in the midsagittal plane, yielding over 42 coordinate points per contour. Each of the three types of intervocalic retroflexes was compared to onset and coda retroflexes at all three temporal points to determine whether their tongue gestures more closely resembled onset or coda patterns.

The 42 x/y coordinate points were edited by discarding those located to the left of the x-coordinate corresponding to the highest y-value. This was done to facilitate a clearer interpretation of tongue tip elevation at the leftmost extent of the tongue. Ultrasound imaging data of the three types of intervocalic retroflexes were analyzed using generalized additive mixed models (GAMMs) implemented in R (R Core Team 2025) with the *mgcv* package (Wood 2017). Generalized Additive Models (GAMs) and Generalized Additive Mixed Models (GAMMs) are statistical techniques that extends linear regression by allowing non-linear relationships between predictors and response variables through the use of smooth functions. This method is particularly useful for analyzing articulatory data, where complex and non-linear patterns often emerge in time-series measurements. In the present study, the data consisted of two-dimensional coordinates extracted from midsagittal ultrasound tongue images, with the X-axis representing the tongue tip anterior-posterior displacement and the Y-axis representing tongue raising. These coordinate data were analyzed to compare the non-linear tongue contours across different types of intervocalic retroflexes. In the statistical analysis, the predictor variable was the normalized temporal dimension (X), and the response variable was the vertical displacement of the tongue in the midsagittal plane (Y). The primary fixed effect of interest was retroflex type, contrasting onset, coda, and each of intervocalic retroflexes. While the GAMMs evaluated non-linearity of the tongue contour, post-hoc pairwise comparisons with estimated marginal means (*emmeans*; Lenth 2024) were used to test whether each intervocalic retroflex differed significantly from onset and coda retroflexes.

To account for variability not attributable to retroflex type, both random intercepts and slopes for lexical items and speakers were evaluated as random effects through stepwise model comparison. Model comparisons were conducted using likelihood ratio tests, and differences in explained variance were examined. This modeling structure enabled the comparison of non-linear tongue contours across retroflex types while appropriately accounting for measures within words and speakers. Nine models (three retroflex types × three timing points) were identified as the best fit, each incorporating random intercepts and slopes for both Speaker and Item. Results are presented from GAMMs that include X (tongue anterior–posterior position) and Word (retroflex type) as fixed effects, with random intercepts and slopes for Speaker and Item.

2.4 Results

Three intervocalic retroflex were compared with onset and coda retroflexes at initial, mid, and final temporal points to examine whether ambisyllabic retroflexes shift from a coda-like to an onset-like configuration over time, and/or whether they maintain an intermediate degree of articulatory gestures between onset and coda throughout articulation. Furthermore, ambisyllabic retroflexes were compared with non-ambisyllabic intervocalic retroflexes to examine whether they exhibited distinct tongue trajectories.

Figure 1 shows GAMM-fitted curves of ambisyllabic, onset, and coda retroflex at the initial temporal point. For the comparison, the generalized additive mixed model (GAMM) analysis and post-hoc pairwise comparisons were executed. The results of the GAMM were presented in Table 2. All smooth terms were highly significant ($p < .001$), indicating systematic differences in tongue trajectories across ambisyllabic, coda, and onset retroflexes. Random effects for both word and speaker were also significant, confirming that lexical variation and, more dominantly, inter-speaker anatomical differences contributed substantially to the model fit. Post-hoc pairwise comparisons using estimated marginal means (emmeans) in R further revealed that ambisyllabic retroflexes showed significantly greater tongue height as a function of X than coda retroflexes (estimate = 0.097, $p < .0001$), but significantly lower tongue height than onset retroflexes (estimate = –0.079, $p = .0002$). Taken together, these results demonstrate that ambisyllabic retroflexes occupy an intermediate articulatory position between onset and coda retroflexes.

Table 2. Results of GAMM (ambisyllabic retroflex at initial point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordambi_ini	7.85	8.65	125.30	< .001 ***
s(X):wordcoda_ini	7.54	8.45	136.40	< .001 ***
s(X):wordonset_ini	7.89	8.66	193.30	< .001 ***
s(Item)	11.89	15.00	16709.00	< .001 ***
s(Speaker)	2.96	3.00	116556.70	< .001 ***
s(X,Speaker)	2.99	3.00	323157.10	< .001 ***
s(X,Item)	12.46	15.00	19989.80	< .001 ***

