Do non-native language listeners perceive Mandarin tone continua categorically?

Chang, Yueh-chin[a], Halle, Pierre[b]; Best, Catherine T.[c]; Abramson, Arthur[d]

[a.] Graduate. Institute of Linguistics., Natl. Tsing Hua Univ., Hsinchu, Taiwan,

[b.] Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/Sorbonne-Nouvelle, Paris, France

[c.] University of Western Sydney, Milperra NSW, Australia

[d.] Haskins Laboratories, New Haven, USA

Abstract

Previous studies of the perception of tone continua found that Mandarin listeners perceive tones more categorically than French listeners do. The latter listeners' performance is essentially psychophysically motivated. In this study, we examined native-language effects on the perception of Mandarin tone continua by Cantonese, Thai, Vietnamese, Japanese, and American English listeners.

Listeners were tested on three Mandarin tone continua constructed from natural utterances and were run on AXB identification and AXB discrimination tests on isolated syllables. The intercepts and steeper slops found in the identification tasks point toward more catergorical perception for the speakers of tone languages than for the listseners of non-tonal languages. Although American and Japanese listeners had better performance in the discrimination tests than the non-native listeners of tone languages did, but they needed more Response Times in the discrimination tasks. That implies that their performance is mostly determined by psychophysical factors rather than being biased by linguistic factors. Thai and Vietnamese listeners' data show that non-native tone contours were perceived in a more psychophysical way, while the native-like tone contours were perceived in a lingusitic way. We also found that Mandarin tone 3, rather than tone 4, is processed by JP listeners in a more linguistic way.

1. Introduction

Literature on the linguistic status of tones in Chinese languages is rather uncontroversial. The main physical correlates of Mandarin tone are fundamental frequency (F0) and amplitude (intensity) contours. Tone information seems to be perceived by native listeners of tone languages in a linguistically contrastive way, presumably just like segments are. (Fox & Unkefer, 1985; Lee, 2000) In our previous study, we found that Chinese speakers who use tonal variations in a phonologically contrastive way at the lexical level categorize tones on-line and pre-lexically or post-lexically, and speakers of non-tonal languages (especially a non-stress language such as French who uses prosodic variations at the sentence level in a non-phonological way) process Chinese tones differently than speakers of Chinese do.

There are a few pioneer studies of the categorical perception of tones. Abramson (1979) claimed that tone perception in Thai is not categorical. But he used a strict criterion to determine the categoricity: the categorical perception involved both steep slopes in categorization functions and marked peaks at the category boundary in discrimination functions. Using an 11 step continuum with the syllable [i], Chan, Chuang, & Wang (1975, reported in Wang, 1976) revealed a typical pattern of categoricity for Mandarin listeners but not for English listeners. Based on these findings, Wang (1976) claimed that tone perception in Mandarin Chinese was clearly categorical.

No further studies using tone continua have reconsidered the issue of tone categorization until our recent endeavor with Taiwanese Mandarin and French listeners (Chang & Hallé, 2000; Hallé et al. 2004). We provided arguments to support the categorical perception of tones based on cross-linguistic data and defined the categorical perception of tones as an increased sensitivity near category boundaries. We concluded that Taiwanese Mandarin listeners perceive tones in a quasi-categorical manner while French listeners perceive them in a more psychophysical fashion and that French listeners are not absolutely "deaf" to tonal variations. They simply fail to perceive tones along the lines of a well-defined and finite set of linguistic categories.

Some studies show that language experience affects tone perception from second language acquisition's perspective. Lee, Vakoch, and Wurm (1996) found that native speakers (Cantonese, Mandarin and English) were more successful at discriminating tones from their own languages for both words and nonwords. The aim of this study is to examine if Mandarin Chinese tone continua will be categorized in a quasi-categorical way by non-native speakers of tone language (such as Cantonese, Thai, Southern Vietnamese), by speakers of pitch accent language (e.g. Japanese), and by speakers of stress language (e.g. English).

Cantonese is a lexical tone language. Cantonese distinguishes six tones: two rising tones (high-rising [24]/[25], low-rising [23]), low-falling [21], and three level tones (high level [55], mid level [33], and low level [22]). Cantonese has two allotones: high falling tone and high level tone. The Cantonese high level, high-rising, low-falling and high falling tones are phonetically similar to Mandarin tones. (Fok, 1974; Yuan, 1983; Zee, 1991) Moreover, Cantonese is one of Chinese languages spoken in southern China. It is expected, thus, that Cantonese listeners categorize the tone continuum contrastively as Mandarin speakers do.

