Journal of Phonetics 64 (2017) 71–89

Contents lists available at ScienceDirect

Journal of Phonetics

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

Special Issue: Mechanisms of regulation in speech, eds. Mücke, Hermes & Cho

Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English

Taehong Cho [a,*], Daejin Kim [a], Sahyang Kim [b]

[a] *Hanyang Phonetics and Psycholinguistics Lab, Department of English Language and Literature, Hanyang University, Seoul, South Korea*
[b] *Department of English Education, Hongik University, Seoul, South Korea*

ARTICLE INFO

Article history:
Received 1 May 2016
Received in revised form
15 November 2016
Accepted 23 December 2016
Available online 18 January 2017

Keywords:
Vowel nasalization
Coarticulation
Coarticulatory resistance
Prosodic strengthening
Prominence
Boundary
Phonetic grammar

ABSTRACT

This study explores the relationship between prosodic strengthening and linguistic contrasts in English by examining temporal realization of nasals (*N*-duration) in CVN# and #NVC, and their coarticulatory influence on vowels (V-nasalization). Results show that different sources of prosodic strengthening bring about different types of linguistic contrasts. Prominence enhances the consonant's [nasality] as reflected in an elongation of *N*-duration, but it enhances the vowel's [orality] (rather than [nasality]) showing coarticulatory resistance to the nasal influence even when the nasal is phonologically focused (e.g., *mob-bob; bomb-bob*). Boundary strength induces different types of enhancement patterns as a function of prosodic position (initial vs. final). In the *domain-initial* position, boundary strength reduces the consonant's [nasality] as evident in a shortening of N-duration and a reduction of V-nasalization, thus enhancing CV contrast. The opposite is true with the *domain-final* nasal in which *N*-duration is lengthened accompanied by greater V-nasalization, showing coarticulatory vulnerability. The systematic coarticulatory variation as a function of prosodic factors indicates that V-nasalization as a coarticulatory process is indeed under speaker control, fine-tuned in a linguistically significant way. In dynamical terms, these results may be seen as coming from differential intergestural coupling relationships that may underlie the difference in V-nasalization in CVN# *vs.* #NVC. It is proposed that the timing initially determined by such coupling relationships must be fine-tuned by prosodic strengthening in a way that reflects the relationship between dynamical underpinnings of speech timing and linguistic contrasts.

© 2017 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

One of the central questions that have vigorously been explored in the field of phonetics and laboratory phonology concerns how the fine phonetic detail of segments at a subphonemic level may serve as phonetic hallmarks of higher-order linguistic structure. Prosodic structure is one such linguistic structure that modulates phonetic shaping of segments at a fine-grained (low) level in accordance with its linguistic functions such as demarcating prosodic junctures and signaling relative salience or prominence of prosodic constituents (e.g., Keating & Shattuck-Hufnagel, 2002; Cho & Keating, 2009; see Fletcher (2010) or Cho (2016) for a review). With a view to understanding the phonetics-prosody interplay, quite a few researchers have explored phonetic manifestations of prosodic structure in terms of *prosodic strengthening* associated with prosodic landmarks such as prosodic domain edges at prosodic junctures and syllables with prominence (e.g., Byrd & Riggs, 2008; Cho, 2004, 2005; Cho & Keating, 2001, 2009; Cho, Lee, & Kim, 2011, 2014; de Jong, 1995, 2004; Cho & McQueen, 2005; Fougeron, 2001; Kuzla, Cho, & Ernestus, 2007; Krivokapić & Byrd, 2012; Mücke, Grice, & Cho, 2014, *inter-alia*). Segments in these prosodic landmarks are generally articulated 'strongly' with an expansion in the spatial and/or temporal dimension, although their actual strengthening patterns may differ depending on the kind of the source of strengthening (e.g., prominence vs. prosodic boundary). Given that the 'strong' articulation is likely to heighten the phonetic clarity of the segments, an important question regarding the phonetics-prosody interplay concerns how the phonetic granularity of prosodic strengthening is related to enhancement of linguistic contrasts (e.g., Cho & Jun, 2000; Cho, 2005; Cho & McQueen, 2005; Cho et al., 2014; de Jong, 1995, 2004; Fougeron & Keating, 1997; Fougeron, 1999).

In the present study, we continue to build on this issue by examining effects of prosodic structure on the phonetic granularity of a coarticulatory process to be reflected in the acoustic realization of nasal consonants and their coarticulatory impacts on the neighboring vowels in CVN# and #NVC contexts in English (where '#' refers to a prosodic boundary such as an

---

* Corresponding author. Fax: +82 2 2220 0741.
E-mail addresses: tcho@hanyang.ac.kr (T. Cho), daejin.kim.hyu@gmail.com (D. Kim), sahyang@hongik.ac.kr (S. Kim).

http://dx.doi.org/10.1016/j.wocn.2016.12.003
0095-4470/© 2017 The Author(s). Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Intonational Phrase boundary). A coarticulatory process is often considered to be an automatic phonetic process as it has its origin from physiological and biomechanic constraints imposed on the human speech production system (cf. Kühnert & Nolan, 1999), but it has been suggested in the literature that the process may be fine-tuned as a function of linguistic and non-linguistic factors, thus being characterized as something that is controllable by the speaker in a linguistically systematic way (see Farnetani & Recasens (1990, 2010), for a review; Beddor (2015), for a related discussion; and Scarborough, Zellou, Mirzayan, and Rood (2015), for a case indicating how the degree of coarticulatory vowel nasalization is modulated in a language-specific way in a language which employs contrastive oral-nasal vowels). The goal of the present study is therefore to explore (1) how the acoustic temporal realization of English nasal consonants and their coarticulatory influence on the neighboring vowel (i.e., V-nasalization as measured by A1-P0) are modulated by two different types of prosodic strengthening (i.e., prominence-related vs. boundary-related) in CV<u>N</u># versus #<u>N</u>VC, and (2) how the prosodically-conditioned coarticulatory process may be understood in relation to linguistic contrasts. It will then be discussed how the prosodically-conditioned fine-tuning of a coarticulatory process may be understood in dynamical terms.

1.1. Theoretical considerations

With methodological advancements in the experimental phonetics in the past several decades, significant progress has been made in our understanding of how a linguistic code is phonetically encoded at a fine-grained level of detail in the sound system of the language. In particular, numerous studies have illuminated the importance of the role that phonetic granularity plays in the grammar of the language (e.g., Browman & Goldstein, 1990, 1992; Cho & Ladefoged, 1999; Goldstein, Byrd, & Saltzman, 2006; Keating, 1984, 1990; Kingston & Diehl, 1994; Mücke et al., 2014; Cho, 2005, *inter alia*). One of the compelling theoretical assumptions that have emerged particularly from studies on the phonetics-prosody interplay is that the phonetic granularity of prosodic strengthening is indeed controlled by the speaker as it is modulated by higher-order linguistic structure, and "so must be specified in a linguistic description of the phonetics-prosody interface as part of the phonetic grammar of the language" (Cho, 2016). In other words, if segments are strengthened as specified in the grammar of a given language, the phonetic granularity of prosodic strengthening must convey some linguistic functions possibly linked to maintenance or maximization of phonological contrasts of the sound system of a given language, which in turn should be eventually exploited by the listener in speech comprehension (e.g., Cho, McQueen, & Cox, 2007; Fougeron & Keating, 1997; Gow, Melvold, & Manuel, 1996; Mitterer, Cho, & Kim, 2016).

An important theoretical consideration of the present study therefore concerns the nature of contrast enhancement that may underlie a prosodically-conditioned coarticulatory process as reflected in the granularity of V-nasalization under prosodic strengthening. Given that boundary and prominence markings are often characterized by different phonetic hallmarks, the nature of linguistic contrasts that may be mediated by prosodic strengthening is also likely to differ (e.g., Fougeron, 1999; Harrington, Fletcher, & Beckman, 2000; Cho & Jun, 2000; de Jong, 1995, 2004; Cho & McQueen, 2005; Cho et al., 2011, 2014; see Cho (2016) for a review). Prominence marking is assumed to enhance *paradigmatic contrast*, which results in a maximization of phonological distinction of contrastive sounds (e.g., de Jong, 1995, 2004). The paradigmatic contrast enhancement has often been characterized as a *localized hyperarticulation* (e.g., de Jong, 1995) as hyperarticulation of this sort involves enhancement of phonemic distinction locally in a stressed syllable rather than globally to the entire utterance in the sense of Hyper- & Hypo-articulation (H&H) theory (Lindblom, 1990). The localized hyperarticulation thus pertains to enhancement of phonetic features of phonemes under prominence, determining the phonetic content of the phonological contrast in the language. For example, in an acoustic study of Dutch stops, Cho and McQueen (2005) showed that the Dutch voiceless stop /t/ was produced with a shorter VOT when the /t/-bearing syllable was accented than when it was unaccented, showing the opposite of the VOT lengthening effect in English, although the voiceless stop in both languages may be specified with the same phonological feature [-voice] (e.g., Keating, 1984, 1990; Kingston & Diehl, 1994). The asymmetrical modulation of VOT under prominence between the two languages may be interpreted as stemming from the use of language-specific phonetic features (e.g., {+/-spread glottis}) which participates in the phonetic coding of phonological contrast. The prominence-induced VOT shortening effect has also been observed in English in a /s/-stop sequence in which the stop gets shortened due to the allophonic rule in English. Cho, Lee and Kim (2014) demonstrated that the shortened VOT in the /s/-stop cluster becomes even shorter in line with enhancement of the relevant phonetic feature (e.g., {-spread glottis}).

As for the boundary effect, Cho and McQueen (2005) also reported that VOT for Dutch voiceless stops may be shortened (rather than lengthened) phrase-initially, but only in a limited way—i.e., the boundary-induced VOT shortening effect was not observed between what they termed "Big" phrase (with longer phrase-final lengthening followed by pause) and "Small" phrase (without pause). Furthermore, Cho et al. (2014) showed that unlike the VOT shortening effect under prominence, the VOT of the unaspirated stop in the /s/-stop cluster did not undergo further shortening at a higher prosodic boundary. These results, together, appear to indicate that the nature of phonological (paradigmatic) contrast driven by boundary marking is not as clear as the case of prominence marking.

Instead, boundary marking is often taken to be structurally motivated, resulting in enhancement of *syntagmatic contrast* between neighboring segments at prosodic junctures, so that the consonant becomes more consonant-like (with a decrease of sonority) and the vowel becomes more vowel-like (with an increase of sonority), enhancing the contrast at or across prosodic boundaries (e.g., Pierrehumbert & Talkin, 1992; Fougeron & Keating, 1997; Fougeron, 1999). For example, based on the frequently observed increase in constriction degree and duration for domain-initial consonants, Fougeron and Keating (1997) and Cho and Keating (2001, 2009) suggested that strengthening of domain-initial consonants is interpretable as heightening the consonantality of the initial segments by virtue of which #CV contrast in the initial syllable or V#C contrast across a prosodic juncture may be enhanced. Frequently observed longer VOTs for the voiceless aspirated stop and more aspiration for /h/ in the domain-initial position (than in the medial position) can also be seen as a heightening of consonantality (the more aspirated, the more consonant-like), thus enhancing CV contrast (e.g., Pierrehumbert & Talkin, 1992; Cho & Keating, 2009). As for the case of the nasal consonant, it has been shown that it is produced with a reduced nasal airflow and a shortened acoustic nasal murmur in the domain-initial than in the domain-medial position across languages—e.g., in English (Fougeron & Keating, 1997; Cho & Keating, 2009), in French (Fougeron, 2001), in Korean (Cho & Keating, 2001) and in Estonian (Gordon, 1996). The reduced nasal murmur has been interpreted as resulting in an increase in consonantality (being more consonant-like) (e.g., Cho & Keating, 2009; Cho, 2001) under the assumption that the consonantality is inversely related to the generally assumed degree of sonority (see Ladefoged (2001), or Parker (2002), for a review of sonority hierarchy). In line with Cho and Keating's interpretation, we will interpret (temporal) reduction in nasality for the nasal consonant which is to be reflected in a shortened nasal murmur as suggesting an increase in the nasal's consonantality (by virtue of reduced sonority) which would in turn effectively augment the CV contrast—making the consonant more consonant-like.

The notion of contrast enhancement can be further considered in connection with coarticulatory resistance. Coarticulatory resistance often refers to the degree to which an articulatory gesture (or a segment) resists its coarticulation with its neighboring segment as a function of the articulatory constraint imposed on the gesture (e.g., Bladon & Nolan, 1977; Recasens, 1987, 1989; see Farnetani & Recasens (2010), for a review). A classic example comes from different degree of coarticulatory resistance of clear vs. dark /l/ in English (e.g., Bladon & Al-Bamerni, 1976). Given that dark /l/ is produced with articulatory constraint involving multiple gestures required for both the alveolar and the dorsopharyngeal constriction, its coarticulation with an adjacent vowel (which also involves the articulatory constriction in the dorsopharyngeal region) is suppressed to a greater degree compared to the clear /l/ which is relatively free from such a constraint. Fowler and Saltzman (1993) elaborates on the notion of coarticulatory resistance in gestural terms—that is, a high degree of strength of gestural blending (i.e., when there are conflicting demands on the same articulator) increases a gesture's resistance to coarticulatory interference coming from neighboring gestures.

