Production and perception of tones by Dutch learners of Mandarin

Zou, T.; Zou T.

Chapter two

Effect of cognitive load on tonal coarticulation: Evidence from native Mandarin speakers and Dutch learners of Mandarin

2.1 Introduction

In Mandarin Chinese, a lexical tone language, pitch movements (cued mainly via fundamental frequency) are used to convey lexical meaning. When produced in isolation, different tones are realized with stable and distinctive pitch contours. However, when produced in connected speech, tones can be influenced by the preceding and following tones and undergo substantial acoustic variation, leading to coarticulated f0 realizations which are different from the canonical contours. Such deviated f0 shapes make it a great challenge for adult non-tonal learners of Mandarin to achieve native-like tone production. It is evident in the literature that such anomalous tonal coarticulation patterns can be the cause of quite a part of the foreign accent of less proficient Mandarin speakers (Hao, 2012; Lee, Vakoch, & Wurm, 1996; Wang, Jongman, & Sereno, 2003).

Previous research on tonal coarticulation has mainly focused on the directionality (carryover or anticipatory), the nature (assimilatory or dissimilatory), and the magnitude of contextual effects on tonal production by native speakers (see Chen, 2012, for a review). In contrast, the underlying mechanisms of tonal coarticulation and the acquisition of coarticulated patterns by second language (L2) learners of Mandarin have remained much less-understood. This study was therefore designed to examine tonal coarticulation by both native and learners of Mandarin, with a particular focus on the effect of cognitive load on tonal coarticulation and the developmental trajectory with regard to the acquisition of tonal coarticulation by learners of non-tonal languages.

2.1.1 Tonal coarticulation in Mandarin Chinese

Coarticulation, the influence of one sound on a neighboring sound in speech production, is an issue that has been extensively studied. The traditional view is that it is a universal phenomenon caused by speech physiology, but it has become clear that both the pattern and the degree of coarticulation can be language-specific (Baumotte & Dogil, 2008; Beddor, Harnsberger, & Lindemann, 2002; Choi & Keating, 1991; Gandour, 1994; Hardcastle & Hewlett, 2006; Manuel, 1990; Oh, 2008).

A speech sound is influenced by both the preceding sound (i.e. carryover effect) and the subsequent sound (i.e. anticipatory effect). Such bidirectional coarticulatory effects have been reported in vowels and consonants, e.g., carryover effect as shown in Recasens (1984) and Beddor, Harnsberger, and Lindemann (2002); and anticipatory effects as shown in Martin and Bunnell (1981) and Grosvald (2009). Whalen (1990) has proposed that the carryover effect is a result of physiological constraints in realizing some motor program, since this effect remains robust when cognitive planning is constrained. The anticipatory effect, on the other hand, would be a reflection of speech planning, in that it decreases when the participants' planning mechanism is inhibited.

Different from vowel and consonant coarticulation, the realization of pitch target in tone production relies on a single articulator, the larynx. Therefore, adjacent tones with opposing pitch targets have to compromise with each other, as overlap in the timing of different gestures, which is common for coarticulation in segments (Browman & Goldstein, 1986), is less feasible in tonal coarticulation (Xu, 1994; DiCanio, 2014). Understanding tonal coarticulation, in comparison to segmental coarticulation, is thus important for research on coarticulation in general.

Most experimental studies on tonal coarticulation have been based on Asian contour-tone languages and the findings, as mentioned earlier, have been mainly in three aspects: the directionality (carryover or anticipatory), the nature (assimilatory or dissimilatory), as well as the magnitude and temporal extent of the articulatory effects. Similar to vowel and consonant coarticulation, previous findings show that tonal coarticulation can also be bidirectional, with assimilatory carryover effect and dissimilatory anticipatory effect. The carryover effect is generally strong and its influence can extend to the first half of or even the entire following syllable. The magnitude and temporal extent of the anticipatory effect is generally smaller compared to the carryover effect (e.g., Thai: Abramson, 1979; Gandour, Potisuk, & Dechongkit, 1994; Gandour, Potisuk, Ponglorpisit, Dechongkit, Khunadorn, & Boongird, 1996; Potisuk, Gandour, & Harper, 1997; Vietnamese: Brunelle, 2009; Han & Kim, 1974), although more recent experimental studies suggest that the anticipatory effect can also be quite salient (see, e.g. Chang & Hsieh, 2012; Li & Chen, 2016).

The patterns of Mandarin tonal coarticulation generally agree with the findings from the above reported patterns for other tone languages. Xu (1997) examined tonal coarticulatory patterns in disyllabic non-words /mama/ with all possible tonal combinations in Standard Chinese produced by native Beijing Mandarin speakers. The results showed that both carryover and anticipatory effects exist in Mandarin. The carryover effect exhibits an assimilatory nature; a high offset of the first tone can raise the onset of the following tone, while a low offset lowers the onset of the following tone. This effect shows a strong influence on the initial and middle part of the final syllable. The anticipatory effect, however, is largely dissimilatory. Xu (1997) showed that in Standard Chinese, the anticipatory effect is mainly on the maximum f0 of the preceding tone. Specifically, the low tone (T3) in the second syllable showed a raising effect on the initial part of the falling tone (T4) and the final part of the rising tone (T2). The magnitude of the anticipatory effect, however, is much smaller compared to the carryover effect. (It is to be noted that in a closely related Mandarin dialect, Tianjin Mandarin, Li & Chen (2016) showed that anticipatory raising may manifest as the raising of the whole tonal contour. More experimental studies are therefore needed to understand the full range of tonal coarticulatory effects in different dialects within the Mandarin dialect family.)

Thus far, what has remained under-investigated is the underlying mechanism and source of the tonal coarticulatory effects. Among the limited number of studies focusing on this issue, Xu (2001) argued that the carryover effect in Mandarin tone is most likely caused by articulatory constrains (i.e., the maximum speed of pitch change). For the anticipatory effect, Tilsen (2009, 2013) reported that native Mandarin speakers tended to dissimilate tones that were planned contemporaneously, which led him to suggest that the dissimilatory effect may result from an inhibitory speech planning mechanism between articulatory targets planned in parallel, with the goal to maintain and maximize phonemic contrasts.

Recently, Franich (2015) took a step further in this line of investigation by introducing the effect of cognitive load. Since motor planning in speech production is believed to recruit central processing resources (Gathercole & Baddeley, 2014; Meyer & Gordon, 1985), increase of cognitive load (Mattys & Wiget, 2011) is therefore expected to introduce the reduction of processing resources for articulatory planning. To introduce cognitive load, a dual-task paradigm was used in Franich's (2015) study. Native Mandarin speakers were asked to read disyllabic Mandarin non-words while being told that they would need to recall the two-digit numbers given before the reading of the non-words. A robust carryover effect was found in both normal and cognitive load conditions. Furthermore, dissimilatory anticipation effect was found to increase under high cognitive load, especially on the high tone (Tone 1) and the low tone (T3). This result is puzzling, however, given the earlier finding that inhibited planning should lead to decrease in anticipatory effect (Whalen, 1990). A possible reason is, as argued by Franich, that anticipatory coarticulation carries important linguistic function (maintaining and maximizing contrasts between phonemic categories) and may therefore have a dedicated cognitive mechanism for its realization even under high cognitive load. This then predicts that native and non-native speakers (especially beginning learners) may show differential effects of cognitive processing constraint on tonal coarticulation, if the cognitive mechanism is developed as a consequence of mastering the native language.

2.1.2 Tonal coarticulation patterns by L2 learners of Mandarin

Thus far, although existing studies on L2 Mandarin learners show that learners are able to correctly produce lexical tone in isolation (e.g., Hao, 2012; Wang, Jongman, & Sereno, 2003), their production of tones in connected speech is greatly challenged, evident in the higher error rates and decreased intelligibility of L2 speech (Hao, 2012; Shen, 1989; Sun, 1998; Yang, 2011, 2016). He and Wayland (2010) found that when producing coarticulated tones in disyllabic words, more experienced American learners of Mandarin are more accurate than less experienced learners. Brengelmann, Cangemi and Grice (2015) examined anticipatory tonal coarticulation in disyllabic sequences in German learners of Mandarin. Compared to native Mandarin speakers, German learners showed that for all four lexical tones, the influence of the following tone was mainly on the final part of the initial syllable. Furthermore, they also produced more f0 variations in the last 20 percent of the tone contours on the first syllable, with much of the variability due to non-native like anticipatory coarticulation. What remains to be learned is to what extent the observed pattern is general among learners of non-tonal languages. Their results motive more systematic studies of the acquisition of tonal coarticulation. A follow-up question is the developmental trajectory and mechanisms of tonal coarticulation that underlies the ultimate attainment of tonal acquisition by non-tonal second language learners.

To address these questions, the current study adapted the paradigm used in Franich (2015) and tapped further into the cognitive mechanisms of tonal coarticulation by both native and non-native Mandarin speakers. We intended to replicate the findings of Franich (2015) for native Mandarin speakers. Furthermore, we also tested beginning and advanced Dutch learners of Mandarin under cognitive load, aiming to reveal the developmental path of tonal coarticulation acquisition, which, we hope, can help to shed further light on the general mechanisms underlying tonal coarticulation.

2.2 Methods

2.2.1 Participants

Twelve Mandarin control participants and 22 Dutch learners of Mandarin participated in the experiment (10 beginning learners and 12 advanced learners). The native Mandarin control group had 3 males and 9 females (age: M = 26.3, SD = 3.0). All were from the Northern part of China and spoke standard Mandarin on a daily basis and fluently. Four were native speakers of Beijing Mandarin and the other eight speakers spoke standard Chinese as their dominant language, but they could speak another northern Mandarin dialect. All Dutch learners of Mandarin received formal Chinese training from the Chinese Studies program at Leiden University. The beginning group consisted of 4 males and 6 females (age: M = 20.6, SD = 2.5). Their Mandarin learning and speaking experience varied between 0.5 and 2 years (mean = 1.2, SD = 0.5), and they had never lived in China. The other 12 participants (4 males and 8 females; age: M = 24.0, SD = 3.6) were advanced Mandarin learners, who had Mandarin experience between 3 and 14 years (M = 4.8, SD = 3.1), and had spent at least one year in China.