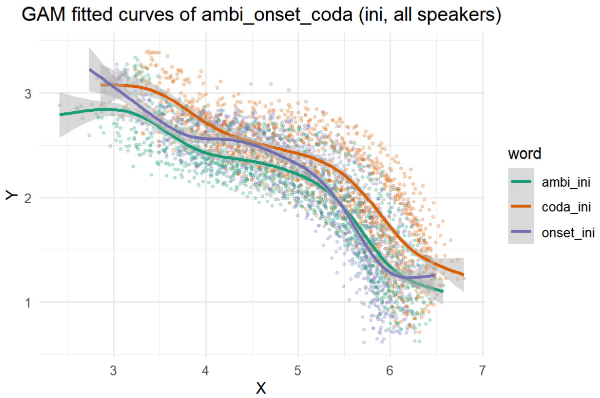


Figure 1. GAM-fitted curves for ambisyllabic retroflexes compared with onset and coda retroflexes at initial time point, collapsed across four speakers

The GAMM fitted curve of ambisyllabic retroflexes at mid point of articulation is compared with those of onset and coda retroflexes in Figure 2. The GAMM analysis revealed in Table 3 that tongue position trajectories varied significantly across retroflex types in mid timing point. Smooth terms for ambisyllabic, onset and coda retroflexes were all highly significant (all $p < .001$), indicating distinct nonlinear tongue contour patterns. The factor *s(Item)*, representing lexical items, did not reach statistical significance ($p = .073$), suggesting that item-level variability contributed little to the overall tongue trajectory differences. By contrast, random smooths for speakers were significant ($p < .001$), confirming substantial between-speaker variability in articulatory patterns. Post-hoc pairwise comparisons revealed that the tongue contours of ambisyllabic retroflexes were significantly higher along the tongue anterior-posterior position than those of onset retroflexes and significantly lower than those of coda retroflexes at their mid point.

Table 3. Results of GAMM (ambisyllabic retroflex at mid point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordambi_mid	8.22	8.84	132.40	< .001 ***
s(X):wordcoda_mid	7.63	8.39	153.90	< .001 ***
s(X):wordonset_mid	8.21	8.81	244.00	< .001 ***
s(Item)	5.49	14.00	813.00	.073 .
s(Speaker)	2.98	3.00	192523.10	< .001 ***
s(X,Speaker)	2.99	3.00	204481.30	< .001 ***
s(X,Item)	12.99	14.00	664.50	< .001 ***

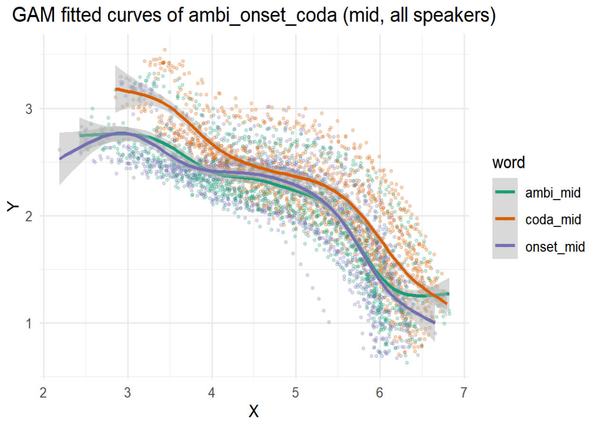


Figure 2. GAM-fitted curves for ambisyllabic retroflexes compared with onset and coda retroflexes at initial time point, collapsed across four speakers.

At the final time point of retroflex articulation, the tongue trajectories of ambisyllabic, onset, and coda retroflexes are illustrated in Figure 3. The results of the GAMM analysis corresponding to this time point are summarized in Table 4. Smooth terms for ambisyllabic, onset and coda retroflexes were all highly significant (all p < .001). Lexical items also contributed significantly (p < .001), alongside strong speaker-level effects (p < .001) and interactions (both p < .001) indicating both inter-speaker and item-specific variability.

Table 4. Results of GAMM (ambisyllabic retroflex at final point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordambi_fin	7.44	8.24	234.40	< .001 ***
s(X):wordcoda_fin	7.92	8.66	203.40	< .001 ***
s(X):wordonset_fin	8.09	8.75	278.30	< .001 ***
s(word2)	10.11	14.00	5225.40	< .001 ***
s(Speaker)	2.98	3.00	274440.40	< .001 ***
s(X,Speaker)	2.99	3.00	310574.20	< .001 ***
s(X,word2)	11.41	14.00	3833.90	< .001 ***

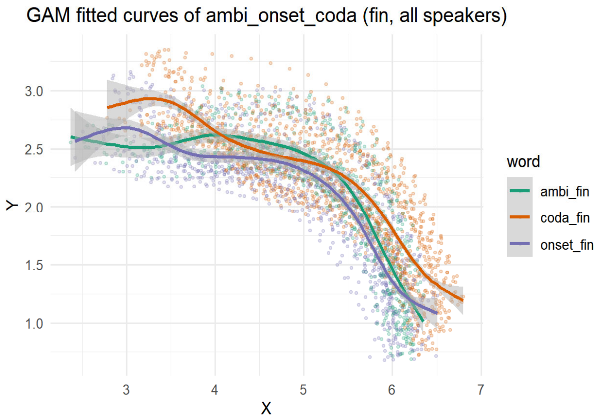


Figure 3. GAM-fitted curves for ambisyllabic retroflexes compared with onset and coda retroflexes at final time point, collapsed across four speakers

