

LANGUAGE-SPECIFIC PROSODIC STRUCTURAL MODULATION OF COARTICULATION VOWEL NASALIZATION IN #NV AND CVN# IN MANDARIN CHINESE

Hongmei Li¹, Sahyang Kim², Taehong Cho¹

Hanyang Institute for Phonetics and Cognitive Sciences of Language (HIPCS), Hanyang University¹;
Hongik University²

lhm82cn@hanyang.ac.kr; sahyang@gmail.com; tcho@hanyang.ac.kr

ABSTRACT

This study investigates prominence and boundary effects on acoustic realization on N-duration and V-nasalization in Mandarin Chinese. In both #NV and VN#, prominence was found to elongate N-related duration (enhancing N's nasality), while V tended to resist coarticulatory influence from N (enhancing V's orality). In phrase-initial position (#NV), boundary strength induced a shortening of N in line with CV contrast enhancement (C being less sonorant hence more consonant-like), while V was nasalized categorically regardless of boundary strength, different from a general coarticulatory pattern. In phrase-final position (CVN#), while there was no clear-cut acoustic distinction between V and N, the nasalized portion showed preboundary lengthening. Crucially, the degree of V-nasalization was not larger phrase-finally, deviating from a general increase in V-nasalization found in other languages in which the nasal-oral contrast in the coda is enhanced phrase-finally. The results are discussed in terms of distributional restrictions in Mandarin Chinese which allows only nasals in the coda which are lenited with no oral counterparts.

Keywords: prosodic structure, vowel nasalization, prominence, boundary, Mandarin Chinese

1. INTRODUCTION

Phonetic variation is one of the properties of continuous speech. One source of the variation is coarticulation caused by an inevitable overlap of multiple articulatory gestures [12]. The process of coarticulation is generally fine-tuned in relation to linguistic structures such as the phonological contrast and the syllable structure of a language [1,2,13]. It has been suggested that the nature of the phonetic enhancement of linguistic contrasts may be echoed in a coarticulatory process in V-nasalization that is modulated by prosodic strengthening that stems from prosodic structure [8,10]. Prosodic structure is assumed to regulate placement of prosodic boundaries and distribution of prominence among phonological units [6,14]. Under the influence of

prominence and boundary strength, segments often undergo prosodic strengthening which effectively increases phonetic clarity and may mediate linguistic contrasts [6].

Previous studies on coarticulatory V-nasalization [8,10] have indeed demonstrated ample evidence that the prosodic-structural fine-tuning of the coarticulatory process leads to linguistically meaningful phonetic outcome. For example, in both American English and Korean, focus-related prominence has been found to elongate N's duration but to induce V's coarticulatory resistance to nasalization in CVN and NVC. These effects are taken to enhance N's nasality and V's orality, respectively, pertaining to an enhancement of each phoneme's paradigmatic contrast. Both languages also showed boundary-related *domain-initial* strengthening effects on #NVC, in such a way that N becomes less nasal (enhancing its consonantality) and V becomes more oral (enhancing its orality), which has been interpreted as a syntagmatic enhancement of CV contrast. Finally, in *domain-final* position (CVN#), both languages showed coarticulatory vulnerability (i.e., more V-nasalization), in line with some kind of articulatory weakening towards the end of a phrase which may result in loosening the velic elevation gesture.

In the present study, we build on this cross-linguistic evidence for the phonetics-prosody interface reflected in the coarticulatory process of V-nasalization by examining the coarticulatory V-nasalization in Mandarin Chinese (henceforth Mandarin). In line with the above-mentioned previous studies, we investigate the extent to which V-nasalization in Mandarin is realized according to the prosodic-structural factors (boundary strength and focus-induced prominence), and how the coarticulatory effect in Mandarin speaks to the language's linguistic contrast systems. The results will then be compared with general coarticulatory patterns observed in other languages, which will enrich our understanding of the coarticulatory process of V-nasalization from both the cross-linguistic and the language-specific perspectives.