2.2.2 Material and procedure

The stimuli, following the design of Xu (1997) on tonal coarticulation, were disyllabic non-word /mama/ with each syllable bearing one of the four Mandarin tones: the high tone (T1); the rising tone (T2), the low tone (T3) and the falling tone (T4) (Chen & Gussenhoven, 2008; Duanmu, 2000). When produced in the second syllable, T3 was expected to show a dipping contour just like its canonical form. It would be realized as a variant with low falling contour preceding T1, T2 and T4. According to the sandhi rule, T3 would be realized with a rising contour, similar to T2 preceding another T3. All 16 possible tonal combinations were tested with four repetitions in three conditions: no-cognitive-load, low-cognitive-load and high-cognitive-load condition. The cognitive-load conditions were manipulated following the paradigm of Lavie, Fockert and Viding (2004), with a minor change of using two-digit numbers as memory material in the low-cognitive-load condition instead of one-digit number. The participants were recorded individually in the Leiden University Phonetics Lab using E-prime (44.1 kHz, 16 bit) with a Sennheiser MKH416T microphone. The three groups of participants were asked to read the sequences given in pinyin, with instructions in their respective native languages (i.e. Chinese for the native Mandarin speakers and Dutch for the learners).

In the control no-cognitive-load (NCL) condition, a fixation point ("+") was first presented on the screen for 2s at the beginning of each trial. After that, a disyllabic pinyin with tone marks appeared on the screen. The participants were asked to simply read them aloud. They had 2.5s for each trial, and after that the next trial proceeded automatically.

For the two cognitive load conditions, the reading task was presented in the retention interval of a short-term memory task. In each trial, the reading task was preceded by memory material, and followed by memory testing material. In the low-cognitive-load (LCL) condition, the memory material was two one-digit numbers and in the high-cognitive-load (HCL) condition, it was six one-digit numbers. For both conditions, each trial started with a 2s presentation of a fixation point ("+") in the center of the screen. After that, a row of two digits were presented equally spaced (horizontally) for 500 ms in the LCL condition. In the HCL condition, a row of six digits were presented for 2s. During the presentation of the memory material, the participants were asked to try their best to remember the digits. Then, the memory digits were replaced by masking arrays with a 500-ms display of two asterisks for the LCL condition, and one 1s display of 6 asterisks for the HCL condition. The masking array was then followed by the presentation of pinyin with tone marks. A time window of 2.5s was provided for participants to read the pinyin aloud. After that, a green digit was presented as the memory testing material. The participants were required to decide whether this digit was present or not in the preceding memory material by pressing "j" (indicated by a green sticker with "yes" on the keyboard) or "k" (indicated by a red sticker with "no" on the keyboard). After the participants responded, the next trial followed automatically.

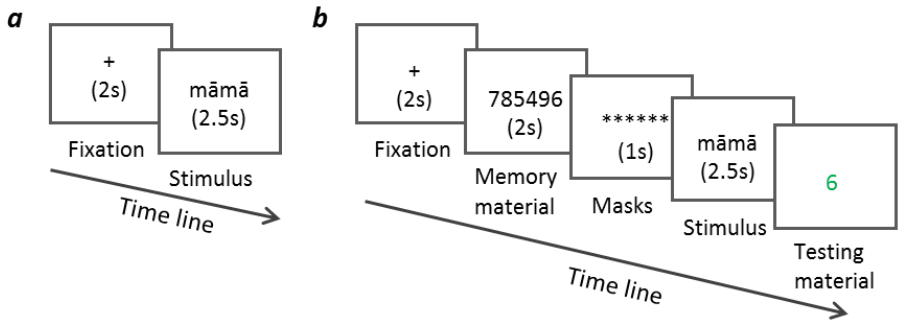

Figure 2.1. *The procedures of the no-cognitive-load condition (panel a) and high-cognitive-load condition (panel b).*

The digits in the memory set were selected from 1 to 8. Each digit was equally likely to appear in each position in the memory set of both conditions. The order of the two digits in the LCL condition was random, and the two digits for the same trial were always different from each other. The order of the six digits in the HCL condition was also random, under the condition that the same digit never appeared more than twice in a trial, and no more than two digits appeared in sequential order. For both conditions, the memory testing digit was equally likely to be present or absent in the memory material. If the digit was present in the memory material, it was equally likely to appear in any possible two or six positions in the memory sequences (see Figure 2.1).

Each condition consisted of four repetitions of the 16 disyllabic non-words. The 64 trials were presented as two blocks in different random orders to each participant. Furthermore, the order of the three conditions was also randomized across participants. In total, there were 192 trials (16 disyllabic combinations × 4 repetitions × 3 conditions) for each participant.

2.2.3 Pre-processing of the data

For all three participant groups, the coarticulation patterns in the LCL condition were not apparently different from the NCL control condition. So, only the results from the NCL control condition and the HCL condition are reported here. In the HCL condition, the error rate in the memory test was low across three groups (2.3% for NM; 7.5% for BL; 5.7% for AL). A total of 238 trials (out of 4,352) in which participants failed to respond accurately in the memory test were excluded. Furthermore, for both conditions, 34 trials were excluded in which participant had failed to read the stimulus within the assigned time window (2 trials for NM; 23 trials for BL; 12 trials for AL). In all, the recordings from 4,077 trials (out of 4,352) were included in the next step.

All 4,077 recordings were evaluated by a native Mandarin speaker in a tone identification task. For native Mandarin speakers, 2.9% trials could not be correctly identified. The production error rate for advanced learners was 7.7% and even higher for the beginning learners (13.8%). Only the recordings that were correctly identified by the native Mandarin listener (3,767 recorded disyllabic sequences) were used for the final f0 analysis.

2.2.4 F0 analysis

The boundaries of the vowels and nasal consonants were manually labelled in Praat (Boersma & Weenink, 2016) using a custom-written script (Chen, 2011).[3] The f0 extraction was also done in Praat (time step = 0.01s; pitch floor = 75 Hz). The f0 contours were obtained by taking 20 equidistant points for vowels, and 10 points for nasal consonants using the same custom-written script. To normalize the individual differences in f0 range, each participant's raw f0 data was transformed to participant-specific z-scores.

[3] Chen, Y. (2011). Generate norm F0.praat (praat script).

2.2.5 Statistical analysis

In order to give a more systematic and detailed report on coarticulatory effect, we examined the overall f0 height, the slope, as well as the steepness of the f0 contour of target tones.

Given the time-varying nature of the tonal contours, we adopted the growth curve analysis (GCA) (Mirman, 2014) with linear mixed model in R. GCA is a multi-level regression method using orthogonal polynomials to fit non-linear time course data. It is powerful in quantifying and analyzing the shapes of time course curves (e.g. f0 data). In the present study, second-order orthogonal polynomials were used with three parameters representing a curve's characteristics, that is, $y = a + bx + cx^2$. The intercept *a* refers to the overall mean of the f0 curves; the linear term *b* indicates the direction of the f0 curves (rise or fall); the quadratic term *c* refers to the steepness of the curvature. Two different f0 contours should expect at least one statistical significance among the three aspects.

Models were built for each target tone in both the first-syllable and second-syllable positions. The fixed effects consisted of the Linear Term, the Quadratic Term, and the experimental conditions, which include the Tonal Context (i.e. the preceding and following tones), Cognitive Load Condition (i.e. NCL and HCL), the Participant Group (i.e. NM, BL and AL), in addition to their interactions. Repetition and interaction between Repetition and Linear and Quadratic Terms were also included as fixed effects. For random effects, we had intercepts for Subjects, as well as by-Subject random slopes for the Cognitive Load Condition and Tonal Context.

Duration of the first and second vowel in the /mama/ sequence were also tested with linear mixed modeling. A model was built with Participant Group, Tone of the 1st syllable, Tone of the 2nd syllable, Cognitive Load, the interactions of these factors, and Repetition as fixed effects. Intercepts for Subjects was used as a random effect.

The significance of main effects in models of f0 contours and vowel durations were obtained via likelihood ratio comparisons with the change in log-likelihood distributed as χ^2. The degrees of freedom equaled the number of parameters added. The results of the main effects are presented in Appendices A1 and A2.

For both models of f0 contours and vowel duration, post-hoc comparisons were conducted using the *glht* function in the Multcomp package with Bonferroni adjustment in R (Hothorn, Bretz & Westgall, 2008). More specifically, for models of f0 contours, we compared the influence of each pair of contextual tones with contrastive offsets or onsets on the target tones for each participant group and cognitive load condition in the post-hoc comparison. Specifically, for the carryover effect, we compared all pairs of high-ending tones versus low-ending tones (T1 vs. T3, T1 vs. T4, T2 vs. T3 and T2 vs. T4). For the anticipatory effect, tonal pairs with contrastive onsets were compared (T1 vs. T2, T1 vs T3, T2 vs. T4 and T3 vs. T4).

2.3 Results

The fixed effect of Tonal Context was significant in all models of target tones in the first- and second-syllable positions (all p values < 0.05). The main effect of Cognitive Load was significant for all the models of tones in the first-syllable position (all p values

< 0.05), while for models of tones in the second syllable, this effect was only significant for T4 [$\chi^2(1) = 7.42$, $p < 0.05$]. The effect of Participant Group was found significant only for T2 and T3, both in the first syllable position [T2: $\chi^2(2) = 32.47$, $p < 0.001$; T3: $\chi^2(2) = 10.64$, $p < 0.01$] and in the second syllable position [$\chi^2(2) = 16.55$, $p < 0.01$; $\chi^2(2) = 15.08$, $p < 0.001$]. The three-way interaction of Tonal Context, Cognitive Load and Participant Group was significant in all models for target tones on both the first and the second syllable positions (all p values < 0.05) (see Appendix A1 for more detailed results).

For models of vowel duration in first- and second-syllable positions, the effects of Participant Group, Tone of the 1st Syllable, Tone of the 2nd Syllable and Cognitive Load were significant (all p values < 0.05). For the model of vowel duration in the first syllable, other significant effects were the interaction between Participant Group and Tone in the 1st Syllable [$\chi^2(6) = 249.96$, $p < 0.001$], the interaction between Participant Group and Tone in the 2nd Syllable [$\chi^2(6) = 27.05$, $p < 0.001$], as well as the three-way interaction of Participant Group, Tone in the 1st Syllable and Cognitive Load [$\chi^2(8) = 31.97$, $p < 0.001$]. For the model of vowel duration in the second syllable, significant interaction was found between Participant Group and Tone in the 2nd Syllable [$\chi^2(6) = 423.95$, $p < 0.001$].

In the following, we will present figures and statistical analyses of the carryover (§ 2.3.1) and anticipatory effects (§ 2.3.2). In both sections, the figures of tonal contours will be presented first. Subsequently, the post-hoc results of interaction of Participant Group, Tonal Context and Cognitive Load in each model will be presented for discussion of the fine-grained details in f0-contour. Finally, the results of the target tones' duration will be reported.