Building on these insights of coarticulation, a number of studies on coarticulation have suggested that the degree of coarticulatory resistance is not only determined by the above-mentioned physical constraints but it may also be further modulated by system constraints in a given language (e.g., Fowler, 1981; Manuel, 1990, 1999; de Jong 1995; Beddor, Harnsberger, & Lindermann, 2002; Cho, 2004; Fletcher, 2004). Coarticulation is thus often expected to be suppressed (or resisted) when it would otherwise result in the blurring of phonetic contrasts in a given language (e.g., Manuel, 1990, 1999), or when the contrast needs to be enhanced either in a communicatively-driven global hyperarticulation context (e.g., Lindblom, 1990) or a localized hyperarticulation context under prominence (e.g., de Jong 1995; Cho, 2004; Fletcher, 2004). de Jong (1995), for example, reported that the coarticulation of /t/ into a following /θ/ in English was reduced when /t/ received prominence (e.g., nuclear pitch accent under focus), showing coarticulatory resistance. Cho (2004) also showed a similar coarticulatory resistance pattern in V-to-V coarticulation under prominence. This type of coarticulatoy resistance was interpreted as being driven by an enhancement of the distinctiveness of the phoneme. In the present study, we will consider the degree of vowel nasalization in terms of coariculatory resistance in relation to enhancement of linguistic contrast.

Another important insight on the nature of coariculation particularly with respect to anticipatory vowel nasalization is found with Beddor's (2009) observation of the inverse relationship between the degree of vowel nasalization and the duration of the following nasal consonant, the coarticulatory source. Based on her own experimental results and the survey on the coarticulatory literature, Beddor hypothesized that the timing of the velum lowering gesture associated with the nasal coda (i.e., in the anticipatory vowel nasalization context) remains more or less constant, so that in a context in which the nasal coda becomes longer, the duration of the vowel nasalization is proportionally shorter. Beddor further noted that although the temporal domain of the velum lowering gesture may be stretched or shrunk depending on various linguistic factors, the inverse relationship tends to remain largely invariant. In the present study, we will consider to what extent this purported relationship accounts for the degree of vowel nasalization that is expected to vary as a function of boundary marking and prominence marking.

1.2. Hypotheses and predictions

What has emerged from previous studies on prosodic structuring is that there is a relatively clear dichotomy in the linguistic functions between the two kinds of prosodic strengthening, so that the culminative function of prominence is linked with paradigmatic contrast, and the delimitative function of prosodic boundary with syntagmatic contrast. With this distinction in mind, we will consider the following hypotheses and related research questions in the present study.

Under the paradigmatic enhancement hypotheses the phonemic distinction of the nasal consonant is expected to be enhanced, so that the duration of the nasal murmur (N-duration) is expected to be lengthened which may be interpreted as an enhancement of [nasality]. In relation to coarticulatory process, segments under prominence (e.g., in stressed syllables) are often assumed to exert more coarticulatory influence on neighboring segments, as has been described as *coarticulatory aggression* in the literature (Fowler & Saltzman, 1993; Cho, 2004). One might therefore expect that the nasal under prominence will also have a stronger coarticulatory impact on the neighboring vowels to be reflected in greater degree of V-nasalization. This is based on the assumption that the velum lowering gesture for the nasal consonant would be strengthened under prominence, hence its greater influence on the neighboring vowel.

To test the prominence effect, three focus conditions were employed: phonological focus (PhonFOC), lexical focus (LexFOC) and no focus. Under phonological (or segmental) focus a specific phoneme is emphasized in contrast with another phoneme (e.g., ‘mob’ *vs.* ‘bob’ in the onset or ‘bomb’ *vs.* ‘bob’ in the coda), whereas under lexical (semantic) focus a lexical item is emphasized in contrast with another lexical item that is semantically related (e.g., ‘bomb’ *vs.* ‘war’). As discussed in de Jong and Zawaydeh (2002) and de Jong (2004), phonological distinctions of particular phonemes are more likely to be maximized in the phonologically focused context than in the lexically focused context, as the speech material in the latter context is affected as a whole without making specific reference to the phonemic content of the speech material. It was therefore important to test phonological vs. lexical focus conditions in connection with prominence—i.e., whether the enhancement of [nasality] may be achieved to a greater degree when focus was realized on the nasality of the consonant (‘mob’ *vs.* ‘bob’ or ‘bomb’ *vs.* ‘bob’) rather than on the lexical (semantic) contrast (‘mob’ *vs.* ‘gang’ or ‘bomb’ *vs.* ‘war’), and whether the hypothesized coarticulatory aggression of the nasal on the vowel may be more robust in the phonologically focused context in which the nasal consonant is emphasized.

An alternative hypothesis can be thought of, however, when we consider how focus is phonetically implemented. Focus, whether phonological or lexical, is in general phonetically realized with prominence in association with pitch accent in English. Given that the vowel is the primary locus of prominence regardless of whether prominence (accent) is linked to phonological or lexical focus, it is likely that the vowel itself may undergo strengthening. If so, the vowel may show coarticulatory resistance to V-nasalization in order to heighten its orality under prominence, rather than being vulnerable to coarticulatory aggression of the nasal even in the phonological focus condition in which nasality is emphasized.

The domain-initial boundary effect, on the other hand, is expected to enhance syntagmatic contrast between the consonant and the neighboring vowels as discussed above. Under this hypothesis, [consonantality] is assumed to be enhanced with a decrease in sonority in #NVC. Thus, being consistent with previous findings (e.g., Fougeron & Keating, 1997; Fougeron, 1999; Cho & Keating, 2001, 2009), *N*-duration is expected to be shortened domain-initially. Fougeron (1999, 2001) interpreted the decrease in the consonant's nasality in the initial position as stemming from the overall augmentation of the ‘oral articulatory force’ that applies to the segments in the initial position. In the case of nasal consonants, the strengthening of the oral articulatory force has an effect of elevating the velum resulting in reduced nasality. More crucially for the present study, given the likelihood of a decrease in nasality for the consonant itself and the syntagmatic driving force for CV contrast enhancement, the nasal's coarticulatory influence on the following vowel in #NVC will also be reduced in the domain-initial position. If this is indeed the case, there will be reduced nasality for the consonant and increased orality for the vowel, which is, taken together, interpretable as an enhancement of CV contrast.

With respect to the domain-final nasals in CVN#, however, a different assumption can be made. The final consonant is generally subject to articulatory weakening rather than strengthening (e.g., Fougeron, 1999; Keating, Wright, & Zhang, 1999), although the final consonant undergoes domain-final lengthening. This phrase-final articulatory weakening is also consistent with Browman and Goldstein's (1992, 1995) observation of word-final reduction of the oral constriction gesture for /n/ (as well as for other consonants such as /l/ and /t/). The consonantal weakening in the coda position may then cause a weakening of the oral articulatory force (cf. Fougeron, 1999, 2001), which would in turn result in a reduction of the velum elevation effect, hence an increase in nasality. If the hypothesized weakening of oral articulatory force would indeed reduce the velum elevation effect in CVN#, one could expect more nasality associated with the nasal consonant, likely to be accompanied by an elongated nasal murmur. The expected increase in nasality for the nasal coda is likely to be reflected in the degree of vowel nasalization as well—i.e., an increase in V-nasalization. Furthermore, the velum lowering gesture is often assumed to be sequentially coupled with the oral constriction (i.e., in anti-phase mode) in CVN#, which would effectively reduce their intergestural cohesiveness as may be reflected in less stable intergestural timing associated with anti-phase mode (e.g., Goldstein, Nam, Saltzman, & Chitoran, 2009; Byrd, Tobin, Bresch, & Narayanan, 2009; see below for more discussion on the gestural coupling relationship). Considering both the weakening of the oral articulatory force and the less stable velic-oral intergestural timing due to the anti-phase coupling relationship, it is reasonable to assume that the degree of V-nasalization is more vulnerable to prosodic modification in the (phrase-final) anticipatory than in the (phrase-initial) carryover coarticulatory context. This hypothesized coarticulatory vulnerability of the vowel in the domain-final position, however, would be difficult to interpret in terms of contrast enhancement at a prosodic boundary, but rather it would be better understood as a byproduct of the weakening of consonantality of the nasal in combination with the anti-phase velic-oral coupling relationship in the domain-final position.

Testing the above-discussed hypotheses is directly germane to another important question regarding the nature of coarticulatory process of V-nasalization. As discussed above, the underlying assumption behind the terms *coarticulatory resistance* or *coarticulatory aggression* (cf. Farnetani & Recasens, 1999, 2010) implies that the effect may be under speaker control. The coarticulatory process as an effect controlled by the speaker, for example, has been discussed in the literature as evident in systematic coarticulatory jaw movements as a function of prominence (Harrington, Fletcher, & Roberts, 1995), and lingual V-to-V coarticulatory patterns under prosodic strengthening in English (Cho, 2004; Fletcher, 2004). The present study continues to explore this issue by examining the time course of V-nasalization during the vowel which allows us to assess the extent to which V-nasalization is physically constrained and the extent to which it is controlled by the speaker. On the one hand, if V-nasalization is driven purely by physiological/biomechanical factors, the coarticulatory effect will be time-locked, being strictly localized to part of the vowel near the nasal consonant. On the other hand, if V-nasalization is modulated by higher order linguistic structure being controlled by the speaker, V-nasalization will extend beyond what a pure phonetic effect might show, thus being more globally pervasive over a larger part of the vowel. These possibilities, especially in the anticipatory coarticulatory context, will also be considered further in connection with Beddor's (2009) observation of the inverse relationship between the duration of the nasal coda consonant and the vowel nasalization. Beddor's observation would lead to a prediction that when the nasal coda is lengthened regardless of the source of lengthening, the duration of vowel nasalization is expected to be shortened. What is particularly relevant for the present study is then to what extent the proposed inverse relationship would explain the degree of vowel nasalization that is expected to vary as a function of prominence marking vs. boundary marking. Exploring this question would allow us to understand the nature of anticipatory vowel coarticulatory process in terms of the constant temporal realization of the velum lowering gesture or linguistic principles that may underlie the paradigmatic vs. syntagmatic functions.

Finally, the obtained data will be further discussed in terms of how the prosodically-conditioned variation of V-nasalization may be understood in dynamical terms. The aforementioned asymmetric position effects on V-nasalization (greater in the anticipatory (domain-final, CVN) context than in the carryover (domain-initial, NVC) context), for example, may be understood as coming from differential intergestural coupling relationships (anti-phase *vs.* in-phase, respectively), which, in theory, determines the timing between the consonantal constriction gesture and the velum lowering gesture (e.g., Goldstein et al., 2009; Byrd et al., 2009; cf. Krakow, 1989). In an effort to understand the asymmetrical coarticulatory effects on V-nasalization between nasals in the onset (initial) vs. the coda (final) position, Krakow (1989) examined the relative timing between the velum lowering gesture and the oral (lip closing) gesture for /m/ (see Krakow, 1999 for a review). She found that in the syllable onset (NVC) context, the time point of the maximum of the velum lowering gesture was aligned roughly with the time point of the target attainment of the oral closing gesture, showing a roughly synchronous timing relationship between the velic and the oral gestures. In the syllable coda (CVN) context, on the other hand, the time point of the velum lowering maximum was largely aligned with the onset of the oral (lip closing) gesture, showing an earlier timing of the velum gesture relative to the oral gesture. The asymmetrical timing patterns were later further instantiated in Byrd et al. (2009) with innovative methodology using real-time MRI for the nasal /n/. The results were discussed in terms of the intrinsic modes of intergestural timing that are hypothesized to underlie human motor behavior. That is, in the intrinsic modes gestures are either synchronously coupled (in in-phase mode) or sequentially coupled (in anti-phase mode) (e.g., Byrd et al., 2009; Goldstein et al., 2009; Nam, Goldstein, & Saltzman, 2009). Thus, the synchronous velum-oral gestural timing relationship is consistent with the in-phase coupling relationship underlying CV structures, and the earlier timing of the velum gesture relative to the oral gesture is consistent with the anti-phase coupling relationship underlying VC structures. Under this assumption, any observed systematic variation in V-nasalization due to prosodic strengthening can be taken to ensue from a fine tuning of the velum-oral gestural timing in an in-phase vs. an anti-phase coupling mode as a function of prosodic structure. It will be discussed how the results of the present study further inform this discussion in connection with the relationship between dynamical underpinnings of speech timing and the enhancement of linguistic contrasts under prosodic strengthening that might underlie systematicity of coarticulatory variation.

2. Methods

2.1. Participants and recording

Fifteen native speakers of American English (8 females and 7 males) participated in the study for pay. All participants were born and raised in the United States and they were all in their 20 s or 30 s (mean age = 25.9), staying in Seoul Korea as exchange students, visitors and English teachers at the time of recording. Acoustic data were recorded in a sound attenuated booth at Hanyang Phonetics and Psycholinguistics Lab with a Tascam DR-680 multi-channel digital recorder and a Shure VP88 condenser microphone at a rate of 24 bit and a sampling rate of 44.1 kHz.

