PROSODIC STRUCTURAL EFFECTS ON COARTICULATORY VOWEL NASALIZATION IN AUSTRALIAN ENGLISH IN COMPARISON TO AMERICAN ENGLISH

Hyunjung Joo[1], Jiyoung Jang[2], Sahyang Kim[3], Taehong Cho[1], and Anne Cutler[4]

Hanyang Institute for Phonetics and Cognitive Sciences of Language (HIPCS), Hanyang University[1], University of California, Santa Barbara[2], Hongik University[3], MARCS Institute, ARC Centre of Excellence for the Dynamics of Language, Western Sydney University[4]
hyunjung.joo.ling@gmail.com, jiyoungljang@gmail.com, sahyang@gmail.com, tcho@hanyang.ac.kr, a.cutler@westernsydney.edu.au

ABSTRACT

This study investigates effects of prosodic factors (prominence, boundary) on coarticulatory V-nasalization in Australian English (AusE) in CVN and NVC in comparison to those in American English (AmE). As in AmE, prominence was found to lengthen N, but to reduce V-nasalization, enhancing N's nasality and V's orality, respectively (paradigmatic contrast enhancement). But the prominence effect in CVN was more robust than that in AmE. Again similar to findings in AmE, boundary induced a reduction of N-duration and V-nasalization phrase-initially (syntagmatic contrast enhancement), and increased the nasality of both C and V phrase-finally. But AusE showed some differences in terms of the magnitude of V nasalization and N duration. The results suggest that the linguistic contrast enhancements underlie prosodic-structure modulation of coarticulatory V-nasalization in comparable ways across dialects, while the fine phonetic detail indicates that the phonetics-prosody interplay is internalized in the individual dialect's phonetic grammar.

Keywords: prosodic structure, coarticulation, vowel nasalization, Australian English, American English.

1. INTRODUCTION

Coarticulation is an inevitable low-level phonetic process that underlies connected speech across languages, entailing cross-linguistic similarities in phonetic implementation [10]. It is, however, also known to be conditioned by various other higher-order linguistic structures, which, as they vary across languages, engender language specificity in fine phonetic detail [1,3]. One such higher-order linguistic structure is prosodic structure which modulates phonetic implementation of speech segments in a language-specific way [4,8]. Recent studies [5,6] have indeed shown that coarticulatory V-nasalization operates in reference to the prosodic structure in which segments occur. For example, results of an acoustic study on V-nasalization in CVN and NVC in American English (AmE) [5] suggest that the seemingly low-level V-coarticulation is fine-tuned by prosodic-structural factors (i.e., boundary strength and prominence), leading to enhancement of different kinds of linguistic contrasts (syntagmatic vs., paradigmatic), depending on the source of prosodic strengthening (boundary vs. prominence).

The AmE results of [5] indicated that under the focus-induced prominence in both CVN and NVC, N duration was lengthened, enhancing N's nasality, but V showed coarticulatory resistance to nasalization, enhancing V's orality. Crucially, the coarticulatory resistance effect was pervasive throughout the vowel, suggesting that it is not a mere outcome of a low-level process, but something controlled by the speaker in reference to the paradigmatic contrast system of the language. As for boundary-related effects, in domain-initial position (#NVC), boundary strength acted to *decrease* N duration and also *reduce* coarticulation with the following vowel. In domain-final position (CVN#), in contrast, N duration was *lengthened* phrase-finally, and at the same time V nasalization *increased* throughout the vowel. The initial effects enhance syntagmatic CV contrast—i.e., the reduced duration of N enhances C's consonanality which, together with reduced nasalization of V, contributes to the CV distinction. The final effects increase coarticulatory propensity and are taken to stem from a general phrase-final articulatory weakening process, which loosens the articulatory linkage of the oral constriction and the velum lowering gestures.

Based on these results in AmE, [5] suggested that the phonetics-prosody interface as reflected in coarticulatory V nasalization must be internalized in the phonetic grammar of each language. This opens the possibility of both cross-linguistic and cross-dialectal variation in the way that coarticulatory processes are instantiated.

As a test of such variation, the present study extends [5]'s study in AmE to another variety of English, Australian English (AusE). By employing similar methods, the present study will deliver results that enable cross-dialectal comparisons across closely matched data sets. Such comparisons will shed light on the relative universality versus language-specificity of the phonetics-prosody interface that underlies coarticulatory nasalization of vowels.