2.3.1 Carryover effect

2.3.1.1 Native Mandarin speakers

Figure 2.2 presents the carryover effect of the preceding tones on the contours of the following tone in /mama/ sequences produced by native Mandarin speakers without cognitive load. At the syllable boundary, the f0 onset of the second syllable was considerably influenced along the same direction by the offset of the first tone. Specifically, the high offset of the preceding tone (T1 and T2) raised the f0 of the initial nasal part of the following tone, and the low offsets (in T3 and T4) lowered the nasal part of the following tone. Furthermore, the influence of the preceding tone decreased over time: the f0 contour varied enormously during the initial nasal part; the f0 contours also differed at the beginning part of the vowel, and remained sizeable at the vowel offset for T1 and T3. In Xu (1997), the target non-words were produced in carrier sentences, therefore T3 in the second syllable was only realized as a low falling contour. Since the non-words were presented in isolation in our study, T3 in the second syllable showed a dipping contour. The general observations of carryover effect in the current study are similar to the results in Xu (1997).

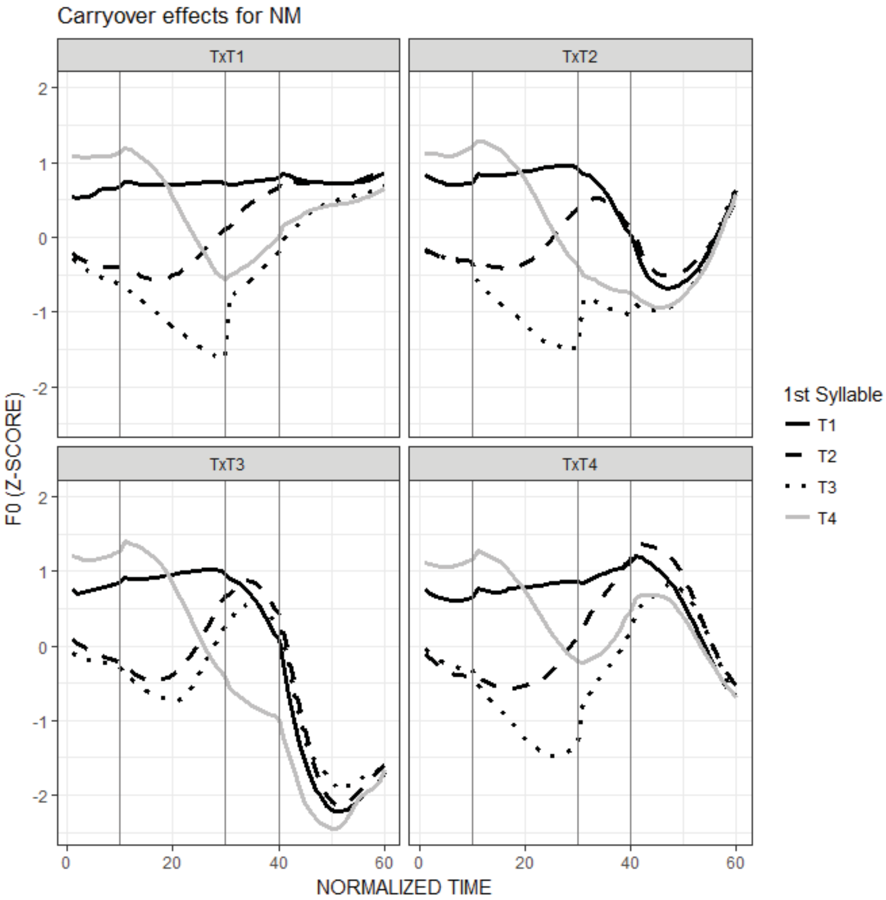

Figure 2.2. *Carryover effects for native Mandarin speakers in the NCL control condition. In each panel, the tone in the second syllable is held constant (T1-T4) and the tone in the initial syllable varies.*

To examine the effect of cognitive load on carryover coarticulation, we plotted in Figure 2.3 the f0 contours over the vowel portion of the second syllable, as a function of different preceding tones under different cognitive load conditions (left: the NCL condition and right: the HCL condition).

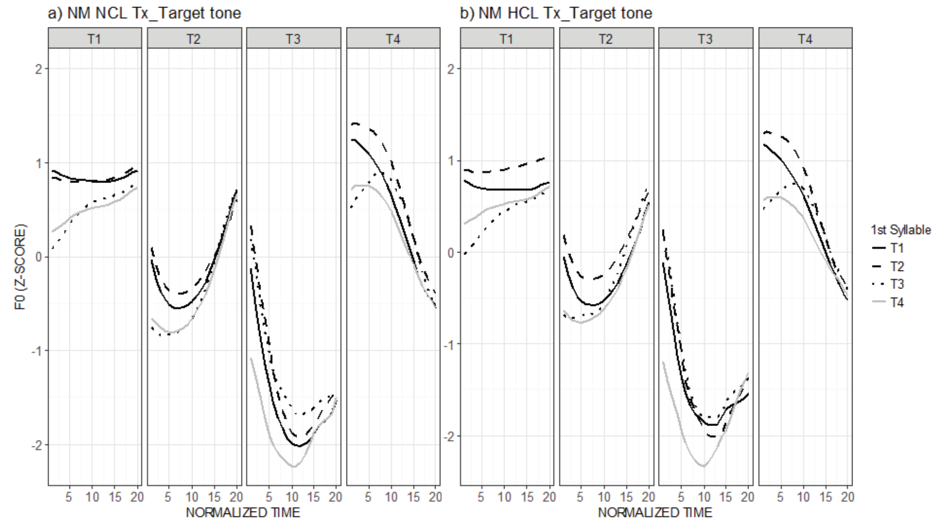

Figure 2.3. *F0 contours (over the vowel part) of the four target tones when preceded by different tones produced by native Mandarin speakers. Normalized f0 contours averaged across participants.*

Figure 2.3*a* shows that in the NCL condition, the whole contour of T1 was lowered by the low offset of the preceding T3 and T4, resulting in an initial rising contour. Both the overall f0 height and f0 slope of the second T1 were significantly different following tones with high offsets vs. low offsets, presenting an assimilatory pattern (Table 2.1*a*). Similar to T1, the initial part of T2 in the second syllable was also significantly affected by the offsets of the preceding tone in the first syllable in the NCL condition (Table 2.1*b*). The contour of T3 (Figure 2.3*a*) in the second syllable was significantly lower when following T4 than following T1 and T2. It should be noted that, according to the phonological rule, when T3 is followed by another T3, the first T3 is realized with a rising contour, similar to the lexical rising tone (T2) with high offset. So the contour of T3 did not show significant difference when following T3 vs. following T1 and T2 (Table 2.1*c*). T4 in the second syllable was affected by the preceding tone (Table 2.1*d*), showing significant difference in at least one parameter in all tone pairs with contrastive offsets (T1 vs. T3 and T4; T2 vs. T3 and T4). Overall, the statistical analysis of the carryover coarticulatory effect for native Mandarin speakers is in line with that in Xu (1997).

In the HCL condition, similar assimilatory carryover effect was found for all tones in the second syllable, as suggested by the similar significant effect of tonal context in the two conditions (Figure 2.3*b*, Table 2.1).

Table 2.1. *Pairwise comparison of results for adapted contours on the vowel part of the second tone due to preceding tones with high offsets (T1 and T2) vs. low offsets (T3 and T4) for native Mandarin speakers.*

***a.* T1 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.32	−4.15	<.001	−0.32	−4.16	<.001	−0.33	−4.22	<.001	−0.33	−4.24	<.001
slope	0.87	8.41	<.001	0.56	5.46	<.001	0.71	6.91	<.001	0.41	3.96	<.01
quadratic	0.07	−5.75	<.001	−0.22	−3.01	<.05	−0.38	−5.29	<.001		n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.30	−3.93	<.01			n.s.	−0.49	−6.31	<.001	−0.39	−5.02	<.001
slope	0.92	8.83	<.001	0.49	4.70	<.001	0.69	6.54	<.001		n.s.	
quadratic	−0.36	−4.86	<.001			n.s.	−0.31	−4.06	<.001		n.s.	

***b.* T2 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept			n.s.			n.s.	−0.26	−3.39	<.05	−0.28	−3.72	<.01
slope	0.88	4.45	<.001	0.62	3.15	<.05	1.16	5.85	<.001	0.90	4.56	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.		−0.26	−3.39	<.05	−0.39	−5.22	<.001
slope	1.03	5.18	<.001	0.61	3.11	<.05	1.09	5.47	<.001	0.68	3.42	<.05
quadratic	−0.32	−3.03	<.05		n.s.			n.s.			n.s.	

Table 2.1. *(Continued)*

***c.* T3 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.		−0.51	−4.52	<.001
slope		n.s.		1.20	3.43	<.05		n.s.		1.76	5.03	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		−0.43	−3.80	<.01		n.s.		−0.48	−4.27	<.001
slope		n.s.		1.47	4.19	<.001		n.s.		1.89	5.33	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	

***d.* T4 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.		−0.56	−4.78	<.001
slope	1.05	5.33	<.001	0.77	3.90	<.01	1.07	5.46	<.001	0.80	4.04	<.001
quadratic	−0.87	−6.91	<.001		n.s.		−0.56	−4.42	<.001		n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.		−0.51	−4.35	<.001
slope	1.14	5.74	<.001	0.89	4.44	<.001	1.28	6.38	<.001	1.02	5.09	<.001
quadratic	−0.57	−4.40	<.001		n.s.			n.s.			n.s.	

Table 2.2. *Mean duration (in ms) of second vowel in* /mama/ *sequence with four tones produced by native Mandarin speakers in the NCL and the HCL conditions.*

Cognitive load	Tone			
	T1	T2	T3	T4
NCL	251	272	273	162
HCL	233	267	264	152

In both NCL and HCL conditions, T4 in the second syllable was the shortest tone, significantly shorter than the other three tones (all p values < 0.01). T1 showed an intermediate duration, while T2 and T3 were the longest tones, significantly longer than T1 and T4 (all p values < 0.01). In the HCL condition, all tones exhibited a reduced duration. T1, T3 and T4 were significantly shorter than that in the NCL condition (all p values < 0.05). This result was different from that in Xu (1997), which found T3 the shortest one in the second syllable, and T1, T2 and T4 comparable in duration. Such divergence may be attributed to different recording procedures. In Xu (1997), the non-words were produced in carrier sentence, and the duration might be influenced by the context. In the current study, the non-words were recorded in isolation. So in the second syllable, the duration of four tones showed a pattern which was similar to that in monosyllabic production.