Thai has five phonemic tones, labeled as high, mid, low, falling and rising. Thai high, rising, low, and falling tones are the 'sister' tones, so to speak, of Mandarin tones 1 to 4. Moreover, the similar patterns of SH activity have been found both in Mandarin tones 2 and 4 and in Thai rising and falling tones. (Hallé, 1994; Sagart et al., 1986) However, the phonetic pitch patterns of high and falling tones in Thai are different from those in Mandarin: the phonetic pattern of high tone is rising in the second part of the syllable. As for the falling tone, regardless of onset pitch value, the pitch pattern fells from high midpoint to low endpoint. (Zsiga et al. 2007) Our intuition is that Mandarin Chinese tones should be categorized phonologically by Thai listeners and the discrimination performance is lower for Thai listeners than for Mandarin and Cantonese listeners.

Southern Vietnamese, also a lexical tone language, has five tones that combine the pitch contours and voice quality: mid level [33], low-falling [21], high-rising [35], glottalized low-falling-rising [312], falling-rising [324]. The Vietnamese level tone is lower than the Mandarin level tone, but high-rising and low-falling tones are similar to those in Mandarin. Since Vietnamese uses voice quality within the tonal system, we could predict that Vietnamese listeners perform worse than Thai and Cantonese listeners do.

Japanese is a pitch accent language. There is one pitch drop (accent) per word. The pitch contrasts are used phonemically as in a tone language. If we agree with Duanmu (2004), Mandarin CV syllables might be considered as having a syllable with two moras. This kind of representation matches perfectly with Japanese two moras words (for instance /kaki/ 'fence'). Japanese pitch accent patterns (LH and HL) are similar to the pitch pattern of Mandarin Tone 2 and Tone 4. Thus, we predict that Japanese listeners are sensitive to category boundaries with the falling-rising continuum, but less sensitive to category boundaries with the high level-rising continuum and that with the high falling-low falling continuum.

American English is a lexical stress language. HL and LH pitch contours could be found in disyllabic words (e.g., digest (n.) vs. digest (v.)). We hypothesize that English speakers might be sensitive to the pitch variations, but they might perceive tone contour differences in a merely psychophysical fashion rather than being biased by linguistic factors.

2. Procedure

Speech materials

The tone continua described here were used in the experiments reported in the previous studies as well as in the present study. We had a male native speaker of Mandarin Chinese (Taiwanese Madarin) pronounce the target syllables /pa/, and /pi/ at each of the four tones, within the carrier sentence "yi ge X zi" ("one character X") where "X" stands for a given target syllable. For each target syllable, three continua were constructed: tone 1-tone 2, tone 2-tone 4, and tone 3-tone 4.

Each continuum proceeded through eight steps from one endpoint to the other. The original sentences used to create the various endpoint syllables were measured for F0 and intensity. For each continuum, the initial endpoint was one of the two original syllables (e.g., /pa2/) on which a stylized F0 contour was imposed (a smoothed version of the original contour), using a pitch-scaling algorithm similar to the time-domain "pitch synchronous overlap add" (PSOLA) method (Moulines & Laroche, 1995). The final endpoint of the continuum was obtained by imposing on the same syllable the stylized and time-adjusted F0 contour of the other endpoint original syllable (e.g., /pa4/). The resulting waveform was further modified by changing its intensity contour to that of the final endpoint original syllable. The remaining six intermediate stimuli were obtained by interpolating, at each time point, both F0 on a log frequency scale and intensity on a decibel scale. For each continuum, one of the two endpoints was chosen as the "starting-point" speech signal from which all the eight steps were derived: t1 for the t1-t2 continuum, t2 for t2-t4, and t3 for t3-t4. The eight F0 contours for each continuum with the syllable /pi/ are shown in Figures 1A-C. The intensity contours for the original /pi/ syllables in the four tones, shown in Figure 2, illustrate that intensity contours are strongly tone dependent and tend to correlate with F0 contours, whereas duration, though tone-specific to some extent, stay within a narrow range of variation.

Participants

Thirteen Cantonese students from City University of Hong Kong (hereafter, CT listeners, aged 19 to 23 years) and 13 Taiwanese Mandarin speakers from National Tsing Hua University (hereafter, TW listeners, aged 20 to 23 years) were paid to participate in three experiments. CT listeners reported have studied Chinese at the age of 8, but English at the age of 6. According to self report, their Chinese level is at 4 (the poorest=1, the best=5). 15 Thai students from Chulalongong University (hereafter, TH listeners, aged 20 to 25 years), 15 Vietnamese students from University of Social Sciences and Humanities (hereafter, VN listeners, aged 20 to 25 years), 15 Japanese students from Nanzan University (hereafter, JP listeners, aged 22 to 28 years), and 14 American students from Yale University and University of Connecticut (hereafter, AM listeners, aged 22 to 28 years) were paid to participate in three experiments. None of the TH and VN students had ever studied Chinese and none of AM and JP participants had ever been exposed to Chinese or any other tone language. All the participants reported normal hearing and speaking.

Participants were run in three sessions: one session per continuum, each involving a discrimination test followed by an identification test.

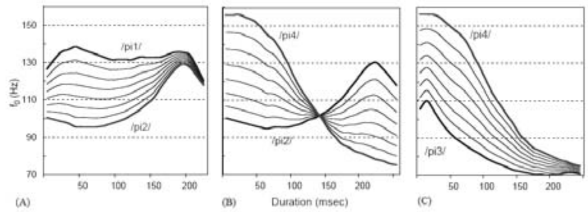

Fig. 1. (a-c) Tone contours for /pi/ in the three continua: (A) in t1-t2; (B) in t2-t4; (3) in t3-t4

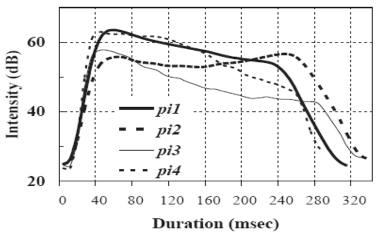

Fig. 2. Intensity contours of the syllable /pi/ in each of the four tones (original stimuli)

Identification tests

In each session, the participants received an AXB identification test, where A and B correspond to the two endpoints of the continuum. For each continuum, participants were presented with 160 trials (8 steps x 2 orders x 2 syllables x 5 repetitions) in quasi-random order, blocked by 16 trials. It was preceded by a training phase of 32 trials. The interstimulus interval was set to 1 s, the intertrial interval to 2.5 s. Participants were instructed (in their native language) to press one of two response buttons labeled "1" and "3" for each trial, according to whether they thought the second stimulus (X) sounded more like A than B or vice-versa. The response times were recorded from the onset of the third stimulus (i.e. B).

AXB discrimination tests

In each session, they received an AXB two-step discrimination test. In the test phase, participants received 144 trials (2 syllables x 6 pairs x 4 combinations x 3 repetitions) presented in quasi-random order, blocked by 18 trials. The preceding training phase comprised 32 trials. The interstimulus interval was set to 1 s, the intertrial interval to 2.5 s. Participants received exactly the same instructions as for the AXB identification task and their response times were recorded from the onset of the third stimulus.

3. Results and discussion

Identification tests

The data for each subject was fitted to a short ogive Gaussian curve (cf. Best & Strange, 1992; Hallé et al., 1999). This yielded individual data of intercept and slope at crossover for each continuum and each syllable type. Table 1 summarizes these data for all the listener groups. The identification curves, for each continuum are shown in Figure 3 (a-c).

	t12		t24		t34		Mean	
	IL	SLOP	IL	SLOP	IL	SLOP	IL	SLOP
AM	3.81	1.26	4.72	1.02	4.96	1.20	4.49	1.16
JP	4.09	1.52	4.70	1.15	4.76	1.20	4.52	1.29
TH	4.38	1.44	4.33	1.15	4.50	1.19	4.40	1.26
VN	3.90	1.68	4.32	1.41	5.18	1.23	4.47	1.44
CT	4.23	1.65	4.55	1.11	4.63	1.53	4.47	1.43
TW	3.82	2.37	4.14	1.87	4.11	1.76	4.02	2.00

Tab. 1 : Intercept and slope of identification data for all the listener groups.

Analyses of variance were conducted on the intercept and on the slope data, with Listener group as between-subject variable and tone continuum (t1-t2, t2-t4, and t3-t4) as within-subject variable. For intercepts, Language group is significant (F(1, 5)=2908, p<0.05). The intercepts, pooled across continua, for the non-native language participants tended to cluster around 4.5, the exact center of an 8 step continuum, whereas those for TW participants fell to the left of the center (4.02). Slopes are actually 1/SD, SD being the standard deviation of the Gaussian functions fitted to the data. Slopes at crossover were steeper for the tone language participants than for non-tonal language participants and were steeper for native language participants than for non-native language participants (TW > TH > VN > CT > JP > AM). The response time data, for each continuum, are illustrated in Figure 4 (a-c).

RTs were much longer in the intermediate region, reaching a peak value between the fourth and fifth stimulus in a continuum, corresponding to the category boundary region. Moreover, RTs were much longer for the non-tonal language participants than for the tone language participants (F(1, 84)=2200, p=0.000). According to Hallé et al. (2004), more RTs are long and more slop is steep, more is the perception psychophysically based, in this study all the RT data as well as the intercept and slope data point toward more categorical perception for the tone language participants, and more psychophysically based perception for the non-tonal language participants.

AXB discrimination tests

In the discrimination test reported in our previous study, Chinese listeners' sensitivity to F0 and intensity contour differences appeared to be biased by the tone category they perceived in any given contour. This should not be obvious for five groups of listeners for whom some of the contours within a continuum have no phonological significance. In order to test this prediction, we used an AXB discrimination procedure.