2.2. Speech materials

Eight test words were used for the study, including four words (*palm*, *bomb*, *ten*, *den*) in the CVN context and four words (*mop, mob, net, Ned*) in the NVC context. Table 1 gives an example set of test words in different conditions for Boundary and Focus. As shown in the table, a test word was embedded in a carrier sentence in a mini discourse situation in which Boundary and Focus were manipulated. As for the boundary conditions, the target nasal /m/ in CVN# (e.g., *BOMB*) was either IP-final (B: *No. I was supposed to write BOMB, wasn't I*?) or phrase-internal Wd-final (B: *No. I wrote* '*say BOMB fast again*'), yielding two boundary conditions (IP-boundary vs. Wd-boundary) in the domain-final context. (Two boundary conditions in the domain-initial (#NVC) context were constructed in a similar way; see Table 1). As for the focus conditions, the test word (e.g,. *BOMB*) was contrastive either phonologically with *BOB* (PhonFOC) or lexically (semantically) with *WAR* (LexFOC), yielding three focus conditions (PhonFOC, LexFOC, NoFOC). Note that words with only non-high vowels (/ɛ, ɑ/) were selected for a practical reason related to measuring vowel nasality in the spectral dimension (to be discussed in Section 2.4.1. below).

2.3. Procedure

In the experiment, the participants were presented with each mini dialogue on a computer screen, and heard a prompt sentence of Speaker A through a loudspeaker. Prompt sentences (i.e., questions) were pre-recorded by a female speaker of American English. (The female speaker read aloud prompt questions several times by placing a pitch accent on focused (accented) words with either L∗ or L∗+H, and one of the recorded utterances with a typical L∗ was chosen for a prompt sentence.) The participants then read the target-bearing sentence (as Speaker B) in response to a prompt sentence presented auditorily as well as visually on the computer screen. As can be seen in Table 1, in order to induce different types of focus, the subjects were asked to make contrast between words in bold in Sentences A and B, so that they made lexically contrastive (narrow) focus (LexFOC, e.g., *BOMB* vs. *WAR*) or phonologically contrastive (narrow) focus on the nasal consonant (PhonFOC, e.g., *BOMB* vs. *BOB*). In the unfocused condition, a contrastive focus fell on another word in the sentence, so that the test word received no focus (e.g., A: *Did you write* '*say bomb FAST again*'? B: *No, I wrote* '*say bomb SLOWLY again*'.). For boundary conditions, an IP boundary was induced by placing a tag question after the test word in the phrase-final condition CVN#

Table 1
An example set of test words with the target nasal consonant in the coda (CVN#) and the onset (#NVC) in mini-discourse contexts with different focus and boundary conditions. Test words are italicized in bold; focused words are in upper case letters; and the phonologically contrasting phonemes are underlined.

Syll. Position	Boundary	Focus	Example sentences
CVN# (e.g., *BOMB*)	#=IP (IP-Final)	PhonFOC	A: Were you supposed to write ***BOB***? B: No. I was supposed to write ***BOMB*** #, wasn't I?
		LexFOC	A: Were you supposed to write ***WAR***? B: No. I was supposed to write ***BOMB*** #, wasn't I?
		NoFOC	A: Were YOU supposed to write ***bomb***? B: No. JOHN was supposed to write ***bomb*** #, wasn't he?
	#=Wd (Wd-Final, IP-Medial)	PhonFOC	A: Did you write 'say ***BOB*** fast again'? B: No. I wrote 'say ***BOMB*** # fast again'.
		LexFOC	A: Did you write 'say ***WAR*** fast again'? B: No. I wrote 'say ***BOMB*** # fast again'.
		NoFOC	A: Did you write 'say ***bomb*** FAST again'? B: No. I wrote 'say ***bomb*** # SLOWLY again'.
#NVC (e.g., *MOB*)	#=IP (IP-Initial)	PhonFOC	A: Did you write '***BOB*** fast again'? B: Not exactly. # '***MOB*** fast again' was what I wrote.
		LexFOC	A: Did you write '***GANG*** fast again'? B: Not exactly. # '***MOB*** fast again' was what I wrote.
		NoFOC	A: Did you write '***mob*** FAST again'? B: Not exactly. # '***Mob*** SLOWLY again' was what I wrote.
	#=Wd (Wd-Initial, IP-Medial)	PhonFOC	A: Did you write 'say ***BOB*** fast again'? B: No. I wrote '*say* # ***MOB*** fast again'.
		LexFOC	A: Did you write 'say ***GANG*** fast again'? B: No. I wrote 'say # ***MOB*** fast again'.
		NoFOC	A: Did you write 'say ***mob*** FAST again'? B: No. I wrote 'say # ***mob*** SLOWLY again'.

Table 2
A list of test words with target phonemes underlined along with phonologically (PhonFOC) and lexically (LexFOC) contrasting words. A phonologically contrastive word forms a minimal pair with a test word (e.g, *palm* vs. *pop*), and a lexically contrastive word is semantically related to a test word (e.g., *palm* vs. *foot*).

CVN# (domain-final)			#NVC (domain-initial)		
Test Words	PhonFOC (Phonologically contrasting words)	LexFOC (Semantically contrasting words)	**Test Words**	PhonFOC (Phonologically contrasting words)	LexFOC (Semantically contrasting words)
palm	*po**p***	*foot*	***mop***	***b**op*	*wash*
bomb	*bo**b***	*war*	***mob***	***b**ob*	*gang*
ten	*Te**d***	*five*	***net***	***d**ebt*	*ball*
den	*deb**t***	*cave*	***Ned***	***d**ead*	*Paul*

(e.g., …*BOMB#, wasn't it?*) or by placing a short utterance before the test word in the phrase-initial condition #NCV (e.g., *Not exactly. #MOB*…). A Wd boundary was induced by placing the test word in the middle of a short phrase (e.g., …'*say BOMB fast again*'… in the final (CVN#) condition or …'*say MOB fast again*'… in the initial (#NVC) condition). (See Table 2 for a list of test words.)

The participants went through a practice session in which they practiced producing a test sentence in response to a prompt sentence at least once in each test condition for each test word. In order to induce production of sentences as naturally as possible in a given context, the participants were asked to think of a hypothetical discourse context in which the participant (as Speaker B in the mini dialogue) had written something on a piece of paper (e.g., '*BOMB fast again*'), but Speaker A (the voice from a loudspeaker) was double-checking with the subject by questioning (e.g., *Did you write* '*BOB fast again*'?). The participants were told that they were supposed to respond to the question by correcting the misunderstood part (e.g., correcting '*BOB*' as '*BOMB*').

Test sentences were presented in a randomized order, and the set of entire sentences were repeated four times. In total, 2880 sentences were collected (i.e., 2 syllable positions (CVN# vs. #NVC) x 4 test words x 3 focus types (LexFOC vs. PhonFOC vs. NoFOC) x 2 boundaries (IP vs. Wd) x 4 repetitions x 15 speakers). Two trained English ToBI transcribers reviewed all recorded data to check whether each token was produced with intended prosodic renditions (as indicated by the presence of pitch accent on a focused word, and the presence of an IP vs. a Wd boundary as predicted by different syntactic structures of the carrier sentences). As a result, 321 tokens were discarded as they were produced with either a wrong placement of pitch accent or a major prosodic boundary before or after a test word in a phrase-internal Wd boundary context. Pitch accent placed on the test word was either H* or L+H*.

2.4. Measurements

2.4.1. A1-P0[1] as an indicator of the degree of V-nasalization

One of the acoustic characteristics of a nasalized vowel is that the nasal murmur is identified within a nasalized vowel in

---

[1] The degree of nasality could also be measured by calculating the difference between the amplitude of the first formant (A1) and the second nasal peak P1 (i.e., A1-P1;

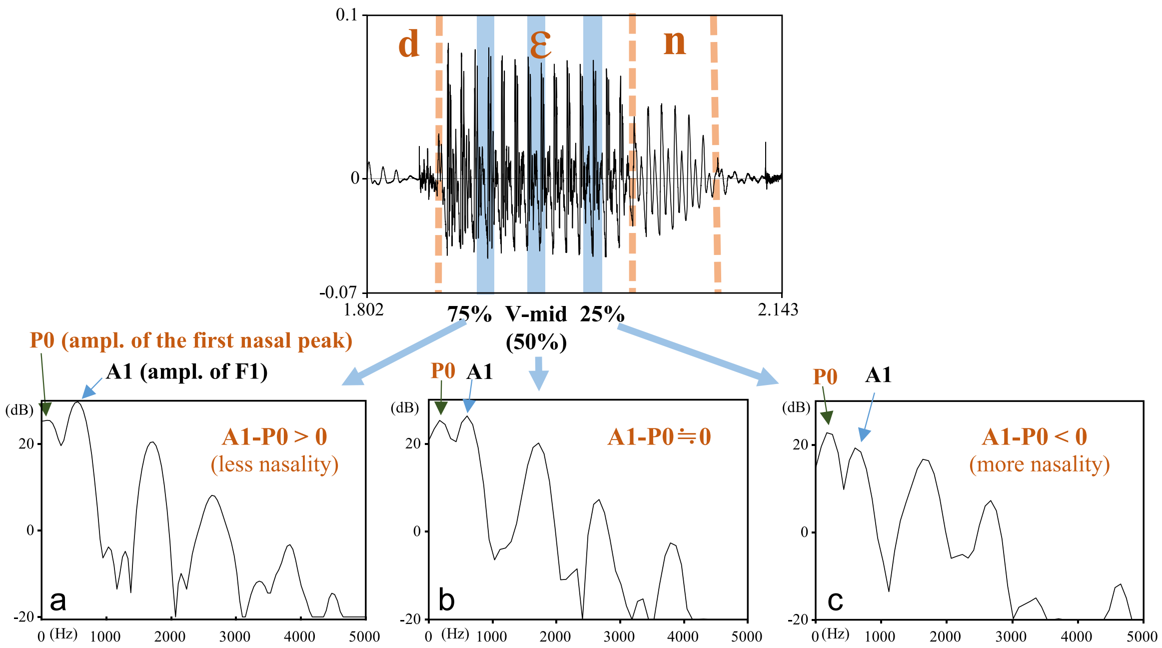

Fig. 1. A schematized process of estimating the degree of V-nasalization by calculating the difference of the amplitude of the first formant (A1) and the first nasal peak (P0).

the vicinity of formants, especially near the first formant (F1) due to oral-nasal coupling which results in a reduction of F1 and an augmentation of the nasal peak (P0) around 250 Hz (cf. Chen, 1996, 1997; see Sampson (1999) for a review). As nasality increases within the vowel, the amplitude of P0 (the nasal peak) increases while the amplitude of A1 (of F1) decreases; hence, crucially for the present study, the more the vowel is nasalized due to the neighboring nasal, the lower the A1-P0 value will be. In English, the degree of vowel nasalization (henceforth V-nasalization) becomes greater as the vowel gets closer to the nasal consonant as a gradient coarticulatory process (e.g., Cohn, 1993). Fig. 1 indeed illustrates the gradient nature of V-nasalization for a token *den* in which A1-P0 gets smaller (thus more nasalized) as a measurement point gets closer to the nasal consonant (i.e., the coarticulatory source of nasalization). At an earlier part of the vowel (i.e., the 75% point away from the nasal consonant; Fig. 1a), P0 is relatively lower than A1, showing the least nasality during the vowel; in the middle of the vowel (the 50% point, Fig. 1b), A1 and P0 are roughly the same in their values, showing an intermediate degree of nasality; and at a later part of the vowel near the nasal consonant (the 25% point; Fig. 1c), P0 is higher than A1, showing the greatest degree of nasality. Note that the measurement points in Fig. 1 are for the anticipatory context (CVN#), and the measurement points in the carryover context (#NVC) are mirror-imaged, so that the 25% point of the vowel in NVC refers to an earlier part of the vowel *near* the preceding nasal consonant and the 75% point of the vowel is a later part of the vowel away from the preceding nasal consonant. In other words, in both the CVN and the NVC context, the 25% point of vowel refers to the near vowel point (i.e., a later vowel part in CVN#, but an earlier vowel part in #NVC), and the 75% point means the distant vowel point (i.e., an early vowel part in CVN#, but a later vowel part in #NVC).

In order to observe how the degree of V-nasalization varies as a function of prosodic factors over the time course of the vowel, A1-P0 values were taken at multiple points during the vowel as in (1) and (2):

(1) *Relative A1-P0 measurement points during the vowel*

a. the distant vowel point (at the 75% point of the vowel away from the nasal)
b. the midpoint (at the 50% of the vowel)
c. the near vowel point (at the 25% of the vowel near the nasal)

(2) *Absolute A1-P0 measurement points during the vowel*

a. the 25ms point away from the nasal
b. the 50ms point away from the nasal
c. the 75ms point away from the nasal
d. the 100ms point away from the nasal

The relative measurement points (75%, 50%, 25% points of the vowel duration away from the nasal consonant) as in (1) will allow us to assess the degree of V-nasalization over the time course of nasalization in a relative term (i.e., relative to the entire vowel duration). We also measured A1-P0 at the absolute measurement points (25 ms, 50 ms, 75 ms, 100 ms away from the onset nasal) as in (2) in order to assess the extent to which the degree of nasality can be accounted for as a time-locked phenomenon—i.e., as a function of the physical distance of the measurement point from the nasal consonant (i.e., the coarticulatory source of nasalization). Note that the tokens which contained the test vowels shorter than 75 ms were not included for the 75 ms and 100 ms measures (37 tokens) and those which contained the vowels whose duration was between 75 ms

---

(*footnote continued*)
Chen, 1996, 1997); however, this method was not used in this study because of obscure characteristic of the second nasal peak (P1) which includes substantial amount of noise and inconsistency. The reader is also referred to recent studies who have employed A1-P0 and other measures for estimating V-nasalization: Zellou and Tamminga, 2014; Scarborough et al., 2015; Styler, 2015; Garellek, Ritchart and Kuang, 2016.

and 100 ms were not included for the 100 ms measures (199 tokens).