2.3.1.2 Beginning Dutch learners of Mandarin

Figure 2.4 presents the carryover effect of the preceding tones on the contours of the following tone in /mama/ sequences without cognitive load produced by beginning Dutch learners of Mandarin. For all tones, the general pattern was also assimilatory, but the magnitude was smaller compared to native speakers. For T1 and T4, the influence of the tonal context was shown on both nasal part and the vowel part. For T2 and T3, however, the assimilatory pattern was not apparent on the vowel part.

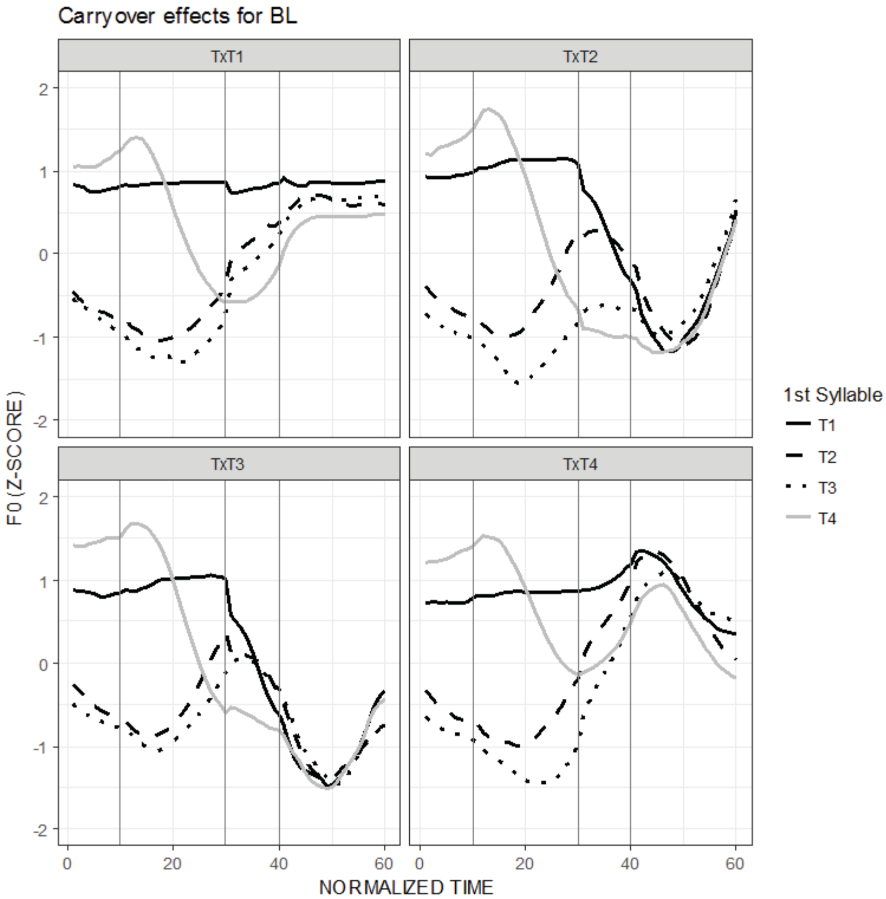

Figure 2.4. *Carryover effects for beginning Dutch learners of Mandarin in the NCL control condition. In each panel, the tone in the second syllable is held constant (T1-T4) and the tone in the initial syllable varies.*

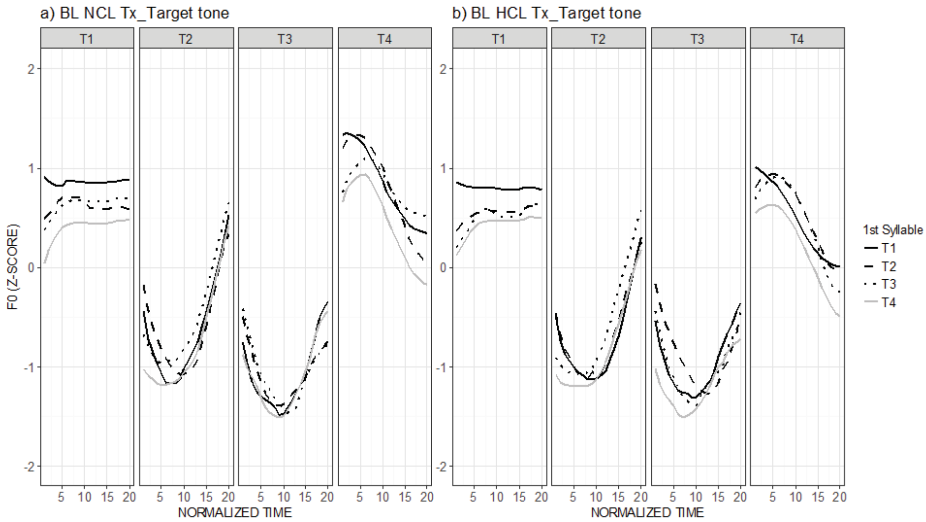

Figure 2.5. *F0 contours (over the vowel part) of the four target tones when preceded by different tones produced by beginning learners of Mandarin. Normalized f0 contours averaged across participants.*

In the NCL condition (Figure 2.5*a*), the f0 contour of target T1 was lower following T4 than the other three tones. Table 2.3*a* shows that overall f0 mean and quadratic term of T1 in the second syllable were significantly different following T1 and T4. As shown in the previous section, for Mandarin native speakers, the contours of target T1 on the second syllable clustered into two groups according to offsets of the initial tones. For beginning learners, this tendency was not clear. In the case of target T2 (Table 2.3*b*), the assimilatory effect on f0 height was not apparent. The influence of the preceding tone was mainly found in the timing of the turning point and sharpness of the contour. When following tones with low offsets (T3 and T4), target T2 was less concave and showed earlier turning point. Statistical significances in slope and sharpness were found when following T2 vs. T3 and T2 vs. T4. For target T3 in the NCL condition (Table 2.3*c*), the assimilatory pattern was not obvious. The contour of T4 (Table 2.3*d*) was influenced along the same direction by the offset of the first tone, showing significant difference in slope or steepness in three tone pairs with contrastive offsets (T1 vs. T3, T1 vs. T4 and T2 vs. T3).

The assimilatory influence from the preceding tone was slightly stronger for T1 in the HCL condition, with additional significant difference after T1 and T3. In case of T2, the influence from the initial tone was mainly on the slope and steepness of the contour, which was similar to that in the NCL condition. For T3, significant difference was found in slope following T2 vs. T4. In terms of T4, significant difference was only found for quadratic terms, showing a slightly weaker assimilatory pattern compared to the NCL condition.

Table 2.3. *Pairwise comparison of results for adapted contours on the vowel part of the second tone due to preceding tones with high offsets (T1 and T2) vs. low offsets (T3 and T4) for beginning learners.*

***a.* T1 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		−0.46	−5.46	<.001		n.s.			n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.		−0.28	−3.50	<.01		n.s.			n.s.	
HCL	**after T1 vs. after T3**			**after T1 vs. after T4**			**after T2 vs. after T3**			**after T2 vs. after T4**		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.35	−4.09	<.001	−0.44	−5.23	<.001		n.s.			n.s.	
slope	0.47	3.98	<.01	0.43	3.66	<.01		n.s.			n.s.	
quadratic		n.s.		−0.27	−3.16	<.05		n.s.			n.s.	

***b.* T2 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.			n.s.		1.01	4.48	<.001	1.15	5.27	<.001
quadratic		n.s.			n.s.		−0.37	−3.05	<.05	−0.51	−4.35	<.001
HCL	**after T1 vs. after T3**			**after T1 vs. after T4**			**after T2 vs. after T3**			**after T2 vs. after T4**		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.		0.67	3.02	<.05	0.78	3.31	<.05	0.82	3.68	<.01
quadratic		n.s.		−0.59	−4.72	<.001		n.s.		−0.46	−3.70	<.01

Table 2.3. *(Continued)*

***c.* T3 in the second syllable**												
NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.			n.s.			n.s.		1.30	3.19	<.05
quadratic		n.s.			n.s.			n.s.			n.s.	

***d.* T4 in the second syllable**												
NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope	0.80	3.56	<.01		n.s.		1.17	4.95	<.001		n.s.	
quadratic		n.s.		−0.53	−3.64	<.01		n.s.			n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic	−0.74	−4.66	<.001	−0.53	−3.53	<.01		n.s.			n.s.	

Table 2.4. *Mean duration (in ms) of second vowel in* /mama/ *sequence with four tones produced by beginning learners in the NCL and the HCL conditions.*

Cognitive load	Tone			
	T1	T2	T3	T4
NCL	305	368	386	177
HCL	293	348	370	172

For beginning learners, T4 was the shortest tone in both NCL and HCL conditions, significantly shorter than the other three tones (all p values < 0.01). T1 showed an intermediate duration, while T2 and T3 were significantly longer than the other two tones (all p values < 0.01). This result was similar to that of native speakers. In the HCL condition, all tones became shorter. T2 and T3 were significantly shorter in the HCL condition compared to the NCL condition.

2.3.1.3 Advanced learners of Mandarin

Figure 2.6 presents the carryover effect of the preceding tones on the contours of the following tone in /mama/ sequences produced by advanced Dutch learners of Mandarin without cognitive load. A similar assimilatory carryover effect was found, but the magnitude was smaller compared to native speakers.