Post-hoc pairwise comparisons for Figure 3 further clarified the status of ambisyllabic retroflexes. Ambisyllabic retroflexes were significantly higher than coda retroflexes ($p < .0001$). However, ambisyllabic retroflexes did not differ significantly from onset retroflexes ($p = .852$), showing convergence with onset articulations at the final temporal point. Unlike at initial and mid time points, ambisyllabic retroflexes are articulated similar to onset retroflexes at final timing point of articulation. Taken together with the results from the initial and medial time points, the tongue contours of ambisyllabic retroflexes appear intermediate between onset and coda retroflexes, but converge with onset retroflexes toward the end of the articulation.

Moving onto non-ambisyllabic retroflexes, recall the non_A type of retroflexes. They are intervocalic retroflexes preceded by a stressed tense vowel as in *siren*, so they have been traditionally assumed to be onset of the following syllable in phonological theory. Their GAMM fitted-curve at initial time point of articulation is compared with those of onset and coda retroflexes in Figure 4. The statistical results of the GAMM analyses are presented in Table 5.

Table 5. Results of GAMM (non_A retroflex at initial point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordnon_A_ini	8.252	8.849	251.5	< .001 ***
s(X):wordcoda_ini	7.918	8.699	182.6	< .001 ***
s(X):wordonset_ini	8.163	8.788	255.2	< .001 ***

s(Item)	11.937	14.000	46604.5	< .001 ***
s(Speaker)	2.986	3.000	257117.8	< .001 ***
s(X,Speaker)	2.994	3.000	835198.9	< .001 ***
s(X,Item)	12.188	14.000	66469.8	< .001***

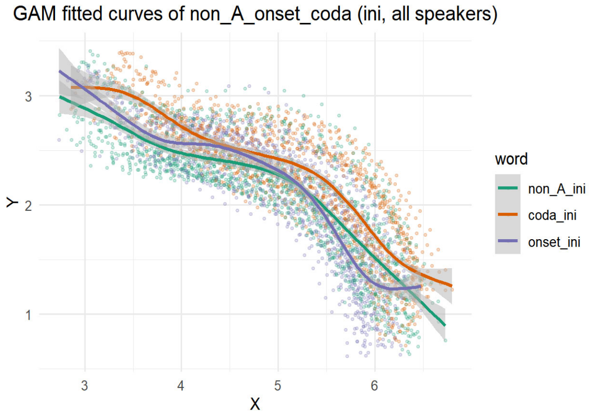


Figure 4. GAM-fitted curves for non_A retroflexes compared with onset and coda retroflexes at initial time point, collapsed across four speakers

The GAMM results demonstrated in Table 5 that tongue position along the anterior –posterior axis (X) varied significantly for all three intervocalic retroflex types (non-ambisyllabic, coda, and onset), with the high estimated degrees of freedom (edf ≈ 8), indicating complex non-linear trajectories. In addition, the random factors of lexical items and speakers and exerted a strong effect, demonstrating that tongue trajectories differed systematically depending on the individual speaker as well as the specific lexical item. Post-hoc pairwise comparisons revealed that the overall tongue height along the horizontal dimension of the tongue of non_A intervocalic retroflexes were significantly higher than those of onset (estimate = -0.071, $p<.001$) and significantly lower that those of coda (estimate = 0.114, $p<.001$).

When the tongue contours of non_A retroflexes at mid timing point were smoothly fitted and compared with onset and coda retroflexes in Figure 5, and the GAMM results are shown in Table 6. The GAMM analyses showed significant non-linear effects of tongue position (X) across all three, non_A, coda, and onset, retroflexes as indicated by the high edf values (≈ 8) and large F-statistics. Both random intercepts and slopes exert significant effects, indicating considerable item-specific variability and inter-speaker differences in how the tongue moved along the anterior–posterior axis.

Post-hoc pairwise comparisons further revealed that non-ambisyllabic retroflexes were significantly lower in tongue position than coda retroflexes (p < .001), but significantly higher than onset retroflexes (p < .001).

