TONAL COARTICULATION IN L2 STANDARD CHINESE

Tingting Brengelmann, Francesco Cangemi, Martine Grice

I*f*L-Phonetik, University of Cologne
wut@smail.uni-koeln.de, fcangemi@uni-koeln.de, martine.grice@uni-koeln.de

ABSTRACT

In this paper we investigated tonal coarticulation in German learners' production of Standard Chinese, as compared to the production of native speakers. Examination of the f0 contour at the end of target syllables revealed that when syllables were uttered in isolation, there was more variability across productions for the learners than for the native speakers. Moreover, this variability increased when the target syllable was followed by a further syllable to form a disyllabic word. Thus, German learners exhibit not only more variability in their production but also more anticipatory coarticulation than their native counterparts.

Keywords: variability, tonal coarticulation, L2 production, anticipatory coarticulation

1. INTRODUCTION

In East Asian tone languages, the phonetic characteristics of tones are influenced considerably by adjacent tones [1][3][6][8][11][13][15][19][20], although the directionality, magnitude and temporal extent of tonal coarticulation vary across languages. Specifically, it has been argued for Standard Chinese that coarticulation is predominantly progressive [20], i.e. a tone is affected by the preceding tone, regressive coarticulation, also referred to as anticipatory coarticulation, being more subtle. This means that in native production the f0 of the initial syllable in a word is relatively similar to the production of this syllable in a monosyllabic word.

Diverse studies on the acquisition of lexical tones in Standard Chinese have reported that advanced English learners, who produced lexical tones in isolation with no difficulty, appeared to have considerable problems producing lexical tones in sequence [4][5][12][17][18][22] e.g. in disyllabic words (the most common structure of words in Standard Chinese). This raises the question as to whether these difficulties are due to deviant tonal coarticulation patterns, specifically with regard to anticipatory coarticulation.

In this study we investigate the f0 in the last 20% of the initial syllable in disyllabic words comprising combinations of all four lexical tones. We compare these productions to productions of the same syllables in monosyllabic words. We hypothesize that the f0 values towards the end of the initial syllable in disyllabic words will be more variable in learners' production than in native production (with the exception of 3rd tone sandhi), indicating that there is more anticipatory coarticulation in the productions of German learners than in the productions of the native speakers.

2. METHOD

2.1. Speech Material

Participants were asked to read aloud three repetitions of target syllables: *wēi, wéi, wěi, wèi, māo, máo, mǎo and mào* in two conditions: (a) as monosyllabic words and (b) as the initial syllable in 16 combinations of two syllables and four tones in disyllabic nonce words of the structure *weimao*, for a total of 24 items × 3 repetitions × 15 speakers =1080 data points.

2.2. Subjects

Five native speakers of Standard Chinese and ten German learners participated in this study. The native speakers were born in Beijing and grew up there. German learners had all been studying Standard Chinese at the University of Cologne for between four and six semesters and had spent a period in China of between one and twelve months.

2.3. Measurements and Analysis

We explored variability in the last portion of the tone contour. The extraction procedure employed, developed by Xu [21], yielded ten equidistant points for each contour and normalization in the frequency domain [14]. We focussed on the average values of the last two points (the last 20%), as a compromise between the potential unreliability of using the final point only, and the risk of capturing variability too far away from the syllable boundary and thus potentially due to factors other than regressive coarticulation (last three points).

3. RESULTS

3.1. F0 contours

For the purposes of illustration Figure 1 shows normalized f0 contours of three repetitions of target syllables for one native speaker (dotted lines) and one learner (solid lines). They are produced as monosyllabic words (left column) and as the initial syllable in disyllabic words (right column). Rows represent productions of Tones 1 to 4.

