Tonal coarticulation in German learners of Standard Chinese
Tingting Brengelmann, Francesco Cangemi, Martine Grice
I/L-Phonetik, University of Cologne
wingting.brengelmann@uni-koeln.de, fcangemi@uni-koeln.de, martine.grice@uni-koeln.de

In East Asian tone languages, tonal realisation is considerably influenced by adjacent tones
(e.g. Han & Kim, 1974 and Brunelle, 2009 on Vietnamese; Abramson, 1979 and Gandour et
al. 1994 on Thai; Peng, 1997 on Taiwanese; Shih, 1988, Shen, 1990 and Xu, 1994 & 1997 on
Standard Chinese), although the directionality, magnitude and temporal extent of tonal
coarticulation vary across languages. Specifically, it has been found for Standard Chinese
(Xu, 1997) that tonal coarticulation is predominantly progressive and assimilatory. This
means that in native production the F0 of the initial syllable in a disyllabic word is largely
unaffected. A number of studies have reported that advanced English learners of Standard
Chinese who are able to produce lexical tones accurately in citation form have considerable
difficulties producing these tones in sequence, e.g. in disyllables (Chen, 2000; Shih, 2010;
Yang, 2011, among others). This opens up the question as to whether these difficulties are due
to deviant tonal coarticulation patterns. Deviation might arise for a variety of reasons, e.g.
because L1 has regressive coarticulation, or because L1 has a sparse tonal specification,
allowing tones to spread onto adjacent syllables. In German, for example, tonal crowding
triggered by right-edge boundary tones leads to the shifting of pitch events to the left. For
German learners, then, the final portion of the f0 contour of a given tone may be affected by
the presence of a following tone. This is not expected to be the case for native speakers.

Five native speakers and ten German learners read aloud three repetitions of 8 syllables, wēi,
wēi, wēi, māo, māo, māo and māo, combining the segmental string weimao and the four
tones in two conditions: (a) syllable in isolation and (b) first syllable of a disyllabic compound
word (24 items × 3 repetitions × 15 speakers = 1080 data points). F0 values were extracted at
10 equidistant points for each syllable. The mean f0 value of the last 20% of each target
syllable 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 (χ2(1)=1.09, p=0.77), but reached
significance when the crucial Group:Condition interaction was dropped (χ2(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. This is
exemplified for tone 2 in Figure 1, in which f0 contours are plotted in normalized time and
height, separately for GROUP (native: dotted lines, L2: solid lines) and CONDITION
(monosyllabic: left, disyllabic: right).

These results provide evidence for regressive coarticulation in L2 production, a pattern that is
very different from the progressive coarticulation reported for native speakers. Whether this
coarticulation is triggered by the presence of regressive assimilation patterns or sparse tonal
specification in L1 (German) remains to be corroborated.
Fig.1 F0 contours of one native speaker (dotted lines) and one German learner (solid lines), the grey bars represent the final portion of the f0 contours that was used for the quantitative analysis. All three repetitions in both conditions are plotted.

References: