SEPTEMBER 14 2020

Prosodic structurally conditioned variation of coarticulatory vowel nasalization in Mandarin Chinese: Its language specificity and cross-linguistic generalizability

Hongmei Li; Sahyang Kim; Taehong Cho

J. Acoust. Soc. Am. 148, EL240–EL246 (2020)
https://doi.org/10.1121/10.0001743

JASA
EXPRESS LETTERS
https://doi.org/10.1121/10.0001743

Prosodic structurally conditioned variation of coarticulatory vowel nasalization in Mandarin Chinese: Its language specificity and cross-linguistic generalizability

Hongmei Li,[1] Sahyang Kim,[2] and Taehong Cho[1,a)]
[1]*Department of English Language and Literature, Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University, Seoul, Korea*
[2]*Department of English Education, Hongik University, Seoul, Korea*
angela.hongmeili@gmail.com, sahyang@hongik.ac.kr, tcho@hanyang.ac.kr

Abstract: This study compares prosodic structural effects on nasal (N) duration and coarticulatory vowel (V) nasalization in NV (Nasal-Vowel) and CVN (Consonant-Vowel-Nasal) sequences in Mandarin Chinese with those found in English and Korean. Focus-induced prominence effects show cross-linguistically applicable coarticulatory resistance that enhances the vowel's phonological features. Boundary effects on the initial NV reduced N's nasality without having a robust effect on V-nasalization, whose direction is comparable to that in English and Korean. Boundary effects on the final CVN showed language specificity of V-nasalization, which could be partly attributable to the ongoing sound change of coda nasal lenition in Mandarin.

[Editor: Rachel M. Theodore] Pages: EL240–EL246

Received: 28 April 2020 **Accepted:** 26 July 2020 **Published Online:** 14 September 2020

1. Introduction

Coarticulation is a physiological, biomechanically driven phonetic process originating from the inevitable overlap of multiple articulatory gestures (Kühnert and Nolan, 1999). This low-level phonetic process, however, can be controlled by the speaker in reference to the higher-order linguistic structures of a given language [e.g., Barlaz *et al.* (2018), Beddor *et al.* (2002), Cho (2004), and Mok (2013)]. Recent studies of English and Korean (Cho *et al.*, 2017; Jang *et al.*, 2018) have shown that the phonetic granularity of coarticulatory vowel nasalization (V-nasalization) in CVN (consonant-vowel-nasal) and NVC (nasal-vowel-consonant) is indeed fine-tuned by prosodic structural factors of prominence and boundary, which can engender linguistically meaningful phonetic outputs, as discussed below [see Fletcher (2010) or Cho (2016) for a related review]. Our purpose in the present study is to extend these previous studies to Mandarin Chinese (henceforth Mandarin) and compare the extent to which coarticulatory V-nasalization in Mandarin can be under the speaker's control in reference to prosodic structural factors with the existing data for American English (henceforth English) and Korean. This comparison will provide insights into cross-linguistic similarities and differences in the coarticulatory process that is assumed to go beyond simple physiological, biomechanical processes. In the rest of this section, we summarize the findings in English and Korean reported in Cho *et al.* (2017) and Jang *et al.* (2018), respectively, and lay out the specific research questions we explored in this study.

English and Korean share comparable focus-induced prominence effects, such that N-duration is longer and the vowel resists the coarticulatory influence of N under focus, indicating an enhancement of N's nasality and the vowel's [oral] feature, respectively. Based on these cross-linguistic similarities, it was suggested that the prominence entails coarticulatory resistance in a cross-linguistically similar way, enhancing phonological contrasts in the language. Our first question is thus how the focus-induced prominence factor modulates the coarticulatory V-nasalization in Mandarin in comparison with the findings in English and Korean and the extent to which the results lend cross-linguistic support to the phonologically informed coarticulatory resistance hypothesis.

English and Korean also show similar boundary effects in the carryover (NV) context: a *domain-initial* shortening of the N-duration that leads to *less* V-nasalization in the following vowel. Such an initial nasal reduction found across languages is generally considered to come from domain-initial articulatory strengthening [e.g., Keating *et al.* (2003)] with a supralaryngeal

[a)]Author to whom correspondence should be addressed, ORCID: 0000-0002-8148-745X.

https://doi.org/10.1121/10.0001743

articulatory force that elevates the velum [e.g., Fougeron (2001)]. Our second question is thus how the boundary strength influences the phonetic realization of NV in Mandarin and the extent to which the obtained boundary effect reflects the domain-initial articulatory strengthening found in English and Korean.

