

# Perception of tones in Mandarin and Dutch adult listeners

Liquan Liu<sup>1,2</sup>, Ao Chen<sup>1</sup> and René Kager<sup>1</sup>

<sup>1</sup>Utrecht University / <sup>2</sup>Western Sydney University

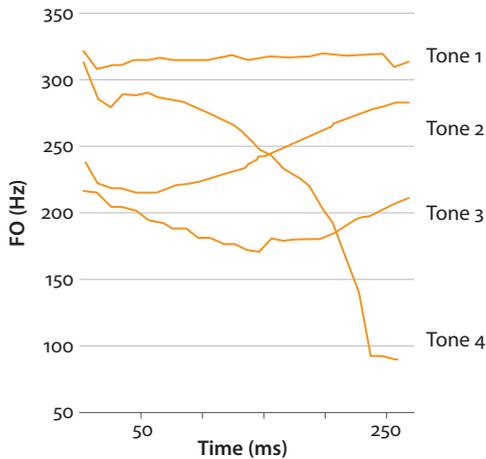
This paper examines the nature of categorical perception (CP) effects in Mandarin and Dutch adult listeners through identification and discrimination tasks using lexical tonal contrasts and through the CP index analysis. In identification tasks, Mandarin listeners identify tones in accordance with their native tonal categories whereas Dutch listeners do so based on acoustic properties. In discrimination tasks, Dutch listeners outperform Mandarin listeners especially in tonal steps on the continuum falling within the Mandarin tonal category boundary, whereas Mandarin listeners display high sensitivity in discrimination of stimuli falling across the native boundary. The CP index analysis shows a higher degree of CP in Mandarin (categorical perception) than in Dutch (psycho-acoustic perception) listeners.

**Keywords:** lexical tone, cross-linguistic perception, CP index analysis, categorical perception, psycho-acoustic perception

## 1. Introduction

In tone languages like Mandarin Chinese, pitch is phonologically contrastive at the word or morpheme level (Figure 1). Tone-language and non-tone-language listeners perceive lexical tones in different ways. On the one hand, Cantonese, Chinese, Thai, Taiwanese, and Vietnamese (tone-language) listeners show categorical perception (CP) effect, the percept invariances in sensory phenomena varied along a continuum, when perceiving lexical tones (e.g. Content & Perwez 2011; Gandour 1978).

On the other hand, CP effect seems lacking in non-tone-language listeners. Hallé et al. (2004) set up 8-step continua for the three segmental sequences /pa/, /pi/, and /kwo/ using three tone contrasts (Mandarin T1-T2; T2-T4; and T3-T4). Taiwanese and French listeners participated in AXB identification and discrimination tasks. The two groups differ qualitatively in categorizing tones: French



**Figure 1.** Tones in Mandarin Chinese [T1: level, T2: rising, T3: dipping, T4: falling]  
Source: Wang, Jongman & Sereno (2001)

listeners display psycho-acoustic but not linguistic tone identification patterns, whereas their Taiwanese peers perceive tones linguistically in accordance with the phonological values in Mandarin. Xu et al. (2006) created a 7-step continuum of a tone contrast (Mandarin T1-T2) to test Chinese and English listeners using a forced choice identification task and an AX discrimination task. CP effect is found only for the Chinese listeners. A category-based interpretation of the perceptual learning of Cantonese lexical tones by Mandarin and English listeners was provided, with listeners' performance not solely to participants' native categories, but also to cross-language mappings including prosodic knowledge of both the native and target languages (Francis et al. 2008). German listeners display broader (psychophysical) identification boundary widths than tone-language listeners' (linguistic) boundaries (Peng et al. 2010).