The overall discrimination performance of each listener group was pooled across participants and syllables (Figures 5a-c). We conducted analyses of variance which included the correct performance of all participant data, with listener Group as between-subject variable, and Continuum (t1-t2, t2-t4, and t3-t4) and Pair (six pairs) as within-subject variables. The results showed a gradient pattern of mean correct discrimination, pooled across continua, among the listener groups: TW (mean 87.9%) > JP (mean 84.9%) > AM (mean 83.7%) > VN (mean 82.2%), CT (mean 81.8%) > TH (mean 80.3%) (F(1,5)=10979, p=0.000). T2-t4 continuum, pooled across participants, was best performed among the three continua (88.9%), followed by t1-t2 continuum (81.9%) and t3-t4 continuum (79.5%) (F(2,158)=26.59, p=0.000). The reaction time data were illustrated in figures 6a-c. A gradient pattern of RTs, pooled across continua, among the listener groups is following: TH (mean 1043 ms.) > AM (mean 1015 ms.), JP (mean 1009 ms.) > TW (mean 975 ms.), VN (mean 957 ms.), HK (mean 927 ms.). Except for the AM, JP, TW, and VN listeners in the t2-t4 continuum, the RTs were generally inversely correlated with correct discrimination level for all the listener groups in three continua: more the discrimination performance was better, shorter were the RTs. The RTs were longest for the t1-t2 continuum (999.8 ms.), followed by t3-t4 continuum (986 ms.) and the t2-4 continuum (973.3 ms.). Moreover, the RTs curves are all U-shaped for the CT and TW listeners. The "troughs" in the RT curves (in the discrimination test) are roughly consistent with the peak locations in the discrimination curves and with the category boundary region found in the identification curves.

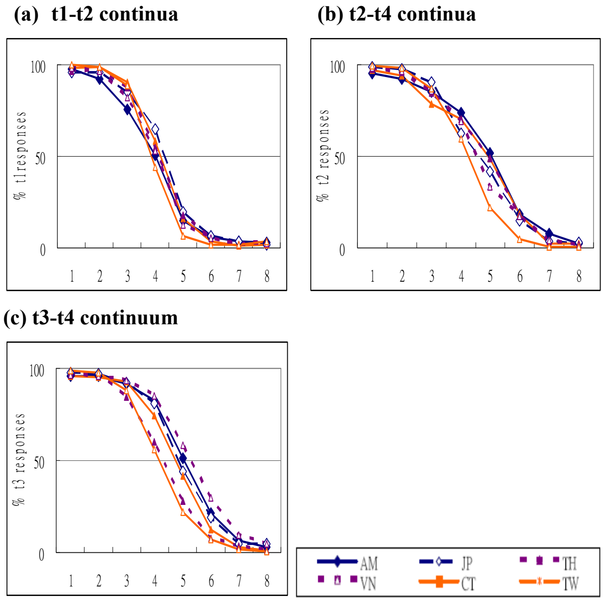

Fig. 3. Identification curves of all the listener groups in the AXB identification task.

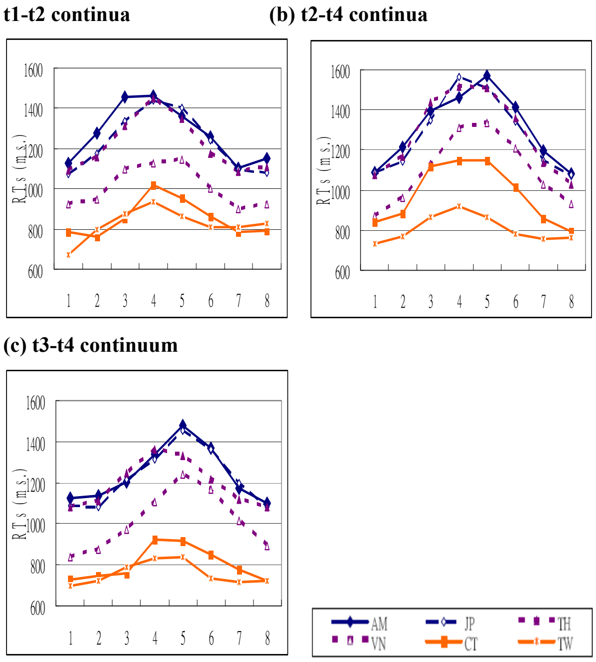

Fig. 4. Response times of all the listener groups in the AXB identification task.

There was a significant Listener Group x Pair interaction, F(25,950)=4.40, p=0.000. The discrimination curves are well-shaped only for the TW and CT listeners (hereafter, listeners of Chinese languages), but 'flat' for the AM and JP listeners (hereafter, listeners of non-tonal languages) and for TH and VN listeners (hereafter, listeners of non-Chinese languages). A detail examination shows that the discrimination level significantly varied across continuum pairs for the JP listeners and the listeners of the tonal languages (*ps* < 0.05), and did not significantly vary across continuum pairs for AM listeners ($F(5,195)=2.331$, $p=0.044$ [marginal]).