Finally, in boundary effects in the anticipatory (VN) context, English and Korean show both similarities and differences. They share a comparable phrase-final coarticulatory vulnerability by which the vowel is *more* nasalized in the phrase-final than in the phrase-medial position. The phrase-final coarticulatory vulnerability has been attributed to a cross-linguistically applicable articulatory weakening process toward the end of a phrase that loosens the velic elevation gesture. On the other hand, English shows a cross-linguistically observed phrase-final lengthening of the N-duration in VN, whereas Korean shows no phrase-final lengthening effect, showing language-specificity in the temporal realization of the N-duration. Our third question is thus how Mandarin modulates V-nasalization and N-duration in the phrase-final CVN context from both cross-linguistic and language-specific perspectives.

Given the available cross-linguistic evidence, one might expect Mandarin to follow the cross-linguistic trend toward the phrase-final coarticulatory vulnerability of the vowel and lengthening of the N-duration in the VN context. But just as Korean showed a language-specific pattern in the boundary effect on N-nasalization in VN (no phrase-final lengthening), so Mandarin could reveal its own language-specific phrase-final patterns. This possibility becomes more plausible when we consider the language's distributional restriction on coda consonants—i.e., only nasal consonants are allowed, being restricted to /n/ and /ŋ/ [e.g., Duanmu (2007)]. Moreover, many Chinese dialects, including Mandarin, undergo a sound change in which the oral constriction of the nasal coda is substantially lenited [e.g., Duanmu (2007)], with a possible place merger of /n/ and /ŋ/ (Chen, 1972; Chen, 2000; Chiu *et al.*, 2019), while the information about nasals can still be preserved in the form of V-nasalization [cf. Chen (1972)]. In other words, V-nasalization in Mandarin might not be characterized merely as a non-contrastive coarticulatory process; it appears to play a phonological role in maintaining phonological contrasts among the nasal consonants. It is thus reasonable to predict that the phrase-final V-nasalization might not simply follow the cross-linguistically observable coarticulatory vulnerability, possibly showing no substantial variation in V-nasalization, so that it can maintain sufficient phonological information from the nasal consonant.

2. Method

2.1 Participants, speech materials, and recording procedure

Sixteen native Mandarin speakers participated in this study. However, four of them produced CVN words with no acoustically discernible, clear-cut division between the vowel and the nasal coda, indicating a weakening of the oral constriction for N that made it difficult to reliably segment the vowel and the coda. These speakers were therefore excluded from further analysis. The 12 remaining speakers reported no language impairments and had lived in Korea for less than 3 years (6 females and 6 males, $M_{age} = 23.4$ years, 20–29 years).

Six monosyllabic target words were recorded in the nasal context: two (/ma/, /na/) in NV and four (/pan/, /tan/, /paŋ, /taŋ/) in CVN. (Note that Mandarin has no NVC words.) The lexical tone was controlled to be tone 1. Two additional CV words (/pa/, /ta/) were separately recorded to be used as the control (oral) context for NV and CVN. Each target word was embedded in a carrier (template) sentence in which the focus and boundary conditions were manipulated. The carrier sentences consisted of a question-answer pair in a mini discourse situation. The NV target word was meant to be placed in either an intonational phrase-initial (IP-initial) or an IP-medial position (with N being word-initial) and to be either focused (in uppercase and bold) with corrective contrast between the target word and a word in the question or unfocused with the contrast elsewhere. The CVN words were included in carrier sentences in a similar way except that the boundary condition for CVN was IP-final (rather than IP-initial) versus IP-medial. See the Supplementary Material[1] for a full set of experimental sentences with NV and CVN words. The full set is also available on the Open Science Framework (Li, 2020).

The discourse situation was intended to be a kind of word game. To obtain natural speech as much as possible, the full carrier (template) sentences were not shown to the speakers in written form. Instead, they were cued by pictures, as shown in Fig. 1. A target word was written on one side of a cube alongside the copular verb ("be") that was to be produced together with the target word. The prompt questions were pre-recorded by a native Mandarin speaker and played back for the participants. The speakers then responded by producing the target words (along with the copular verb) in the carrier (template) sentences according to the cues given in the picture. For example, as shown in Fig. 1(a), two cubes appeared on the screen, and the left cube had the target word *ma* ("mom") written in Chinese alongside the copular verb *shi* (*ma-shi*). The pre-recorded voice asked whether the word on the left cube was *pa-shi* ("eight-be") as in ***pa1***

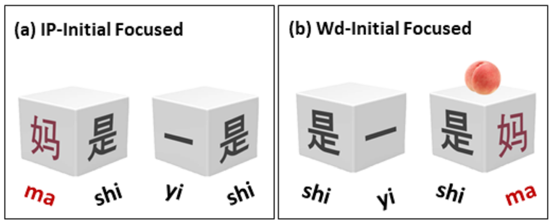

Fig. 1. (Color online) An example of the visual prompts used in the designed word game. In (a), the left cube contains the target word *ma* ("mom") alongside the copular verb *shi*, used to induce an IP boundary condition under focus. In (b), the right cube contains the target word *ma* alongside *shi* (with a fruit on top of it), used to induce a word boundary under focus. (See the text for further explanation.)

ʂi4 tsai4 ts^{w}o3 p^{j}ɛn1 ma? ("Is '**EIGHT**-be' on the left?"). The speaker would say that the left one was *ma*-*shi* (not *pa*-*shi*), inducing a corrective contrastive focus on the target word as in *pu2 ʂi4,* ***ma1*** *ʂi4 tsai4 ts^{w}o3 p^{j}ɛn1* ("No. '**MOM**-be' is on the left."). In this condition, the target word was preceded by another (short) phrase (e.g., [pu2 ʂi4] "No") which created an IP boundary between the target and the preceding phrase. Figure 1(b) shows a visual prompt for a focused condition with a Wd boundary. The target word was cued by a fruit placed on top of it, and as exemplified in *pu2 ʂi4. t^{h}au2 tsai4 ʂi4* ***ma1*** *ʂɑŋ4 p^{j}ɛn1* ("No. The peach is on the top of 'be-**MOM**'."). It was meant to be embedded in the middle of a phrase. There was a practice session of about 30 min, including the instructions for the word game, so that the speakers were comfortable producing response sentences based on the visual prompts.

Acoustic data were recorded in a sound-attenuated booth using a Tascam HP-D2 digital recorder and a SHURE KSM44 microphone at a sampling rate of 44 kHz. Each dialogue was repeated three times. In total, 1440 sentence tokens were collected—10 items (6 target words and 4 control words) × 2 boundaries (IP, Wd) × 2 focus conditions (focused, unfocused) × 3 repetitions × 12 speakers. The obtained tokens were checked for their prosodic conditions by all three authors, and 19 tokens were excluded because they were produced with either a clearly wrong placement of focus or a prosodic boundary before or after a test word in the context of a Wd boundary condition.

2.2 Measurements and statistical analysis

The acoustic duration of the nasal consonant (N-duration) was measured on the spectrogram along with the waveform. It was taken from the onset of the nasal murmur to the onset of the vowel in NV and from the offset of the vowel to the offset of the nasal murmur in CVN. V-nasalization was assessed by A1-P0, following previous studies [e.g., Chen (1997), Zellou (2017), Cho *et al.* (2017), and Jang *et al.* (2018)]. In A1-P0, A1 refers to the amplitude of F1 and P0 to the amplitude of the nasal peak at 250–450 Hz, with lower A1-P0 values indicating greater nasalization. A1-P0 values were taken at both absolute and relative timepoints in the vowel, as in Cho *et al.* (2017) and Jang *et al.* (2018). The absolute measures were taken at 25, 50, and 75 ms from the V onset in NV and before the V offset in CVN to examine the extent to which V-nasalization is related to the physical distance from N as a time-locked low-level process. The relative measures were taken at 25%, 50%, and 75% relative to the entire vowel duration to examine the extent to which the coarticulatory influence pervaded into the vowel as a process beyond a low-level phonetic effect. The A1-P0 values were extracted by using a PRAAT script (Styler, 2015). Following Zellou (2017) and Jang *et al.* (2018), the A1-P0 values were normalized for each participant based on the maximum and minimum nasality values (A1-P0) obtained from each (near) minimal pair of oral-nasal contexts (NV-CV; CVN-CV). The normalized values in proportion (%) indicate how much each A1-P0 measure of the target word is considered to be nasalized relative to the obtained range.