Non-tone-language listeners tend to illustrate psycho-acoustic perceptual patterns while hearing tones (Van Lancker & Fromkin 1973; Wang et al. 2004; but see Luo et al. 2006), and process prosodic variation at the sentence level (e.g. as intonation) (Wang 1973, 1976; Hallé et al. 2004; Xu et al. 2006; Wu & Lin 2008; Huang & Johnson 2010; Sun & Huang 2012), whereas tonal categories function to distinguish between segmentally identical words among tone-language listeners (Lee 2007). Neuro-imaging studies report that tones are perceived in the left hemisphere just as other speech segments for tone-language listeners, and in the right hemisphere for non-tone-language listeners (Hsieh et al. 2000; Gandour et al. 1998; 2001; Klein et al. 2001; Wang et al. 2001; Brown-Schmidt & Canseco-Gonzalez 2004; Kaan et al. 2008). These results strongly indicate a linguistic

perception of tones by tone-language listeners and an acoustic perception of tones by non-tone-language listeners.

While up to date, copious research on the CP of tones has been conducted, the relationship between identification and discrimination has seldom been addressed. Although identification tasks result in a CP pattern only for tone-language listeners, the outcomes from discrimination tasks often suggest that tone-language and non-tone-language listeners do not show much difference in their performance (Abramson 1979). Indeed, non-tone-language listeners display a degree of sensitivity to other supra-segmental information such as stress and intonation (Henderson & Nelms 1980; Repp & Lin 1990; Lee & Nusbaum 1993). If listeners perceive stimuli in a categorical fashion, their identification and discrimination performances should predict each other. Xu et al. (2006) examine the correlation between the predicted (from identification data) and obtained discrimination scores using a two-way ANOVA: English listeners display significantly lower correlation than Chinese listeners in a speech discrimination task.

The interactional perspective on CP calls for a better integration and interpretation of experimental results. To capture this interaction, we adopt a method that is new in the tonal domain, the CP index analysis (van Hessen & Schouten 1999). This method was first used in consonant and vowel studies to test the identification and discrimination of stop consonant and vowel contrasts embedded in CVC words. It was found that the perception of stop-consonant is highly categorical, whereas this is much less so in vowel perception. Consonant perception illustrates a more coherent relationship between identification and discrimination results than the perception of vowel contrasts. CP index analysis seizes the relationship between contrast identification and discrimination, in line with the definition of CP which involves within-category compression and between-category separation (Harnad 1987).

In van Hessen & Schouten (1999), the relationship between identification and discrimination is quantified as such: the stronger the correlation between identification and discrimination, the higher the degree of CP. In other words, the better discrimination performance is predicted by identification performance, the stronger CP becomes (Schouten & van Hessen 1992). Assuming this interpretation, rather than a classical view according to which CP exists in tone-language but not in non-tone-language listeners, it can be understood that tone-language listeners display a stronger CP than non-tone-language listeners. Conversely, the notion 'degree of CP' also becomes applicable to non-tone-language listeners, which is crucial for an understanding of native influence on non-native contrast perception. Non-tone-language listeners from different language backgrounds may show different sensitivities to pitch variation in a non-native language (Soto-Faraco et al. 2001). For example, the tone discrimination peak of French listeners is

relatively smooth (Hallé et al. 2004), whereas English listeners discriminate tones as accurately as Chinese listeners (Wu & Lin 2008).

The 'degree of CP' may also be reflected with contrasts differing in acoustic salience. From a developmental perspective, tone-language-infants show initial and increased sensitivity to lexical tones (Harrison 2000; Mattock & Burnham 2006; Mattock et al. 2008; Yeung et al. 2013), the level of which is likely to be related to the salience of the tonal contrasts. Non-tone-learning infants, on the other hand, experience perceptual decline to lexical tones in the first year after birth, and their sensitivity rebounds in the second year as a function of contrast acoustic salience (Liu & Kager 2014, 2016a). Specifically, non-tone-language (Dutch) infants' tonal sensitivity at the stage of perceptual rebound is argued to be more acoustic than linguistic, reflecting non-tone-learning (Dutch) adults' tonal perceptual pattern. Nevertheless, the correspondent adults' perceptual patterns have not been examined. This study is among the first to match the stimuli used in infant and adult tone perception, aiming to shed light on tone perception of non-tone language listeners across ages. The current study concerns two research questions: the comparison of lexical tone perception by tone and non-tone language listeners by means of identification and discrimination tasks, as well as the comparison of tone perception by non-tone language infant and adult listeners across studies. These questions provide a comprehensive view on cross-linguistic tone perception.