Table 6. Results of GAMM (non_A retroflex at mid point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordnon_A_ini	8.252	8.849	251.5	< .001 ***
s(X):wordcoda_ini	7.918	8.699	182.6	< .001 ***
s(X):wordonset_ini	8.163	8.788	255.2	< .001 ***
s(Item)	11.937	14.000	46604.5	< .001 ***
s(Speaker)	2.986	3.000	257117.8	< .001 ***
s(X,Speaker)	2.994	3.000	835198.9	<. 001 ***
s(X,Item)	12.188	14.000	66469.8	< .001 ***

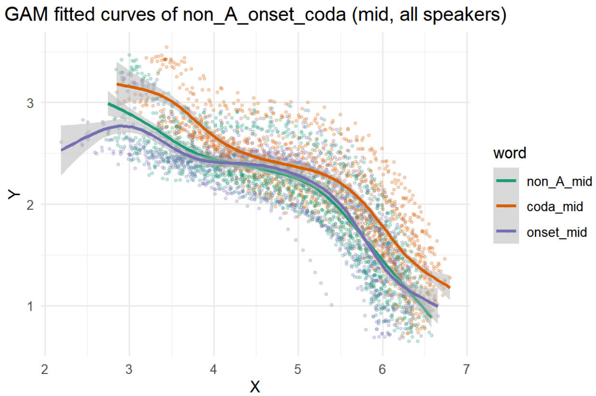


Figure 5. GAM-fitted curves for non_A retroflexes compared with onset and coda retroflexes at mid time point, collapsed across four speakers

Next, the GAMM-fitted curves of non_A, onset, and coda retroflexes at the final time point are displayed in Figure 6 and their statistical results are reported in Table 7. The GAMM results at the final time point demonstrated that tongue position along the anterior–posterior axis (X) varied significantly for all three retroflex types (non-ambisyllabic, coda, and onset), each showing complex non-linear trajectories (edf ≈ 8, p < .001). Random factors such as lexical items and speakers exerted significant effects on tongue trajectories. According to post-hoc pairwise comparisons, non_A intervocalic retroflexes were significantly lower that coda retroflexes (p < .001), but did not differ significantly from onset retroflexes (p = .157). Their tongue movements were intermediate between those of onset and coda retroflexes at the initial and medial stages of articulation, but shifted toward an onset-like configuration at the final stage.

Table 7. Results of GAMM (non_A retroflex at final point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordnon_A_fin	7.972	8.729	253.2	< .001 ***
s(X):wordcoda_fin	8.167	8.805	253.6	< .001 ***
s(X):wordonset_fin	8.344	8.858	342.2	< .001 ***
s(Item)	12.387	14.000	105720.0	< .001 ***
s(Speaker)	2.987	3.000	597787.2	< .001 ***
s(X,Speaker)	2.994	3.000	874838.4	< .001 ***
s(X,Item)	12.608	14.000	112095.2	< .001 ***

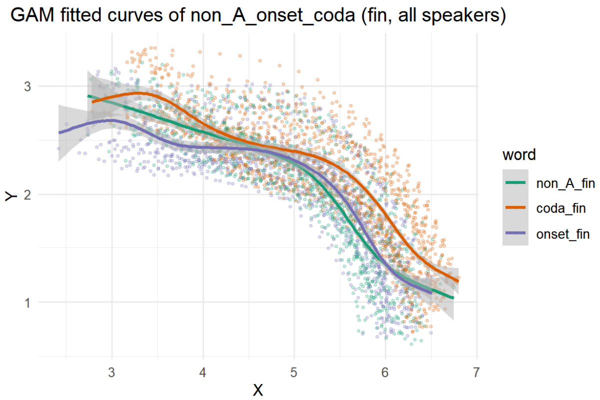


Figure 6. GAM-fitted curves for non_A retroflexes compared with onset and coda retroflexes at final time point, collapsed across four speakers

The last retroflex type that was observed in this experiment was non_B intervocalic retroflexes, where retroflexes were preceded by an unstressed vowel and followed by a stressed vowel as in *arrive*. Figure 7 shows the GAMM-fitted tongue contours of non_B retroflexes, compared with those of onset and coda retroflexes. The GAMM results at the initial time point showed that tongue height (Y) varied significantly with tongue anterior-posterior position (X) across non-ambisyllabic, coda, and onset retroflexes (edf ≈ 8, $p < .001$) as presented in Table 8 . Significant effects were also found for lexical items and speakers, indicating robust variability across words and individuals. Post-hoc pairwise comparisons further demonstrated that non_B retroflexes were significantly lower than coda retroflexes (estimate = -.124, $p < .001$) and significantly higher than onset retroflexes (estimate = .062, p = 0.002). This indicates that non_B intervocalic retroflexes do not align with either onset or coda retroflexes at the initial point of articulation, but instead exhibit intermediate tongue movement between onset and coda.