The A1-P0 values at the specified timepoints were obtained by using a Praat script employed by Styler (2015), which calculated the amplitudes of the first formant (A1) and the nasal peak (P0) from the spectrum of one full cycle near a specified timepoint (as shown in Fig. 1). Given that P0 is located around 250 to 450 Hz, it is often the case that P0 and A1 (peak amplitude of F1) are not easily separable for high vowels whose F1 is below 500 Hz. We therefore included only non-high vowels /ɛ,ɑ/ in the present study to increase the measurement accuracy. The obtained A1-P0 values were then normalized into *z*-scores which were used for statistical analyses. Following Styler (2015), some values in the data were discarded in the following cases: (i) when the amplitudes of the first and second harmonics calculated were erroneously similar, (ii) when the pitch was erroneously detected at less than 85 Hz or more than 300 Hz, (iii) when the Praat script itself failed to find an accurate value of pitch and harmonic structure.

2.4.2. *Duration of the nasal consonant (N-duration)*

The duration of the nasal consonant (*N-duration*) in #NVC and CVN# was measured to examine the effects of prosodic factors in the temporal realization of the nasal consonant along with the investigation of the time course of V-nasalization. *N-duration* was taken from the onset and the offset of nasal energy (murmur) and nasal zeros (weakened formant structures) displayed on the spectrogram.

2.4.3. *Duration of the vowel (V-duration)*

The duration of the vowel in #NVC and CVN# was also measured to confirm whether the durational variation of the vowel would be consistent with expected lengthening effects as a function of the prosodic factors under investigation. As it turned out, *V*-duration was found to be longer when the vowel was accented (both phonologically and lexically focused) than when it was unaccented (receiving no focus) regardless of whether the vowel was in CVN# and #NVC; it was also found to be longer IP-finally than IP-medially in CVN#, while no difference was observed in #NVC between the IP-initial and the IP-medial condition. Given that these durational effects were largely in line with prosodically-conditioned lengthening patterns previously reported in the literature, and given that *V*-duration was not the critical measure for the purpose of the present study, the basic results will be reported only in Appendix for the sake of completeness. Note that possible temporal effects on V-nasalization was already reflected in the relative timepoint measures.

2.5. *Statistical Analysis*

Effects of prosodic factors on the dependent variables *N-duration* and *V-nasality* (A1-P0) were evaluated in Repeated Measures Analysis of Variance (RM ANOVA) separately for the CVN (anticipatory) and the NVC (carryover) context, with two (within-subject) prosodic factors, Focus (LexFOC vs. PhonFOC vs. NoFOC) and Boundary (IP vs. Wd). For the assessment of the time course of vowel nasalization to be reflected in V-nasalization (A1-P0), Timepoint was added as an additional within-subject factor in two different analyses with (1) *the relative time points* (the near (25%), the midway (50%), the distant (75%) points of the vowel from the nasal) and (2) *the absolute time points* (25ms, 50ms, 75ms, 100ms away from the nasal). Each speaker's data were averaged over repetitions per condition across items in order to obtain each speaker's representative value per condition. When there was an interaction between factors, Bonferroni-corrected pairwise comparisons were made by running separate one-way ANOVAs to assess within-factor effects in a particular condition. The main effects and interactions with a *p*-value less than .05 were considered significant while η_p^2 (partial *eta*-squared) values and mean differences were given to estimate the effect size (Sheskin, 2003). Statistical analysis was done using IBM SPSS version 21.0.

3. Results

3.1. *CVN# (domain-final anticipatory context)*

3.1.1. *N-duration in CVN#*

Results of RM ANOVAs on N-duration in the CVN# (domain-final/coda) context are summarized in Fig. 2. As can be seen in the figure, there was a significant main effect of Focus on N-duration in CVN# ($F[2,28] = 58$, $p < .001$, $\eta_p^2 = .81$), so that *N*-duration was longer in the focused conditions (PhonFOC, LexFOC) than in the unfocused condition (NoFOC). In addition, between focused conditions, N-duration was longer in PhonFOC than in LexFOC, although the difference was not as large as the difference between focused and unfocused conditions. There was also a significant main effect of Boundary ($F[1,14] = 43.26$, $p < .001$, $\eta_p^2 = .76$), so that N-duration in CVN# was longer in the IP-final than in the Wd-final position, showing domain-final lengthening (in Fig. 2b).

There was, however, a significant interaction between Focus and Boundary ($F[2,28] = 9.09$, $p < 0.01$, $\eta_p^2 = .39$). As can be inferred from Fig. 2c, the interaction was in part due to the fact that the focus effect was *less* robust in the IP-final context ($\eta_p^2 = .57$) than in the Wd-final context ($\eta_p^2 = .87$). The interaction was also in part due to the fact that the boundary effect was *less* robust in the focused conditions (PhonFOC, $\eta_p^2 = .6$, LexFOC, $\eta_p^2 = .6$) than in the unfocused condition ($\eta_p^2 = .89$). In other words, the focus-induced temporal expansion was constrained in the IP context in which *N*-duration was already substantially elongated; and the boundary-induced temporal expansion was constrained in the focused conditions in which *N*-duration was already substantially elongated.

3.1.2. *V-nasalization in CVN#*

Results of RM ANOVAs on A1-P0 (V-nasalization) in the CVN# (anticipatory) context are summarized in Table 3, and visualized in Figs. 3 and 4. There was a significant main effect of the both prosodic strengthening factors, Focus and Boundary, on A1-P0, but the directionality of the effects differed. As can be seen from Fig. 3a and b, the vowel in CVN# was *less* nasalized (i.e., with larger A1-P0) when focused than unfocused, showing *coarticulatory resistance* or reduction of V-

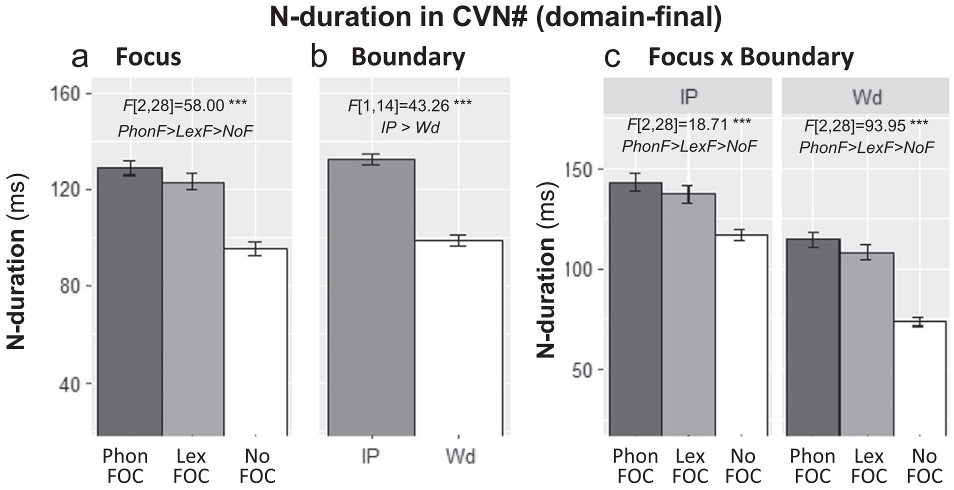

Fig. 2. *N*-duration in CVN#: (a) The main effect of Focus; (b) The main effect of Boundary; (c) The interaction between Focus and Boundary (***: $p<.001$; '>' indicates a difference at $p<.05$ in posthoc tests).

Table 3
A summary of RM ANOVAs for A1-P0 in CVN#: Anticipatory V-nasalization.

Factors	A1-P0 (relative timepoints)		A1-P0 (absolute timepoints)	
Focus	$F[2, 28]=21.85^{**}$	$\eta_p^2=.61$	$F[2, 28]=18.88^{**}$	$\eta_p^2=.38$
Boundary	$F[1, 14]=17.44^{**}$	$\eta_p^2=.56$	$F[1, 14]= 21.56^{**}$	$\eta_p^2=.61$
Timepoint	$F[2, 28]=36.06^{**}$	$\eta_p^2=.72$	$F[3, 42]=8.63^{**}$	$\eta_p^2=.57$
Focus × Boundary	$F[2, 28]=15.66^{**}$	$\eta_p^2=.53$	$F[3, 42]=9.89^{**}$	$\eta_p^2=.41$
Focus × Timepoint	$F[4, 56]<1^{n.s.}$	$\eta_p^2=.04$	$F[6, 84]<1^{n.s.}$	$\eta_p^2=.04$
Boundary × Timepoint	$F[2, 28]=4.66^{*}$	$\eta_p^2=.25$	$F[3, 42]=1.62^{n.s.}$	$\eta_p^2=.10$
Focus x Boundary × Timepoint	$F[4, 56]=2.2^{n.s.}$	$\eta_p^2=.14$	$F[6, 84]=1.18^{n.s.}$	$\eta_p^2=.08$

$^{n.s}$: $p>.1$; **: $p<.01$; *: $p<.05$.

nasalization in a prosodically strong context—i.e., under prominence due to focus at both the relative and the absolute timepoints (relative timepoints, F[2,28] = 21.85, $p<.001$; absolute timepoints, F[2,28] = 18.88, $p<.001$), while no difference was found between the phonologically-focused (PhonFOC) and the lexically-focused (LexFOC) conditions. The boundary effect, on the other hand, indicated an opposite pattern, showing *coarticulatory vulnerability* in a prosodically strong context—i.e., at a higher prosodic boundary. As can be seen in Fig. 3c and d, the vowel was *more* nasalized (i.e., with smaller A1-P0) in the IP-final position than in the Wd-final position at both the relative and the absolute timepoints. (Note that the relative timepoints in this case of anticipatory coarticulation work backwards from the nasal, so that, for example, the 25% point is near the nasal and the 75% point is farther away from it.)

Results also showed that both the focus and the boundary effects were pervasive across timepoints over the vowel. There was no Focus × Timepoint interaction at both the relative and the absolute timepoints (see Table 3). This means that the focus-induced *reduction* of V-nasalization (coarticulatory resistance) was not physically time-locked (or localized) to a later part of the vowel near the nasal in CVN#, but that the effect was robust even at a distant point away from the nasal (e.g., at the 75% point of the vowel, and at a timepoint 100 ms away from the nasal), as can be seen in Fig. 3a and b. As a mirror image, the boundary-induced *increase* in V-nasalization (coarticulatory vulnerability) was also found to be robust across timepoints over the vowel, especially as evident in no Boundary × Timepoint interaction at the absolute timepoints (see Table 3). This indicates that, like the case with the focus effect, the boundary effect was not strictly time-locked, but pervasive over the vowel at least up to a timepoint 100 ms away from the nasal in CVN#, as shown in Fig. 3d. The spreading of the boundary effect was further evident at relative timepoints. Although there was a significant interaction between Boundary and Timepoint at relative timepoints (see Table 3), the interaction was due to the fact that the boundary effect disappeared at the distant vowel point (i.e., the 75% point of the vowel), while the effect was still evident at the midpoint (50%) of the vowel, as shown in Fig. 3c.