## 2. Experiment

### 2.1 Participants

In total, 20 Chinese (mean age: 27 years, standard deviation (SD): 4 years, 6 male) and 20 Dutch (mean age: 24 years, SD: 5 years, 5 male) listeners joined in the study. All participants reported normal hearing without language impairment. None of the Dutch participants had knowledge of Mandarin Chinese or other tone / pitch-accent languages, nor had they been exposed to a Mandarin speaking environment. All Chinese speakers speak Mandarin Chinese or a northern dialect of Mandarin Chinese as their native language. The average musical training is less than 3 years across Chinese and Dutch participants.

### 2.2 Stimuli

In this experiment, we did not use Mandarin T3 to avoid potential tone sandhi. Mandarin T1-T2 and T2-T4 has been examined in various studies (e.g. Hallé et al. 2004). Less has been tested for Mandarin T1-T4 (level vs. falling). The acoustic

salience of this contrast is among the highest as judged by tone language listeners (Whalen & Xu 1992; Huang & Johnson 2010). The tone-bearing syllable was /ta/. Both /taT1/ ‘build’ and /taT4/ ‘big’ are legal words in Mandarin. The productions of a Chinese female phonetician were recorded using the computer program Audacity<sup>1</sup> via an active speaker Genelec 1029A recording system in a sound-proof booth of the Utrecht University phonetics lab. For each sound, four natural T1-T4 pairs were recorded to create within-speaker variation. The distances (in Hz) between temporally aligned points of four natural realizations of the contrast were divided into seven equal steps each, at four points in time (0%, 33%, 67% and 100%) using the PRAAT software (Boersma & Weenink 2012; see Appendix). Then each of the in-between points was connected by simple interpolation to produce new pitch contours (Figure 2). In this way, eight stimuli including the endpoint contours were created for one continuum from stimulus 1 (/taT1/) to stimulus 8 (/taT4/), and 32 stimuli were generated in total for the four continua as multiple tokens (mean duration: 412ms, see Appendix). The average sound intensity level was 70dB. Five native speakers report that the stimuli sounded speech-like.

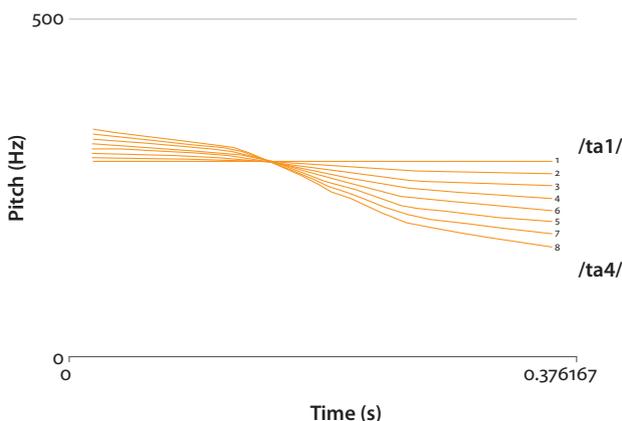


Figure 2. An 8-steps along a T1-T4 continuum

### 2.3 Procedure

All participants were tested on four CP tasks, in the sequence of forced choice identification, AX discrimination, open choice identification, and AXB discrimination. It was not random that multiple tasks were selected to test adult non-tone listeners' identification and discrimination ability. It is insufficient to claim a CP effect solely based on one type of task (Abramson 1979; Francis et al. 2003). The two forms of