2. METHOD

2.1. Participants and Recording

Fourteen native speakers of Australian English (10F and 4F) participated in this study. All were born and raised in Australia, aged from 20 to 30. 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 kHz. AmE data was based on [5] for the comparison with AusE data. Both AmE and AusE data were collected with an identical experimental setup.

2.2. Speech Materials

Eight test words were used, in a CVN context (*pal**m***, *bo**mb***, *te**n***, *de**n***) or NVC (***m**op*, ***m**ob*, ***n**et*, ***N**ed*) context. They included only non-high vowels such as /ɛ, ɑ/ to avoid overlap between the first formants (F1) of the vowel and the nasal peak (P0). A further 16 words in an oral (CVC) context were included for an oral baseline condition as well as to induce different contrastive focus conditions (phonological focus, lexical focus and no focus) in the prompt sentences (see below). These words were embedded in carrier sentences in which Boundary (IP/Wd) and Focus (LexFOC, PhonFOC, and UnFOC) were systematically manipulated, as laid out in Table 1.

Table 1: Example sentences with a CVN word for each condition of boundary and focus. Targets are underlined and focused words are in bold.

Condition		Example sentences
#= IP	Phon FOC	A: Were you supposed to write **BOB**? B: No. I was supposed to write **BOMB** #, wasn't I?
	Lex FOC	A: Were you supposed to write **WAR**? B: No. I was supposed to write **BOMB** #, wasn't I?
	No FOC	A: Were YOU supposed to write **bomb**? B: No. JOHN was supposed to write **bomb #**, wasn't he?
#= Wd	Phon FOC	A: Did you write 'say **BOB** fast again'? B: No. I write 'say **BOMB** # fast again.'
	Lex FOC	A: Did you write 'say **WAR** fast again'? B: No. I write 'say **BOMB** # fast again.'
	No FOC	A: Did you write 'say **bomb** FAST again'? B: No. I write 'say **bomb** # SLOWLY again.'

2.3. Procedure

The participants (Speaker B) were instructed to read out the second sentence (see Table 1) in a mini dialogue in response to the pre-recorded prompt sentences of a native AusE female speaker (Speaker A). Speech rate of the prompt sentences was comparable in both AusE and AmE data, which helped maintaining similar speech rates of the experimental sentences across the languages. To obtain different types of focus, speakers were asked to make contrast between words in bold in Sentences A and B, which induced corrective lexical contrastive focus (e.g., *WAR vs. BOMB*) or corrective phonological contrastive focus on N (e.g., *BOB vs. BOMB*). To obtain different boundary types, an IP boundary after a test word was obtained with a following tag question as in Table 1; and an IP boundary before a test word was induced by placing the words "Not exactly" before the test word. Finally, the Wd boundary was induced by placing the test word midway in a short phrase (e.g., 'say TARGET fast again'). The test sentences were given in a randomized order with 4 repetitions. A total of 2688 sentences were recorded: 2 syllable positions (#NVC vs. CVN#) x 4 test words x 3 focus types (PhonFOC vs. LexFOC vs. NoFOC) x 2 boundaries (IP vs. Wd) x 4 repetitions x 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. Measurements

N duration was measured from the onset to the offset of nasal energy (murmur) as observed in the spectrogram. In the case of V-nasalization, A1-P0 values (A1=amplitude of F1; P0=nasal peak) were extracted using a Praat script [11]; the lower the A1-P0 value, the greater the nasalization. A1-P0 values were obtained at two absolute timepoints (25ms and 50ms from N into the vowel) and at three relative timepoints (25%, 50%, 75% of the vowel). The absolute measures were to examine whether the coarticulatory process would be a time-locked phenomenon, and the relative measures to examine to what extent the coarticulatory process would be pervasive throughout the vowel as a process that may be considered to be under the speaker control.

2.5. Statistical Analysis

Repeated Measures Analysis of Variance (RM ANOVAs) were conducted to examine the effects of prosodic factors on two dependent variables: N-duration and V-nasality (A1-P0 z-score). Within-subject factors were Focus (PhonFOC vs. LexFOC vs. NoFOC), Boundary (IP vs. Wd) and Time (Relative: 25%, 50%, 75%; Absolute: 25ms, 50ms). In addition, for the comparison of AusE data with AmE data, a between-subject factor, Dialect Group (AusE vs. AmE) factor was included. When interactions were observed among factors, one-way ANOVAs with

Bonferroni-corrected pairwise comparison were carried out separately for each within-factor effect. Statistical analysis was performed with IBM SPSS version 23.0.