T1-t2 continuum

For the t1-t2 continuum, the discrimination curves are bell-shaped for the listeners of the Chinese languages (a clear peak of discrimination at pair 3-5 for TW listeners and a fuzzy maximum at pairs 3-5 and 4-6 for CT listeners), but "flat" for AM, JP, TH, and VN listeners. For the latter groups of listeners, there was a trend for better performance toward the tone 1 end of the continuum. A similar trend had been found in the French listeners' discrimination performance in our previous study. In line with Wang et al. (1976), we concluded that French listeners categorized the tone continuum in a psychophysical way, because it is easier to distinguish the level contour (such as tone1) from the rising contour than the two different rising contours (such as tone2) on the psychophysical basis. To understand if AM, JP, TH, and VN listeners processed the tone continuum in a psychophysical fashion, we did a further examination on the discrimination level across continuum. The result showed that the discrimination level significantly varied across continuum pairs for all the four group listeners ($ps<0.05$). However, if we ignore the pair 1-3, where the highest level is reached, the curves are "flat" for AM and JP listeners ($ps>0.05$), but "falling" for TH and VN listeners ($ps<0.05$). The identification data showed that the intercept tended to cluster around stimulus 4 for AM and JP listeners. Thus, the peak at the pair 1-3 could not be considered as the pair that straddled the category boundary for the latter listeners. Moreover, as can be seen in the RTs curves in the fig. 6a, RTs were longer for AM and JP listeners than for TW and CT listeners. It means that AM and JP listeners had less confident to decide whether X is similar to A or to B in the discrimination test. This lack of confidence can be reflected by the longer reaction times. In contrast, TW and CT listeners, the discrimination level was significantly higher at the peak than the level reached in the adjacent pairs and the peaks of discrimination were consistent with the category location found in the identification test. All the data point to noncategorical processing of the t1-t2 continuum by AM and JP listeners, and to categorical processing by TW and CT listeners.

If we take into consideration the boundary location found in the identification for the t1-t2 continuum for all the listener groups (i.e. average at 4.06), the between-boundary was the pair 3-5, within-category pairs were the pairs 1-3 and 2-4 for the tone 1 category and were the pairs 4-6, 5-7, and 6-8 for the tone 2 category. Since Thai and Vietnamese have a rising tone, similar to Mandarin tone 2, in their phonological system, it was expected that the speakers of those two languages will ignore the irrelevant tonal variations in order to efficiently categorize some of the continuum pairs into rising tone, and the within-category discrimination level is expected to be poorer than the between-category discrimination level. Although, no peak had been found in the discrimination curves for VN and TH listeners, the discrimination performance for the within-category pairs of tone 2 category was significantly lower than that for the between-category pair (i.e. the pair 3-5) ($ps<0.05$). Moreover, the VN and TH listeners had significantly poorer discrimination performance than the AM and JP listeners did in the within-category pairs of tone 2 category ($F(1,175)=4.38$, $p=0.037$), but no significant difference had been found in the within-category performance of tone 1 category for those four listener group. In other words, as compared to AM and JP listeners, TH and VN listeners seemed to tend to ignore irrelevant tonal variations toward to the tone 2 category (i.e. more linguistic processing) and to be more sensitive to the tonal variations toward to the tone 1 category (i.e. more psychophysical processing).

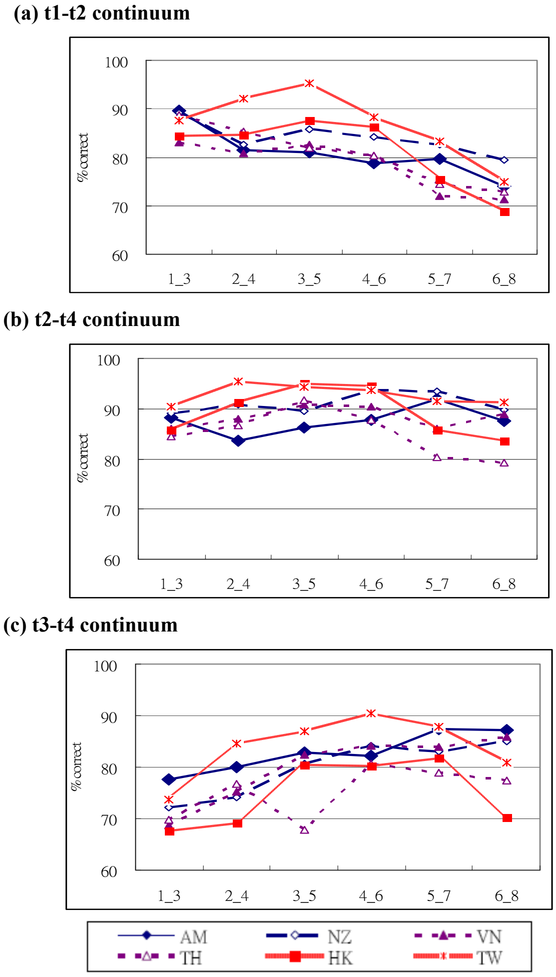

Fig. 5. Two-step discrimination curves in three continua for all listener groups.