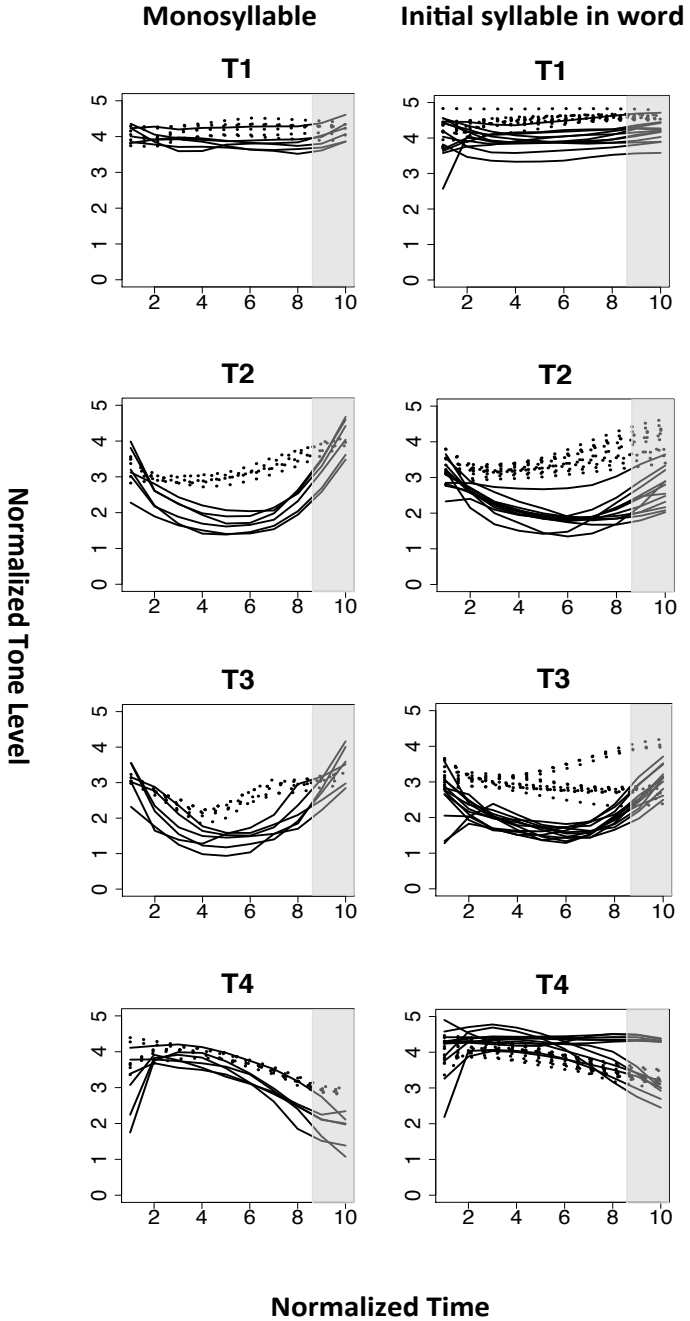

Figure1: F0 trajectory of three repetitions of target syllables of one speaker from each group as monosyllables (left column) and as the initial syllable in disyllables (right column). The grey bars represent the final portion of the f0 contours that was used for the quantitative analysis

For Tone 1, we can see that for this German learner there is more variation in the overall height in the disyllabic condition than in the monosyllabic condition. For this tone, the contours produced by learners and natives did not differ considerably. For Tone 2 (second row), variation in the learner's production was not restricted to the final part of the contour. Looking at the region halfway into the syllable in the disyllabic condition (right column) we can detect a great amount of variation but not the clear low characteristic of Tone 2 and visible in the monosyllabic condition [20]. The picture for Tone 3 is more complex. In the monosyllabic condition, this native speaker shows different patterns. In the disyllabic condition, we can see that this native speaker produce two general patterns, rising when followed by another Tone 3 and low when followed by other tones. This is due to the tone sandhi rule, according to which Tone 3 resembles Tone 2 when followed by another Tone 3. This is not the case in the learner's production. Instead, we cannot see a great amount of variation in comparison to the monosyllabic condition. For Tone 4 the patterns look similar across the two conditions for both speakers, although learner's productions are comparatively more variable than those of the native speaker, especially in the disyllabic condition towards the end of the target syllable.

Generally speaking, comparing the f0 contours in the left and right columns of Figure 1, we can observe that f0 in the final portion of the contours in the German production (solid lines) shows greater variation in the disyllabic condition (right column) than in the monosyllabic condition (left column).

3.2. Statistical Analyses

Figure 2 shows the amount of within speaker Variability in f0 in the final portion of the target syllable (y-axis) in the production of each Tone (four panels), split by Group (lines) and Condition (as monosyllabic words and disyllabic words, x-axis). Variability is expressed as the average of within-speaker standard deviations, as calculated over the final 20% of the f0 contour. For Tone 1, Tone 2 and Tone 4, standard deviations in the disyllabic condition are averaged across the four upcoming tones. For Tone 3 they were first calculated separately for the sandhi and non-sandhi conditions, then averaged. Unsurprisingly, the trends in Figure 2 show that productions of learners are overall more variable than those of native speakers, for all four tones and in both conditions. Crucially, the difference in variability between learners and natives appears larger in the disyllabic condition, most notably for Tone 1 and

Tone 4. Such a trend is consistent with our hypothesis, according to which learners are more prone to anticipatory tonal coarticulation than native speakers.