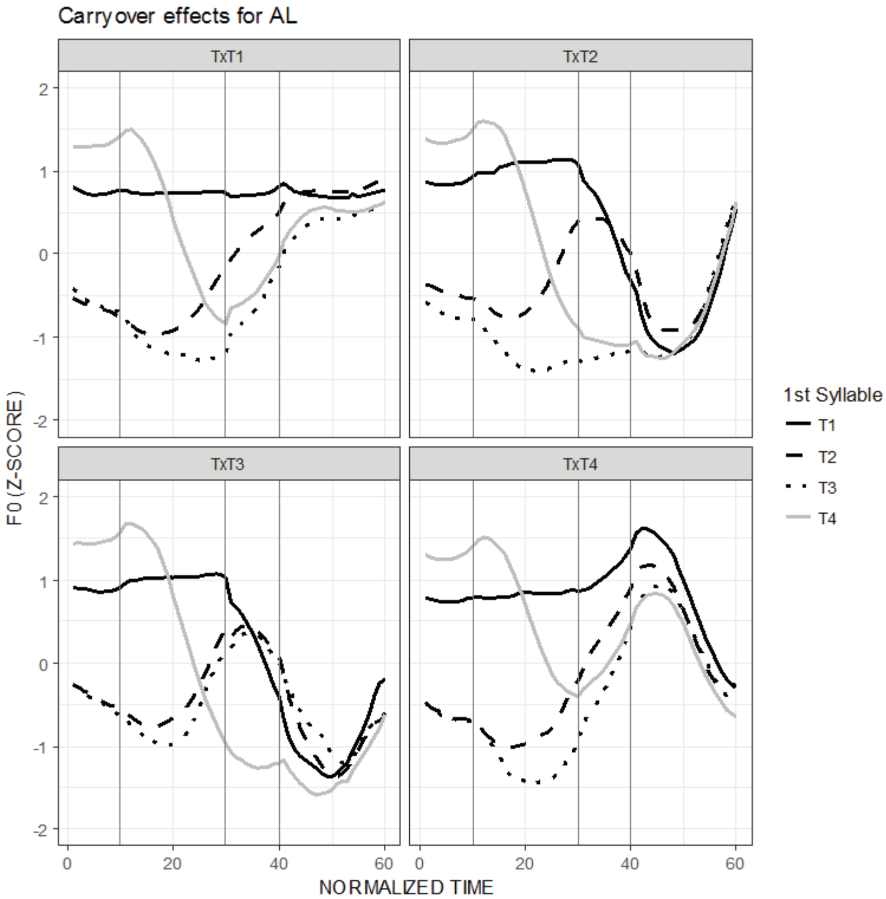

Figure 2.6. *Carryover effects for advanced Dutch learners of Mandarin in the NCL control condition. In each panel, the tone in the second syllable is held constant (T1-T4) and the tone in the initial syllable varies.*

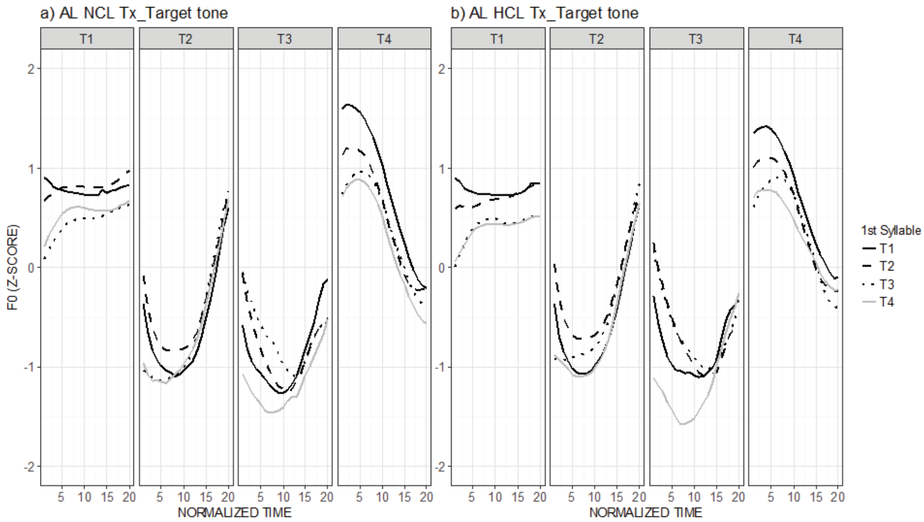

Figure 2.7. *F0 contours (over the vowel part) of the four target tones when preceded by different tones produced by advanced learners of Mandarin. Normalized f0 contours averaged across participants.*

In the NCL condition (Figure 2.7*a*), T1 showed divergent f0 contours following tones with contrastive offsets, presenting an assimilatory pattern. The T1 contours following all tone pairs with contrastive offsets showed statistical significance in at least one parameter (Table 2.5*a*), which was in line with the pattern of native speakers. T2 in the NCL condition also demonstrated assimilatory (Table 2.5*b*) carryover effect exerted by the offsets of the preceding tones, showing significant differences when following all tone pairs with contrastive offsets. A clear assimilatory tendency was also found for T3 (Figure 2.7*a*), however, the significant differences was only found for slope following T1 vs. T3 and T2 vs. T4 (Table 2.5*c*). The significant difference exerted by T1 vs. T3 in the first syllable was not expected. According to the tone sandhi rule, the offset of initial T3 was high when followed by another T3, and therefore T1 and T3 in the initial syllable (both with high offsets) should exert similar carryover effect on the following T3. In terms of T4, its contours was raised or lowered by the high or low offsets of the initial tones, showing significant difference when following all tone pairs with contrastive offsets except for T2 vs. T4 (Table 2.5*d*).

In the HCL condition, the assimilatory effect was maintained for target T1. For T2 in the HCL condition, the assimilatory effect became weaker. The assimilatory carryover effect was generally maintained in the HCL condition for T3 and T4.

Table 2.5. *Pairwise comparison of results for adapted contours on the vowel part of the second tone due to preceding tones with high offsets (T1 and T2) vs. low offsets (T3 and T4) for advanced learners.*

***a.* T1 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.32	−4.18	<.001	−0.23	−3.04	<.05	−0.30	−3.49	<.01		n.s.	
Slope	0.62	6.07	<.001	0.42	4.10	<.001	0.36	3.09	<.05		n.s.	
quadratic	−0.28	−5.29	<.001	−0.41	−5.71	<.001		n.s.		−0.26	−3.18	<.05
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.36	−4.59	<.001	−0.36	−4.69	<.001	−0.29	−3.41	<.05	−0.30	−3.50	<.01
slope	0.46	4.37	<.001	0.45	4.31	<.001		n.s.			n.s.	
quadratic	−0.54	−7.18	<.001	−0.45	−6.04	<.001	−0.35	−3.92	<.01		n.s.	

***b.* T2 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope	0.90	4.53	<.001	0.74	3.77	<.01	1.13	5.72	<.001	0.98	4.96	<.001
quadratic	−0.38	−3.63	<.01	−0.47	−4.51	<.001	−0.37	−3.51	<.01	−0.46	−4.40	<.001
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.		−0.33	−4.37	<.001
slope		n.s.			n.s.		0.88	4.41	<.001	0.85	4.26	<.001
quadratic	−0.46	−4.17	<.001		n.s.		−0.43	−3.91	<.01		n.s.	

Table 2.5. *(Continued)*

***c.* T3 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope	−1.39	−3.92	<.01		n.s.			n.s.		1.08	3.06	<.05
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		−0.36	−3.20	<.05		n.s.		−0.48	−4.16	<.001
slope		n.s.			n.s.			n.s.		2.09	5.75	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	

***d.* T4 in the second syllable**

NCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.42	−3.57	<.01	−0.53	−4.54	<.001		n.s.			n.s.	
slope	0.79	3.95	<.01	0.71	3.60	<.01		n.s.			n.s.	
quadratic		n.s.			n.s.		−0.40	−3.00	<.05		n.s.	
HCL	after T1 vs. after T3			after T1 vs. after T4			after T2 vs. after T3			after T2 vs. after T4		
	Est.	z	*P*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	−0.38	−3.19	<.05	−0.43	−3.61	<.01		n.s.			n.s.	
slope		n.s.		0.73	3.68	<.01		n.s.			n.s.	
quadratic	−0.55	−4.05	<.001		n.s.		−0.44	−3.02	<.05		n.s.	

Table 2.6. *Mean duration (in ms) of second vowel in* /mama/ *sequence with four tones produced by advanced learners in the NCL and the HCL conditions.*

Cognitive load	Tone			
	T1	T2	T3	T4
NCL	278	308	267	176
HCL	257	288	252	168

Different from native speakers, in both NCL and HCL conditions, T2 was the longest one, significantly longer than the other three tones (all p values < 0.01); T1 was the second longest tone, significantly longer than T3 and T4 (all p values < 0.01). T4 was the shortest one. In the HCL condition, all tones except T4 showed significantly shorter duration (all p values < 0.05).

2.3.1.4 Summary of carryover effect

For native Mandarin speakers, a strong and robust assimilatory carryover effect was found. The increase of cognitive load (Mattys & Wiget, 2011) is expected to introduce a reduction of processing resources for articulatory planning. However, the carryover effect was not obviously influenced by the high cognitive load, which indicated a lack in speech planning for this effect. This result was in line with previous findings (Franich, 2015; Whalen, 1990).

For beginning learners, the assimilatory carryover effect was found for T1, T2 and T4 in the NCL condition. The influence of cognitive load was weak with a slight increase in assimilatory effect for T1, T2 and T3, indicating the instability of this effect. Although these learners may be able to produce a lexical tone that fall into one of the four categories, correctly identified by native speakers, their assimilatory carryover effect was weaker than native speakers.

The advanced learners showed stronger carryover effect than beginners, exhibiting similar pattern with native Mandarin speakers. For T1, T2 and T4, strong assimilatory effect was found in the NCL condition. The assimilatory effect was weaker for T3. The carryover effect was generally maintained in the HCL condition.

2.3.2 Anticipatory effect

2.3.2.1 Native Mandarin speakers

Figure 2.8 presents variations in f0 contour of the first tones when followed by different tones in /mama/ sequences produced without cognitive load by native Mandarin speakers. Compared to the strong assimilatory effect exerted by the first ones on the second tones, the contours of the first tones showed less variability when followed by different tones. Moreover, the anticipatory was dissimilatory: tonal contours were higher when they were followed by tones with high onsets (T1 and T4) than followed by tones with low onsets (T2 and T3). The effect can be observed clearly for T1, T2 and T4 in the first syllable. For T1, the whole contour was raised by following tones with low onsets. For T2 and T4, this raising effect was strongest on the maximum f0 value in the tonal contours. The strongest effect was exerted by T3 in the following syllable, the maximum f0 of a tone preceding T3 was always higher than preceding other three tones. T2 in the following syllable also showed a similar raising effect for initial T1, T2 and T4. This general pattern was comparable to the findings of Xu's study (Xu, 1997), except that the whole contour of initial T1 in our study was clearly raised by tones with low onsets.