Table 8. Results of GAMM (non_B retroflex at initial point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordnon_B_ini	8.375	8.894	292.1	< .001 ***
s(X):wordcoda_ini	7.835	8.645	170.0	< .001 ***
s(X):wordonset_ini	8.082	8.749	239.5	< .001 ***
s(Item)	11.480	14.000	19304.8	< .001 ***
s(Speaker)	2.985	3.000	157119.1	< .001 ***
s(X,Speaker)	2.993	3.000	413250.0	< .001 ***
s(X,Item)	11.745	14.000	21310.1	< .001 ***

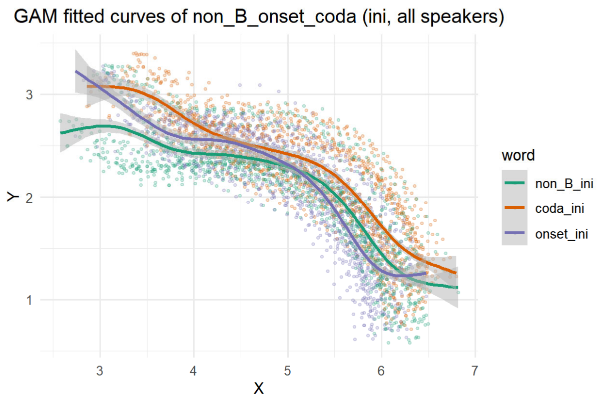


Figure 7. GAM-fitted curves for non_B retroflexes compared with onset and coda retroflexes at initial time point, collapsed across four speakers

Figure 8 illustrates the GAMM-fitted smoothing curves of non-B retroflexes in comparison with onset and coda retroflexes, while Table 9 summarizes the corresponding statistical results. At the medial time point, tongue height (Y) again varied significantly as a function of X (tongue anterior_posterior position) across all three retroflex types (edf ≈ 8, $p < .001$), with complex non-linear trajectories. Lexical items (edf ≈ 9, $p < .001$) and speakers (edf ≈ 3, $p < .001$) also exerted strong effects, indicating systematic differences across individuals and lexical contexts. Post-hoc pairwise comparisons indicated that non_B retroflexes were lower than coda ($p < .001$), but did not differ from onset retroflexes ($p = .0194$) at mid timing point of articulation.

Table 9. Results of GAMM (non_B retroflex at mid point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordnon_B_mid	8.256	8.852	312.6	< .001 ***
s(X):wordcoda_mid	7.887	8.564	202.0	< .001 ***
s(X):wordonset_mid	8.392	8.887	322.7	< .001 ***
s(Item)	9.493	14.000	8947.1	< .001 **
s(Speaker)	2.990	3.000	379795.9	< .001 ***
s(X,Speaker)	2.994	3.000	337888.1	< .001 ***
s(X,Item)	12.724	14.000	4789.5	< .001 ***

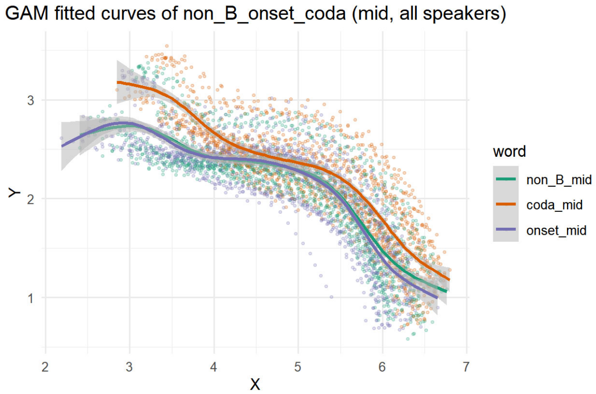


Figure 8. GAM-fitted curves for non_B retroflexes compared with onset and coda retroflexes at mid time point, collapsed across four speakers

Finally, the smoothed tongue contours of non_B retroflexes, compared with those of onset and coda retroflexes, are shown in Figure 10, while Table 9 reports the results of the GAMM analysis. At the final time point, the GAMM results confirmed that tongue height (Y) varied significantly as X for non-A, coda, and onset retroflexes (edf ≈ 8, $p < .001$). Substantial variability was also found across lexical items and speakers. According to post-hoc pairwise comparisons, non_B retroflexes were higher than coda retroflexes ($p < .001$), but did not differ from onset retroflexes ($p = .757$). Overall, the tongue configurations of non_B retroflexes were similar to those of onset retroflexes at both the medial and final time points. This pattern contrasts with ambisyllabic and non_A retroflexes, which shifted toward an onset-like tongue position only at the final stage of articulation.