Only for Tone 2 is there an increase in variability from the monosyllabic condition to the disyllabic condition for both natives and learners. This is in line with the results reported in [20] on variation in Tone 2.

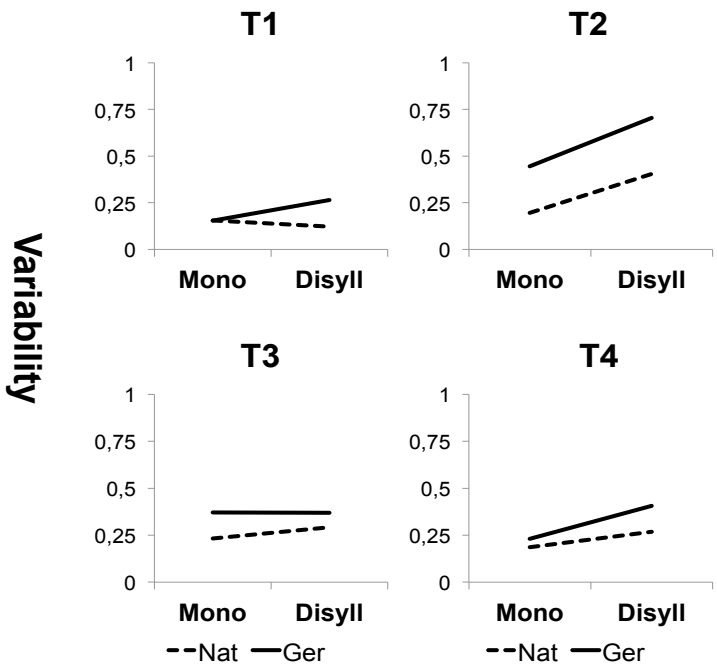

Figure 2: Variability in the final portion of the f0 contour. Averaged within-speaker standard deviations (y-axis), split by Tone (panels), Group (lines: dashed for natives vs. solid for German learners) and Condition (x-axis: as monosyllabic word vs. as first syllable of disyllabic word).

Given the unbalanced number of subjects in the two Groups (5 native speakers vs. 10 German learners), we refrain from a quantitative evaluation of the impact on variability of the interaction between Group and Condition. In line with the preliminary results of [23], however, the final portion of the contours is indeed affected by the Group:Condition interaction. The average between the two final f0 values of each contour was predicted with a mixed-effects model featuring TONE {1,2,3,4},GROUP {native, learner}, CONDITION {monosyllabic, disyllabic} and their two- and three-way interactions as fixed factors, and intercepts for SPEAKER {1,...15} and REPETITION {1,2,3} as random effects. Likelihood Ratio Tests did not reach significance when comparing the full model to a reduced model without the three-way interaction ($\chi^2(1)=1.09$, $p=0.77$), but reached significance when the crucial Group:Condition interaction was dropped ($\chi^2(1)=25.18$, $p<0.0001$). Taken together, the results can be interpreted as suggesting that the differences between the production of tones by learners and native speakers are stronger when an upcoming tone is present, a finding that is consistent with the hypothesis that learners are more likely than native speakers to produce anticipatory tonal coarticulation.

4. DISCUSION & CONCLUSION

For native production it has been reported that the anticipatory effect from the upcoming tone can be seen in the f0 maximum of the preceding tone and not in its final portion, if there is an anticipatory effect at all [20]. This would mean that the most visible anticipatory effect would not be seen in the final portion of the f0 contour (except for Tone 2), as Tone 2 is a rising tone and its maximum is near the syllable boundary.

Our data, however, showed great variability in L2 production in the last 20% of the preceding syllable. If there was an anticipatory effect, there were different patterns used by learners than by native speakers. A possible motivation for this difference is that the location of an accentual H peak in German is highly variable and subject to influences such as proximity to other tones and prosodic edges. Furthermore, H peaks associated with a stressed syllable can occur outside this syllable, either before or after it, depending on the phonological specification (i.e. which pitch accent it belongs to) [7][9][10]. Such variation is possible in German, as, unlike Chinese, it has a sparse tonal specification, allowing tones to spread onto adjacent syllables. From their L1 background, tonal realization of German learners is thus more likely to be affected by tonal coarticulation than native speakers.