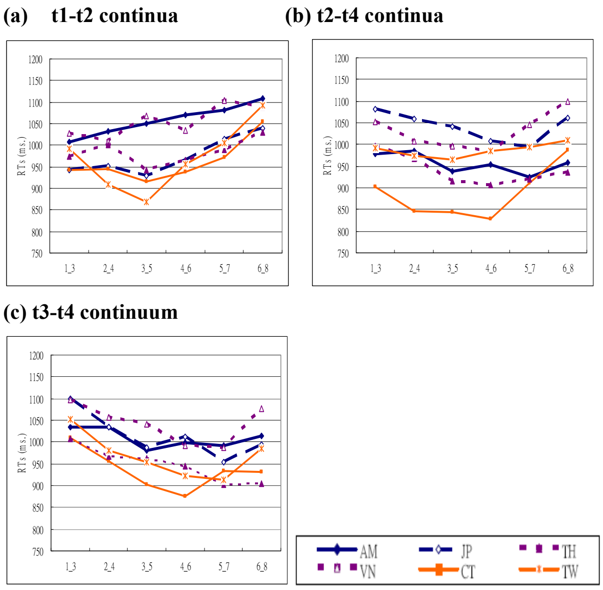

Fig. 6. Response times of all the listener groups in the AXB discrimination task.

T3-T4 continuum

A similar outcome obtains for the t3-t4 continuum, that is, for both the TW and CT listeners, there was a fuzzy maximum at pairs 3-5, 4-6, and 5-7, significantly higher than the adjacent pairs 2-4 and 6-8, (ps<0.05). The 'peaks' of discrimination were consistent with the boundary region found in the identification test (4.11 and 4.63 respectively). Steep slop found in categorization (1.76 for the TW listeners and 1.53 for the CT listeners) and clear peaks at the category boundary found in discrimination showed that the TW and CT listeners categorize the tone continuum in a linguistic way. But, both the intercept and slop data point toward more categorical perception for TW listeners than CT listeners. There was a trend for better discrimination toward the tone 4 end of the continuum for the listeners of non-Chinese languages. One-way ANOVA showed that the discrimination level significantly vary across continuum pairs for JP, TH and VN listeners (ps<0.05), but not for AM listeners ($F(5,78)=1.36$, $p=0.25$). For the latter listeners, moreover, the highest level was at the pairs 5-7 and 6-8, far from the boundary location (i.e.4.96) found in the identification test. Longer RTs in the discrimination task and gentle slop found in the categorization (1.20) also indicated that AM listeners do not perceive the t3-t4 continuum in a linguistic way, as compared to TW and CT listeners.

VN listeners' data demonstrated that there was a trend for better performance toward the tone 4 end. Recall that Vietnamese does not have a high falling tone but has a low falling tone similar to Mandarin Chinese tone 3. For VN listeners, the within-category pairs of tone 3 category were the pairs 1-3, 2-4 and 3-5, based on the boundary location found in the identification, which is on the right side of the center of the continuum (i.e. 4.84). The discrimination data showed that the within-category discrimination level was significantly lower than the between-category discrimination performance (i.e. at the pair 4-6) and the within-category pairs' performance in tone 3 category is significantly lower for VN listeners than for AM listeners (p<0.05). In other words, VN listeners had more difficulty in distinguishing the difference between within-category pairs of tone 3 category (mean 72%), as compared to AM listeners (mean 80.2%). They tended to ignore irrelevant tonal variations toward tone 3 category (i.e. more linguistic processing) and to be more sensitive to the tonal variations toward tone 4 category (i.e. more psychophysical processing).

As for the TH listeners, the discrimination curve is M-shaped. The discrimination performance was lowest at pair 3-5. However, the highest level was at pair 4-6, which is roughly consistent with the category boundary found in the identification test and is significantly higher than the level reached in the adjacent pairs 1-3, 2-4 and 3-5 (i.e. the within-category pairs of tone 3 category), but unsignificantly higher than the level reached in the adjacent pairs 5-7 and 6-8 (i.e. the within-category pairs of tone 4 category). It was expected that TH listeners processed the t3-t4 continuum categorically, since phonemically, Thai has falling and low tones, similar to Mandarin Chinese tones 3 and 4. However, the data showed that TH listeners perceived the within-category pairs of tone 3 category in a more linguistic way, but in a more psychophysical way for the within-category pairs of tone 4 category. It might be due to the phonetic pattern difference between Mandarin falling tone and Thai falling tone. The falling tone is actually realized as rise-fall contour in Thai. (Zsiga, E. & Nitisaroi, R. 2007)