Table 10. Results of GAMM (non_B retroflex at final point)

Smooth term	edf	Ref.df	F	p-value
s(X):wordnon_A_fin	7.940	8.712	255.2	< . 001 ***
s(X):wordcoda_fin	8.146	8.796	255.2	< . 001 ***
s(X):wordonset_fin	8.320	8.848	343.6	< . 001 ***
s(Item)	12.384	14.000	105195.9	< . 001 ***
s(Speaker)	2.985	3.000	594666.2	< . 001 ***
s(X,Speaker)	2.992	3.000	869122.5	< . 001 ***
s(X,Item)	12.607	14.000	110798.7	< . 001 ***

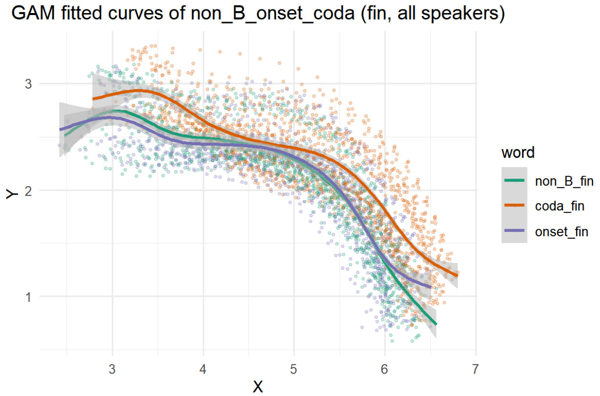


Figure 9. GAM-fitted curves for non_B retroflexes compared with onset and coda retroflexes at final time point, collapsed across four speakers

2.5 Discussion

Three intervocalic retroflexes have been compared with onset and coda retroflexes in their tongue contours and are summarized in Table 11. Ambisyllabic retroflexes which preceded by a stressed lax vowel were articulated with the overall tongue position intermediate between onset and coda retroflexes at initial and midpoint of articulation, but ended with onset-like a tongue contour. The same articulatory shift of the tongue contours occurred to the non_A type (non-ambisyllabic intervocalic retroflexes which preceded by a stressed tense vowel). This means that the overall tongue height along the tongue front-back position was higher than onset and lower than coda retroflexes until the mid time point of articulation. The tongue height, however, lowered to that of onset by the time the retroflexes ended and the following vowels were about to begin. The non_B type, on the other hand, showed distinct patterns of tongue movements. When the intervocalic retroflexes are followed by a stressed vowel, the tongue position intermediate between onset and coda retroflexes was initiated, but changed onset-like as early as at mid point. The tongue contour similar to onset continued to the final stage of articulation.

Table 11. Summary of tongue contours

Types	time point	tongue contour comparisons	tongue contour
Ambi	initial	onset < ambi < coda	intermediate
	mid	onset < ambi < coda	intermediate
	final	onset ≒ ambi < coda	onset-like
Non_A	initial	onset < ambi < coda	intermediate
	mid	onset < ambi < coda	intermediate
	final	onset ≒ ambi < coda	onset-like
Non-B	initial	onset < ambi < coda	intermediate
	mid	onset ≒ ambi < coda	onset-like
	final	onset ≒ ambi < coda	onset-like

Recall the two research questions posed earlier: (1) Are ambisyllabic retroflexes intermediate in temporal and/or spatial dimensions? and (2) Do ambisyllabic retroflexes, unlike non-ambisyllabic intervocalic retroflexes, exhibit a temporal progression from coda-like to onset-like tongue gestures, or do they instead maintain an intermediate configuration throughout articulation? The present findings provide evidence for the former. Ambisyllabic retroflexes demonstrated a temporal shift in tongue posture, progressing from an intermediate configuration toward an onset-like articulation, thereby ruling out the possibility of a stable intermediate position. With respect to the first research question, however, the results revealed that such temporal shifts were not unique to ambisyllabic retroflexes. Both non-ambisyllabic categories A and B displayed comparable gestural transitions, moving from intermediate to onset-like tongue shapes. In particular, non_A retroflexes patterned nearly identically with ambisyllabic retroflexes, exhibiting blended onset- and coda-like characteristics up to the midpoint of articulation. These results challenge the phonological account of ambisyllabicity for English consonants as proposed by Giegerich (1992), since no exclusive articulatory signature was observed for ambisyllabic retroflexes. Rather, the current findings align with Lee (2024), who reported a lack of phonetic evidence for ambisyllabicity in American English laterals based on ultrasound tongue imaging.

Turning to the non_B type of intervocalic retroflexes, the tongue contour exhibited intermediate behavior only at the initial time point, whereas onset-like articulatory gestures predominated at both the medial and final stages of articulation. This pattern contrasts with the temporal dynamics observed for ambisyllabic and non_A retroflexes. Recall Kahn's (1976) assertion that a consonant may be considered ambisyllabic when it is followed by an unstressed vowel. If this account holds, one would expect ambisyllabic and non_A retroflexes to display parallel retroflexion patterns, insofar as both occur in an identical prosodic context: a stressed vowel preceding and an unstressed vowel following. The present results are consistent with this prediction. While the study was initially designed to critically assess Giegerich's (1992) phonological framework of ambisyllabicity, the articulatory evidence observed here lends stronger support to Kahn's (1976) conception of ambisyllabicity.