3. RESULTS

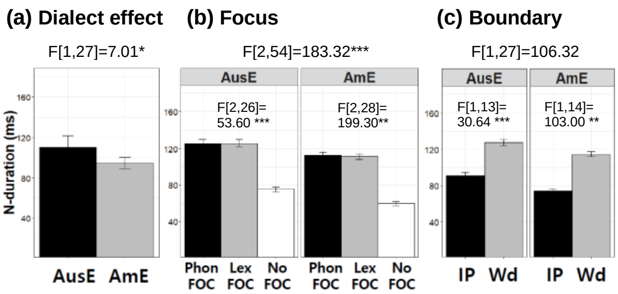
3.1. #NVC (domain-initial effects)

3.1.1. Initial N duration

There was a main effect of Dialect on N duration (F[1,27]=7.01, p<.05), indicating that N duration in NVC was generally longer in AusE than in AmE (Fig.1**a**). There was also a main effect of Focus on N duration as shown in Fig.1**b**, showing that N duration was longer in the focused than in the unfocused condition in both dialects, which augmented N' nasality under prominence. (Note that lexical focus and phonological focus did not differ on any measure, so we will not report the difference between the two for the remainder of the paper.) Crucially, there was no interaction between Focus and Dialect, suggesting that the focus effect on N is comparable between the dialects.

As shown in Fig.1**c**, Boundary also showed a main effect on N duration with no interaction with Dialect, again showing a comparable cross-dialectal effect. But counter to the lengthening of N under focus, the Boundary factor induced a *shortening* of N in IP-initial position, increasing N's consonantality (rather than its sonority).

Figure 1: N duration: Focus and Boundary effects on **#N**VC in AusE and AmE. (AmE data from [5].)



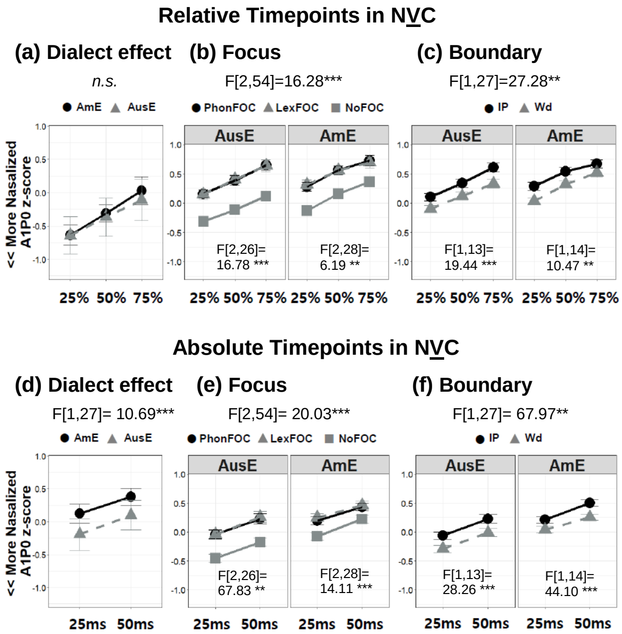
3.1.2. V nasalization in NVC (carryover effect)

There was a main effect of Dialect on A1-P0 in the absolute measure (F[1,27]=10.69, p<.05), indicating that V nasalization in the carryover direction was generally larger in AusE than in AmE when the vowel's physical distance from the coarticulatory source (N) was exactly the same (i.e., fixed at absolute timepoints) across the two dialects (Fig.2**d**). But such cross-dialectal difference disappeared in the relative measure, revealing no Dialect effect (Fig.2**a**).

Focus showed a main effect in both relative and absolute measures. As shown in Fig. 2**b** and **e**, A1-P0 was greater in the focused than in the unfocused condition, and crucially the effect was pervasive through the vowel (no interaction with Time), indicating V's *coarticulatory resistance* to nasalization under prominence. There was no interaction Focus x Dialect interaction, either, suggesting that the focus effect is comparable between the two dialects.