Our findings suggest that the presence of an upcoming tone has a different impact on the final portion of the f0 contours in natives' and learners' productions. Specifically, upcoming tones seem to induce more variability in the final portion of L2 productions, and thus more anticipatory tonal coarticulation. A number of studies have, in fact, reported the finding that learners have difficulties in producing tones in connected speech. This paper explicitly links these difficulties to divergence in tonal coarticulation patterns between learners and native speakers. It is expected that the analysis of differential variability in production proposed here might also prove useful in the exploration of other aspects of tonal coarticulation in L2 production, such as the comparison between learners and native speakers with respect to carry-over effects, i.e. progressive coarticulation.

5. REFERENCES

[1] Abramson, A. (1979). The coarticulation of tones: An acoustic study of Thai. In T.L. Thongkum, P. Kullavanijaya, V. Panupong and T. L. Tingsabadh (eds.), *Studies in Tai and Mon-Khmer Phonetics and Phonology. In Honour of Eugenie J. A. Henderson*. 1-9. Bangkok, Chulalongkorn University Press.

[2] Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3.48) [Computer software]. Retrieved from http://www.praat.org.

[3] Brunelle, M. (2009) Northern and Southern Vietnamese Tone Coarticulation: A Comparative Case Study, *Journal of the Southeast Asian Linguistics Society*, 1: 49-62.

[4] Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *Journal of Chinese Language Teachers Association*, 32, 1, 21-39.

[5] Chen, Q. (2000). *An Analysis of Mandarin Tonal Errors in Connected Speech by English - Speaking American Adult Learners: A Study at and above the Word Level*. Ph.D dissertation. Brigham Young University.

[6] Gandour, J., S. Potisuk and S. Dechongkit (1994). Tonal Coarticulation in Thai. *Journal of Phonetics* 22(4): 477-492.

[7] Grice, M., S. Baumann & R. Benzmüller (2005). German Intonation in Autosegmental-Metrical Phonology. In: Jun, Sun-Ah (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press.

[8] Han, M. and K.-O. Kim (1974). Phonetic variation of Vietnamese tones in disyllabic utterances. *Journal of Phonetics* 2: 223-232.

[9] Ladd, D.R. (2008). *Intonational Phonology* (second edition). Cambridge: Cambridge University Press.

[10] Niemann, H., M. Grice & D. Mücke (2014). Segmental and positional effects in tonal alignment: An articulatory approach. Talk at International Seminar on Speech Production, 5-8 May, Cologne, Germany

[11] Peng, S.-h. (1997). Production and perception in Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics* 25: 371-400.

[12] Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association*, 24, 3, 27-47.

[13] Shen, X. S. (1990). Tonal Coarticulation in Mandarin. *Journal of Phonetics* 18(2): 281- 295.

[14] Shi, F. (1986). The bi-syllabic tone in Tianjin dialect. *Language Study,* 1, 71-90

[15] Shih, C. (1988). Tone and Intonation in Mandarin. *Working Papers of the Cornell Linguistics Laboratory* 3: 83-109.

[16] Shih C. & H. D. Lu (2010). Prosody Transfer and Suppression: Stages of Tone Acquisition. *Proceedings of Speech Prosody*, May 11-14, Chicago.

[17] Sun, S. H. (1998). *The Development of a Lexical Tone Phonology in American Adult Learners of Standard Mandarin Chinese*. University of Hawaii Press.

[18] White, C. (1981). Tonal pronunciation errors and interference from English intonation. *Journal of Chinese Language Teachers Association*, 16, 2, 27-56.

[19] Xu, Y. (1994). Production and Perception of Coarticulated Tones. *Journal of the Acoustical Society of America* 95(4): 2240-2253.

[20] Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics* 25: 61-83.

[21] Xu, Y. (2013). ProsodyPro - A Tool for Large-scale Systematic Prosody Analysis. In Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France. 7-10.

[22] Yang, C. (2011). *The Acquisition of Mandarin Prosody by American Learners of Chineseas a Foreign Language (CFL)*. PhD. diss. Ohio University

[23] Brengelmann, T., F. Cangemi & M. Grice (accepted). Tonal coarticulation in German learners of Standard Chinese. Poster Präsentation . Phonetics and Phonology in Europe 2015 (PaPE), Cambridge, UK.