A series of linear mixed-effects models were fitted to each measure by using the *lme4* (Bates *et al.*, 2015) packages in R (R Core Team, 2019) to investigate the prosodic influences of boundary and focus on the two dependent variables: N-duration and V-nasality (A1-P0 proportion). The fixed effects were focus (focused, unfocused), boundary (IP, Wd), and timepoint (Absolute: 25, 50, 75 ms; Relative: 25%, 50%, 75%) and their interactions. Focus and boundary were contrast-coded, and timepoint was treatment-coded. The random-effect structure for the A1-P0 model in the final (CVN) context includes random intercepts and random slopes for participant and item. Four items (/pan/, /tan/, /paŋ/, /taŋ/) were included in CVN. For NV, there were only two items (/ma, /na/) and the A1-P0 model with relative timepoints converged with both random intercepts and slopes for item. But the A1-P0 model with absolute timepoints did not converge with random slopes for item, so that only random intercepts for item were included. Random slopes for timepoint were not included in the A1-P0 models in both CVN and NV

https://doi.org/10.1121/10.0001743

because the models failed to converge with them. As timepoint was included in the model fits mainly to examine how effects of prosodic factors (focus and boundary) would interact with timepoints, we excluded random slopes for timepoint for the sake of model convergence rather than those for the two critical prosodic factors. (Note also that although timepoint had three levels, direct comparisons among them were not made for the purpose of the present study, so that there was no need for family-wise error correction.) Finally, models for N-duration in both NV and CVN converged with random intercepts and slopes for both participant and item. The R syntaxes used for the analyses are given in the Supplementary Material[1] and on the Open Science Framework (Li, 2020).

The data (including all the measured values for dependent variables in each condition) used for model fits and R scripts are available on the Open Science Framework (Li, 2020).

3. Results

3.1 Initial NV (carryover) context

Focus ($\beta = 17.73$, $t = 6.95$, $p < 0.001$) and boundary ($\beta = -31.53$, $t = -5.48$, $p = 0.027$) had significant effects on N-duration in NV, with N being *longer* in the focused than in the unfocused condition [Fig. 2(a)] and *shorter* in the IP-initial than in the Wd-initial position [Fig. 2(b)]. There was also a significant interaction effect between focus and boundary ($\beta = -22.31$, $t = -5.02$, $p = 0.031$), with the focus-induced lengthening effect being smaller in the IP than in the Wd condition [Fig. 2(c)]. For V-nasalization, focus had a significant effect on A1-P0 (normalized) at both the absolute and relative timepoints (absolute: $\beta = -16.05$, $t = -3.39$, $p = 0.004$; relative: $\beta = -15.58$, $t = -2.38$, $p = 0.035$)—i.e., vowels were nasalized *less* in the focused than the unfocused condition, showing a focus-induced coarticulatory reduction [Fig. 2(d) and 2(f)]. The effect of boundary was not significant in the absolute measurement ($\beta = -5.68$, $t = -1.47$, $p = 0.160$) [Fig. 2(e)], but there was a trend in the relative measurement ($\beta = -8.22$, $t = -1.47$, $p = 0.077$) [Fig. 2(g)]. Focus and boundary did not interact with each other, nor did they interact with timepoint, indicating that the observed patterns were consistent across the timepoints and were thus pervasive in the vowel.

3.2 Final CVN (anticipatory) context

Focus and boundary had significant effects on N-duration in CVN, with N being longer in the focused than in the unfocused condition ($\beta = 9.36$, $t = 5.32$, $p < 0.001$) [Fig. 3(a)] and longer in the IP-final than in the Wd-final condition ($\beta = 28.09$, $t = 5.46$, $p = 0.005$), consistent with a general phrase-final lengthening effect [Fig. 3(b)]. An interaction between focus and boundary was observed ($\beta = -18.91$, $t = -3.5$, $p = 0.013$) [Fig. 3(c)]: the focus-induced lengthening effect was reliable only in the Wd-final condition, disappearing in the IP-final condition in which the N-duration was already substantially lengthened, showing a kind of ceiling effect caused by phrase-final lengthening.

