Interspeech 2022
18-22 September 2022, Incheon, Korea

Mandarin Nasal Place Assimilation Revisited: An Acoustic Study

Luo Mingqiong

Shanghai International Studies University (SISU), Shanghai, China
`joanlmq@shisu.edu.cn`

Abstract

Two types of nasal place assimilation (NPA) have been proved to take place in Mandarin Chinese (MC) disyllabic words: categorical NPA and gradient NPA. The former mainly happens in /n.C/ clusters and the latter /ŋ.C/ clusters. Previous studies show that in fast speech, the nasal coda may be completely lost, indicating a possible effect of speech rate on NPA patterns. Previous studies on NPA in other languages show that not only does the consonant have anticipatory effect on the preceding nasal, but the nasal also has carryover effect on the following consonant. However, no agreement has been reached concerning speech-rate effect on NPA patterns. This study aims to examine whether the nasal coda in MC N.C clusters has carryover effect on C and how speech rate affects N.C coarticulatory patterns. A production experiment was carried out and results show that (i) coda /n/ does not have carryover effect on Cs, but the fronted /a/ does; (ii) coda /ŋ/ does have carryover effect on the consonantal onset; (iii) speech rate does not have significant effect on the overall N.C coarticulatory patterns, but does for certain Cs. The results were explained in light of the degree of articulatory constraint (DAC) model.

Index terms: nasal place assimilation; coarticulatory resistance; speech rate; the DAC model

1. Introduction

Mandarin Chinese (MC) has two nasal codas, specifically, the coronal nasal /n/ and the velar nasal /ŋ/. On the other hand, MC boasts a variety of syllable onsets in terms of both place and manner of articulation, including the bilabial stops /p, p^{h}/, the labiodental fricative /f/, the dental fricative /s/ and affricates /ts, ts^{h}/, the alveolar stops /t, t^{h}/, the palatal fricative /ɕ/ and affricates /tɕ, tɕ^{h}/, the velar stops /k, k^{h}/ and the retroflex fricative /ʂ/ and affricates /ʈʂ, ʈʂ^{h}/.

A number of recent studies confirmed that NPA did happen in MC V$_1$N.CV$_2$ disyllabic words in connected speech (e.g. [1][2][3]). Specifically, in slow- and medium-rate speech, the place of articulation of coda /n/ assimilates completely to that of the following consonantal onset. For example, /p^{h}an.pa/ > [p^{h}am.pa] (‘climb’), and /kan.ka/ > [kaŋ.ka] (‘embarrassed’). On the other hand, nasal coda /ŋ/ usually undergoes gradient assimilation, that is, the closure gesture for /ŋ/ overlaps with that for the following consonantal onset. For example, /taŋ.pa/ > [taŋ͡m.pa] (‘be a father’), and /paŋ.ta/ > [paŋ͡n.ta] (brand name of a medicine). In fast-rate speech, however, both nasal codas may undergo complete deletion, leaving the preceding vowel heavily nasalized. That is to say, these studies show that speech rate can be an important factor affecting NPA patterns in MC.

However, recent studies on NPA patterns in other languages (e.g. [4][5]) discovered that clear speech, although produced in fast rate, actually showed similar assimilatory patterns as that produced in normal rate. Casual speech, on the other hand, does have coarticulatory patterns significantly different from those in formal speech. That is to say, speech rate may not be as important a factor in affecting NPA patterns as speech style is.

Moreover, [4]&[5] on Spanish NPA found that the impact of assimilation may be bi-directional, i.e. not only does C affect N place of articulation, but N also affects that of C. This finding agrees with a number of recent articulatory research findings on gestural overlap in consonant clusters, the effect of which is confirmed to be bi-directional [6][7][8]. However, previous research on MC NPA paid little (if not no) attention to the carryover effect in NC sequence. Therefore, the present study aims to tackle two (sets of) questions:

(1) Does N place have carryover effect on C place? If yes, what are the patterns like, and how to account for them?

(2) Does variation in speech rate affect NPA patterns (i.e., categorical NPA and gradient NPA) in MC clear speech?