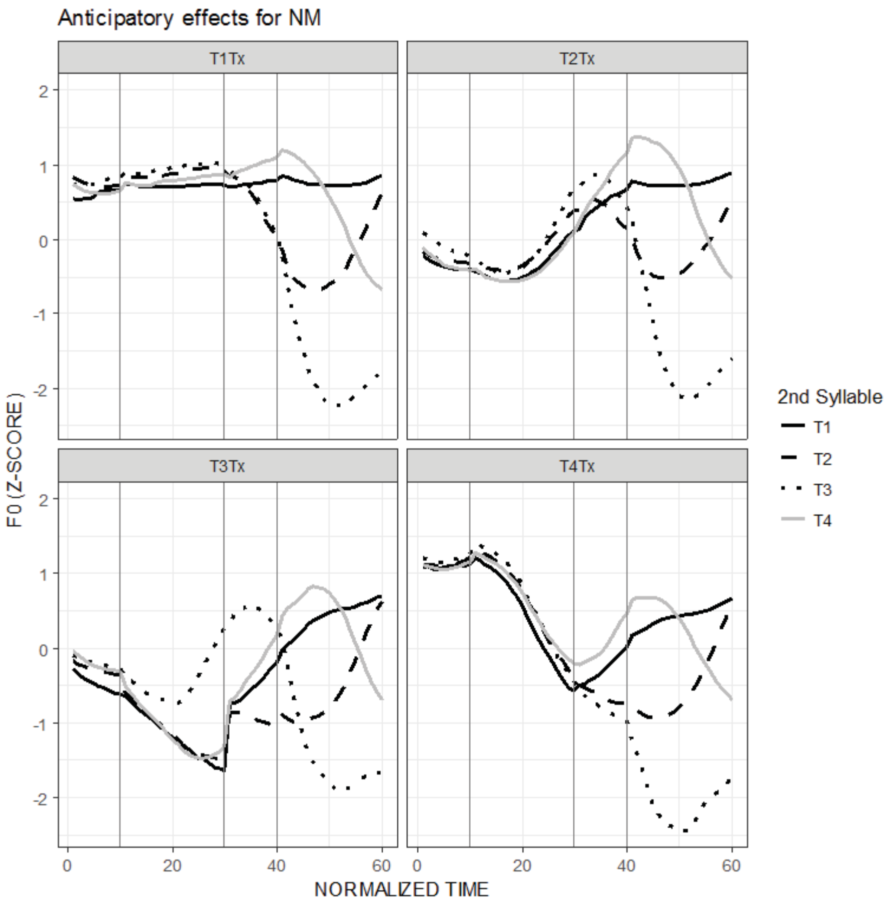

Figure 2.8. *Anticipatory effects for native Mandarin speakers in the NCL control condition. In each panel, the tone in the first syllable is held constant (T1-T4) and the tone in the second syllable varies.*

To examine the effect of cognitive load on anticipatory coarticulation, we plotted in Figure 2.9 the f0 contours over the vowel portion of the first syllable, as a function of different following tones under the NCL condition (panel *a*) and the HCL condition (panel *b*).

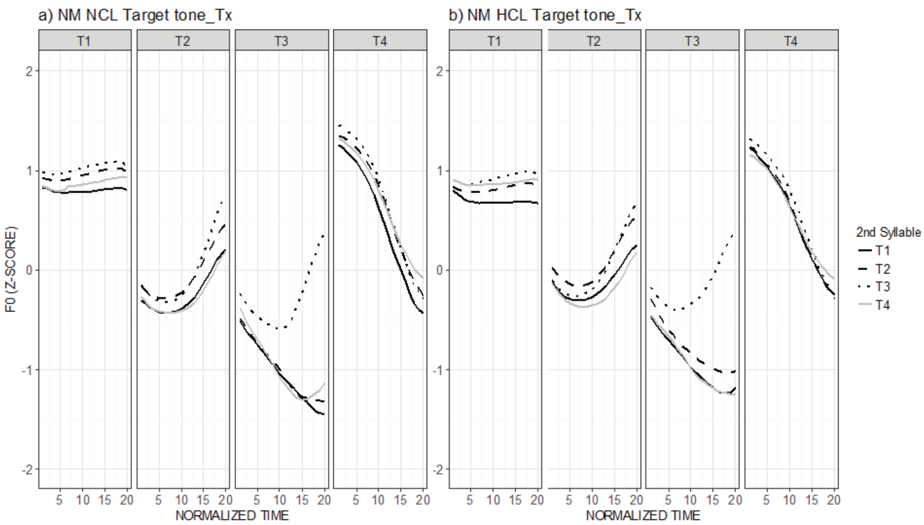

Figure 2.9. *F0 contours (over the vowel part) of the four target tones when followed by different tones produced by native Mandarin speakers. Normalized f0 contours averaged across participants.*

Figure 2.9*a* shows that in the NCL condition, the whole contour of T1 was raised by the low onsets of the following tones with significant differences when followed by all tone pairs with contrastive onsets except T2 vs. T4 (Table 2.7*a*). For T2, the whole contour was raised by the low onsets of the following tone. Significant difference was found in at least one parameter when T2 was followed by all tone pairs with contrastive onsets except T1 vs. T2 (Table 2.7*b*). According to the phonological rule, when T3 is followed by another T3, the first T3 is realized with a rising contour, similar to the lexical rising tone (T2). So the contour of T3 before another T3 shows different contour compared to T3 followed by the other tones. Other than this, the contours of T3 did not vary significantly when followed by tone pairs with contrastive onsets (when followed by T1 vs. T2 and followed by T2 vs. T4). This result was in line with the finding of Xu (1997), which also presented that the contour of T3 showed no statistical difference whether the following tone had a high offset of a low offset. The potential reason could be that the anticipatory raising effect mainly exerted on the maximum value of the initial tone. There was no high target in T3 therefore its contour was not sensitive to the anticipatory effect. In the case of T4, its contour was higher when followed by tones with low onsets than followed by tones with high onsets. The initial portion was raised the most, which also suggested that the anticipatory raising effect mainly affected the maximum f0 value. The comparison was significant when T4 was followed by T1 vs. T2 and T1 vs. T3 (Table 2.7*d*).

Table 2.7. *Pairwise comparison of results for adapted contours on the vowel part of the first tone due to following tones with high onsets (T1 and T4) vs. low onsets (T2 and T3) by native Mandarin speakers.*

***a.* T1 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	Z	*P*	Est.	z	*p*	Est.	z	*p*
intercept	0.16	3.10	<.05	0.23	4.39	<.001		n.s.		−0.16	−3.02	<.05
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	Z	*P*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		0.24	4.50	<.001		n.s.			n.s.	
slope		n.s.		0.25	3.41	<.05		n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	

***b.* T2 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	Z	*P*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		0.23	3.64	<.01	−0.20	−3.13	<.05	−0.25	−4.17	<.001
slope		n.s.		0.56	3.53	<.01		n.s.		−0.66	−4.18	<.001
quadratic		n.s.		0.31	4.82	<.001		n.s.		−0.26	−3.99	<.001
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	Z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.		−0.25	−3.91	<.01		n.s.	
slope		n.s.		0.60	3.74	<.01		n.s.		−0.75	−4.73	<.001
quadratic		n.s.		0.21	3.10	<.05		n.s.			n.s.	

Table 2.7. *(Continued)*

***c.* T3 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		0.71	7.14	<.001		n.s.		−0.66	−6.62	<.001
slope		n.s.		2.18	8.35	<.001		n.s.		−2.08	−8.17	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		0.77	7.77	<.001		n.s.		−0.83	−8.27	<.001
slope		n.s.		1.92	7.38	<.001		n.s.		−2.02	−7.68	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	

***d.* T4 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.19	3.00	<.05	0.23	3.69	<.01		n.s.			n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	

In the HCL condition (Figure 2.9*b*), the dissimilatory anticipatory effect decreased for T1, with a significant difference only before T1 vs. T3. For T2 and T3, the dissimilatory anticipatory effect was maintained compared to the NCL condition. The anticipatory effect decreased for T4, with no significant difference found before tone pairs with contrastive onsets.

Table 2.8. *Mean duration (in ms) of first vowel in* /mama/ *sequence with four tones produced by native Mandarin speakers in the NCL and the HCL conditions.*

Cognitive load	Tone			
	T1	T2	T3	T4
NCL	176	188	184	176
HCL	174	187	188	171

In the NCL condition T2 was significantly longer than T1 and T4 (all p values < 0.05). In the HCL condition, the duration of T2 and T3 was significantly longer than T1 and T4 (all p values < 0.01). Compared to the NCL condition, the duration of the four tones in the HCL condition did not change significantly.

2.3.2.2 Beginning Dutch learners of Mandarin

As plotted in Figure 2.10, the tonal contours in the first syllable in the NCL condition also varied when followed by different tones for beginning Dutch learners of Mandarin. For most tones, the general pattern of the anticipatory effect was also dissimilatory.

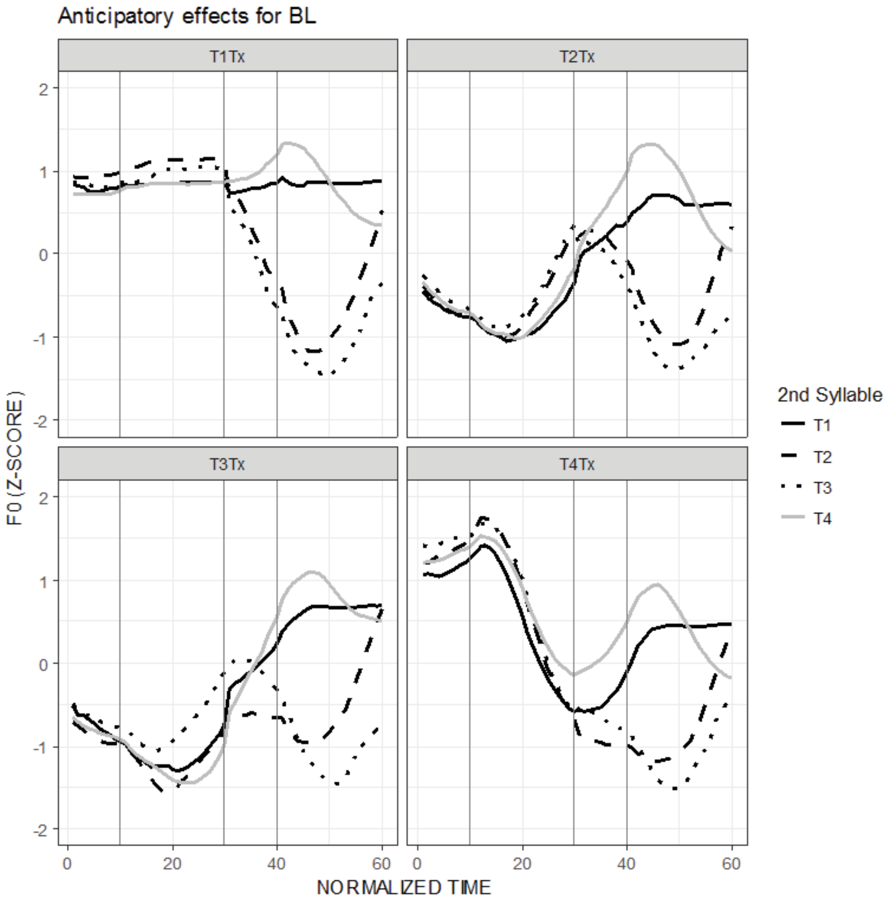

Figure 2.10. *Anticipatory effects for beginning learners of Mandarin in the NCL control condition. In each panel, the tone in the first syllable is held constant (T1-T4) and the tone in the second syllable varies.*