1. Audacity open resource: “<http://audacity.sourceforge.net>”

tasks for each type of ability, inherently tapping into the same function, create performance referents within the same participants and provided more accurate insights of participants' perceptual patterns. The data from this 2 x 2 design help form multiple correlates for the CP index analysis. The total task duration was approximately 25 minutes. Proper instructions were provided at the beginning of each task on how to conduct the experiments. The instruction included a brief introduction to tone language in order to encourage Dutch listeners to a language perception mode. Prior to the actual tasks, participants received two auditory examples as practice trials, listed in the individual tasks below. The experiment was conducted in the Phonetics Laboratory of Utrecht University through the Flexible Experiment Program (FEP, Veenker 1998). Participants sat in front of a monitor in a sound-attenuated booth and stimuli were presented at a sound level adequate to the listeners. All trials were presented in random order in each task, and response data were collected.

### *Forced choice identification*

Participants heard one stimulus per trial, and were required to respond as accurately and quickly as possible by mouse-clicking one of the two buttons on the screen, labelled "T1" and "T4" (for Chinese participants) or "flat" and "falling" (for Dutch participants),<sup>2</sup> upon hearing each stimulus. After each click, the next trial was started. All 32 stimuli along the 8 steps along the T1-T4 continua were played twice in random order, adding up to 64 trials for each participant. Prior to the task, two auditory examples, T1 and T4, were played as practice trials.

### *Open choice identification*

In this task, eight buttons, labelled from "1" to "8", were provided correspondent to the 8 steps along the continua. Participants were instructed that 1 referred to T1 (flat) and 8 to T4 (falling), and were required to choose a number that they felt was the best match for the stimuli from the T1-T4 continua. Compared to the aforementioned forced choice identification task or the AXB identification task in which two choices were provided, a wider range of choices provided a more accurate examination of participants' perceptual patterns. Prior to the task, two auditory examples, T1 and T4, were played as practice trials.

---

2. It was not plausible to ask Dutch participants who had no tone-language background to choose between T1 and T4 labels. Introducing labels representing non-native tones may introduce a certain degree of confusion in identification. Nevertheless, Dutch participants were aware that they were listening to Chinese tones through instructions, reducing any potential influence introduced by the labels. In addition, labeling does not appear to affect perception of Dutch listeners. For instance, Figure 3 illustrates that a substantial amount of choices were given to T1 even when Dutch listeners were hearing T4, whereas almost no choice of T1 was given by Chinese listeners.

### *AX discrimination*

Participants heard one pair of stimuli (A and X) per trial, and were required to respond as accurately and quickly as possible by mouse-clicking one of the two buttons on the screen, labelled “same” and “different”, upon hearing each pair. After each click, the next trial was presented. Participants’ responses to 2-step pairs (the difference between A and B was two steps apart: 1–3, 2–4, 3–5, 4–6, 5–7 and 6–8) were recorded, adding up to 48 trials for each participant. Additionally, 48 trials with evident different pairs (more than 4 steps apart, e.g. Steps 1–6) and 32 trials with identical tokens (e.g. Steps 1–1) were used as filler trials to prevent experiment-induced bias. It remained unclear how listeners would judge the test trials before the test, hence it was hard to predict whether identical and different trials were equal on a perceptual level. Preliminary tests showed that participants across language backgrounds discriminated these filler trails with ceiling performances, reducing the likelihood of participant bias. The results of the filler trails were thus automatically excluded by the FEP program, leaving only relevant data for analysis. Prior to the task, two auditory examples, T1-T1 and T1-T4, were played as practice trials. The inter stimulus interval between tokens in each pair was set at 500 ms.