There was also a main effect of Boundary (domain-initial effect) on V nasalization. As shown in Fig.2**c** and **f**, this consisted of a *reduction* of V nasalization (greater A1-P0) in IP-initial position on both the relative and absolute measures. Again, Boundary did not interact with Dialect (showing cross-dialectal comparability) or Time (showing the pervasiveness of the effect into the vowel).

Figure 2: V nasalization: Focus and Boundary effects. A1P0 z-score at relative and absolute timepoints in **#N**VC in AusE and AmE. (AmE data from [5].)



3.2. CVN# (domain-final effects)

3.2.1. Final N duration

Unlike the case with initial N (longer in AusE than in AmE), there was no main effect of Dialect on final N (Fig.3**a**). But Focus showed a main effect on N duration in CVN. As shown in Fig.3**b**, N duration was significantly longer in the focused than unfocused conditions, with no interaction with Dialect. Boundary also showed a main effect on N duration (Fig.3**c**), such that N was longer in IP-final than Wd-final conditions, showing a general phrase-final lengthening effect. Again there was no interaction between Boundary and Dialect.

3.2.2. V nasalization in CVN (anticipatory effect)

Similar to the effect on NVC, there was a main effect of Dialect on the absolute measure in CVN but in an opposite direction (Fig.4**d**)—i.e., whereas NVC (carryover effect) showed more V nasalization in

AusE than AmE, CVN (anticipatory effect) revealed *less* V nasalization in AusE than AmE.

CVN also showed a main effect of Focus on both the relative and absolute measures, as shown in Fig.4**b** and **e**, which was again pervasive throughout the vowel (no interaction with Time). This indicates a general coarticulatory reduction (*resistance*) under focus. CVN also showed a main effect of Boundary on both the relative and absolute measures as shown in Fig.4**c** and **f**. But unlike the focus-induced coarticulatory resistance effect, the boundary-induced effect revealed a coarticulatory vulnerability in IP-final position, as evident in an increase in the degree of V nasalization.

Figure 3: N duration: Focus and Boundary effects in CV**N#** in AusE and AmE. (AmE data from [5].)

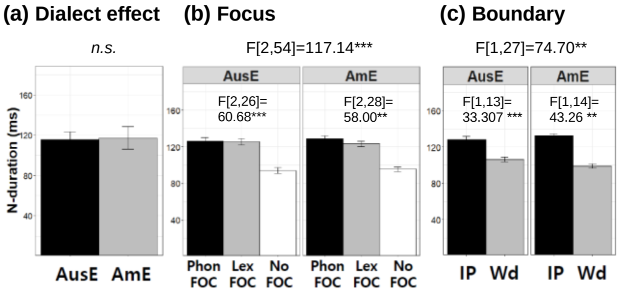
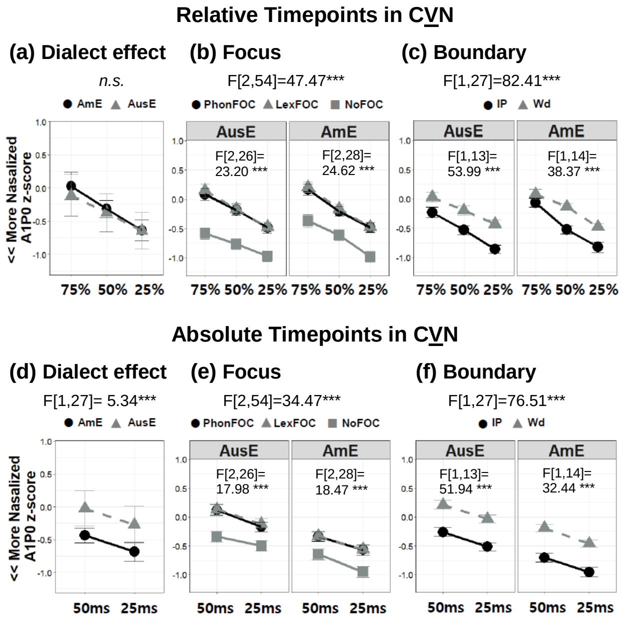


Figure 4: V nasalization: Focus and Boundary effects. AusE and AmE A1P0 z-score at Relative (75%, 50%, 25%) and Absolute timepoints (50ms, 25ms) in CV**N#.**



4. DISCUSSION AND CONCLUSION

Comparisons of the results between AusE and AmE revealed interesting cross-dialectal differences. N duration in the onset of NVC was generally longer in AusE than in AmE, but no such dialectal difference was observed for N in the coda of CVN. As for V nasalization, there was an asymmetry between the two dialects in that AusE showed *less* V nasalization in the NVC (carryover) context, but *more* V nasalization in the CVN (anticipatory) context, as compared with V nasalization in AmE. These cross-dialectal differences suggest that the low-level phonetic coarticulatory process is indeed regulated differently even across dialects of the same language, extending the general view that a non-contrastive phonetic process is internalized in the phonetic grammar at an individual language level [2,3,7,9,12] to a dialectal level.