While one of the most conspicuous differences between Mandarin and English/Korean lies in the fact

that Mandarin employs a lexical tone system, a particular property of Mandarin that we focus on in the present study is its distributional restrictions on the occurrence of consonants in the coda position. Mandarin allows only /n/ and /ŋ/ in the coda position [3] with no oral counterparts, and the nasals’ oral constriction is often incomplete or substantially lenited [5,9]. In fact, many Chinese dialects have been reported to be undergoing a nasal elision process in the coda [4], but in Mandarin the two nasal consonants whose oral place features may be delinked in the coda are considered to be still preserved, leaving their phonetic trace in the form of V-nasalization [4]. A specific question that arises is then how Mandarin V-nasalization manifests itself in such a nasal lenition (VN) context and how it is conditioned by prosodic strengthening factors. Given that V-nasalization in VN carries some phonological function of preserving the phonemic information of the following nasal, prosodic-structurally conditioned variation in V-nasalization in Mandarin may show different effects as compared to those found in other languages which do not employ similar restrictions on the occurrence of consonants in the coda.

2. METHOD

2.1. Participants and speech materials

Sixteen native speakers (8F,8M) of Mandarin who had lived for less than 3 years in Korea (except for 1 speaker who resided in Korea for 5 years) participated in the experiment for pay. They were students at Hanyang University, who were in their 20s and 30s.

Speech materials included six monosyllabic target words in the nasal context with two (/ma/, /na/) in #NV and four (/pan/, /tan/, /paŋ/, /taŋ/) in CVN#. Note that Mandarin has no NVC words. For the purpose of testing V-nasalization, the lexical tone was controlled to be Tone 1. Two additional CV (/pa/, /ta/) words were also collected to be used for the control (oral) context to estimate V-nasalization relative to the oral context.

Each target word was embedded in a carrier sentence which was an answer to a question in a mini discourse situation, where the Boundary (IP vs. Wd) and Focus (contrastive focused vs. unfocused) conditions varied. The discourse situation was designed as a kind of a word game (as shown in Fig. 1.) to induce focus and prosodic boundary conditions. In order to facilitate as much spontaneous speech as possible, the full carrier sentences were not written, but were induced by pictures as shown in Fig. 1 (see below for details).

2.2. Procedure

The speakers were presented with a picture where a target word was displayed on a computer screen, and

the prompt question (Speaker A, pre-recorded by a native Mandarin speaker) was played back from a loudspeaker. After having seen the picture and heard the question, the speakers answered the question. The answers served as the experimental sentences that contained target words. For example, as shown in Fig. 1, two cubes appeared on the screen and a monosyllabic test word /pan/ was displayed on the cube on the left. The pre-recorded voice asked the subject whether the word on the left cube was /pa/ (“eight”). After having seen the displayed word /pan/ on the cube, the speaker was supposed to say that the left one was /pan/ (“class”) (not /pa/), which induced a corrective contrastive focus on the target word. In this way, the target words were either focused or unfocused. For the unfocused condition, a contrastive focus was placed on an adjacent word in a carrier sentence, leaving the target unfocused. For the boundary conditions, as exemplified in Table 1, the target word was immediately followed (or preceded) by another short phrase, which created an IP boundary. For the Wd-boundary condition, the target word was embedded in the middle of a phrase.

Figure 1: An example of the visual cues used in a designed word game. Displayed on the left cube in each figure is a target word (*pan*, ‘class’) for the IP condition (a), and the *be*-verb + target word (*shi pan*, ‘is class’) for the Wd condition (b).

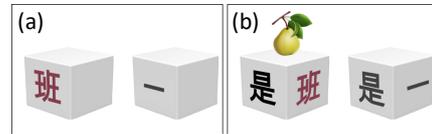


Table 1: Examples of CVN in carrier sentences, in which the target words are underlined and in *italic*, while the focused words are in **bold**. The number after each word indicates lexical tone (e.g. 1 represents Tone 1). Examples of NV are not listed due to the space limit.

Boundary	Focus	Example sentences
IP-final	Focused	A: [tsuo3 pjen1 ši4 pa1 ma] ? Is the left one EIGHT ?
		B: [pu2 ši4]. IP [tsuo3 pjen1 ši4 <u>pan1</u>]. IP [ši4 pu]? No. The left one is <i>CLASS</i> , isn't it?
	Unfocused	A: [tsuo3 pjen1 ši4 pan1 ma] ? Is the LEFT one CLASS?
		B: [pu2 ši4]. IP [joo4 pjen1 ši4 <u>pan1</u>]. IP [ši4 pu]? No. The RIGHT one is <i>CLASS</i> , isn't it?
Wd-final	Focused	A: [li2 tsar4 ši4 pa1 ŋan4 mjen4 ma]? Is the pear on the top of IS EIGHT ?
		B: [pu2 ši4]. IP [li2 tsar4 ši4 <u>pan1</u> ŋan4 mjen4]. No. The pear is on the top of <i>IS CLASS</i> .
	Unfocused	A: [tʰuo2 tsar4 ši4 pan1 ŋan4 mjen4 ma] ? Is the PEACH on the top of <i>IS CLASS</i> ?
		B: [pu2 ši4]. IP [li2 tsar4 ši4 <u>pan1</u> ŋan4 mjen4]. No. The PEAR is on the top of <i>IS CLASS</i> .