### *AXB discrimination*

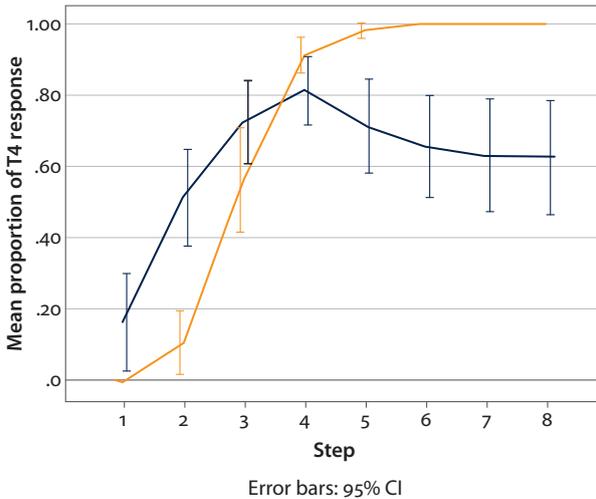
Participants heard a triplet of stimuli (A, X, B) per trial, and were required to respond to whether the second stimulus (X) sounded the same as the first (A) or the third stimulus (B) as accurately and quickly as possible by clicking one of the two buttons on the screen, labelled “first” and “third”, upon hearing each triplet. After each click, the next trial was presented. The AXB trials had four possible combinations (AAB, ABB, BAA, and BBA). Similar to the AX discrimination task, responses were recorded in which the difference between A and B was two steps apart, adding up to 96 trials for each participant. Additionally, 96 trials with evident different pairs (more than 4 steps apart, e.g. Steps 1-1-6) were used as filler trials. The filler trails were automatically excluded by the FEP program, leaving only relevant data for analysis. Prior to the task, two auditory examples, T1-T1-T4 and T1-T4-T4, were played as practice trials. The inter-stimulus interval between each token in a triplet was set at 500 ms.

## 2.4 Results and discussion

### *Forced choice identification*

A repeated measures analysis of variance (RM ANOVA) was conducted with language group (Chinese and Dutch) as the between-subject factor and mean proportion of T4 responses at each step as the within-subject variable. Corrections for

violations of sphericity were made using the Greenhouse-Geisser method where appropriate. Tukey's HSD post hoc test was applied to make pairwise comparison when necessary. The effect of step was significant,  $F(7, 32) = 69.522$ ,  $p < .001$ , indicating different responses across steps. The interaction between language group and step was also significant,  $F(7, 32) = 8.906$ ,  $p < .001$ , marking the difference between Chinese and Dutch listeners along the continua (Figure 3). Test of between-subjects effects showed that the two language groups differed significantly,  $F(1, 38) = 4.534$ ,  $p = .040$ .



**Figure 3.** T1-T4 forced choice identification curves for Chinese (solid line) and Dutch (dash line) listeners

Chinese participants consistently separated the tonal continuum into two categories, revealing a categorical perceptual pattern. Figure 3 illustrates that this tendency was stronger in the right-end of the continuum starting from Step 3 (pairwise comparisons showing that Step 3 was significantly different from any other step,  $ps < .002$ ), around which the category boundary seemed to reside in. On the other hand, no clear CP pattern can be observed among Dutch listeners, whose performance displayed high variability in almost every step along the T1-T4 continuum. At the left-end, Dutch listeners were able to separate a flat pitch (Step 1) from a slightly falling one (Step 2), differing from Chinese listeners.

Following previous studies, we conducted a PROBIT analysis calculating the category boundary location and slope of the Gaussian distribution functions that fitted the identification data (Best & Strange 1992; Hallé et al. 2004; Wu & Lin 2008). The mean response value for each participant was taken out to fit into an Ogive curve (Gaussian function). Subsequently, the analysis calculated two key

values: the intercept, the estimated 50% response position indicating the categorical boundary location, and the slope, the estimated steepness value by 1/SD of the Gaussian function fitted to the data indicating the sharpness of the boundary.

The mean intercepts (categorical boundary locations) and slopes (potential degree of CP) were summarized in Table 1. The intercept differed significantly between groups:  $F(1, 22) = 22.61, p = .015$ , with the category boundary (the 50% point) at around step 3 for Chinese listeners. The slope was significantly sharper for the Chinese group:  $F(1, 38) = 2.8, p < .001$ . The results of Chinese listeners resembled a CP function more than those of the Dutch group.