Focus had a significant effect on V-nasalization (normalized A1-P0) in both the absolute and relative measures (absolute: $\beta = -24.49$, $t = -6.02$, $p < 0.001$; relative: $\beta = -24.75$, $t = -5.97$, $p < 0.001$). V-nasalization was less in the focused than in the unfocused condition, showing a reduction of coarticulatory vowel nasalization under focus [Figs. 3(d) and 3(f)]. Although focus and timepoint showed some statistical interactions in the absolute measures (at 50 ms, $\beta = 3.72$, $t = 2.88$, $p = 0.004$; at 75 ms, $\beta = 5.82$, $t = 4.45$, $p < 0.001$), no interaction was found in the relative measures. In fact, as can be seen in Figs. 3(d) and 3(f), the focus effect was largely consistent across timepoints in the vowel in both the absolute and relative measures, although the effect appears to be slightly more robust near the nasal consonant, which could account for the focus × timepoint interactions. Unlike focus, boundary showed no effect on V-nasalization in the absolute measure ($\beta = -1.08$, $t = -0.62$, $p = 0.547$) [Fig. 3(e)]. The relative measure revealed a

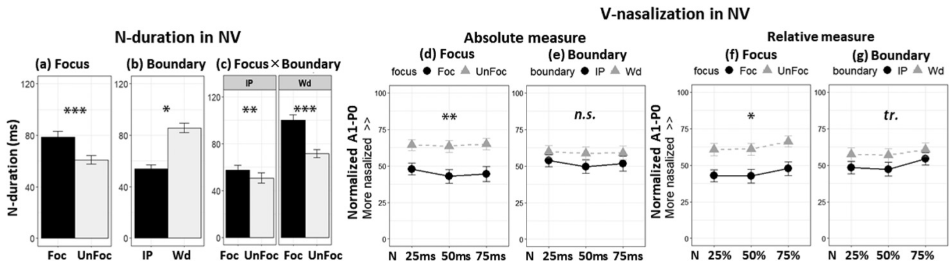

Fig. 2. Effects of focus and boundary on initial N-duration [(a), (b), (c)] and V-nasalization at the absolute (d), (e) and relative (f), (g) timepoints in the carryover (NV) context. Not significant (*n.s.*), p > 0.1; *, p < 0.05; **, p < 0.01; ***, p < 0.001.

https://doi.org/10.1121/10.0001743

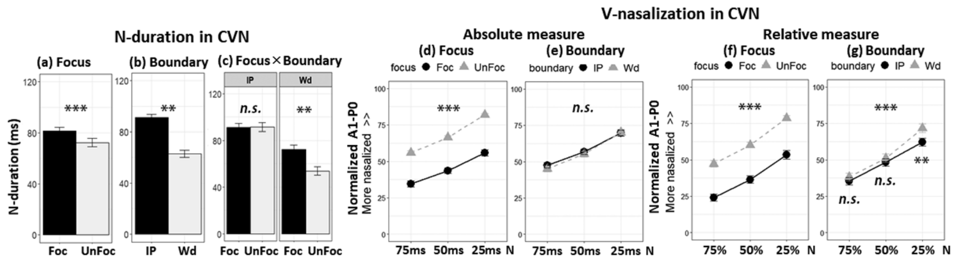

Fig. 3. Effects of focus and boundary on final N-duration [(a), (b), (c)] and V-nasalization at the absolute (d), (e) and relative (f), (g) timepoints in the anticipatory (CVN) context. *n.s.*, $p > 0.1$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

significant effect of boundary ($\beta = -9.48$, $t = -5.29$, $p < 0.001$), but, as shown in Fig. 3(g), the effect interacted with timepoint in such a way that only the timepoint at 25% (nearest the nasal source) revealed a significant boundary effect ($\beta = -9.83$, $t = -4.26$, $p = 0.002$).

