Phonetic cue-weighting in the production of Mandarin rising [T2] and low [T3] tones by Japanese learners

Yufei Niu[1] & Ricky Chan[2]

[12]The University of Hong Kong (HK)
yufeiniu@connect.hku.hk, rickykwc@hku.hk

L2 tone acquisition can be challenging and previous research has shown that T2/T3 confusion is common for L2 Mandarin learners [1]. For native Mandarin speakers, in addition to the primary F0 cue, voice quality (T3 allophonic creakiness) and duration can serve as important secondary cues in distinguishing T2 and T3 [2, 3]. However, the relative importance of these cues for L2 learners remains unclear. Similar to Mandarin, Japanese contrasts some words with F0 cues and it is interesting to determine if Japanese learners of Mandarin will use similar acoustic cues in lexical tone production. We also explore whether there is a shift towards a more native-like cue weighting pattern with increasing Mandarin proficiency.

We recruited two groups of Japanese learners of Mandarin (JAPH: high Mandarin proficiency and JAPL: low Mandarin proficiency) as the target groups and one group of native Mandarin speakers (MAN) as the control group. This contribution reports preliminary results from five speakers (2 from JAPH, 2 from JAPL, and 1 from MAN). Participants read aloud monosyllabic words in a Mandarin carrier phrase “wo du X zhe ge zi” (‘I read the word X’) in a quiet room, with 1000 tokens in total. 10 equidistant F0 measurements were obtained and F0 tone contours were modelled with quadratic polynomials ($y = a + bx + cx^2$), in which ‘a’ refers to intercept, ‘b’ to slope, and ‘c’ to curvature. For the secondary cues, we measured the duration and extracted the minimum values of six voice quality parameters (H1*-H2*, H2*-H4*, H1*- A1*, H1*-A2*, H1*-A3*, and CPP) from VoiceSauce. We then conducted the Linear Discriminant Analysis on T2 and T3 for each participant to examine the weighting of each acoustic parameter for distinguishing the two tones (Table 1). In addition, to explore the extent of creakiness in T3 production, we performed mixed-effects linear regression across language groups on each voice quality cue and the result of Tukey’s post-hoc test is shown in Table 2.

Figure 1 shows that the T3 contour of the JAPL group substantially rises in the later part (even reaching the height of T1 at the endpoint), probably because at the beginning stage, T3 contour is the focus of teaching and L2 learners were taught to produce a very complete T3. Then, with the improvement of Mandarin proficiency, L2 learners gradually got a better grasp of the T2/T3 distinction and their T3 contours converged towards that of native speakers (only a small rise in the second half). For the cue-weighting pattern, results in Table 1 suggest that F0 cues are the most robust, accounting for the top three weights across language groups (except for JAPL2). Duration plays an essential role for low-proficiency learners and relatively speaking, it has more weight in the JAPL group compared to JAPH and MAN group, i.e., when F0 cannot clearly distinguish T2 and T3 for beginners, secondary cues such as duration comes into play. However, as Mandarin proficiency improves, L2 learners tend to weigh more on voice quality than duration, which is consistent with native speakers’ cue-weighting pattern. For the specific voice quality cues, there are between-group differences in most of them except for H2*-H4* and H1*-A1*. All the other parameters were significantly lower for the MAN group than for the L2 learners, and the JAPH group had significantly lower values for H1*- H2* and CPP than the JAPL group (Table 2). Since lower values suggest a higher degree of creakiness, we may infer that high-proficiency learners produce creakier T3 than their low counterparts, indicating a native-like tendency. In sum, this preliminary study suggests that the cue-weighting pattern in Japanese learners’ T2/T3 distinction skew towards native speakers as their Mandarin proficiency increases.