Interestingly enough, the pattern of discrimination data obtained for the JP listeners (mean 79.9%, ranging 72.3%~85.0%) was similar to that for the VN listeners (mean, 80.0%, ranging 68.6%~85.8%) ($F(1,168)=0.008$, $p=0.93$), that is, the highest level reached at pair 6-8 and a poor discrimination performance toward to the tone 3 end. Japanese pitch accent patterns (LH and HL) are similar to those of Mandarin Tone 2 (LH) and Tone 4 (HL). We expected that Japanese listeners had a better discrimination performance toward the tone 3 category than toward the tone 4 category. However, the data showed that the JP listeners had more difficulty in discriminating the continuum pairs in the tone 3 category. However, the discrimination level reaches 75.7% in the within-category pairs of tone 3 category, lower than 84.0% in those of tone 4 category, 85.8% in those of tone 1 category, and 82% in those of tone 2 category obtained in the t1-t2 discrimination test. It implied that Japanese listeners who are sensitive to the pitch drop categorize some of continuum pairs (i.e. in tone 3 category) more linguistically. They seem to assimilate the pitch drop to a low tone, rather than to a falling tone or to a rising tone. This also provides an explanation of why the JP listeners were not able to categorize some of continuum pairs into a rising tone in the t1-t2 continuum as VN and TH listeners did.

T2-T4 continuum

As for t2-t4 continuum, one-way ANOVA showed that the discrimination performance significantly varied across continuum pairs only for TH and CT listeners. Their discrimination curves are bell-shaped (a clear peak of discrimination at pair 3-5 for TH listeners and a fuzzy maximum at pairs 3-5 and 4-6 for CT listeners). The peak of discrimination is significantly higher than the discrimination level reached in the adjacent pairs for CT listeners (ps<0.05), but not for TH listeners (ps>0.05). The peaks of discrimination were consistent with the boundary locations (4.33 and 4.55) found in the identification test. Thai has a falling tone and a rising tone, and Cantonese has a high rising tone and phonetically a falling tone, allotone of level tone. The discrimination data showed that TH and CT listeners' sensitivity to the F0 contour differences seemed to be biased by their native tone categories..

The correct discrimination is highest at the pair 5-7 for AM listeners (91.8%) and at the pairs 4-6 and 5-7 for JP listeners (93.6% and 93% respectively), but unsignificantly higher than the discrimination level reached in the adjacent pairs for both the listener groups (ps>0.05). That is, the discrimination level did not significantly vary across continuum pairs. The peaks in the discrimination curves were consistent with the intercept locations found in the identification test for JP listeners (4.70), but not for AM listeners (4.96). The gentle slops found in categorization (1.02 for the AM listeners and 1.15 for the JP listeners) and small peaks found in discrimination indicated that AM and JP listeners did not process the t2-t4 continuum as categorically as TH and CT listeners did.

A shallow and broad peak had been found at pairs 3-5 and 4-6 for VN listeners, but the discrimination level did not significantly vary across continuum pairs. The peak locations tended to be consistent with the boundary location (i.e. 4.32) found in the identification tests. Recall that Vietnamese has a rising tone, and in the t1-t2 continuum, the data showed that the VN listeners' discrimination performance is biased by native tone category. However, the data obtained in the t2-t4 continuum did not show categorical processing rather than a noncategorical processing. Similar outcome had been found in the TW listeners' data. For the latter listeners, the discrimination curve is "flat": the discrimination performance did not significantly vary across continuum pairs. The result obtained in this study is different from that obtained in the previous study, that is, a fuzzy maximum is reached at the pairs 3-5 and 4-6, significantly higher than the level reached in the adjacent pairs. Why the TW listeners did not categorize the t2-t4 continuum linguistically in this study? Recall that the t2-t4 continuum is a dynamic-to-dynamic continuum whose endpoint contours are dramatically opposed. The tonal variations (i.e. contour direction and contour slop) between stimulus pairs are more salient, as compared to those in the t1-t2 and t3-t4 continua, both static-to-dynamic continua. The participants of this study might pay attention to tonal variations in within-category pairs rather than ignoring them. Moreover, some of TW listeners reported that they heard some of stimuli as level tone in the t2-t4 continuum session. Then, the TW listeners, as well as VN listeners, might perceive the t2-t4 continuum in a mix way between 'psychophysical' and 'linguistic'.

4. General discussion

In this study, using the definition of categorical perception claimed in the previous study (i.e. increased sensitivity near category boundaries), we show that the native listeners outperform the non native listener in both the identification accuracy and between-category discrimination performance. The listeners of tone languages, such as CT, VN, and TH listeners, process Mandarin tone continuum more categorically, when the tones are part of the native phonological system; otherwise, they perceive them in a psychophysical way. The non-tonal listeners' data point to noncategorical processing of Mandarin tones by AM and JP listeners, even the LH and HL pitch contours could be found in English and Japanese words. This indicates that LH and HL pitch contours do not have linguistically contrastive significance for AM listeners, and surprisingly also for JP listeners. In contrast, Japanese listeners tend to assimilate the pitch drop (accent) to Mandarin low tone. Japanese pitch accent pattern (HL/LH) contrast thus could be reinterpreted as the presence/absence of low tone in the second mora.