4. General discussion

Mandarin showed focus-induced prominence effects on N-duration and V-nasalization that were largely consistent with those observed in English (Cho *et al.*, 2017) and Korean (Jang *et al.*, 2018). N-duration was elongated under focus in both NV and CVN, showing an enhancement of N's nasality. The vowel in both the carryover (NV) and anticipatory (CVN) contexts was nasalized *less* under focus across the timepoints in the vowel. As in English and Korean, such across-the-board coarticulatory reduction in Mandarin indicates that the effect goes beyond what could be normally observed with a localized low-level phonetic process to show an enhancement of the vowel's phonological [oral] features [cf. de Jong (2004) and Cho (2016); see Kim *et al.* (2018) for related discussion on enhancement of stop voicing features in English]. This confirms that prominence modulates the low-level coarticulatory process in relation to the sound system in much the same way across languages, lending support to the prominence-induced coarticulatory resistance hypothesis.

The boundary effect in the carryover (NV) context in Mandarin also demonstrated cross-linguistically comparable patterns. N-duration was shortened IP-initially more than Wd-initially, as in English and Korean [see also Cho and Keating (2001, 2009) for related results]. The domain-initial reduction of N (the nasal murmur) builds on the cross-linguistic tendency for nasal reduction to decrease the onset consonant's sonority. This can be interpreted as showing an enhancement of its consonantality in NV (the less sonorant, the more consonant-like) and hence an enhancement of the syntagmatic CV contrast [cf. Fougeron and Keating (1997) and Keating *et al.* (2003)], albeit the effect might also be attributed to an articulatory force that applies to the vocal tract, elevating the velum (Fougeron, 2001). Interestingly, however, despite the robust phrase-initial reduction of N's nasality, there was only a weak trend towards a phrase-initial reduction of V-nasalization. This weak effect differs substantially from the robust phrase-initial reduction of V-nasalization found in English and Korean. At the moment, we do not have a definite explanation for this potential cross-linguistic difference, but the trend found in Mandarin is at least in the same direction as in English and Korean, suggesting that all three languages have similar phonetic underpinnings that underlie the boundary effect in the carryover (NV) context, although the fine phonetic detail appears to be determined in a language-specific way.

Finally, and most importantly, the boundary effect in the anticipatory (CVN) context demonstrated both cross-linguistic similarity and difference. N-duration in CVN was elongated IP-finally, showing a generally observed phrase-final lengthening effect, as was found in English. (Note that Korean showed no phrase-final lengthening effect on the final N-duration.) Crucially, however, no substantial boundary effect was observed on V-nasalization in either the absolute or relative measure. There was only one timepoint (25%) nearest the nasal source in CVN in which the vowel was more nasalized IP-finally; no other timepoint (absolute or relative) showed any boundary effect. These results differ categorically from the robust phrase-final effects across timepoints found in both English and Korean. The almost null boundary effect on V-nasalization in CVN in Mandarin, therefore, stands in sharp contrast with the general pattern of phrase-final coarticulatory vulnerability that can arise with phrase-final articulatory weakening, which attenuates the oral articulatory force to close off the velopharyngeal port (Cho *et al.*, 2017; Jang *et al.*, 2018).

One way of understanding this cross-linguistic discrepancy is to simply attribute it to a language-specific phonetic effect internalized in the phonetic grammar of Mandarin [e.g., Keating (1984), Cho and Ladefoged (1999), and Cho *et al.* (2019)]. In other words, Mandarin might simply not follow the presumably biomechanically driven universal phonetic process understood to underlie the phrase-final coarticulatory vulnerability. The cross-linguistic variation in V-nasalization (as observed in Mandarin vs English and Korean) can then be understood in terms