2. Literature review

2.1 Coarticulation in consonant clusters (CC) and DAC

In recent Articulatory Phonology [9][10][11], categorical NPA and gradient NPA are viewed as different processes in connected speech, the former being assimilation and the latter gestural overlap. Both are typical of CC coarticulatory patterns and attested in many languages across the world. Many variables have been found to affect CC coarticulation patterns. Major language-internal variables include the inherent coarticulatory properties of the consonants, prosodic positions (e.g. syllable onset *vs.* coda, syllable boundary, phrase boundary) and features (e.g. stressed *vs.* unstressed) of the consonants, and major language-external variables include speech rate and speaker idiosyncrasies.

A number of theories and models have been proposed to account for coarticulation patterns (e.g. [12][13][14][15][16]), among which the most recent one is the degree of articulatory constraint (DAC) model. DAC derives from the notion “coarticulatory resistance”, which, in simple terms, refers to the likelihood of a segment to affect and be affected by other segments [17]. The DAC model proposes that the more strictly a specific region of the tongue (especially the tongue dorsum) is required to be involved in forming the constriction, the higher the DAC value is for the consonant. Thus, for example, bilabials have low DAC values since they do not require any tongue involvement in forming the constriction. Dentoalveolars have higher DAC values than bilabials because they require that tongue front be involved in forming the constriction. Velars exhibit systematic variations because although they require that tongue dorsum (i.e. highly restricted tongue region according to the model) be involved in forming the constriction, the constriction location may vary considerably according to neighboring vowels, and therefore, the DAC value can be high or low. Palatals have the highest DAC value because they have

10.21437/Interspeech.2022-89

very strict requirements on the involvement of tongue dorsum in forming the constriction. In addition, the DAC model proposes that manner of articulation also affects the DAC specification of consonants. Thus, fricatives (and affricates) have higher DAC values than stops with the same places of articulation because the former require the formation of a narrow central groove for the passage of airflow. Nasal consonants have lower DAC values than non-nasal ones with the same places of articulation because the former have looser requirements on the closure gesture for oral constriction than the latter do [18][19][20][21][22].

According to the DAC model, consonants with high DAC values are more likely to influence neighboring segments and resist coarticulatory effect from them, whereas those with low DAC values are less resistant to coarticulatory effect from neighboring segments and more likely to be affected. Thus, consonants can be listed according to their capacity in coarticulatory resistance as follows:

(1) bilabials < dentoalveolars, velars < palatals

(2) stops < fricatives, affricates

(3) nasals < non-nasals

These lists have been proved to be valid in V_1CV_2 sequences in many languages across the world, such as English, Catalan and Spanish. Moreover, recent studies show that these lists are also valid for MC [23][24]. Based on (1, 2, 3) and the fact that categorical NPA happens in MC /n.C/ sequence so that V_1 have direct interactions with C place, we make the following prediction:

(4) Consonant /p/ in V_1n.CV_2 sequence is least resistant to carryover effect from V_1, and /t, k/ are more resistant. Affricate /ts/ is more resistant to V_1 carryover effect than /t/, but less so than /tɕ, ʈʂ/, which are most resistant to V_1 carryover effect.

When two consonants with different DAC values come in contact, it is not always the case that the consonant with higher DAC value exerts more coarticulatory effects on that with lower DAC value than the other way round. Order of consonants and gestural antagonism also come into play. Gestural antagonism refers to the conflict between C_1 and C_2 in the requirements for the involvement of specific articulators in closure gesture. Coarticulation is expected to be more prominent for CCs with no conflict in articulator involvement than for those with conflict. Thus, we can make the following predictions:

(5) Carryover effect from /ŋ/ to /p, t, ts/ in /ŋ.p, ŋ.t, ŋ.ts/ clusters should be more prominent than that from /ŋ/ to /tɕ, ʈʂ/ in /ŋ.tɕ, ŋ.ʈʂ/ clusters because gestural antagonism in the former group is greater than that for the latter group, and /tɕ, ʈʂ/ have higher DAC values than /p, t, ts/.

(6) No carryover effect from /ŋ/ to /k/ is expected in /ŋ.k/ cluster because the two consonants share major articulators for closure gesture, but carryover effect from V_1 to /k/ is expected because /ŋ/ and /k/ are homorganic.

2.2 Speech-rate effect on coarticulatory patterns

Previous studies have reported inconsistent results on speech-rate effect on CC coarticulatory patterns. Some studies showed that faster speech rate did cause segmental shortening in duration and more gestural overlap in CCs, which in turn induced more prominent coarticulatory effects [25][26][27][28]. Other studies, however, showed that CC gestural overlap and speech rate was not indicative of a direct relationship. In other words, faster speech rate does not necessarily cause greater gestural overlap, and thus coarticulatory patterns do not vary significantly [29][30][31].