The pervasiveness of prosodic effects across timepoints, however, does not mean that there was no gradient phonetic effect as a function of physical proximity to the source of coarticulation, the nasal consonant in the coda. In fact, there was a main effect of Timepoint (for both the relative and the absolute timepoints; see Table 3), indicating that A1-P0 decreases (i.e., V-nasalization increases) progressively as the

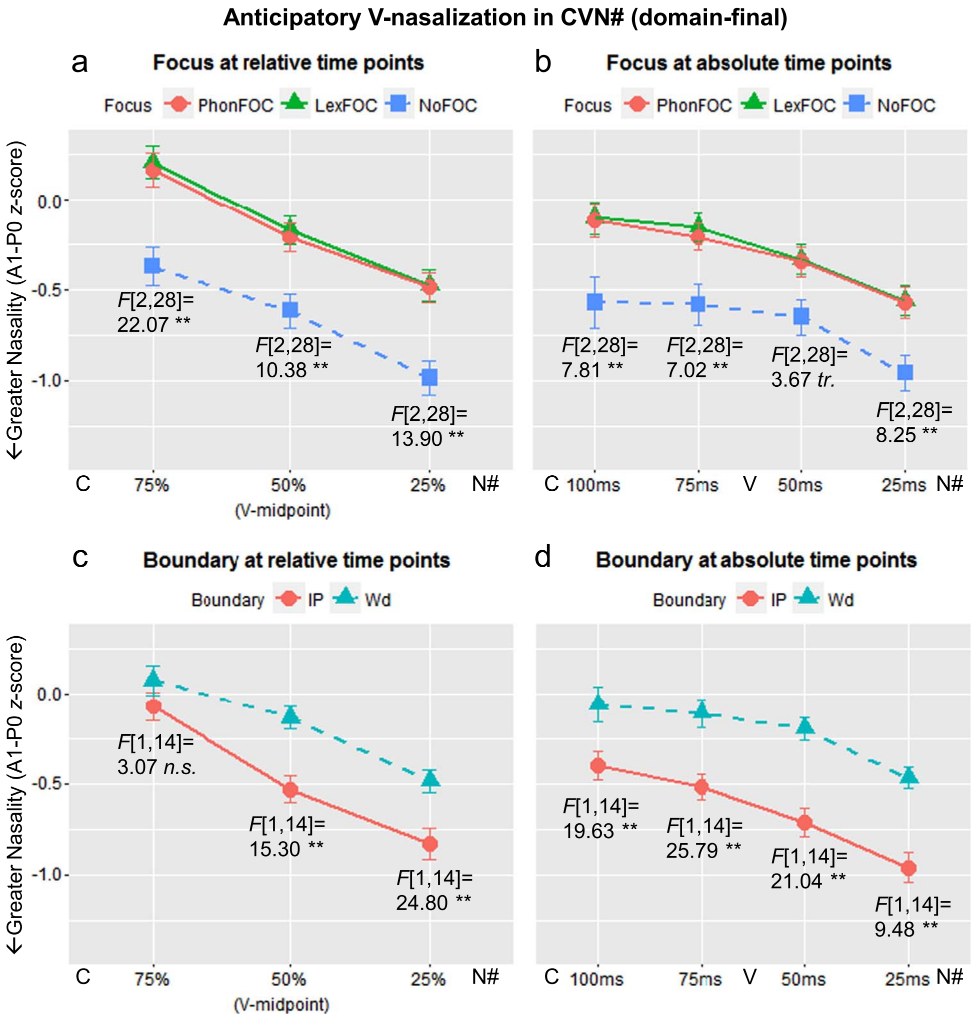

Fig. 3. Effects of Focus and Boundary on V-nasalization in CVN#: (a) The effect of Focus at the relative time points; (b) The effect of Focus at the absolute time points; (c) The effect of Boundary at the relative time points; (d) The effect of Boundary at absolute time points (**: $p<.01$, *tr.* $.05<p<.08$, *n.s*: $p>.08$).

timepoint gets physically closer to the nasal consonant as shown in Fig. 3a–d with an overall rightward declination of A1-P0 towards the nasal. Nevertheless, the fact that prosodic factors did not interact with Timepoint (see Table 3) indicates that the phonetically-driven coarticulatory effects are further modulated by higher-order prosodic factors (i.e., the focus-driven coarticulatory resistance and the boundary-induced coarticulatory vulnerability).

Finally, there was a significant two-way interaction between Focus and Boundary while there was no three way (Focus x Boundary × Timepoint) interaction (see Table 3). As can be seen in Fig. 4, the interaction was in part due to differential focus effects as a function of Boundary—i.e., the focus effect (coarticulatory resistance) was much more robust in the IP-final condition (relative, $F[2,28]=21.92$, $p<.01$; absolute, $F[2,28]=11.9$, $p<.01$) than in the Wd-final condition (relative, $F[2,28]=8.5$, $p<.01$, absolute, $F[2,28]=0.93$, $p>.1$). In other words, the Focus factor induced a greater coarticulatory reduction (suppression) of V-nasalization in a context (i.e., the IP-final condition) in which the vowel is more prone to nasalization due to boundary, thus leaving much room for further reduction. Seen from a different angle, the interaction was also in part attributable to differential boundary effects as a function of Focus—i.e., the boundary effect (coarticulatory vulnerability) was far greater in the unfocused condition (NoFOC, relative, $F[2,28]=20.8$, $p<.01$, $\eta_p^2=.35$; absolute, $F[2,28]=19.17$, $p<.01$, $\eta_p^2=.58$) than in the focused conditions (PhonFOC, relative, $F[2,28]=6.08$, $p<.05$, $\eta_p^2=.3$, absolute, $F[2,28]=12.52$, $p<.01$, $\eta_p^2=.47$; LexFOC, relative, $F[2,28]=1.76$, $p>.1$, $\eta_p^2=.11$, absolute, $F[2,28]=7.07$, $p<.05$, $\eta_p^2=.34$).

As for the boundary effect, on the other hand, V-nasalization increased more IP-finally than Wd-finally, especially in the unfocused context (i.e., NoFOC) in which the increase of V-nasalization was not heavily counteracted by focus.

3.2. #NVC (domain-initial carryover context)

3.2.1. N-duration in #NVC

Results of RM ANOVAs on *N*-duration in the #NVC (domain-initial/onset) context are summarized in Fig. 5. As can be seen in Fig. 5a, there was a significant main effect of Focus on N-duration in #NVC ($F[2, 28]=199.3$, $p<.001$, $\eta_p^2=.93$), so that *N*-duration was longer in the focused conditions (PhonFOC, LexFOC) than in the unfocused condition (NoFOC), while no difference was observed between focused conditions (PhonFOC vs. LexFOC). There was also a significant main effect of Boundary ($F[1, 14]=103.01$, $p<.001$, $\eta_p^2=.88$), but unlike the

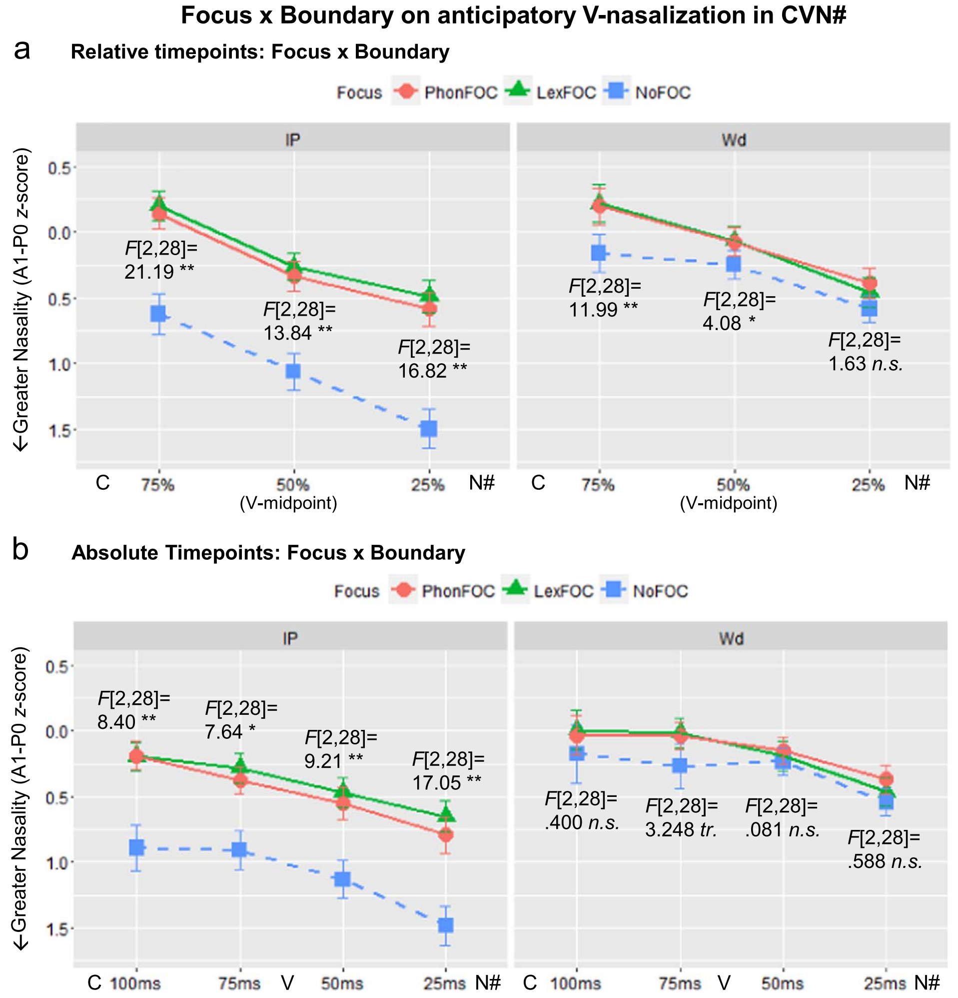

Fig. 4. Focus × Boundary interaction in CVN#: (a) at the relative time points and (b) at the absolute time points (**: $p<.01$, *: $p<.05$, *tr.* $.05<p<.08$, *n.s*: $p>.08$). (Note that although there was no three-way interaction involving Timepoint, the data were plotted at different timepoints in order to provide further information on the detailed time course of V-nasalization.).

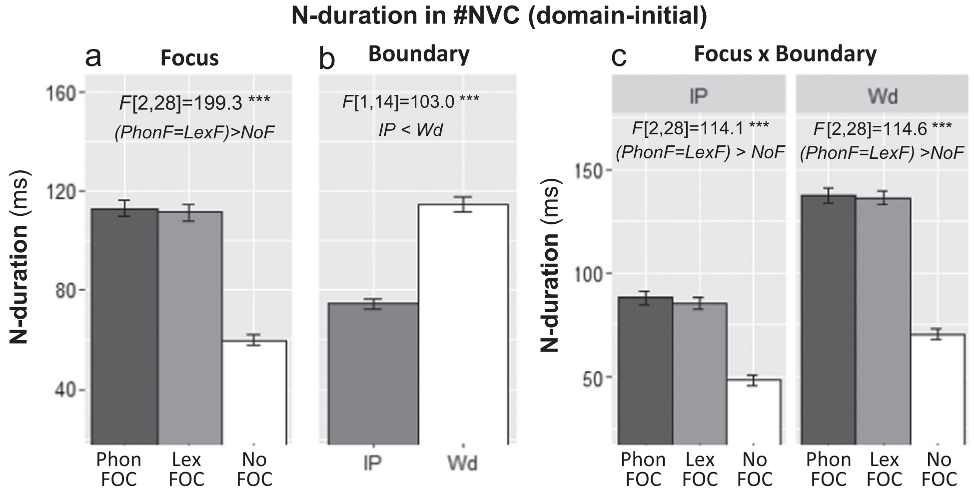

Fig. 5. *N*-duration in #NVC: (a) The main effect of Focus; (b) The main effect of Boundary; (c) The interaction between Focus and Boundary (***: $p<.001$; '>' indicates a difference at $p<.05$ in posthoc tests.).

domain-final case which showed longer N-duration at an IP boundary, the domain-initial *N*-duration was *shorter* in the IP-initial than in the Wd-initial position, as shown in Fig. 5b. There was also a significant interaction between Focus and Boundary (F[2,28] = 25.31, $p<.001$, $\eta_p^2 = .64$). The interaction was in part due to the fact that the focus-induced lengthening effect was attenuated in the IP-initial context in which there was a shortening force on *N*-duration.

Table 4
A summary of RM ANOVAs for A1-P0 in NVC: Carryover V-nasalization.

Factors	A1-P0 (relative timepoints)		A1-P0 (absolute timepoints)	
Focus	$F[2, 28]=16.92^{**}$	$\eta_p^2=.55$	$F[2, 28]=35.72^{**}$	$\eta_p^2=.51$
Boundary	$F[1, 14]=11.81^{**}$	$\eta_p^2=.46$	$F[1, 14]=15.49^{**}$	$\eta_p^2=.53$
Timepoint	$F[2, 28]=17.43^{**}$	$\eta_p^2=.56$	$F[3, 42]=14.49^{**}$	$\eta_p^2=.72$
Focus x Boundary	$F[2, 28]=.32^{n.s.}$	$\eta_p^2=.16$	$F[3, 42]=2.17^{n.s.}$	$\eta_p^2=.13$
Focus x Timepoint	$F[4, 56]=.64^{n.s.}$	$\eta_p^2=.04$	$F[6, 84]=3.09^{n.s.}$	$\eta_p^2=.18$
Boundary × Timepoint	$F[2, 28]=2.65^{n.s.}$	$\eta_p^2=.02$	$F[3, 42]=1.3^{n.s.}$	$\eta_p^2=.09$
Focus x Boundary × Timepoint	$F[4, 56]=1.11^{n.s.}$	$\eta_p^2=.07$	$F[6, 84]=.77^{n.s.}$	$\eta_p^2=.05$

$^{n.s.}$: $p>.1$; **: $p<.01$.