There was a 30-minute practice session including the instructions of the game. After the practice, participants reported that they could produce the carrier sentences spontaneously as intended while playing the role of Speaker B by answering the questions as guided by the pictures.

Acoustic data were collected in a sound-attenuated booth, using a Tascam HC-P2 digital recorder and a SHURE KSN44 condenser microphone at a sampling rate of 44kHz. Each dialogue was repeated three times. In total, 1920 sentence tokens were collected (6 items x 1 Tone type x 2 boundaries x 2 Focus types + 16 control words) x 3 repetitions x 16 speakers). 37 tokens were discarded, which were produced with either a wrong placement of focus or a major prosodic boundary before or after a test word in an intended phrase-internal Wd boundary context.

2.3. Measurements

For NV, N-duration (nasal murmur) and A1-P0 (where A1 refers to the amplitude of F1 and P0 to the amplitude of nasal peak which is formed usually below 400 Hz). A1-P0 values were taken at relative time points during the vowel. The relative time points were 25%, 50%, and 75% points relative to the entire vowel duration. The relative measures were expected to show to what extent V-nasalization occurs in reference to the temporal structure of the vowel. (Note that the absolute measures were also taken to see how the V-nasalization is directly conditioned by the physical distance from N, but they are not reported in this paper due to the space limit.)

The coda N, however, did not clearly distinguish itself from the preceding vowel as discussed above. In fact, as discussed in [4], the coda N was often produced without clear acoustic evidence for its oral constriction, so that it was often impossible to detect the acoustic onset of the coda N. Thus, we measured the whole VN-duration rather than N-duration to examine the overall durational effect in CVN#. Again, A1-P0 was taken from the relative time points. In other words, the entire VN-duration was divided into 10 portions by an increment of 10%, and nine points were taken: from 10% point (near the end of VN) to 90% point (near the vowel onset). The A1-P0 values were extracted by using a Praat script [15].

2.4. Statistical Analysis

A series of linear mixed effects models were conducted. Fixed effects were Focus (UnFoc vs. Foc), and Boundary (Wd vs. IP) with contrast coding. The relative Timepoint (nominal factor) was added as another fixed effect with three time points in NV. In VN context, the relative time points were measured in three phases separately (Phase 1: 90%~70%; Phase 2: 60%~40%; Phase 3: 30%~10%) in which three time points were used for each Timepoint (nominal) factor. In order to examine how the prosodic factors interacted with the time points across the entire VN sequence in the coda context, an additional model was run with Timepoint as a continuous factor (with 9 time points). The maximal random effects structure

was employed as long as the model converged (i.e. by-subject intercept and slopes for all test variables).

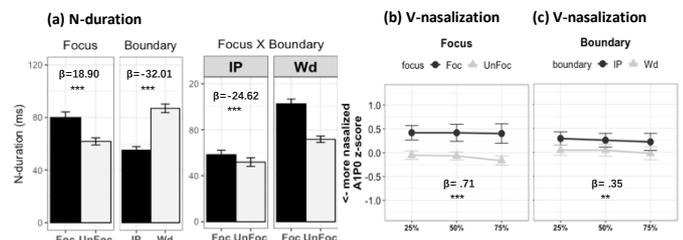
3. RESULTS

3.1 #NV (phrase-initial) context

N-duration in #NV [Fig.2a] showed a main effect of Focus ($\beta=18.90, p<.001$), with N being longer in the focused than in the unfocused condition. A main effect of Boundary was found as well ($\beta=-32.01, p<.001$), showing N being longer in the Wd-initial position than in the IP-initial position. There was an interaction between Focus and Boundary ($\beta= -24.62, p<.001$), indicating that the focus effect was suppressed in the IP condition presumably due to the general shortening trend for IP-initial N.