In summary, in the present study, we are going to test Hypotheses (4, 5, 6) listed in Section 2.1 and then examine whether and how speech rate affects NC coarticulation patterns in MC in both spatial and temporal extent, i.e. the magnitude of coarticulatory effect and temporal extent of gestural overlap.

3. A production experiment

3.1 Materials

A disyllabic word list containing $V_1N.CV_2$ sequences was designed for this experiment. V_1 and V_2 were both set to be /a/. The consonantal onset of the first syllable is set to be unaspirated stops in Mandarin, i.e. either /p/, /t/ or /k/. The nasal coda is either /n/ or /ŋ/. The consonantal onset following N is one of /p, d, k, ts, tɕ, ʈʂ/, all of which have a definite constriction location along the vocal tract during the closure phase. So altogether the material contains 1×2×6 = 12 disyllabic words.

In order to observe the carryover effect of N place on C place, a list of $V_1.CV_2$ disyllabic words in Mandarin was also prepared. V_1, V_2 and C in $V_1.CV_2$ sequence were exactly the same as those in the $V_1N.CV_2$ sequence. Altogether this material contains 1×2×6 = 12 disyllabic words. The assumption is that if N place has carryover effect on C place, the carryover pattern of in $V_1N.CV_2$ sequence should vary as a function of N place; on the other hand, if there is no N-to-C effect, the carryover pattern of N.C sequences remains the same despite N place variation.

3.2 Subjects

Altogether 10 subjects participated in this experiment. All were native Mandarin speakers with a specific northern dialect as their other native languages. They were all second-year college students at the time of recording. None of them reported articulation- or auditory-related impairments. The subjects participated in a training session of about 15 minutes on speech-rate variation before recording. and received a moderate amount of payment after the experiment.

3.3 Procedures and measurements

The experiment was carried out in a sound-proof booth in an audio-visual classroom affiliated with a university in Shanghai, China. The recording was made with a RODE-mini microphone connected to the computer. The software used for recording was Audacity. The subjects were instructed to read the words on the materials in slow, medium and fast rate, respectively, for three times. Each time they were instructed to read first at their normal speech rate (i.e., medium rate), and then at perceptually slower or faster rates, respectively. If speech rate distinction is perceptually unclear in one reading, the subject was required to read again with speech rate adjusted. Altogether (12×3×3+12×3×3) × 10 = 1440 tokens of disyllabic words were used for later analysis.

The segmentation was made with Praat [32] on the computer by using a cohort of acoustic cues in the waveform and spectrogram, including abrupt changes in the acoustic wave form and spectral energy display, formant tracking and voicing bar intensity as shown in the spectrogram. Audio verification was used when necessary.

The annotation was done manually in textgrid with Praat and the measurement of formants was done manually by placing the cursor at the midpoint and at the boundary and reading out the value provided by Praat directly. The sampling rate was 44100Hz and the formant frequency was set at 5000Hz for male subjects and 5500Hz for female ones. The formant number was 5. Moderate adjustments were made on the parameter settings when necessary.

The difference between second vowel formant F2 at the V_1-C, V_1-N and C-V_2 transitions and that at the mid-point of V_1 and V_2, i.e. ΔV_1F2 and ΔV_2F2, were used as the phonetic correlates for the anticipatory and the carryover effect, respectively, following [6]. The F2 frequency of V_1 and V_2 at the mid-point was obtained by averaging across all $V_1.CV_2$ and $V_1N.CV_2$ tokens, respectively, for each subject.

The ratio between the duration of N and that from N closure onset to C release onset in $V_1N.CV_2$ sequence was used as the phonetic correlate for the temporal extent of N.C gestural overlap. Our assumption is that the greater the ratio, the greater the extent of N.C gestural overlap may be.

3.4 Statistical results

The statistics were run in R using the lme4 package [33].

3.4.1 Carryover effect of N place on C place

To analyze the carryover effect of N place on C place, a linear mixed-effects regression models were fit to the data with ΔV_2F2 value as the dependent variable and N place as the independent variable. Subject was set as the random effect (Note that gender is not included because R shows that it is a singular factor). V.CV sequence (zero coda) was the baseline of comparison. Table 1 below shows the statistical results.