https://doi.org/10.1121/10.0001743

of how languages can differ in the gestural magnitude (and the temporal extent) of the velum lowering gesture for a nasal consonant and its alignment timing with the oral constriction gesture. Following the insight that variation in V-nasalization is related to the alignment of the velum lowering gesture with the oral constriction gesture of Beddor (2009), Jang *et al.* (2018) offered a gestural account for a cross-linguistic difference between English and Korean. Recall that Korean did not show a phrase-final lengthening of N-duration, whereas both languages showed a similar phrase-final coarticulatory vulnerability of V-nasalization. The null effect on N-duration in Korean could be accounted for by a smaller velum lowering gesture in Korean than in English. The assumed smaller velum lowering gesture in Korean appears to be aligned *earlier* (shifted to the left) relative to the oral constriction gesture in VN in the phrase-final position. As a result, N-duration in VN in Korean might not undergo a phrase-final lengthening (because the velum lowering gesture has been shifted to the left, thus overlapping less with the oral constriction gesture for N), and the preceding vowel overlaps more with the velum lowering gesture, showing more V-nasalization in the phrase-final position. We suggest that a similar mechanism for modulating inter-gestural timing could apply to Mandarin but in a language-specific way, showing an exact opposite pattern. The phrase-final velum lowering gesture in Mandarin could be aligned *later* (rather than earlier as in Korean) so that it is shifted to the right relative to the oral constriction gesture for N. The resulting increase in overlap between the velum lowering and oral constriction gestures could account for phrase-final lengthening of the N-duration to some extent. Such a later alignment would in turn have caused *less* V-nasalization, if the magnitude of the velum lowering gesture remained largely the same as a function of boundary strength, but the degree of V-nasalization may also remain more or less the same, if the velum lowering gesture indeed becomes larger at an IP boundary, thus overlapping more with the consonantal constriction. This inter-gestural timing account for the cross-linguistic variation in V-nasalization is reminiscent of the inter-gestural timing account proposed by Cho and Ladefoged (1999) for cross-linguistic variation in VOT—i.e., it is assumed to arise as a consequence of the language's arbitrary use of differential inter-gestural timing between the oral release gesture and the glottal adduction gesture.

This language-specific effect on CVN could also be related to a language-specific distributional restriction on the occurrence of consonants in the coda position and a possible sound change currently in progress in Mandarin. Recall that in many Chinese dialects, the oral constriction for the coda nasal can be substantially lenited (Duanmu, 2007), possibly leading to a sound change in which the oral constriction is being elided with a merger between /n/ and /ŋ/ (Chen, 1972; Chen, 2000; Chiu *et al.*, 2019). Chen (1972) also suggested that although the oral place features for the two nasal consonants could be delinked in the coda position in Mandarin, they could still be preserved with their phonetic trace left in the form of V-nasalization. This is presumably why four of the original 16 speakers who participated in this study did not show a clear-cut division between the nasalized vowel and the nasal consonant. From the articulatory point of view, the velum lowering gesture could be more loosely aligned (with less gestural cohesion) with the oral constriction gesture for C$\underline{\text{VN}}$ in Mandarin than in English, which could eventually have led to a weakening of oral constriction while the V-nasalization carries sufficient information about the presence of the upcoming nasal coda. The results of the present study thus lend support to the view that V-nasalization in Mandarin is no longer a mere coarticulatory process but plays a phonological role in preserving information about the coda. Thus, it is less subject to the putatively universal biomechanically driven weakening effect in the phrase-final position. It is, however, worth pointing out that the sound change in Mandarin appears to be still at its embryonic stage. Twelve out of 16 speakers still showed a relatively clear acoustic division between the vowel and the nasal murmur for the consonant. Moreover, if the V-nasalization was fully phonologized, one might expect a focus-induced enhancement of the phonological nasal feature, showing more nasalization under focus. But the nasalized vowel in Mandarin was *less* nasalized under focus, indicating that the nasal feature is not yet phonologically linked with the preceding vowel.

In conclusion, we have here demonstrated that although Mandarin Chinese follows cross-linguistically applicable coarticulatory patterns, some of the effects are attributable to the language's specific linguistic structure. The observed cross-linguistic generalizability and language-specificity support the general view that low-level phonetic processes can be under speaker control in reference to prosodic structure across languages, but they can also be further fine-tuned in language-specific ways that depend on the language's internal phonological and prosodic structures. The results of this study, however, are based on a limited tonal context. Further research is needed to generalize the current finding by examining broader tonal contexts and other Chinese dialects.

Acknowledgments

We would like to thank the Mandarin speakers who participated in our experiment. Preliminary results based on the subset of the data were presented at Hanyang International Symposium on

Phonetics and Cognitive Sciences (HISPhonCog) 2019, and appeared in the proceedings of International Congress on Phonetic Sciences (ICPhS) 2019. This work was supported by research funds from Hanyang University (HY-2019) awarded to T. Cho.