**Table 1.** The mean intercept and slope values for Chinese and Dutch listeners' forced choice identification task

/taT1/-/taT4/	Intercept		Slope	
	Chinese	Dutch	Chinese	Dutch
	2.927	4.117	2.276	0.811

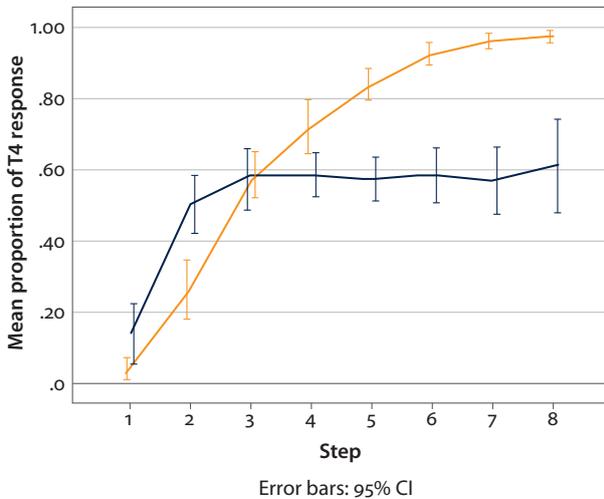
### *Open choice identification*

The ordinal data (responses 1 to 8) were proportionately restructured to fit a 0–100% continuum (i.e. the proportion of response 8 was equal to 100%).<sup>3</sup> The same analysis and criteria as in the forced choice identification task was conducted. The effect of step was significant,  $F(7, 32) = 58.158, p < .001$ , showing different responses across steps. The interaction between language group and step was also significant,  $F(7, 32) = 9.352, p < .001$ , revealing the difference between Chinese and Dutch listeners in various steps (Figure 4). Test of between-subjects effects showed that the two language groups differed significantly,  $F(1, 38) = 4.202, p = .047$ . The slope is less steep in this task than the previous task probably due to task-induced effect: As the name goes, the open choice task offers more variations rather than a sharp binary distinction in the forced choice one.

Chinese participants' performance revealed a slope between the two ends of the continuum, whereas Dutch listeners' mean responses were moderately above 50%, with the exception of step 1 (T1, flat). Note that even step 1 was falsely identified as T4 across many Dutch listeners.

Data of Chinese and Dutch participants were fitted to the PROBIT analyses. The mean intercepts and slopes were summarized in Table 2. The intercept differed significantly between Chinese and Dutch listeners,  $F(1, 23) = 32.64, p = .007$ . The

3. The actual transformation between number and proportion was: step 1 = 0%, step 2 = 14.2857%, step 3 = 28.5714%, step 4 = 42.8571%, step 5 = 54.1429%, step 6 = 71.4286%, step 7 = 85.7143%, step 8 = 100%.



**Figure 4.** T1-T4 open choice identification curves for Chinese (solid line) and Dutch (dash line) listeners

slopes were significantly steeper for the Chinese group,  $F(1, 19) = 9.22$ ,  $p = .012$ , indicating a higher degree of CP.

**Table 2.** The mean intercept and slope values for Chinese and Dutch listeners' open choice identification task

/taT1/-/taT4/	Intercept		Slope	
	Chinese	Dutch	Chinese	Dutch
	2.860	4.189	0.834	0.253

Chinese and Dutch participants' intercepts and slopes between forced and open choice identification tasks were compared separately using a One-way ANOVA. The intercepts did not differ significantly between the two identification tasks for Chinese ( $F(1, 38) = 0.13$ ,  $p = .721$ ) and for Dutch ( $F(1, 38) = 0.14$ ,  $p = .907$ ) listeners. The slopes were not significant between the two identification tasks for Dutch listeners ( $F(1, 38) = 1.60$ ,  $p = .214$ ), but they were significant for Chinese participants ( $F(1, 38) = 28.56$ ,  $p < .001$ ). This is not surprising since a binary measure in the forced choice identification task is likely to result in a steeper curve than the gradient measures in the open choice identification task when participants have well-established tonal categories.