Table 1: *Carryover effect in N.C clusters across Cs*

	n	**ŋ**
Slow	+***	−
Mid	+**	−**
Fast	+***	−

Table 1 shows that in speeches of all rates, ΔV_2F2 values in /an.Ca/ sequences are significantly higher (+) than the baseline, indicating prominent carryover effect from V_1. On the other hand, ΔV_2F2 value in /aŋ.Ca/ sequences in medium-rate speech is significantly lower (−) than the baseline, but those in slow- and fast-rate speech are not significantly different, indicating prominent carryover effect from /ŋ/ in the former but not in the latter.

To further explore the carryover effect of N place on specific C place, subsets of the data for each individual C in slow-, medium- and fast-rate speech, respectively, was drawn from the general data. A linear mixed-effect regression model was fit to the data using N place as the major effect and subject as the random effect. V.CV sequence was used as the baseline for comparison. Table 2 below displays the statistical results.

Table 2: *Carryover effect in N.C clusters for each C*

		p	**t**	**k**	**ts**	**tɕ**	**ʈʂ**
n	*slow*	+**	+***	+	+**	+	+*
	mid	+***	+	+	+***	+	+
	fast	+*	+***	+**	+***	+*	+*
ŋ	*slow*	−	+	−**	−	+	−
	mid	−	−*	−**	−*	−	−
	fast	−	+	−	−*	+	+

Table 2 shows that in /n.C/ clusters, ΔV_2F2 value for /p/ is significantly higher than the baseline across speech rates, indicating very prominent carryover effect from V_1. This is also true with /ts/, on which the effect seems more prominent. ΔV_2F2 values for /t/ are significantly higher than the baseline in slow- and fast-rate speech but not in medium-rate speech, indicating very prominent carryover effect from V_1 in the former contexts but not in the latter. This is also true with /ʈʂ/, although the effect seems not as prominent as that on /t/. ΔV_2F2 values for /k, tɕ/ are not significantly different from the baseline in slow- and medium-rate speech but is in fast-rate speech, indicating no prominent effect from V_1 in the former contexts but very prominent effect in the latter.

In /ŋ.C/ clusters, ΔV_2F2 value for /p, tɕ, ʈʂ/ are not significantly different from the baseline across speech rates, indicating no carryover effect from the back /a/ or coda /ŋ/ on C place. In addition, ΔV_2F2 values for consonants /t, ts/ in medium-rate speech and that for /ts/ in fast-rate speech are significantly lower than the baseline, indicating prominent carryover effects from coda /ŋ/. In addition, ΔV_2F2 value for /k/ is significantly lower in slow- and medium-rate speech, but not in fast-rate speech, indicating prominent carryover effect from the back /a/ in the former contexts but not in the latter.

Overall, results show that in /n.C/ clusters across speech rates, /p, ts/ are least resistant to carryover effect from V_1, and /k, tɕ/ are most resistant, with the effects of /t, ʈʂ/ at a medium-level. In /ŋ.C/ clusters, /p, tɕ, ʈʂ/ are more resistant to carryover effect from coda /ŋ/, /t/ is less resistant and /ts/ is the least resistant.

3.4.2 Speech-rate effect on coarticulatory patterns in MC NPA

Since N.C anticipatory patterns (i.e., NPA patterns) differ for different nasal codas, we drew subsets for codas /n/ and /ŋ/, respectively, and fit a linear mixed-effects regression model to each. The formula was set using ΔV_1F2 as the dependent variable and the interaction between C and speech rate as the independent variable. Subject was set as the random effect. The baseline of comparison for /n/ was consonant /t/ in medium-rate speech, and that for /ŋ/ was /k/ in medium-rate speech.

Statistical results on both /n.C/ sequences and /ŋ.C/ sequences show that ΔV_1F2 values for all interactions between C and speech rate are not significantly different from the baseline. That is to say, speech rate has no significant effect on C-to-N anticipatory patterns for all Cs examined in this study.

To explore speech-rate effect on N-C carryover patterns, we drew subsets from the general data for each C and fit a linear mixed-effects regression model to each subset. The formula was set with ΔV_2F2 as the dependent variable and with the interaction between N and speech rate as the independent variable. Subject was set as the random effect. The baseline of comparison was $V_1.CV_2$ sequences in medium-rate speech.