When a retroflex consonant is followed by a stressed vowel, the Onset Maximization Principle, which is assumed to apply universally across languages, tends to align the preceding consonant with the onset position, thereby producing onset-like articulation by the midpoint of the gesture. In contrast, when the following vowel is unstressed, the intervocalic retroflex appears to resist such alignment. Rather than fully patterning with onset retroflexes, it remains intermediate until the midpoint of articulation. Consonant–vowel (CV) bondage, understood here as an implementation of the Onset Maximization Principle, appears to be fully achieved only under conditions of stress. This suggests that the Onset Maximization Principle is weakened in the context of an unstressed vowel, which lacks the prosodic strength necessary to trigger full onset alignment.

The results indicate that ambisyllabic and non_A retroflexes are more likely to occupy an intermediate position between onset and coda retroflexes, whereas non_B retroflexes are more strongly onset-like, with their midpoint reflecting the temporal target at which gestures are achieved. These findings converge with Lee and Seo (2019), who examined two types of intervocalic consonants (ambisyllabic and non_A) and compared their acoustic features with those of word-initial onsets and word-final codas. Their results similarly demonstrated that ambisyllabic consonants do not possess exclusive acoustic properties. Instead, both ambisyllabic and non_A consonants exhibited an intermediate status between onset and coda categories, suggesting that they may be regarded as constituting an independent allophonic category distinct from onset and coda allophones.

Taken together, the present findings, along with those of Lee and Seo (2019) and

Lee (2024), suggest that ambisyllabicity lacks consistent phonetic substance and may function primarily as a theoretical construct. Historically, the notion of ambisyllabicity was motivated by the Weight–Stress Principle (Giegerich 1992), according to which an intervocalic consonant following a stressed lax vowel is parsed into both syllables so as to render the stressed syllable heavy. However, the ultrasound evidence reported here revealed no detectable effect of vowel weight and stress interaction on the articulation of intervocalic retroflexes. While ambisyllabicity may continue to serve as a useful analytical construct in phonological theory, its phonetic grounding appears empirically unsupported. Moreover, the present results indicate that intervocalic retroflexes may be regarded as constituting an independent allophone, articulatorily distinct from onset and coda allophones, when a retroflex is preceded by a stressed vowel and followed by an unstressed vowel.

3. Conclusion

This study set out to examine whether ambisyllabic retroflexes in American English occupy an intermediate articulatory position between onset and coda realizations and whether they exhibit distinct articulatory dynamics compared to non-ambisyllabic retroflexes. Ultrasound evidence demonstrated that ambisyllabic retroflexes, along with non_A retroflexes, consistently began with tongue configurations intermediate between onset and coda positions but shifted toward onset-like gestures by the final stage of articulation, while non_B retroflexes displayed an earlier onset-like alignment at the midpoint. These findings indicate that the intermediate status of ambisyllabic retroflexes is not unique, but shared with other intervocalic retroflexes, thereby undermining claims of a stable phonetic basis for ambisyllabicity as proposed in Giegerich's (1992) phonological framework. Instead, the results provide stronger support for Kahn's (1976) view that prosodic context—particularly the presence of an unstressed following vowel—shapes the articulatory realization of intervocalic consonants. Taken together, the present findings, in line with Lee and Seo (2019) and Lee (2024), suggest that ambisyllabicity functions primarily as a theoretical construct without consistent phonetic grounding, and further point to the possibility that intervocalic retroflexes preceded by a stressed vowel and followed by an unstressed vowel may be regarded as constituting an independent allophone distinct from onset and coda allophones.

References

Articulate Instruments Ltd. 2008. Ultrasound Stabilisation Headset Users.

Articulate Instruments Ltd. 2012. Articulate Assistant Advanced User Guide: Version 2.14. Edinburgh, UK: Articulate Instruments Ltd. Manual: Revision 1.4. Edinburgh, UK: Articulate Instruments Ltd.

Delattre, Pierre and Donald C. Freeman. 1968. A dialect study of American r's by x-ray motion picture. *Linguistics* 44(6): 29-68.

Derwing, Bruce. 1992. A 'pause-break' task for eliciting syllable boundary judgments from literate and illiterate speakers: Preliminary results for five diverse languages. *Language and Speech* 35(1-2): 219-235.

Duanmu, San. 2010. Onset and the weight-stress principle in English. In *Linguistic essays in honor of professor Tsu-Lin Mei* on his 80th Birthday, 1-31. Commercial Press, Beijing.

Durvasula, Karthik and Ho-Hsin Huang. 2017. Word-internal "ambisyllabic" consonants are not multiply-linked in American English. *Language Sciences* 62: 17-36.