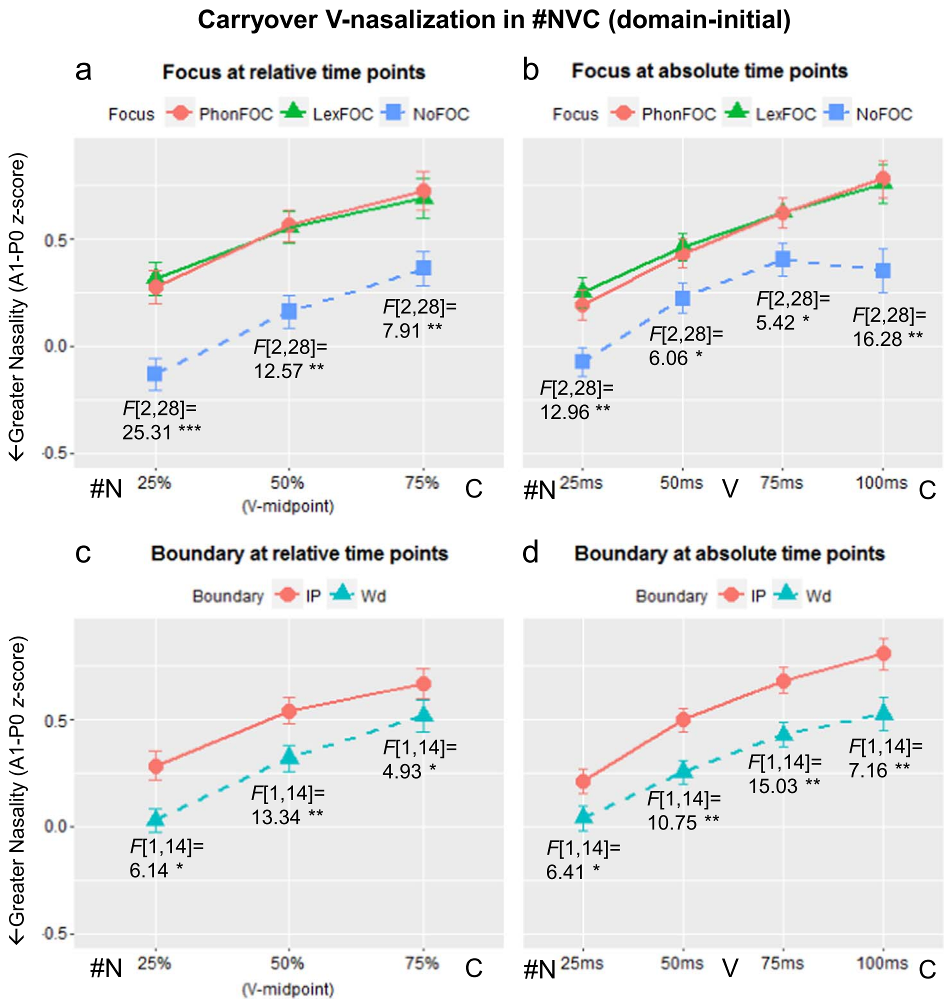

Fig. 6. Effects of Focus and Boundary on V-nasalization in #NVC: (a) The effect of Focus at the relative time points; (b) The effect of Focus at the absolute time points; (c) The effect of Boundary at the relative time points; (d) The effect of Boundary at absolute time points (***: $p<.005$, **: $p<.01$, *: $p<.05$).

3.2.2. *V-nasalization in #NVC*

Results of RM ANOVAs on A1-P0 (V-nasalization) in the #NVC (carryover) context are summarized in Table 4, and important findings are visualized in Fig. 6. There was a significant main effect of Focus, Boundary and Timepoint on A1-P0 at both the relative and the absolute timepoints. The focus effect on the carryover V-nasalization (in #NVC) indicated that the vowel was nasalized *less* (i.e., with larger A1-P0) when focused than unfocused with no difference between PhonFOC and LexFOC (p>.08 in all time points). This focus-induced coarticulatory reduction was similar to the one found in the domain-final anticipatory (CVN) context. But the boundary effect in #NVC (in the domain-initial carryover context) stood in sharp contrast with the boundary effect in the domain-final anticipatory context in CVN#. That is, in contrast with the vowel's coarticulatory *vulnerability* at a higher prosodic boundary in CVN# (in the domain-final context), the vowel in the domain-initial (#NVC) context showed coarticulatory *reduction* at a higher prosodic boundary.[2] Furthermore, unlike the case in CVN#, #NVC showed no Focus x Boundary interaction (see Table 4), indicating that the boundary effect was independent of focus conditions.

As for the main effect of Timepoint in #NVC, the effect was significant at both the relative and the absolute timepoints. As was the anticipatory case in CVN#, there was a gradient coarticulatory phonetic effect in the carryover (#NVC) context as well—i.e., V-nasalization was gradually and progressively attenuated (or A1-P0 increases) as a timepoint in the vowel was getting farther way from the nasal. Crucially, neither Focus nor Boundary interacted with Timepoint at both the relative and the absolute timepoints (see Table 4), indicating that prosodic effects were robust across timepoints even at a distant vowel point away from the onset nasal (i.e., at the 75% point of the vowel and at a point 100ms away from the onset nasal). This again indicates that V-nasalization as a phonetic coarticulatory process was further modulated by higher-order prosodic factors.

4. General discussion

4.1. Prominence-induced effects under focus

One of the basic findings of the present study regarding the focus effects was that the nasal consonant was temporally elongated (as shown in longer *N*-duration) in both CVN and NVC under focus, while a significant difference between focus conditions was observed only in CVN# (i.e., longer in the phonological than in the lexical focus condition). From the viewpoint of contrast enhancement, the increase in nasality for the consonant as reflected in longer N-duration may be interpreted as a paradigmatic enhancement of the [nasal] feature consistent with the localized hyperarticulation hypothesis (e.g., de Jong, 1995, 2004). Importantly, however, despite the temporal expansion of the nasal in the focused conditions, its coarticulatory impact on the neighboring vowel was not augmented but reduced in both the anticipatory (CVN) and the carryover (NVC) context. The reduction of its coarticulatory impact on the neighboring vowel therefore does not resonate with an enhancement of the [nasal] feature, and it shows no coarticulatory aggression effect although coarticulatory impacts of some phonetic features are expected to be greater when the triggering segment is stressed (or more broadly when accented) (e.g., Fowler & Saltzman, 1993; Cho, 2004; cf. Farnetani, 1990). (See below for further discussion on this point in relation to Beddor's (2009) hypothesis.)

The observed coarticulatory reduction of V-nasalization under focus can be interpreted as the vowel's *coarticulatory resistance* to encroachment of the nasal which is consistent with findings in the literature on coarticulation. For example, coarticulatory resistance as observed in V-to-V coarticulation may be viewed as another type of prosodic strengthening, contributing to marking prosodic structure (Cho, 2004) in correlation with degree of accentuation (Fletcher, 2004). Likewise, the reduction of V-nasalization under focus found in the present study demonstrates another case of coarticulatory resistance in V-nasalization. A question then arises as to why the vowel resists coarticulation with the nasal consonant even in the phonologically focused condition in which the nasal consonant was emphasized. An answer may be found in relation to the finding that there was no difference between phonological focus and lexical focus. Recall that the nasal consonant was phonologically contrasted with the oral consonant in the phonological focus condition (e.g., ‘mob’ *vs.* ‘bob’ in the onset or ‘bomb’ *vs.* ‘bob’ in the coda), while in the lexically focused condition, the target bearing word was contrastive with a semantically related word (e.g., ‘bomb’ vs. ‘war’). The fact that there was no difference between the two conditions suggested that both types of focus were mediated by prominence (i.e., pitch accent), yielding commensurable effects regardless of focus types. Given that the vowel is a stress-bearing unit and a locus of accentuation in prosodic structure in English (e.g., Turk & White, 1999), the vowel's coarticulatory resistance to nasalization can be taken as a prominence enhancement strategy for the vowel, which may still be seen as paradigmatic enhancement in a sense that the vowel's [orality] feature is enhanced.

4.2. Boundary-induced effects: domain-final vs. domain-initial case

The boundary effects on N-duration and V-nasalization were found to be asymmetrical as a function of whether the nasal was in the final, anticipatory context (CVN#) or in the initial, carryover context (#NVC). In CVN#, both *N*-duration and V-nasalization *increased* in the domain-final position, whereas in #NVC, both *N*-duration and V-nasalization *decreased* in the domain-initial positon.

The shortening of initial *N*-duration and the reduction of carryover V-nasalization in the domain-initial position are interpretable in terms of syntagmatic enhancement. At the outset of the paper, it was hypothesized that the boundary-induced prosodic effect would be motivated structurally, so that the effect could give rise to an enhancement of syntagmatic contrast which would maximize CV contrast in the initial position—i.e., the consonant would become more consonant-like (or less sonorous) by virtue of which the contrast between the consonant and the following vowel would be enhanced (e.g., Pierrehumbert & Talkin, 1992; Fougeron & Keating, 1997; Fougeron 1999; Cho, 2004, 2016). The shortening of initial *N*-duration, which was also observed in previous studies (e.g., Fougeron & Keating, 1997; Cho & Keating, 2001), may be interpreted as an enhancement of the [consonantality] feature for the nasal which participates in maximizing CV contrast. The reduction of V-nasalization in #NVC also goes hand in hand with the shortening of *N*-duration, interpretable as increasing the vowel's [orality] feature which may contribute to an enhancement of CV contrast.

The reduction of V-nasalization due to boundary, however, cannot be characterized as coarticulatoy resistance which was assumed to underlie the prominence-induced reduction of V-nasalization. As a boundary effect, the initial nasal consonant itself was reduced in its nasality, which in turn has a reduced coarticulatory influence on the following vowel (i.e., less V-nasalization). In other words, it is likely that the following vowel is nasalized less in the IP-initial context not because the vowel resisted the nasal's coarticulatory influence but because

---

[2] As a reviewer suggested, the vowel in the initial context (in #NVC), all else being equal, may be taken to be more salient than the one in the final context (in CVN#), hence the greater boundary-induced coarticulatory reistance (reduction) in the former.

the nasality of the consonant that triggered coarticulation was weakened in that context

Unlike shorter *N*-duration and less V-nasalization in the domain-initial (#NVC) position, *N*-duration in the domain-final (CVN#) position was longer and V-nasalization was greater before an IP than before a Wd boundary. The longer *N*-duration, which was in line with frequently observed phrase-final lengthening in the literature, may have exerted greater coarticulatory impact on anticipatory V-nasalization. The boundary effect in CVN# was similar to the prominence effect in that N-duration was elongated, but it differed sharply from the prominence effect as the former induced coarticulatory vulnerability, and the latter coarticulatory resistance. The results are consistent with the articulatory weakening hypothesis that pertains to domain-final consonants as was discussed in the introduction (e.g., Byrd, 1996; Fougeron, 1999; Keating et al., 1999). From an articulatory point of view, it seems reasonable to assume that a weakening of consonantality allows for a loosening of the articulatory linkage of the oral constriction and the velum lowering gesture, especially given that they are assumed to be anti-phase coupled, thus being less stable compared to the in-phase coupling assumed for the initial NVC case (See Section 4.4 for further discussion on the intergestural coupling relationship.). To the extent that the hypothesized loosening of the articulatory linkage works, it would increase coarticulatory propensity, accounting for more nasal coarticulation in the domain-final position. (At this point, it is worth noting a reviewer's point that the assumed articulatory linkages may not necessarily increase coarticulatory propensity, but rather simply increase coarticulatory fluctuations in both directions, causing more or less coarticulation. We, however, propose that when there is a natural coarticulatory force presumably driven by the principle of ease of articulation and efficiency arising with parallel transmission of phonetic information (e.g., Mattingly 1981), the loosened articulatory linkage, all else being equal, works in a direction to accommodate the coarticulatory force rather than to suppress it.) In the acoustic dimension, such a weakening could reduce the nasal's [consonantality] as reflected in an increase in both *N*-duration and V-nasalization (i.e., the more the nasality, the less consonant-like (or the more sonorous) the nasal).

4.3. V-nasalization as a coarticulatory process under the control of the speaker

The results of the present study confirmed the low-level characteristic of V-nasalization by showing that the degree of V-nasalization gradually increased as a measurement point in the vowel became physically closer to the triggering nasal consonant in both the anticipatory (CVN) and the carryover (NVC) context. However, just because a low-level phonetic process is physiologically and/or biomechanically driven in origin does not mean that the process is automatic (e.g., Keating, 1985; Cho & Ladefoged, 1999). A close examination of vowel nasalization patterns found across languages reveals that languages may vary in the degree of vowel nasalization (e.g., Clumeck, 1976; Cohn, 1990, 1993; Solé, 1992, 1995, 2007; Krakow, 1989, 1999; Farnetani & Recasens, 2010; see Zellou & Tamminga (2014), for a systematic difference among age groups within a dialect in English). A recent study by Scarborough et al. (2015), for example, showed that in Lakhota which employs the oral-nasal vowel contrast, although there is no phonetic reason for the underlying nasal vowel (which is produced already with substantial nasality) to be more nasalized, nasal vowels in the nasal consonant context (e.g., VN) were produced with even more nasality not necessarily near the source of coarticulation (the nasal consonant) but at an earlier point during the vowel. The authors interpreted this result as suggesting that the degree of coarticulatory vowel nasalization is modulated by the phonetic grammar and/or by the phonological contrast in a language specific way. This supports the view that coarticulatory patterns that may arise as a low-level phonetic process for mechanical reasons are in fact controlled or tuned by speakers in language-specific ways, and therefore must be specified in the phonetic grammars of individual languages. In a similar vein, the results of the present study indicate that such a low-level coarticulatory process is further fine-tuned even within a language driven by the phonetics-prosody interplay. Recall that the coarticulatory effect as reflected in V-nasalization was counteracted by the prominence system of the language. A similar reasoning may also apply to the boundary-related effect. Although the directionality of the boundary effect differed as a function of position (initial vs. final), what emerged was that the boundary-related effect on V-nasalization was also maintained over a large portion of the vowel rather than as a time-locked process in which case V-nasalization would have been localized to the vicinity of the nasal.