References and links

[1]See supplementary material at https://doi.org/10.1121/10.0001743 for a full set of experimental sentences with NV and CVN words and the R syntaxes used for the statistical analyses (linear mixed effects modeling).

Barlaz, M., Shosted, R., Fu, M., and Sutton, B. (**2018**). "Oropharygneal articulation of phonemic and phonetic nasalization in Brazilian Portuguese," J. Phon. **71**, 81–97.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Soft. **67**(1), 1–48.

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

Beddor, P. S., Harnsberger, J. D., and Lindemann, S. (**2002**). "Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates," J. Phon. **30**, 591–627.

Chen, M. (**1972**). "Nasals and nasalization in Chinese: Explorations in phonological universals," Ph.D. dissertation, University of California, Berkeley, CA.

Chen, M. Y. (**1997**). "Acoustic correlates of English and French nasalized vowels," J. Acoust. Soc. Am. **102**(4), 2360–2370.

Chen, M. Y. (**2000**). "Acoustic analysis of simple vowels preceding a nasal in Standard Chinese," J. Phon. **28**, 43–67.

Cho, T. (**2004**). "Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English," J. Phon. **32**(2), 141–176.

Cho, T. (**2016**). "Prosodic boundary strengthening in the phonetics—Prosody interface," Lang. Ling. Compass **10**(3), 120–141.

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

Cho, T., and Keating, P. (**2009**). "Effects of initial position versus prominence in English," J. Phon. **37**, 466–485.

Cho, T., Kim, D., and Kim, S. (**2017**). "Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English," J. Phon. **64**, 71–89.

Cho, T., and Ladefoged, P. (**1999**). "Variation and universals in VOT: Evidence from 18 languages," J. Phon. **27**, 207–229.

Cho, T., Whalen, D. H., and Docherty, G. (**2019**). "Voice onset time and beyond: Exploring laryngeal contrast in 19 languages," J. Phon. **72**, 52–65.

Chiu, C., Lu, Y., Weng, Y., Jin, S., Weng, W., and Yang, T. (**2019**). "Uncovering syllable-final nasal merging in Taiwan Mandarin: An ultrasonographic investigation of tongue postures and degrees of nasalization," in *Proceedings of the 19th International Congress of Phonetic Sciences*.

De Jong, K. (**2004**). "Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration," J. Phon. **32**(4), 493–516.

Duanmu, S. (**2007**). *The Phonology of Standard Chinese* (Oxford University Press, New York).

Fletcher, J. (**2010**). "The prosody of speech: Timing and rhythm," in *The Handbook of Phonetic Sciences*, 2nd ed., edited by W. J. Hardcastle, J. Laver, and F. E. Gibbon (Blackwell, Oxford, UK), pp. 523–602.

Fougeron, C. (**2001**). "Articulatory properties of initial segments in several prosodic constituents in French," J. Phon. **29**, 109–135.

Fougeron, C., and Keating, P. A. (**1997**). "Articulatory strengthening at edges of prosodic domains," J. Acoust. Soc. Am. **101**(6), 3728–3740.

Jang, J., Kim, S., and Cho, T. (**2018**). "Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences," J. Acoust. Soc. Am. **144**(1), EL33–EL39.

Keating, P., Cho, T., Fougeron, C., and Hsu, C. (**2003**). "Domain-initial strengthening in four languages," in *Papers in Laboratory Phonology VI: Phonetic Interpretations*, pp. 145–163.

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

Kim, S., Kim, J., and Cho, T. (**2018**). "Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English," J. Phon. **71**, 65–80.

Kühnert, B., and Nolan, F. (**1999**). "The origin of coarticulation," in *Coarticulation: Theory, Data and Techniques*, edited by W. J. Hardcastle and N. Newlett (Cambridge University Press, Cambridge), pp. 7–30.

Li, H. (**2020**). "Prosodically-conditioned vowel nasalization in Mandarin Chinese," Open Science Framework, https://osf.io/sjteq/ (Last viewed 22 August 2020).

Mok, P. (**2013**). "Does vowel inventory density affect vowel-to-vowel coarticulation?," Lang. Speech **56**, 191–209.

R Core Team. (2019). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/ (Last viewed 22 August 2020).

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

Zellou, G. (**2017**). "Individual differences in the production of nasal coarticulation and perceptual compensation," J. Phon. **61**, 13–29.