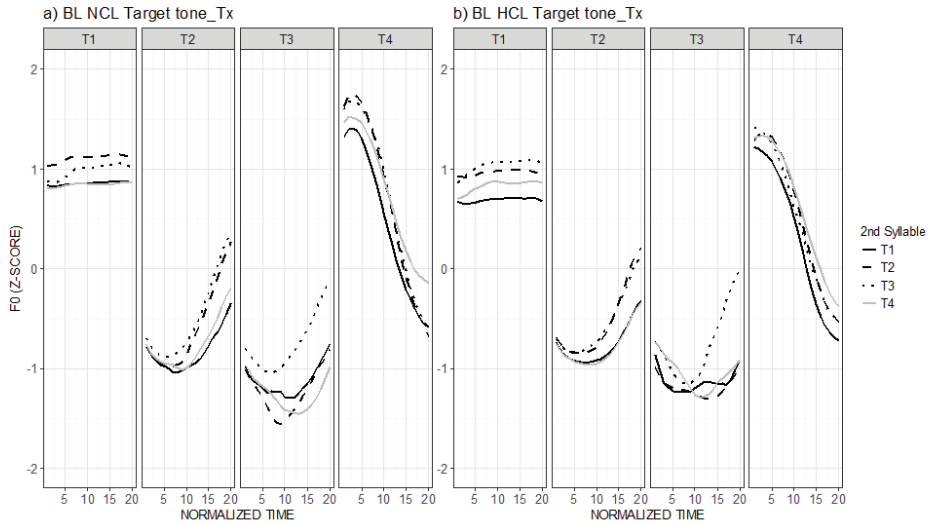

Figure 2.11. *F0 contours (over the vowel part) of the four target tones when followed by different tones produced by beginning learners of Mandarin. Normalized f0 contours averaged across participants.*

The contours of T1 were raised by the following tones with low onsets in the NCL condition (Figure 2.11*a*, Table 2.9*a*), showing significant difference before T1 vs. T2 and T2 vs. T4. T2 showed the strongest anticipatory effect among the four tones with significant differences when followed by all tone pairs with contrastive onsets. T3 in the NCL condition exhibited significantly different f0 contours when followed by T1 vs. T3 and T3 vs. T4 due to the phonological sandhi rule (see 2.3.1.1). The initial portion of T4 was raised by tones with low onsets, showing a dissimilatory pattern. The T4 contours also differed significantly followed by all tone pairs with contrastive onsets.

In the HCL condition, a similar pattern was found for T1, T2 and T3. The anticipatory effect became weaker in the HCL condition for T4, with only significant difference found in the overall f0 height when followed by T1 vs. T2.

Table 2.9. *Pairwise comparison of results for adapted contours on the vowel part of the first tone due to following tones with high onsets (T1 and T4) vs. low onsets (T2 and T3) by beginning learners.*

***a.* T1 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.27	4.57	<.001		n.s.		−0.26	−4.40	<.001		n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	**before T1 vs. before T2**			**before T1 vs. before T3**			**before T2 vs. before T4**			**before T3 vs. before T4**		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.24	4.06	<.001	0.31	5.30	<.001		n.s.		−0.19	−3.29	<.05
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	

***b.* T2 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.29	4.09	<.001	0.40	5.58	<.001		n.s.		−0.26	−3.63	<.01
slope	1.22	6.66	<.001	1.10	5.98	<.001	−0.90	−4.92	<.001	−0.78	−4.25	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	**before T1 vs. before T2**			**before T1 vs. before T3**			**before T2 vs. before T4**			**before T3 vs. before T4**		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.26	3.59	<.01	0.27	3.67	<.01	−0.22	−3.09	<.05	−0.23	−3.17	<.05
slope	1.03	5.37	<.001	0.94	4.90	<.001	−0.77	−4.20	<.001	−0.68	−3.72	<.01
quadratic		n.s.			n.s.			n.s.			n.s.	

Table 2.9. *(Continued)*

***c.* T3 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		0.35	3.22	<.05		n.s.		−0.41	−3.71	<.01
slope		n.s.			n.s.			n.s.		−1.01	−3.44	<.05
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept		n.s.		0.39	3.55	<.01		n.s.			n.s.	
slope		n.s.		1.04	3.52	<.01		n.s.		−1.11	−3.65	<.01
quadratic		n.s.			n.s.			n.s.			n.s.	

***d.* T4 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.02	2.98	0.05		n.s.			n.s.			n.s.	
slope		n.s.			n.s.		1.18	5.59	<.001	1.06	5.06	<.001
quadratic		n.s.		−0.34	−3.28	<.05		n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.21	3.08	<.05		n.s.			n.s.			n.s.	
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	

Table 2.10. *Mean duration (in ms) of first vowel in* /mama/ *sequence with four tones produced by beginning learners of Mandarin in the NCL and HCL conditions.*

Cognitive load	Tone			
	T1	T2	T3	T4
NCL	272	274	294	215
HCL	256	265	274	209

In both conditions, T3 was the longest tone, significantly longer than the other three tones (all *p* values < 0.01). T1 and T2 had similar intermediate duration, while T4 was the shortest one, significantly shorter than the other three tones (all *p* values < 0.01). In the HCL condition, the duration of T1, T3 and T4 was significantly shorter than the NCL condition.

2.3.2.3 Advanced learners of Mandarin

The varied tonal contours in the first syllable due to different following tones produced without cognitive load by advanced learners of Mandarin in the NCL condition is plotted in Figure 2.12. The contours of the target tones were high when followed by tones with low onsets, showing a dissimilatory anticipatory pattern comparable to that of native Mandarin speakers.

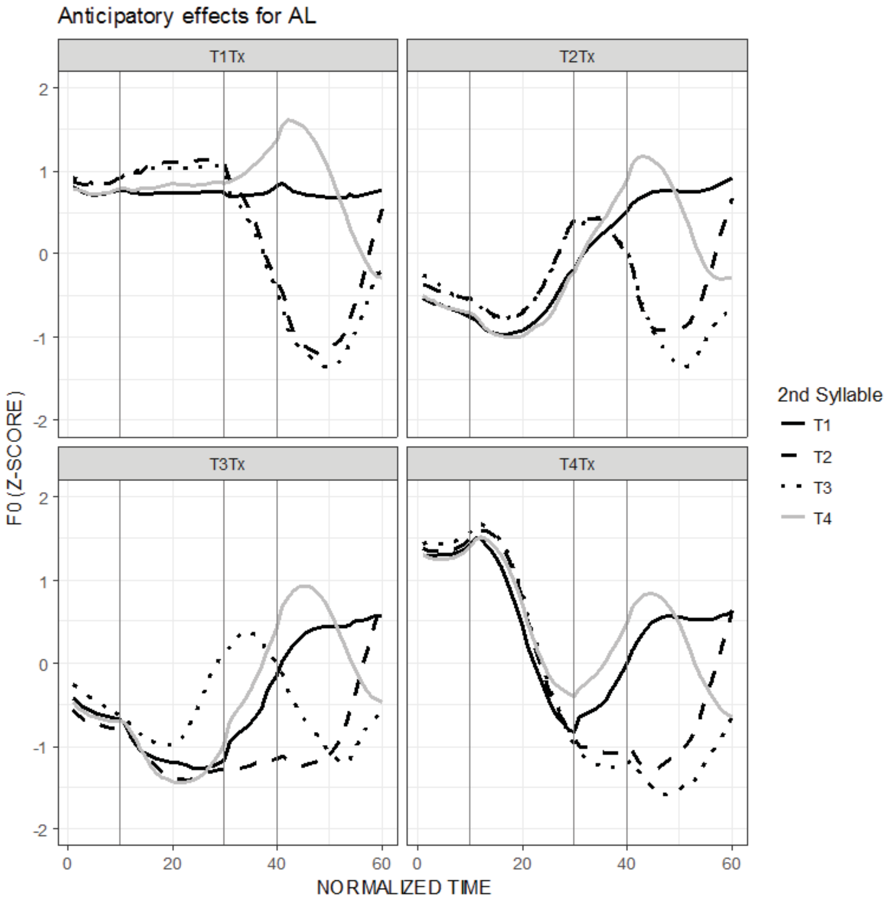

Figure 2.12. *Anticipatory effects for advanced Dutch learners of Mandarin in the NCL control condition. In each panel, the tone in the first syllable is held constant (T1-T4) and the tone in the second syllable varies.*

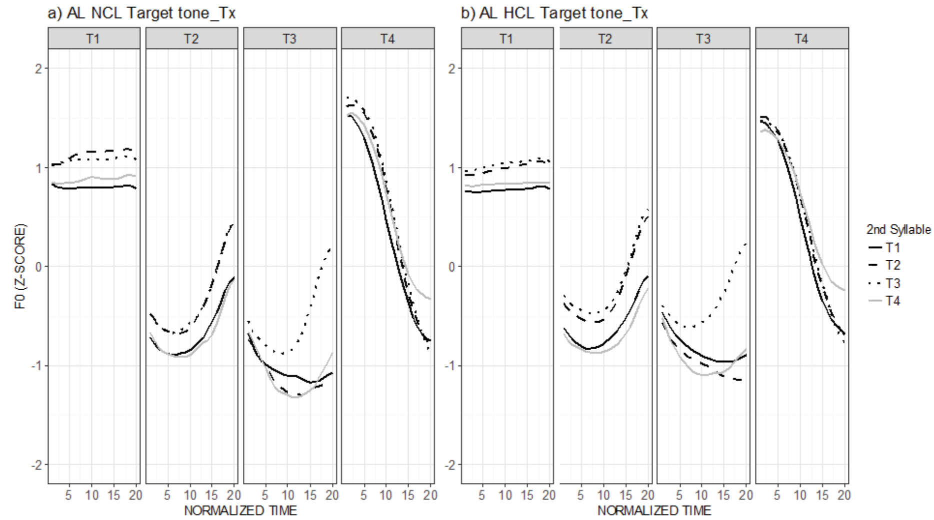

Figure 2.13. *F0 contours (over the vowel part) of the four target tones when followed by different tones produced by advanced learners of Mandarin. Normalized f0 contours averaged across participants.*

For advanced learners, the anticipatory effect for T1 in the NCL condition was strong (Figure 2.13*a*, Table 2.11*a*). When followed by tones with low onsets, the overall f0 of initial T1 was significantly higher than when it followed by tones with high onsets. Strong anticipatory effect was also found on initial T2. Contours of target T2 clustered into two groups according to the onsets of the second tone. T3 exhibited significantly different f0 contours when followed by T3 vs. T1 and T4 due to the phonological sandhi rule, which was similar to the performance of native Mandarin speakers. When followed by tones with contrastive onsets, the contours of T3 showed more variation than native Mandarin speakers, but the differences were not significant. The dissimilatory effect on initial T4 was also strong. The initial portion in the contours of the target tone was significantly raised by the low onsets of the following tones.