Eddington, David and Dirk Elzinga. 2008. The phonetic context of flapping in American English: Quantitative evidence. *Language and Speech* 51(3): 245-266.

Espy-Wilson, Carol, Suzanne Boyce, Michel Jackson, Shrikanth Narayanan, and Abeer Alwan. 2000. Acoustic modeling of American English /r/. *Journal of the Acoustical Society of America* 108(1): 343-356.

Espy-Wilson, Carol and Suzanne Boyce. 1994. Acoustic differences between "bunched"and "retroflex" variants of American English /r/. *The Journal of the Acoustical Society of America 95.5_Supplement*: 2823-2823.

Fallows, Deborah. 1981. Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics* 17(2): 309-317.

Gick, Bryan. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In John Local, Richard Ogden, and Rosalind Temple (eds.), *Phonetic interpretation: Papers in laboratory phonology VI*. Cambridge, Cambridge University Press.

Giegerich, Heinz. 1992. *English phonology*. Cambridge: Cambridge University Press.

Guenther, Frank, Carol Espy-Wilson, Suzanne Boyce, Melanie Matthies, Majid Zandipour, and Joseph Perkell. 1999. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *Journal of the Acoustical Society of America* 105(5): 2539–2541.

Gussenhoven, Carlos. 1986. English plosive allophones and ambisyllabicity. *Gramma* 10(2): 119-141.

Hagiwara, Robert. 1995. Acoustic realizations of American /ɹ/ as produced by women and men.

UCLA Phonetics Laboratory Working Papers 90: 2-187.

Hayes, Bruce. 2011. *Introductory phonology*. West sussex, John Wiley & Sons.

Ishikawa, Keiichi. 2002. Syllabification of intervocalic consonants by English and Japanese Speakers. *Language and Speech* 45(4): 355-385.

Kahn, Daniel. 1976. *Syllable-based generalizations in English phonology*. PhD Dissertation. Messachusetts Institute of Technology.

Lee, Joo-Kyeong. 2024. An articulatory examination of ambisyllabicity: Focusing on ultrasound images of American English laterals. *Studies in Phonetics, Phonology, and Morphology* 31(1): 91-110.

Lee, Joo-Kyeong and Yuhyeon Seo. 2019. A phonetic examination of phonological ambisyllabicity: Focusing on temporal and spectral characteristics. *Linguistic Research* 36(1): 91-110.

Lee-Kim, Sang-Im, Lisa Davidson, and Sangjin Hwang. 2013. Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology* 4(2): 475-511.

Lenth, Russell. 2024. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.10.0. https://CRAN.R-project.org/package=emmeans.

Mielke, Jeff, Adam Baker, and Diana Archangeli. 2016. Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Linguistic Society of America* 91(1): 101-140.

Nesbitt, Monica and Karthik Durvasula. 2016. Ambisyllabic consonants are codas: Evidence from a syllable tracking task. Paper presented at Linguistic Society of America 90, Washington, January 7-10.

Ong, Darryl and Maureen Stone. 1998. Three-dimensional vocal tract shapes in /r /and /l/: A study of MRI, ultrasound, electropalatography, and acoustics. *Phonoscope* 1(1): 1-13.

R Core Team. 2025. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Sproat, Richard and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21(3): 291-311.

Treiman, Rebecca and Andrea Zukowski. 1990. Toward an understanding of English syllabification. *Journal of Memory and Language* 29(1): 66-85.

Treiman, Rebecca and Catalina Danis. 1988. Syllabification of intervocalic consonants. *Journal of Memory and Language* 27(1): 87-104.

Treman, Rebecca, Judith Bowey, and Derrick Bourassa. 2002. Segmentation of spoken words into syllables by English-speaking children. *Journal of Experimental Child Psychology* 83(3): 213-238.

Westbury, John R., Michiko Hashi, and Mary J. Lindstrom. 1998. Differences among speakers in lingual articulation for American English /ɹ/. *Speech Communication* 26(3): 203-226.

Wood, Simon N. 2017. *Generalized additive models: An introduction with R*. 2nd edn. Boca Raton, FL: Chapman & Hall/CRC. https://doi.org/10.1201/9781315370279.

Zamuner, Taniaand and Diane Ohala. 1999. Preliterate children's syllabification of intervocalic consonants. In the *Proceedings of the 23rd annual Boston Conference on Language Development*, 753-763. Somerville MA: Cascadilla Press.

Joo-Kyeong Lee
Professor
Department of English Language and Literature
University of Seoul
163 Seoulsiripdaero, Dongdaemun-gu,
Seoul, 02504, Korea
E-mail: jookyeong@uos.ac.kr

Received: 2025. 08. 11.
Revised: 2025. 09. 26.
Accepted: 2025. 10. 20.