Table 3: *Interactions between N and speech rate*

	p	**t**	**k**	**ts**	**tɕ**	**ʈʂ**
*n*slow*	−	+	+	−	−	+
*ŋ*slow*	+	+*	+	+	+	+
*n*fast*	−	+**	+	+	+	+
*ŋ*fast*	+	+	+	−	+	+

Table 3 shows that ΔV_2F2 values for all interactions between N and speech rate are not significantly different from the baseline except for /ŋ.t/ sequence in slow-rate speech ($p < 0.05$) and /n.t/ sequence in fast-rate speech ($p < 0.01$). That is to say, speech rate has no significant effect on N-to-C carryover patterns for all Cs except /t/.

In order to explore the speech-rate effect on the temporal extent of gestural overlap in /ŋ.C/ sequences, a linear mixed-effects regression model was fit to the data with the ratio between N duration and (N + C release onset) duration as the dependent variable, and the interaction between C and speech rate as the independent variable. Subject was set as the random effect. The baseline of comparison was consonant /k/ in medium-rate speech.

Statistical results show that C place has significant effect on the temporal extent of /ŋ.C/ gestural overlap. Specifically, gestural overlap between /ŋ/ and /ts, tɕ, ʈʂ/ is significantly greater in temporal extent than that in /ŋ.k/ sequence (refer to Table 4 for specific values of gestural overlap for each C). However, the interaction between C and speech rate does not show any significant effect. That is to say, speech rate does not have significant effect on the temporal extent of /ŋ.C/ gestural overlap. However, N/(N+C) ratio differences in /ŋ.p, ŋ.t/ clusters in fast speech are significantly different from the baseline, indicating a possible significant effect of fast speech on N.C gestural overlap for certain Cs.

4. Discussion

4.1 DAC and carryover effects in MC N.C clusters

Statistical results in Section 3 on N-to-C carryover effect in MC $V_1N.CV_2$ sequence prove that Hypothesis (4, 5, 6) in Section 2 are true with two exceptions, i.e. consonants /ts/ and /p/.

To be more specific, our hypothesis for /ts/ is that affricates are more resistant to coarticulation than stops, but experimental results show that /ts/ is least resistant to the carryover effect from both front /a/ and coda /ŋ/ among all Cs examined in this study. Consonant /p/, on the other hand, is expected to be the least resistant to coarticulation according to the DAC model, but experimental results show that it is as resistant to carryover effect from /ŋ/ as /tɕ/ is, which is most resistant to coarticulation.

Our explanation is that since /ts/ in MC only has requirements on the tongue tip in forming the constriction and leaves tongue blade and tongue dorsum relatively free for coarticulation, carryover effects at tongue blade and tongue dorsum can be prominent. Since /ts/ has no conflict with front /a/ or coda /ŋ/ in major articulators, it is expected that carryover effect from the latter are prominent on the former. That is to say, gestural antagonism may override manner of articulation and become the dominant factor in affecting coarticulatory patterns.

However, gestural antagonism can not explain why /p/ can resist the carryover effect from coda /ŋ/. Specifically, /p/ has no conflict in closure gesture with /ŋ/ for any major articulator, and gestural overlap between the two consonants is clear and definite, but no carryover effect from /ŋ/ to /p/ is found in speech of any rate. That is to say, the bilabial closure gesture for /p/ is forceful enough to resist the carryover effect of tongue-back retraction movement in the closure gesture for /ŋ/, whatever the length the latter has. This finding quite unexpected under the present version of the DAC model, and requires further research.

4.2 Speech-rate effect on coarticulatory patterns

Experimental results in this study are supportive of previous studies which found that speech rate does not have significant effects on CC coarticulation patterns in clear speech. It is worth noting that the two exceptions, i.e. carryover effects in /n.t/ and /ŋ.t/ clusters, may not be completely accidental. They show that /t/ may be more sensitive than other consonants to speech-rate variation in coarticulation. In addition, the results also show that speech rate has different effects on the temporal extent of gestural overlap in /ŋ.C/ clusters for different Cs.