Despite the cross-dialectal differences, however, the present results also reveal remarkable cross-dialectal similarities. Both dialects showed a boundary-induced shortening of N, accompanied by *less* V nasalization in #NVC, but the reverse was true in CVN# in which N was lengthened and V was *more* nasalized. Furthermore, both dialects showed a prominence-induced lengthening of N in both #NVC and CVN#, while V was nasalized *less* under prominence, a pattern interpretable as coarticulatory *resistance*. These results reinforce a view of the phonetics-prosody interface in which phonetic realization of segments is fine-tuned by the prosodic structure in which segments occur (cf. [3]).

Crucially, both dialects also showed that these effects were not limited to the vicinity of the source of nasalization (N) which might otherwise signal a low-level phonetic effect. Instead, the effects were pervasive throughout the entire vowel, as V nasalization was extended beyond the physiological/biomechanical time-locked effect. This also implies that the coarticulatory process is controlled by the speaker with reference to higher-order prosodic structure. More importantly, the cross-dialectal similarities are grounded on linguistic contrasts that may underlie the phonetics-prosody interface. For example, both dialects showed a boundary-induced enhancement of syntagmatic (CV) contrast in domain-initial position. This was evident in a combination of a shortening of nasal murmur for N (which increases N's consonantality) and a reduction of V nasalization (which increases V's orality). Furthermore, both dialects showed a prominence-induced enhancement of paradigmatic contrast, such that the nasal murmur of N was lengthened, enhancing N's nasality feature whereas V showed coarticulatory resistance to nasalization.

These results add to the fast-growing body of literature on the phonetics-prosody interface. They demonstrate that although AusE and AmE differ in the magnitude of coarticulatory nasalization in carryover *vs.* anticipatory contexts, such seemingly different coarticulatory propensities across dialects operate in much the same way by making reference to linguistic contrasts in universally applicable ways.

5. ACKNOWLEDGEMENT

We thank the Australian English speakers for their participation in the experiment and Yewon Hong for her assistance with acoustic measurements. We also thank the reviewers for useful comments. This work was supported by Global Research Network program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A2A2912410).

6. REFERENCES

[1] Barlaz, M., Shosted, R., Fu, M., Sutton, B. 2018. Oropharygneal articulation of phonemic and phonetic nasalization in Brazilian Portuguese. *Journal of Phonetics*, 71, 81-97.

[2] Cho, T., Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27(2), 207-229.

[3] Cho, T. 2015. Language effects on timing at the segmental and suprasegmental levels. In: Redford, M. A. (ed), *The Handbook of Speech Production*. Hoboken, NJ: Wiley-Blackwell, 505-529.

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

[5] Cho, T., Kim, D., Kim, S. 2017. Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, 64, 71-89.

[6] 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.

[7] Keating, P. A. 1990. Phonetic representations in a generative grammar. *Journal of Phonetics*, 18(3), 321-334.

[8] Kim, S., Kim, J., Cho, T. 2018. Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics*, 71, 65-80.

[9] Kingston, J. 2007. The phonetics- phonology interface. In: P. de Lacy. (ed), *The Cambridge Handbook of Phonology*. Cambridge University Press, Cambridge, 401-434.

[10] Kühnert, B., Nolan, F. 1999. The origin of coarticulation. In: Hardcastle, W., Hewlett, N. (eds), *Coarticulation: Theory, Data and Techniques*. Cambridge University Press, Cambridge, 7-30.

[11] Styler, W. 2015. *On the acoustical and perceptual features of vowel nasalization.* Ph.D. dissertation. University of Colorado.

[12] Yu, J., Zahner, K. 2018. Truncation and compression in Southern German and Australian English. *Proc. Interspeech 2018*, 1833-1837.