As for V-nasalization, there was a main effect of Focus ($\beta= .71, p<.001$) [Fig.2b], showing that vowels were more nasalized in the unfocused condition. The Boundary factor showed a main effect ($\beta= .35, p<.01$) [Fig.2c], indicating a reduction of V-nasalization IP-initially. A significant downtrend was observed at 75% point ($\beta= -.19, p<.001$). No interaction among Focus, Boundary and Timepoint was observed.

Figure 2: N-duration and V-nasalization (A1-P0) in #NV



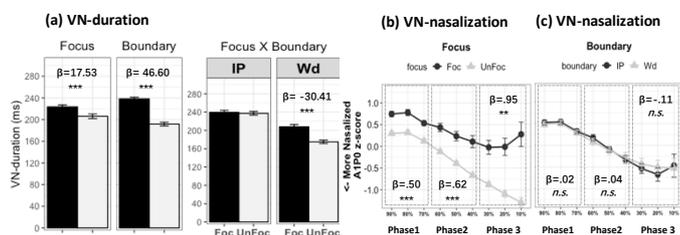
3.2 CVN# (phrase-final) context

As can be seen in Fig.3a, there was a significant main effect of Focus on VN-duration in CVN# ($\beta=17.53, p<.001$), showing a focus-induced lengthening of VN-duration. VN-duration also showed a main effect of Boundary ($\beta= 46.60, p<.001$), being longer IP-finally than Wd-finally, showing a general final lengthening effect. An interaction between Focus and Boundary ($\beta=-30.41, p<.001$) was observed as well.

Turning to V-nasalization in CVN#, as seen in Fig.3b, vowels were less nasalized in the focused than unfocused condition in all three phases ($\beta=.50, p<.001$ in Phase 1; $\beta=.62, p<.001$ in Phase 2; $\beta=.95, p<.01$ in Phase 3) throughout the vowel. The Timepoint effect was significant at 70% point in Phase 1, at 50% and 40% points in Phase 2, and at 20% and 10% points at Phase 3. The output of the model with Timepoint as a continuous factor indicated that there was a significant main effect of Focus ($\beta=.69, p<.001$) and Timepoint ($\beta=.01, p<.001$), and there was an interaction effect between the two ($\beta=.01,$

$p < .001$). The interaction, as can be seen in Fig. 3, was due to the fact that the focus-induced difference in V-nasalization became progressively larger towards the end of the phrase.

Figure 3: VN-duration and V-nasalization (A1-P0) in CVN#.



Crucially, there was no main effect of Boundary in the models of three Phases and in the model with Timepoint as a continuous factor. As shown in Fig. 3c, V was not more nasalized IP-finally. An interaction between Boundary and Timepoint was found only for Phase 3 ($\beta = -.17$, $p < .05$). Separate (posthoc) analyses indicated that there was a significant Boundary effect only at 20% time point ($\beta = -.22$, $p < .05$). As can be seen in Fig. 3c, the interaction was due to the trend towards the boundary-induced difference becoming clearer towards the end of the VN.

4. SUMMARY AND DISCUSSION

One of the basic findings regarding the focus-induced prominence effect was that N-duration (in the case of #NV) and the VN-duration (in the case of CVN#) were temporally augmented in both #NV and CVN#, indicating a strengthening of N's nasality in the contrastively focused context. Despite the augmented nasality for N under focus, its coarticulatory influence on the neighboring vowel was not reinforced but, if anything, V-nasality in the vowel tended to be reduced in both #NV and CVN# context. In other words, the strengthening of N did not show coarticulatory aggression effect on the following V (cf. [6]), so that V's orality was either maintained (in the case of #NV) or enhanced (in the case of CVN#). The fact that the V's coarticulatory resistance spreads to all time points in the vowel suggests that the effect is more than a simple low-level phonetic effect. This effect pertains to linguistic contrasts, consistent with previously reported cross-linguistic patterns found in American English and Korean, all showing a kind of paradigmatic contrast enhancement [8,10].

As for boundary effects, we observed a general shortening of N in IP-initial position in line with the cross-linguistic observations that domain-initial strengthening increases the consonantality by decreasing its sonority (nasality). The following vowel was also less nasalized, showing that

boundary-induced domain-initial strengthening gives rise to an enhancement of syntagmatic CV contrast with an increase in consonantality and a heightening of V's phonetic (oral) clarity.