In the HCL condition, the strong effect on T1 and T2 maintained. Similar pattern was also found for T3 in the HCL condition. The effect also remained robust for T4, except for the insignificance when followed by T1 vs. T2.

Table 2.11. *Pairwise comparison of results for adapted contours on the vowel part of the first tone due to following tones with high onsets (T1 and T4) vs. low onsets (T2 and T3) by advanced learners.*

***a.* T1 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	P	Est.	z	*p*	Est.	z	*p*
intercept	0.33	6.25	<.001	0.28	5.26	<.001	−0.25	−4.74	<.001	−0.20	−3.75	<.01
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.24	4.56	<.001	0.28	5.17	<.001	−0.18	−3.38	<.05	−0.21	−3.97	<.01
slope		n.s.			n.s.			n.s.			n.s.	
quadratic		n.s.			n.s.			n.s.			n.s.	

***b.* T2 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.36	5.21	<.001	0.35	5.10	<.001	−0.38	−5.85	<.001	−0.37	−5.72	<.001
slope	0.65	3.70	<.01	0.74	4.17	<.001	−0.68	−4.17	<.001	−0.77	−4.68	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	*p*	Est.	z	*p*	Est.	z	*p*
intercept	0.35	5.09	<.001	0.44	6.26	<.001	−0.41	−6.31	<.001	−0.49	−7.55	<.001
slope	0.69	3.86	<.01	0.66	3.64	<.01	−0.81	−4.83	<.001	−0.78	−4.57	<.001
quadratic	0.24	3.05	<.05		n.s.		−0.27	−3.68	<.01	−0.24	−3.06	<.05

Table 2.11. *(Continued)*

***c.* T3 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	p	Est.	z	p	Est.	z	p
intercept		n.s.		0.48	4.78	<.001		n.s.		−0.53	−5.26	<.001
slope		n.s.		1.65	6.31	<.001		n.s.		−1.47	−5.58	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	p	Est.	z	p	Est.	z	p
intercept		n.s.		0.47	4.67	<.001		n.s.		−0.57	−5.70	<.001
slope		n.s.		1.47	5.59	<.001		n.s.		−1.20	−4.49	<.001
quadratic		n.s.			n.s.			n.s.			n.s.	

***d.* T4 in the first syllable**

NCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	p	Est.	z	p	Est.	z	p
intercept		n.s.		0.23	3.65	<.01		n.s.			n.s.	
slope		n.s.			n.s.		0.80	4.25	<.001	0.90	4.79	<.001
quadratic	−0.43	−4.62	<.001	−0.52	−5.53	<.001	0.36	3.88	<.01	0.45	4.79	<.001
HCL	before T1 vs. before T2			before T1 vs. before T3			before T2 vs. before T4			before T3 vs. before T4		
	Est.	z	P	Est.	z	p	Est.	z	p	Est.	z	p
intercept		n.s.			n.s.			n.s.			n.s.	
slope		n.s.			n.s.		0.89	4.70	<.001	0.88	4.63	<.001
quadratic		n.s.		−0.43	−4.48	<.001		n.s.			n.s.	

Table 2.12. *Mean duration (in ms) of first vowel in* /mama/ *sequence with four tones produced by advanced learners of Mandarin in the NCL and the HCL conditions.*

Cognitive load	Tone			
	T1	T2	T3	T4
NCL	215	218	207	190
HCL	195	198	192	179

In both cognitive-load conditions, T2 was the longest tone, significantly longer than T3 and T4 (all p values < 0.01). T3 was realized with an intermediate duration, while T4 was the shortest. In the HCL condition, the duration of all four tones became significantly shorter (all p values < 0.01).

2.3.2.4 Summary of anticipatory effect

Compared to the carryover effect, the dissimilatory anticipatory effect was found in T1, T2 and T4 with smaller magnitude for native Mandarin speakers. This effect decreased under the influence of cognitive load. Compared to native speakers, the magnitude of dissimilatory effect was smaller for beginning learners, which decreased in the HCL condition (especially for T4). The advanced learners exhibited a dissimilatory effect similar to native Mandarin speakers in the NCL condition which remained robust in the HCL condition.

2.4 Discussion

In this study, we reexamined the directionality, the nature and the magnitude of tonal coarticulation for native Mandarin speakers using disyllabic non-words following the design in Xu (1997), and more important, we tapped into the underlying mechanism and source of the tonal coarticulatory effects by introducing the effect of cognitive load. Since L2 acquisition has been remained less investigated, we further tested the beginning and advanced Dutch learners of Mandarin to reveal the developmental trajectory and mechanisms of tonal coarticulation that underlies the ultimate attainment of tonal acquisition in by non-tonal second language learners.

2.4.1 Tonal coarticulation for native Mandarin speakers

Our results showed that, for native speakers, tonal coarticulation was bidirectional, with both carryover and anticipatory effects. The carryover effect exerted by the offset of the initial tones exhibited an assimilatory nature, which replicated the findings in Xu (1997). In the current study we examined the variability in target tones exerted by all pairs of initial tones with contrastive offsets. The assimilatory effect was found for all four tones when preceded by all pairs of tones with contrastive offsets. Although, the carryover effect decreased over time, the influence can still be seen at least two third into the vowel. The anticipatory effect exerted by the following tones on the tones in the initial position showed a dissimilatory nature and had a smaller magnitude compared to the carryover effect. These findings were also in line with Xu (1997). For T1, T2 and T4, the dissimilatory effect was found when followed by most pairs of tones with contrastive onsets. It should be noted that, different form Xu's (1997) finding, the raising effect of the following low onsets were not constrained to the maximum f0. The whole contour of T1 and T2 were raised by following tones with low onsets (T2 and T3) in the current experiment. There was a lack of anticipatory effect for initial T3, except for the phonological change when it was followed by another T3.

We further investigated the effect of cognitive load on the tonal coarticulatory effects for native speakers of Mandarin. The carryover effect was robust and was not affected by high cognitive load. This is in line with previous findings that carryover effect does not involve advance planning (Whalen, 1990), thereby supporting the view that it is mainly caused by physiological constraints (Xu, 2011). The anticipatory effect decreased with high cognitive load, which was in contrast to the findings of Franich (2015). Significant dissimilatory anticipatory effect was found for all tones except T3 in the NCL condition, while in the HCL condition, this effect on T1 and T4 became weaker. This result lends support to the view that anticipatory coarticulation is planned, and diminishes under the influence of concurrent mnemonic processing. This finding can be potentially accounted for by the model proposed by Tilsen (2009, 2013), which argued that an inhibitory speech planning mechanism was used for contemporaneously planned articulatory targets to maintain and maximize the contrasts of different phonemes. In the NCL condition of the present study, the inhibitory mechanism functioned well and led to a clear dissimilatory anticipatory effect, maximizing the contrast of the adjacent tones. In the HCL condition, however, such inhibitory mechanism was constrained and resulted in a decreased dissimilatory anticipatory effect.

The overlap of different articulatory gestures, which happens in vowel and consonant coarticulation, is less feasible in tonal coarticulation. However, the involvement of cognitive planning in anticipatory tonal coarticulation found in the current study is compatible with the planned anticipatory effect found in vowel and consonant coarticulation (Katsika, Whalen, Tiede, & King, 2015; Whalen, 1990) and may reflect a characteristic of coarticulation in general.

2.4.2 Tonal coarticulation for Dutch learners of Mandarin

For both groups of learners, tonal coarticulation was also bi-directional. For beginning learners, assimilatory carryover effect was found for T1, T2 and T4 in the second syllable in the NCL condition. The magnitude of this effect was smaller compared to that of native Mandarin speakers. The influence of cognitive load was weak. The carryover effect for advanced learners was also assimilatory in nature and substantial for T1, T2 and T4, with magnitude similar to that of native speakers. Comparing the patterns of carryover effect of the beginning vs. advanced learners of Mandarin, we observed a clear developmental path. These interesting and new findings suggest that although the carryover effect does not include advance planning and is mainly bases on physiological constraints in articulation, its acquisition is still a gradual learning process. Fine-tuned motor skills are required for native-like production of tonal sequences.

Different from the carryover effect, the anticipatory coarticulation was strong for Dutch learners of Mandarin. For beginning learners, the dissimilatory anticipatory effect was found on T1, T2 and T4 in the first syllable. This effect was robust for T1 and T2 with high cognitive load. Our advanced learners also showed anticipatory effect in the NCL condition like that of native Mandarin speakers. For T1, T2 and T4, the dissimilatory effect was found when followed by all pairs of tones with contrastive onsets. This strong effect could be seen for the whole contour for T1 and T2, and for the maximum f0 value for T4. It did not show obvious decrease for all tones under the influence of high cognitive load for advanced learners, showing a more robust pattern than native Mandarin speakers. That is, as suggested by Tilsen's model, the inhibitory mechanism was acquired by beginning and advanced learners as an effective way to maintain the contrast and ensure the perceptibility of different tonal categories in sequence.

Brengelmann, Cangemi and Grice (2015) reported results which revealed greater variability in L2 production in the final portion of the initial tone in disyllabic sequences than native speakers. More specifically, German learners of Mandarin were more likely to produce anticipatory coarticulation. However, only the final part of the initial tone was examined in that study and therefore it is not clear whether there is an anticipatory effect on the whole contour or maximum f0 value of the initial tone for German learners of Mandarin. More studies examining learners with different L1s are needed to shed light on the general pattern and language-specific characteristics among learners of non-tonal languages.

2.4.3 Conclusion

The results of the current study show that for native Mandarin speakers, the carryover effect in tonal coarticulation was assimilatory in nature and did not involve speech planning. The anticipatory coarticulation, on the other hand, was dissimilatory in nature and was planned. The carryover effect could be acquired gradually by L2 learners, suggested by a developmental path found in beginning and advanced Dutch learners of Mandarin. The anticipatory effect was strong for both beginning and advanced learners. The advanced learners showed a more robust anticipatory effect compared to native Mandarin speakers, since for them, this effect was not reduced in high cognitive load condition. The anticipatory effect was adopted by L2 learners as an effective way in maintaining and maximizing contrast of tonal categories.