These results taken together consolidate insights into the nature of the phonetics-prosody interplay which is assumed to regulate the human motor behavior. The phonetic granularity of vowel nasalization fine-tuned by high-order prosodic structure is therefore characterized as being cognitive rather than automatic. A prosodically-driven fine-tuning of a low-level process must be specified in the grammar of the language in reference to linguistic principles such as paradigmatic and syntagmatic contrast enhancements. Such a cognitive phonetic process, however, does not belong to the realm of phonology in a traditional sense. The vowel in English is often assumed by phonologists to be unspecified for its nasality (e.g., Cohn, 1990, 1993), and therefore vowel nasalization in English is thought to operate outside the phonological component of the grammar. How could then such a cognitive aspect of speech production that has traditionally been thought to be outside the realm of phonology be adequately captured in the grammar of the language?

The 'phonetic' component of the grammar as has been substantiated by Keating (1984, 1985, 1996; see Cho and Ladefoged (1999) for a related discussion) may do the job by operating 'phonetic' rules that fine-tune the input that is fed down from the phonological component of the grammar. One way of implementing a fine-tuning of speech production is proposed in Keating's (1990) window model (see also Byrd (1996), Cho (2004), Cohn (1990), Docherty (1992) and Keating (1996); see, for example, Guenther, 1995; Guenther, Hampson, & Johnson, 1998, for an independent development of a window model). In the model, a window is specified as a range of articulatory movement (or acoustic variation) allowed for each segment. Crucially, the range of articulations may be expanded or shrunk (Keating & Shattuck-Hufnagel, 2002), so that an expanded window allows for more variation than a shrunk one, which may capture coarticulatory variation that arises as a function of prosodic factors. Developing this idea further, Cho

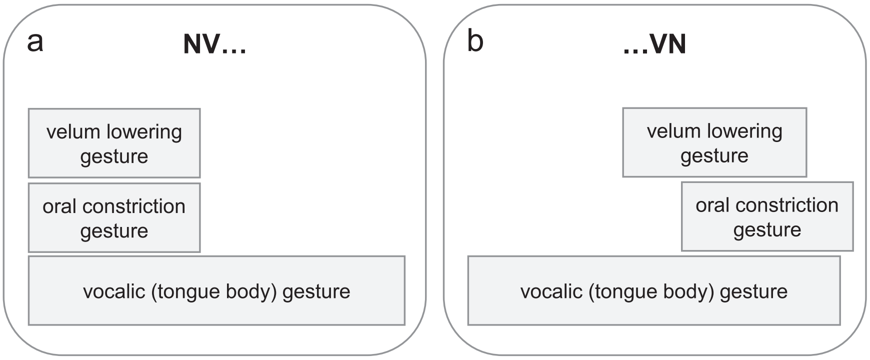

Fig. 7. Schematics of assumed gestural scores for the timing between the velum lowering and the oral constriction formation gestures in NV vs. VN (cf. Brownman and Goldstein, 1995).

(2004) proposed that the size of the window be invariably determined for a segment in line with Keating's (1990) earlier assumption, but prosodic factors could specify an operating target region within the fixed window. The model therefore provides a way of capturing systematic phonetic variation in a linguistically intuitive way, so that the phonological category of a segment may be expressed by an invariable window size, but the contextually-driven variation is accommodated within that window. For example, in V-to-V coarticulation, a prominence factor may narrow an operating target region within a window for a vowel, which effectively reduces its vulnerability to coarticulatory influence, accounting for coarticulatory resistance, whereas a widely specified operating region allows for a wider range of variation, showing greater degree of coarticulation as needed. (Here, by variation we refer to a systematic change as a function of linguistic context rather than numerical variability in distribution of phonetic values in a particular context.) However, the window model was developed primarily for coarticulatory variation in the spatial dimension that involves similar articulators (e.g., the tongue body for V-to-V coarticulation), but the model has not been explicit about coarticulatory variation that involves temporal overlap of different articulatory gestures such as the oral and the velum gestures that determines the nasalization of the vowel. (But see Byrd (1996) who proposed a phase window model which specifies a range of possible values in which intergestural overlap is implemented in the temporal dimension). It therefore remains to be seen how this interpretation of windows can be further developed in a computationally implementable way, capturing various types of coarticulatory variation in both the spatial and the temporal dimension that are modulated by prosodic factors.

Finally, results of the present study have implications for Beddor's (2009) hypothesis as laid out in the introduction. The prominence-driven lengthening of the nasal consonant in the coda position and the accompanied shortening of vowel nasalization is indeed compatible with the prediction that follows from the assumption that the temporal domain of the velum lowering gesture associated with the nasal coda remains constant, so that the longer the nasal murmur in the coda, the shorter the duration of vowel nasalization. This would lead to an alternative account for the present finding under prominence at least in the anticipatory context, undermining or at least being compatible with the enhancement-driven coarticulatory resistance account as discussed above. Crucially, however, the invariant velum lowering account does not apply to the boundary-driven vowel nasalization effect observed in the present study. Recall that the duration of vowel nasalization in CVN# was found to be longer IP-finally than IP-medially, despite the fact that the duration of the nasal consonant was also lengthened in that context. It remains to be seen how Beddor's model would account for the asymmetrical effect due to prominence marking vs. boundary marking, especially for the boundary-induced temporal expansion of both the nasal murmur during the consonantal closure and the extended duration of vowel nasalization in the CVN# context.

4.4. Dynamical accounts

Another way of capturing a cognitive aspect of the phonetic granularity of speech production that has traditionally been thought to be outside the realm of phonology may be considered by resorting to the notion of Articulatory Phonology in which phonetics and phonology are integrated with a set of unified formal mechanisms without referring to features—i.e., phonological contrasts are directly expressed by coordination (timing) of articulatory gestures in temporal and spatial dimensions (e.g., Browman & Goldstein, 1990, 1992; Goldstein & Fowler, 2003; Goldstein et al., 2006). In Articulatory Phonology, the intergestural timing is assumed to be specified in the gestural score, which is in principle expected to vary across languages as the timing information is assumed to be stored in the (language-specific) lexicon. In this framework, the degree of vowel nasalization can be defined in terms of the relative timing between the vowel gesture and the velic lowering gesture, which overlap in time and space. In recent development of a task-dynamic model of speech production (e.g., Saltzman, Nam, Krivokapić & Goldstein, 2008), it is assumed that phonological systems prefer to employ intrinsically stable modes of intergestural timing relationship which is either an in-phase mode (that characterizes CV structures in which gestures are coordinated synchronously) or an anti-phase mode model (that characterizes VC structures in which gestures are coordinated sequentially).

The asymmetrical V-nasalization in the initial (carryover) vs. the final (anticipatory) context that has been found in English (and across many other languages) may be accounted for by differential intergestural coupling relationships, in-phase vs. anti-phase (e.g., Goldstein et al., 2009; Byrd et al., 2009). As schematized in Fig. 7, the assumed in-phase coupling relationship between the velum lowering and the oral constriction (closing) gesture for the nasal in the onset indicates that the gestural movements initiates roughly simultaneously (Fig. 7a), while the anti-phase coupling relationship indicates the velum lowering gesture is timed earlier relative to the oral constriction gesture for the nasal in the coda (Fig. 7b). The overall reduced amount of V-nasalization in the initial (carryover) context is thus consistent with the in-phase (synchronous) coupling relationship of the velum-oral gestural timing that is hypothesized to underlie CV structures. On the other hand, the greater amount of V-nasalization in the final (anticipatory) context is consistent with the anti-phase (sequential) coupling relationship underlying VC structures (e.g., Goldstein et al., 2009; Byrd et al., 2009). Furthermore, the assumed differential coupling relationships are hypothesized to differ in terms of phasing stability—i.e., they are more stable in an in-phase mode than in an anti-phase mode. In connection with this, the hypothesized loosening of the articulatory linkage of the oral-velum gestures in the final (anticipatory) context, which induces more nasal coarticulation domain-finally than domain-initially, is indeed consistent with the assumptions, as discussed in the introduction, that the intergestural timing is less stable in an anti-phase mode than in an in-phase mode and that the oral articulatory force (which would suppress the velum lowering gesture as discussed by Fougeron (1999, 2001)) is reduced, allowing for more coarticulatory flexibility in the final (anticipatory, CVN#) context.

The assumption about differential gestural coupling relationships thus provide insights into how asymmetrical patterns in the domain-initial (carryover) and the domain-final (anticipatory) context may have emerged in a dynamical system. However, it does not directly answer the question as to how the intergestural timing, whether in-phase or anti-phase, is further modulated by boundary strength or by prominence. Saltzman et al. (2008) indeed suggested that prosodically-conditioned variation in speech production can be adequately incorporated into the model by invoking a new component into a task-dynamic model, *modulation gestures* which have a function of fine-tuning the spatial and temporal properties of constriction gestures that are concurrently being activated. There are two types of modulation gestures. Temporal modulation gestures (μ_T-gestures) modulate the rate of utterance timeflow by smoothly changing all frequency parameters of the planning oscillator ensemble. Spatial modulation gestures (μ_S-gestures) magnify or reduce the motions of constriction gestures by smoothly changing the spatial target parameters of these constriction gestures. For example, a spatial modulation gesture (a μ_S-gesture) is assumed to operate directly on constriction gestures in a stressed syllable accounting for frequently observed spatial expansion of gestures under prominence. Similarly, a temporal modulation gesture (a μ_T-gesture) is assumed to operate in connection with constriction gestures at prosodic junctures, which regulates a local slowing down in the vicinity of prosodic boundary (e.g., phrase-final lengthening). The temporal modulation gesture that operates at prosodic juncture is called "a prosodic gesture" (a π-gesture) (Byrd, 2000, 2006; Byrd, Kaun, Narayanan, & Saltzman, 2000; Byrd & Saltzman, 2003; Byrd, Krivokapić, & Lee, 2006). The model, however, has not been fully explicit on the question regarding whether and how these modulation gestures operate in a way to account for different types of prosodically-conditioned coarticulatory variation such as coarticulatory resistance vs. coarticulatory vulnerability as found in the present study. The results of the present study therefore have implications for this question that will be useful for developing a task-dynamic model.

First, although the timing between the oral and velic gestures which determines degree of V-nasalization is assumed to differ between CVN and NVC (with anti-phase *vs.* in-phase couplings), prominence (focus) was found to influence the intergestural timing in a unified way—i.e., by coarticulatory resistance as reflected in reduction of V-nasalization regardless of directionality (CVN/NVC) and focus type (lexical/phonological). The degree of coarticulatory resistance indicates that intergestural timing needs to be directly modified to reduce overlap between the velum and the oral gesture. However, a direct modification of intergestural timing does not seem to be the province of modulation gestures as they serve primarily to modulate the spatio-temporal properties of all concurrently active constriction gestures. It may be possible that spatial/temporal modulations of concurrently active constriction gestures give rise to a change in intergestural timing on the surface (e.g., Byrd & Saltzman, 2003), but mechanisms for directly modifying intergestural timing of constricting gestures postlexically would be a useful device in the model in order to illuminate the systematic coarticulatory resistance under the control of the speaker.

Second, the asymmetry found in the anticipatory CVN# vs. the carryover #NVC context has an implication for the theory of π-gesture that postulates a local slowing down of articulatory movement at a prosodic juncture. Crucially, preboundary lengthening of final *N*-duration vs. postboundary shortening of initial *N*-duration, and their accompanying increase vs. decrease in V-nasalization indicate that the model needs to devise separate ways to account for effects on the intergestural timing in CVN# (with anti-phase coupling) *vs.* #NVC (with in-phase coupling)—i.e., the velum lowing gesture seems to *start* earlier before a higher than a lower prosodic boundary in CVN#, while the velum lowering gesture seems to *end* earlier after a higher than a lower prosodic boundary in #NVC. In simulations of the kinematic consequences of π-gestures, Byrd & Saltzman (2003) in fact demonstrated a possibility that intergestural timing may be modified as a result of operation of π-gesture, but again, the asymmetrical boundary effect on the realization of the nasal consonant and its coarticulatory impact on the neighboring vowel need to be specified separately and incorporated in a dynamical model of boundary-related speech production.

Finally, the results of the present study suggest that fine-tuning of intergestural timing occurs as a function of prosodic structure in a way to enhance linguistic contrast. Modulation of the spatial and the temporal properties of constriction gestures thus needs to be determined in a dynamical model by making reference to the linguistic contrast system of the language in which different types of contrast enhancement (*paradigmatic* vs.