The most interesting finding, however, pertains to the boundary-related *domain-final* effect on CVN#. While Mandarin does not show a clear-cut division between V and N in the acoustic signal (cf. [5]), in most cases (8 out of 9 time points) there was no significant boundary effect observed in our data, except for only one time point near the end (20%) which showed more V-nasalization in IP-final than in IP-medial position. The overall lack of boundary effects stands in sharp contrast with phrase-final V-nasalization patterns found in American English and Korean. Both languages showed more nasalization in IP-final than IP-medial position generally across the board, which indicated a universal propensity of V-nasalization possibly due to phrase-final articulatory weakening which attenuates the oral articulatory force to close off the velopharyngeal port ([8,10]). This then implies that Mandarin does not follow the general, presumably biomechanically-driven, cross-linguistic pattern. While more data is required to confirm this pattern, our interim interpretation is that this unique effect may be driven by distributional restrictions on the occurrence of consonants in the coda position in Mandarin. Given that Mandarin allows only nasal consonants in the coda position, it does not employ a rich phonological manner contrast in that position. This is also presumably why nasals in the coda are often produced without oral constriction as its phonological function may be sufficiently signaled by the presence of [nasal] feature only (i.e., vowel nasalization). One could then further assume that in the absence of the nasal murmur cue during stop closure due to a delink of place feature, V-nasalization carries information about the nasal consonant, thus being less vulnerable to the biomechanically-driven articulatory weakening effect.

In conclusion, the present study has indicated that while Mandarin shows cross-linguistically applicable coarticulatory effects, some of the effects should be attributable to the language's specific linguistic structure. The observed cross-linguistic generalizability and language-specificity supports the general view that the low-level phonetic process operates across languages but it is fine-tuned by the language's internal structure such as the language-specific phonological and prosodic structure (e.g., [6,7,8,10, 11]). The results of the present study, however, are based on a limited tonal context. We are in the processing of analyzing data in different tonal contexts, which will better illuminate the nature of V-nasalization in relation to tonal structures with an increased generalizability.

5. ACKNOWLEDGEMENT

We thank the Mandarin Chinese speakers for their participation and the reviewers for very helpful comments. This work was supported by Global Research Network program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A2A2912410) awarded to TC.

6. REFERENCES

- [1] Barlaz, M., Shosted, R., Fu, M., Sutton, B. 2018. Oropharyngeal articulation of phonemic and phonetic nasalization in Brazilian Portuguese. *Journal of Phonetics*, 71, 81-97.
- [2] Beddor, P.S., Harnsberger, J.D., Lindermann, S. 2002. Language-specific patterns of vowel-to vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics* 30, 591-627.
- [3] Chao, Y.R. 1948. *Mandarin Primer*. Cambridge: Harvard University Press.
- [4] Chen, M. 1972. *Nasals and nasalization in Chinese: an essay in natural phonology*. Unpublished doctoral dissertation. University of California Berkeley.
- [5] Chen, M. Y. 2000. Acoustic analysis of simple vowels preceding a nasal in standard Chinese. *Journal of Phonetics* 28, 43-67.
- [6] Cho, T. 2016. Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass*, 10(3), 120-141.
- [7] Cho, T., James McQueen. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2) 121-157
- [8] Cho, T., Kim, D., Kim, S. 2017. Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics* 64, 71-89.
- [9] Duanmu, San. 2000. *The Phonology of Standard Chinese*. New York: Oxford University Press.
- [10] Jang, J., Kim, S., Cho, T. 2018. Focus and boundary effects on coarticulatory vowel nasalization in Korean with implications for cross-linguistic similarities and differences. *The Journal of the Acoustical Society of America* 144, EL33-39.
- [11] Kim, S., Kim, J., Cho, T. 2018. Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics*, 71, 65-80.
- [12] Kühnert, B & Nolan, F. 1999. "The origin of coarticulation" in *Coarticulation: Theory, Data and Techniques*, edited by W.J. Hardcastle and N. Newlett. Cambridge: Cambridge University Press, 7-30.
- [13] Scarborough, R., Zelou, G., Mirzayan, A., and Rood, D.S. 2015. Phonetic and phonological patterns of nasality in Lakota vowels. *Journal of the International Phonetic Association*, 45(3), 289-309.
- [14] Shattuck-Hufnagel, S., and A. E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193-247.
- [15] Styler, W. 2015. *On the acoustical and perceptual features of vowel nasalization*. Ph.D. dissertation. University of Colorado.