Table 4: *Specific values of the temporal extent of /ŋ.C/ gestural overlap and statistical results on the anticipatory and carryover effects of /ŋ.C/ clusters for each C in three speech-rate contexts*

		p	**t**	**k**	**ts**	**tɕ**	**ʈʂ**
slow	*ge.ov*	0.63	0.65	0.67	0.74**	0.77	0.74
	anti.	–***	–	(bsl.)	–**	–	–
	caov.	–	+	–**	–	+	–
mid	*ge.ov*	0.64	0.65	0.67	0.79	0.77	0.78
	anti.	–***	–	(bsl.)	–	+	–
	caov.	–	–	–**	–*	–	–
fast	*ge.ov*	0.70**	0.71*	0.65	0.83	0.78	0.74
	anti.	–***	–***	(bsl.)	–***	+	–*
	caov.	–	+	–	–*	+	+

Table 4 shows that the temporal extent of gestural overlap in /ŋ.p/ cluster is smallest in slow- and medium-rate speech, but the anticipatory effect is the most prominent among all /ŋ.C/ clusters. Moreover, gestural overlap in /ŋ.p/ sequence is steadily increasing in temporal extent, but neither anticipatory nor carryover effect shows correlative changes. In contrast, gestural overlap in /ŋ.tɕ, ŋ.ʈʂ/ clusters are larger in temporal extent than when C is a stop across speech rates, but few significant anticipatory or carryover effect is found. In addition, gestural overlap in /ŋ.ʈʂ/ cluster is smallest in fast-rate speech as compared to that in slow- and medium-rate speech, but the anticipatory effect is most prominent.

All the evidence shows that the relationship between the temporal extent of gestural overlap and spatial extent of coarticulatory effects is non-linear.

5. Conclusion

This study uses acoustic measures to explore N.C coarticulatory patterns in MC disyllabic words, focusing on the carryover effect of N place on C place and the speed effect, and manages to provide explanations in light of the DAC model. We found that in /n.C/ clusters, /n/ place has no carryover effect on C place, but /a/ before /n/ has significant carryover effect on C place. On the other hand, in /ŋ.C/ clusters, /ŋ/ place has significant effect on some consonants but not others. In both cases, we find that the spatial extent of coarticulatory effects differs for different Cs. The DAC model can help explain MC N.C coarticulatory patterns, but explorations on /N.p, N.ts/ clusters should be made further.

This study also finds that speech rate has little effect on either anticipatory or carryover effect of N.C coarticulation, provided that the disyllabic words are produced in a clear way. Moreover, generally speaking, gestural overlap increases as speech rate increases, but the temporal extent changes differ for different Cs. Besides, no correlative relation between the temporal extent of gestural overlap and the spatial extent of coarticulation is found.

6. References

[1] J. Liang. “On the Factors Leading to the Loss of Nasal Rhyme in Connected Speech,” *Proceedings of National Conference on Modern Phonetic,* pp126-129, 2001.

[2] Q. Fang. “Loss of nasal coda in connected speech,” *Speech Research Report*, pp. 91-97, 2004.

[3] Y. Zheng and J. Liu. “The Acoustic Mode of Mandarin N1C2 (C#C) Coarticulatory Patterns,” *Proceedings of National Seminar on Human-machine Communication System*, pp. 61-66, 2005.

[4] A. Kochetov and L. Colantoni, “Spanish nasal assimilation revisited: A cross-dialect electropalatographic study,” *Laboratory Phonology*, vol. 2, no. 2, pp. 487-523, 2011.

[5] D. Honorof, *Articulatory gestures and Spanish nasal assimilation*. Yale University, 1999.

[6] G. Modarresi, H. Sussman and B. Lindblom, et al. “An acoustic analysis of the bidirectionality of coarticulation in VCV utterances”, *Journal of Phonetics*, vol. 32, no. 3, pp. 291-312, 2004.

[7] F. Gibbon, W. Hardcastle and K. Nicolaidis, “Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: a cross-linguistic investigation,” *Language and Speech*, vol. 36, pp. 261-277, 1993.

[8] W. Hardcastle, “Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences,” *Speech Communication*, vol. 4, pp. 247-263, 1985.

[9] C. Browman and L. Goldstein, “Articulatory gestures as phonological units,” *Phonology*, vol. 6, pp.201-251, 1989.

[10] C. Browman and L. Goldstein, “Gestural specification using dynamically-defined articulatory structures,” *Journal of Phonetics*, vol. 18, pp. 299-320, 1990.