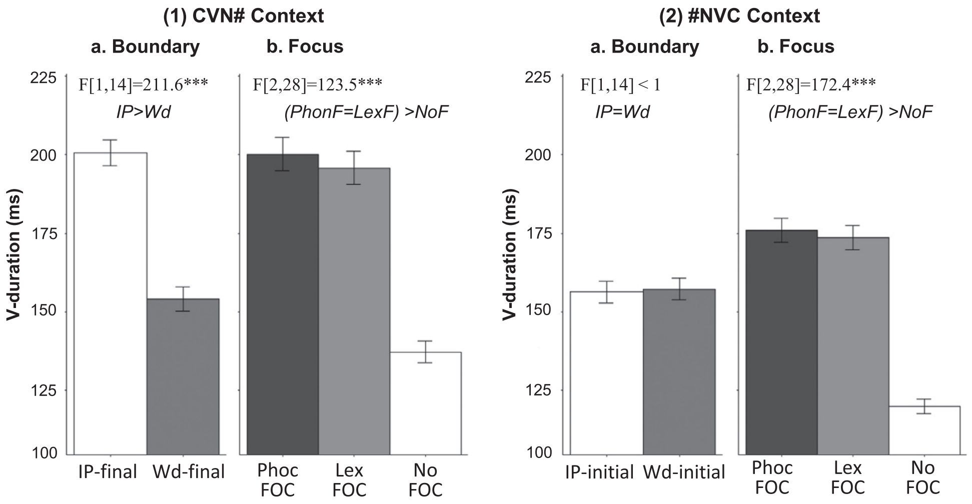

Fig. 8. *V*-duration: (1a) The main effect of Boundary in CVN#; (1b) The main effect of Focus in CVN#; (2a) The main effect of Boundary in #NVC; and (2b) The main effect of Focus in #NVC. (***: *p* < .001; '>' indicates a difference at *p* < .05 in posthoc tests.).

syntagmatic) yields different timing patterns in the surface. Another challenge for the model is to devise a way to accommodate the interaction between prominence and boundary factors. However fine-tuning is implemented in a task-dynamic model, it needs to take into account not only influences of individual higher-order factors, but also their interactions in a linguistically informed way to reflect their communicative functions.

5. Conclusion

The present study has demonstrated that effects of prosodic structure are differentially reflected in N-duration and V-nasalization, depending on its source. The prominence-induced coarticulatory resistance indicates that prominence enhances the [orality] feature of the vowel (rather than [nasality] of the consonant) even when focus fell on the nasal, showing paradigmatic contrast enhancement of the vowel. The boundary-induced domain-initial effect, on the other hand, increases the nasal's [consonantality] through its shortening and reduction of its coarticulatory exertion on the following vowel, showing syntagmatic (CV) contrast enhancement. Yet, this boundary effect on *initial* nasals stands in sharp contrast with the boundary effect on *final* nasals the latter of which is better characterized as a weakening of the nasal's consonantality. The weakening may have led to a loosening of the articulatory linkage of the oral constriction and the velum lowering gesture, increasing coarticulatory propensity. These results also indicate that the low-level phonetic phenomenon which may have originated from physiological and biomechanical properties of speech may indeed be fine-tuned under the control of the speaker, which needs to be integrated into the cognitive system of the individual language and therefore to be linguistically specified in the grammar of the language. It is proposed that while a fine-tuning of the phonetic granularity as a function of prosodic structure may be captured in a computationally feasible way in a task dynamic model, exact mechanisms of prosodically-conditioned fine-tuning of intergestural timing should be determined by making reference to the relationship between dynamical underpinnings of speech timing and linguistic contrasts.

Acknowledgement

We would like to thank Will Styler and Rebecca Scarborough for letting us use their Praat scripts for extracting A1-P0 values, and Jiyoung Jang for her assistance. A preliminary version of this paper (with a smaller number of subjects) was presented at the 18th International Congress of Phonetic Sciences 2015, Glasgow, UK (Cho, Kim, & Kim, 2015). Part of the data reported here was used in the second author's MA thesis at Hanyang University (Seoul, Korea). 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).

Appendix. Results on *V*-duration

The results of RM ANOVAs on *V*-duration in CVN# indicated that there was a main effect of both Boundary and Focus on *V*-duration in CVN#, so that *V*-duration was longer in the IP-final than in the Wd-final (IP-medial) condition (Fig. 8.1a), and it was longer when the vowel was focused than when it was unfocused (Fig. 8.1b). No difference was found due to focus type (phonological vs. lexical). In #NVC, on the other hand, there was no significant Boundary effect on *V*-duration (Fig. 8.2a), but as was the case in CVN#, Focus yielded a significant main effect on *V*-duration with no difference between the phonological and the lexical focus condition. In both CVN# and #NVC, there was no significant interaction between Boundary and Focus (CVN#, F[2,28] < 1, *p* > .1; #NVC; F[2,28] = 1.44, *p* > .1).

References

Beddor, P. S. (2009). A coarticulatory path to sound change. *Language*, *85.4*, 785–821.

Beddor, P. S. (2015). The relation between langauge users' perception and production repertoires. *Proceedings of the 18th International Congress of Phonetic Sciences.* Glasgow, UK. Paper number 1041.1-9 retrieved from ⟨http://www.icphs2015.info/pdfs/Papers/ICPHS1041.pdf⟩.

Beddor, P. S., Harnsberger, J. D., and Lindemann, S., Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates, *Journal of Phonetics* ***30***, 2002, 591-627.

Browman, C., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 341–376). Cambridge: Cambridge University Press.

Browman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

Browman, C., & Goldstein, L. (1995). Gestural syllable position effects in American English. In F. Bell-Berti, & L. J. Raphael (Eds.), *Producing speech: Contemporary issues for Katherine Safford Harris* (pp. 19–33). New York: AIP Press.

Bladon, A., & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics*, *4*, 137–150.

Bladon, R. A. W., & Nolan, F. (1977). A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics*, *5*, 185–193.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*(2), 209–244.

Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, *57*(1), 3–16.

Byrd, D. (2006). Relating prosody and dynamic events: Commentary on the papers by Cho, Navas, and Smiljanic'. In L. Goldstein, D. H. Whalen, & C. Best (Eds.), *Papers in laboratory phonology 8* (pp. 549–561). Mouton de Gruyter.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe, & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 70–87). Cambridge: Cambridge University Press.

Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effect. *The Journal of the Acoustical Society of America*, *120*(3), 1589–1599.

Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association*, *38*, 187–202.

Byrd, D., & Saltzman, E., The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening, *Journal of Phonetics*, ***31***, 2003, 149–180.

Byrd D.,Tobin S., Bresch E. and Narayanan S., Timing effects of syllable structure and stress on nasals: a real-time MRI examination, *Journal of Phonetics* ***37***(1), 2009 97-110.

Chen, M. (1996). *Acoustic correlates of nasality in speech* Ph.D. dissertation. Massachusetts Institute of Technology.

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

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: evidence from Korean. *Phonetica*, *58*, 129–162.

Cho, T. (2004). Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*, 141–176.

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English. *Journal of the Acoustical Society of America*, *117*(6), 3867–3878.

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

Cho, T., & Jun, S. (2000). Domain-initial strengthening as featural enhancement: aerodynamic evidence from Korean. *Chicago Linguistics Society*, *36*, 31–44.

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, *29*, 155–190.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, *37*, 466–485.

Cho, T., Kim, D. and Kim, S. (2015). Prosodic strengthening on consonnatal nasality and its asymmetric coarticulatory influence on vowel nasalization in CVN# and #NVC in English. In*Proceedings of the 18th international congress of phonetic sciences.* Glasgow, UK. Paper number 0608.1-4 retrieved from <http://www.icphs2015.info/pdfs/Papers/ICPHS0608.pdf>.

Cho T. and Ladefoged L., Variation and universals in VOT: evidence from 18 languages, *Journal of Phonetics* ***27***, 1999, 207-229.

Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics*, *39*(3), 344–361.

Cho, T., Lee, Y., & Kim, S. (2014). Prosodic strengthening on the /s/-stop cluster and the phonetic implementation of an allophonic rule in English. *Journal of Phonetics*, *46*, 128–146.

Cho, T., & McQueen, J. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*, 121–157.

Cho, T., McQueen, J., & Cox, E. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243.

Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, *4*, 337–351.

Cohn, A. (1990). Phonetics and phonological rules of nasalization. Ph.D. disseration. UCLA (Also published as UCLA Working Papers in Phonetics, No. 76.).

Cohn, A. (1993). Nasalization in English: phonology or phonetics. *Phonology*, *10*(01), 43.

de Jong, K. (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.

de Jong, K. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, *32*, 493–516.

de Jong, K., & Zawaydeh, B. A. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variationin Arabic vowel duration. *Journal of Phonetics*, *30*, 53–75.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*, Vol. 9. Walter de Gruyter.

Farnetani, E. (1990). VCV lingual coarticulation and its spatiotemporal domain. In W. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 93–130). Springer.

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.

Farnetani, E., & Recasens, D. (2010). Coarticulation and connected speech processes. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd edition). Oxford: Blackwell.

Fletcher, J. (2004). An EMA/EPG study of vowel-to-vowel articulation across velars in Southern British English. *Clinical Linguistics & Phonetics*, *18*(6-8), 577–592.

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 edition). Oxford: Blackwell.

Fougeron, C. (1999). Prosodically conditioned articulatory variations: A review. *UCLA Working Papers in Phonetics*, *97*, 1–74.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*(2), 109–135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *106*, 3728–3740.

Fowler, C. A., Production and perception of coarticulation among stressed and unstressed vowels, *Journal of Speech and Hearing Research*, ***46***, 1981, 127–139.

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

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

Gordon, M. (1996). The effect of stress and prosodic phrasing on duration, acoustic amplitude and airflow of nasal consnants in Estonian. *UCLA Working Papers in Phonetics*, *92*, 151–159.

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.), *Action to Language via the Mirror Neuron System* (pp. 215–249). New York: Cambridge University Press.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N. O. Schiller, & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159–207). Mouton de Gruyter.

Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled oscillator planning model of speech timing and syllable structure. *Frontiers in Phonetics and Speech Science*, 239–250.

Gow, D. W., J. Melvold, & S. Manuel. (1996). How word onsets drive lexical access and segmentation: evidence from acoustics, phonology, and processing. In *Proceedings of ICSLP (CD-ROM)*. Philadelphia, USA.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*(3), 594.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, *105*(4), 611.

Harrington, J., Fletcher, J., & Beckman, M. (2000). Manner and place conflicts in the articulation of accent in Australian English. In M. B. Broe, & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 40–51). Cambridge: Cambridge University Press.

Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics*, *23*(3), 305–322.

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: Academic Press.

Keating, P. A. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 451–470). Cambridge: Cambridge University Press.

Keating, P. A. (1996). The phonology–phonetics interface Studia grammatica 41. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 262–278). Berlin: Akademie Verlag.

Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics*, *101*, 112–156.

Keating, P., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, *97*, 157–173.

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

Krakow, R. A. (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures* Ph.D. Dissertation. Yale Univeresity.

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

Krivokapić, J., & Byrd, D. (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics*, *40*, 430–442.

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

Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, *35*, 301–320.

Ladefoged, P. (2001). *A course in phonetics*. Fort Worth: Harcourt College Publishers.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–440.). Dordrecht: Kluwer Academic Publishers.

Manuel, S., The role of contrast in limiting vowel-to-vowel coarticulation in different languages, *Journal of the Acoustical Society of America*, ***88***(3), 1990, 1286–1298.

Manuel, S., Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques,* 1999, Cambridge, UK: Cambridge University Press, 179-198.

Mattingly, I. G., Phonetic representation and speech synthesis by rule. In T. Myers, J. Laver, & J. Anderson (Eds.), *The Cognitive representation of Speech*, 2009, Amsterdam: North-Holland, 415-420.

Mitterer H., Cho T. and Kim S., How does prosody influence speech categorization?, *Journal of Phonetics* 54 ,2016, 68-79.

Mücke, D., Grice, M., & Cho, T. (2014). More than a magic moment–Paving the way for dynamics of articulation and prosodic structure. *Journal of Phonetics*, *44*, 1–7.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marsico, I. Chitoran, & C. Coupé (Eds.), *Approaches to phonological complexity* (pp. 299–328). Mouton de Gruyter.

Parker, S.G., Quantifying the Sonority Hierarchy, (Ph.D. Dissertation), 2002, University of Massachusetts Amherst.

Pierrehumbert J. and Talkin D., Lenition of /h/ and glottal stop, In: Docherty G.J. and Ladd D.R.,(Eds.), Papers in laboratory phonology II: Gesture, segment, prosody, 1992, Cambridge University Press; Cambridge, 90-117.

Recasens, D., An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish V-C-V sequences, 1987, *Journal of Phonetics, **15***, 299-312.

Recasens, D., Long range coarticulation effects for tongue dorsum contact in VCVCV sequences, 1989, *Speech Communication*, 8, 293-307.

Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa, S. Madureira, and C. Reis (Eds.), *Proceedings of the Speech Prosody 2008 Conference*. Campinas, Brazil, 175-184.

Sampson, R. (1999). Nasal vowels and vowel nasalization: preliminaries. In R. Sampson (Ed.), *Nasal vowel evolution in Romance* (pp. 1–31). New York: Oxford University Press.

Scarborough, R., Zellou, G., Mirzayan, A., & Rood, D. (2015). Phonetic and phonological patterns of nasality in Lakota vowels. *Journal of the International Phonetic Association*, *45*(3), 289–309.

Sheskin, D. (2003). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton: Chapman & Hall/CRC.

Solé, M. J. (1992). Phonetic and phonological processes: The case of nasalization. *Language and Speech*, *35*(1-2), 29–43.

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. In M.-J. Solé, P. S. Beddor, & M. Ohala (Eds.), *Experimental approaches to phonology*. Oxford: Oxford University Press 302-321.

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

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, *27*(2), 171–206.

Zellou, G., & Tamminga, M. (2014). Nasal coarticulation changes over time in Philadelphia English. *Journal of Phonetics*, *47*, 18–35.