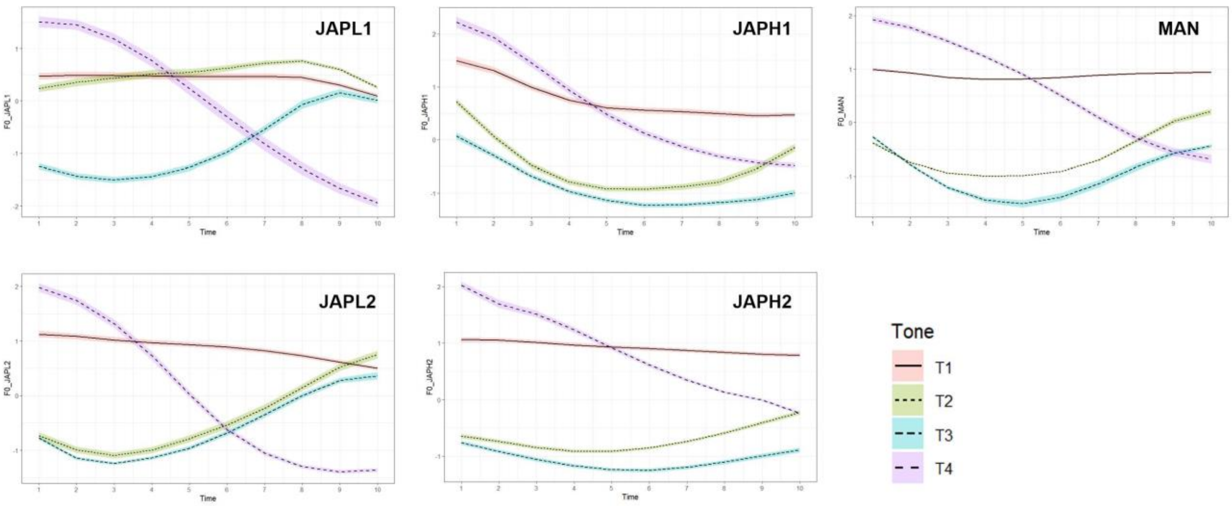

Figure 1. Individual tone contour patterns across language groups

Table 1. Individual LDA results of Mandarin T2/T3 distinction (The five highest weighted cues are bolded)

	JAPH1	JAPH2	JAPL1	JAPL2	MAN
a	**-1.4804**	**-1.6223**	**-1.5180**	**-0.6855**	**-0.6327**
b	**-3.0432**	**-3.7727**	**-0.6433**	**-2.1609**	**-5.2126**
c	**-3.1628**	**-2.6683**	**0.9427**	**-1.8471**	**-4.0688**
duration	0.1973	**0.2145**	**0.5627**	**1.1038**	**0.2647**
H1*-H2*	-0.1852	**-0.2747**	-0.0396	**0.4441**	**-0.6252**
H2*-H4*	-0.2758	-0.0106	0.1759	0.0027	-0.2041
H1*-A1*	**0.3064**	-0.1218	**-0.4124**	-0.2205	0.1051
H1*-A2*	**-0.3992**	-0.0010	-0.1829	-0.1391	0.1549
H1*-A3*	0.2061	-0.0210	0.3415	-0.0298	-0.2032
CPP	-0.1629	-0.1472	0.1290	-0.2125	0.0715

Table 2. Results of Tukey's post-hoc comparisons of voice quality parameters in T3 production

	Language	estimate	SE	*p*.value
H1*-H2*	JAPH - JAPL	-3.11	0.53	<.0001
	JAPH - MAN	2.88	0.83	0.0050
	JAPL - MAN	5.99	0.83	<.0001
H1*-A2*	JAPH - JAPL	0.80	1.37	0.8283
	JAPH - MAN	7.08	1.85	0.0005
	JAPL - MAN	6.28	1.88	0.0029
H1*-A3*	JAPH - JAPL	-1.87	1.33	0.3395
	JAPH - MAN	6.32	2.05	0.0127
	JAPL - MAN	8.18	2.06	0.0011
CPP	JAPH - JAPL	-1.29	0.16	<.0001
	JAPH - MAN	1.94	0.20	<.0001
	JAPL - MAN	3.23	0.21	<.0001

References

[1] Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonallanguage speakers. *Journal of Phonetics, 40*(2), 269–279.

[2] Van Way J (2014) Effect of creaky voice simulation on third-tone perception in Mandarin Chinese. *University of Hawai'i at Mānoa: Working Papers in Linguistics* 45(3): 1-13.

[3] Yang, J., Zhang, Y., Li, A., & Li, X. (2017). On the duration of Mandarin tones. *Proceedings of theAnnual Conference of the International Speech Communication Association*, 1407–1411.