[11] C. Browman and L. Goldstein, “Articulatory phonology: An overview,” *Phonetica*, vol. 49, pp. 155-180, 1992.

[12] S. E. Öhman, “Coarticulation in VCV utterances: Spectrographic measurements,” *Journal of the Acoustical Society of America*, vol. 39, pp. 151–168, 1966.

[13] C. A. Fowler, “Coarticulation and theories of extrinsic timing,” *Journal of Phonetics*, vol. 8, pp. 113–133, 1980.

[14] F. Bell-Berti and K. Harris, “A temporal model of speech production,” *Phonetica*, vol. 38, pp. 9–20, 1981.

[15] P. A. Keating, “The window model of coarticulation: Articulatory evidence,” *UCLA Working Papers in Phonetics*, vol. 69, pp. 3–29, 1988.

[16] B. Lindblom, “Explaining phonetic variation: A sketch of the H & H theory,” In W. J. Hardcastle & A. Marchal (eds.), *Speech Production and Speech Modelling*, pp. 403–440, 1990.

[17] R. Bladon and A. Al-Bamerni, “Coarticulation resistance in English /l/,” *J. Phonetics,* vo. 4, pp. 137–150, 1976.

[18] D. Recasens, M. Pallarès and J. Fontdevila, “A model of lingual coarticulation based on articulatory constraints,” *J. Acoust. Soc. Am.* Vol. 102, pp. 544–561, 1997.

[19] D Recasens and MD Pallarès, “Coarticulation, assimilation and blending in Catalan consonant clusters,” *Journal of Phonetics*, vol. 29, no. 3, pp. 273-301, 2001.

[20] D. Recasens, “Vowel-to-vowel coarticulation in Catalan VCV sequences,” *The Journal of the Acoustical Society of America*, vol. 76, no. 6, pp. 1624-1635, 1984.

[21] D. Recasens and A. Espinosa, “An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan,” *Journal of the Acoustical Society of America,* vol. 125, no. 4, pp. 2288-2298, 2009.

[22] D. Recasens, “The Effect of Stress and Speech Rate on Vowel Coarticulation in Catalan Vowel-Consonant-Vowel Sequences,” *Journal of Speech Language & Hearing Research*, vol. 58, no. 5, pp. 1407-1425, 2015.

[23] X. Chen, “A Study on Coarticulation in Mandarin Chinese (MC),” *Studies of the Chinese Language*, vol. 5, pp. 345-351, 1997.

[24] L. Zhang, *Coarticulation in Mandarin Syllables: An Acoustic Study,* Doctoral Dissertation, East China Normal University, 2012.

[25] K. Munhall and L. Anders, “Gestural aggregation in speech-laryngeal gestures,” *Journal of Phonetics,* vol. 20, pp. 111–126, 1992.

[26] D. Byrd and C. Tan, “Saying consonant clusters quickly,” *Journal of Phonetics,* vol. 24, pp. 263–282, 1996.

[27] K. Tjaden and W. Gary, “Speaking-rate-induced variability in F2 trajectories,” *Journal of Speech, Language, and Hearing Research,* vol. 41, pp.976–989, 1998.

[28] E. Fosler-Lussier and M. Nelson, “Effects of speaking rate and word frequency on pronunciations in conversational speech,” *Speech Communication*, vol. 29, pp. 137–158, 1999.

[29] E. Zsiga, “Acoustic evidence for gestural overlap in consonant sequences,” *Journal of Phonetics*, vol. 22, pp. 121-140, 1994.

[30] R. Dixit, R. Prakash and J. Flege, “Vowel context, rate and loudness effects of linguopalatal contact patterns in Hindi retroflex /ʈ/,” *Journal of Phonetics*, vol. 19, pp. 213–229, 1991.

[31] A. Kochetov, P. Marianne and S. Minjun, “Cross-language differences in overlap and assimilation patterns in Korean and Russian,” *Proceedings of the 16th ICPhS*, pp. 1361–1364. Saarbrücken, 2007.

[32] P. Boersma and D. Weenink (1992–2022): Praat: doing phonetics by computer [Computer program]. Version 6.2.06, retrieved 23 January 2022 from https://www.praat.org.

[33] D. Bates, M. Mächler, B. Bolker et al. “Fitting Linear Mixed-Effects Models Using lme4,” *Statistics & Computing,* arXiv:1406(1), pp. 133-199, 2014.