Moreover, the discrimination performance is higher for the AM and JP listeners than for the CT, VN, and TH listeners. Although, LH and LH pitch contours have no phonological significance for AM listeners and the duration is strongest correlate and cue to the word stress in English (Sluijter *et al.*, 1995), high fundamental frequency and greater intensity still play important roles in distinguishing accented primary stressed syllables from the neighboring non-primary syllables. Also, JP listeners are sensitive to low pitch. Both the listener groups are supposed to be very sensitive to tonal and/or intensity variations in the continuum pairs based on the psychophysical properties. However, the CT, VN, and TH listeners tend to ignore the irrelevant tonal variations in the within-category continuum pairs, but be sensitive to the tonal variations in the between-category continuum pairs. Therefore, their discrimination performances were lower AM and JP listeners'.

Another interesting finding is that TH listeners' discrimination performances are poorest and the RTs in the discrimination are longest among all the listener groups. This might be due to phonemic similarities and phonetic differences between Thai and Mandarin tones might affect TH listeners' assimilating Mandarin tone stimuli pairs to Thai tone categories.

REFERENCES

Abramson, A. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech*, *21*, 319-325.

Abramson, A. (1979). The noncategorical perception of tone categories in Thai. In B. Lindblom & S. Ohman (Eds.), *Frontiers of Speech Communication* (pp. 127-134). London: Academic Press.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, *20*, 305-330.

Chan, S. W., Chuang, C.-K., & Wang, W. S.-Y. (1975). Cross-linguistic study of categorical perception for lexical tone. *Journal of the Acoustical Society of America*, *58*, S119.

Chang, Y.-C., & Hallé, P. (2000). Taiwan Huayu shengdiao fanchou ganzhi [Categorical perception of Taiwanese Mandarin tones]. *Tsing Hua Journal of Chinese Studies*, *new series XXX*, *1*, 51-65.

Duanmu, S., 2004, Tone and non-tone languages: An alternative to language typology and parameters. *Language and Linguistics*. *5.4,* 891-923.

Fok, C.Y.C. (1974) *A Perceptual Study of Tones in Cantonese*. Hong Kong: University of Hong Kong Press.

Fox, R., & Qi, Y.-Y. (1990). Context effects in the perception of lexical tone. *Journal of Chinese Linguistics*, *18*, 261-283.

Fox, R., & Unkefer, J. (1985). The effect of lexical status on the perception of tone. *Journal of Chinese Linguistics*, *13*, 69-90.

Gandour, J. (1981). Perceptual dimensions of tone: evidence from Cantonese. *Journal of Chinese Linguistics*, *9*, 20-36.

Gandour, J. (1983). Tone perception in far eastern languages. *Journal of Phonetics, 11*, 149-175.

Gandour, J., & Harshman, R. (1978). Crosslanguage differences in tone perception: a mulltidimensional scaling investigation. *Language and Speech*, *21*, 1-33.

Hallé, P. (1994). Evidence for tone-specific activity of the sternohyoid muscle in Modern Standard Chinese. *Language and Speech*, *37*, 103-124.

Hallé, P., Best, C., & Levitt, A. (1999). Phonetic versus phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics, 27*, 281-306.

Hallé, P., Chang, Y-C., & Best, C. (2004). Categorical perception of Taiwan Mandarin Chinese tones by Chinese versus French native speakers. *Journal of Phonetics*, *32*, 395-421.

Lee, C.-Y. (2000). Lexical tone in spoken word recognition: A view from Mandarin Chinese. Unpublished doctoral dissertation, Brown University. Providence, RI.

Lee, Yuh-Shiow , Douglas A. Vakoch and Lee H. Wurm. (1996).Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research* 25 (5), 527 - 542

Moulines, E. & Laroche, J. (1995). Non-parametric techniques for pitch-scale and time-scale modification of speech. Speech Communication, 16, 175-205.

Sagart, L., Hallé, P., de Boysson-Bardies, B., Arabia- Guidet, C. (1986). Tone Production in modern standard Chinese: an electro-myographic investigation. *Cahiers de Linguistique Asie Orientale*, *15*, 205-220.

Sluijter, A.M.C., Shattuck-Hufnagel, S., Stevens, K.N., and van Heuven, V.J. (1997). Supralaryngeal resonance and glottal pulse shape as correlates of stress and accent in English. *Proceeding of the 13th International Conference of Phonetic Sciences, Stockholm, 2,* 630-633.

Wang, W. S.-Y. (1976). Language change. *Annals of the New York Academy of Sciences*, *28*, 61-72.

Yang, Y.-F. (1991). Ear differences in distinguishing consonant features and lexical tones. *Acta Psychologica Sinica*, *2*, 131-137.

Yuan (1983) *Essentials of Chinese Dialectology*. Beijing: Wenzi Gaige Publisher.

Zee, E. (1991) Chinese (Hong Kong Cantonese). *Journal of the International Phonetic Association* 21 (1), 46–8.

Zsiga, E. & Nitisaroi, R. (2007) Tone Features, Tone Perception, and Peak Alignment in Thai. *Language and Speech, 50 (3),* 343-383.