Participants' proportion of correct responses (i.e. choosing 4 when hearing step 4) along each step was pooled in Figure 5. Chinese listeners were relatively accurate in identifying the endpoints along the continuum. In contrast, only correct response for step 1 (T1, flat) was above chance in Dutch participants. Though

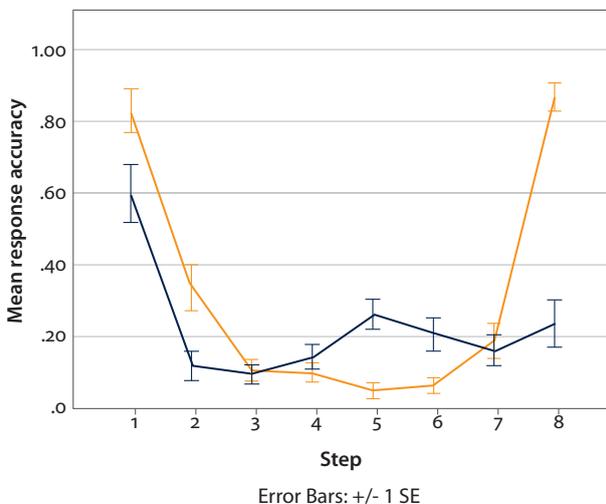


Figure 5. T1-T4 open choice identification correct responses for Chinese (solid line) and Dutch (dashed line) listeners

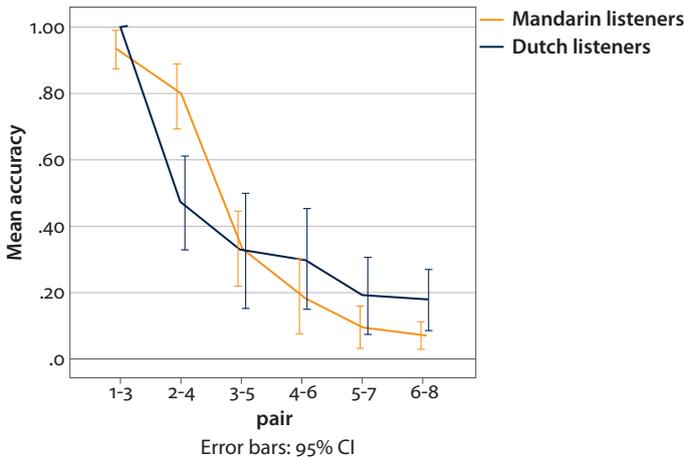
having no representations to lexical tones, Dutch listeners were able to detect the acoustic difference between flat and falling.

The intercept differed significantly between Chinese and Dutch listeners in both identification tasks, and the boundary locations were identical across tasks. Chinese participants were biased by their knowledge of native tonal categories and preferred choices closer to the category prototype. Meanwhile, without previous linguistic experience of tones, Dutch listeners identified tones psycho-acoustically.

#### *AX discrimination*

Similar to forced choice identification, an RM ANOVA was conducted with language group (Chinese and Dutch) as the between-subject factor and mean proportion of response accuracy (i.e. choosing “same” when  $A = X$ ) at each pair (6-level, 1–3, 2–4, 3–5, 4–6, 5–7 and 6–8) as the within-subject variable. The effect of pair was highly significant,  $F(5, 34) = 253.000$ ,  $p < .001$ , suggesting different responses across pairs. So was the interaction between language group and pair,  $F(5, 34) = 9.034$ ,  $p < .001$ , reflecting the cross-pair difference between Chinese and Dutch listeners (Figure 6). Nevertheless, test of between-subjects effects illustrated that language group was not a significant factor,  $F(1, 38) = .033$ ,  $p = .857$ .

The performance of Chinese and Dutch listeners was both high at the left end (T1), and fell sharply in the middle and at the right end (T4) of the continuum. For both groups, pair 1–3 was the most salient contrast pair (Mean accuracy: Chinese: 93%; Dutch: 100%), while pair 6–8 being the most difficult contrast pair for both groups (Mean accuracy: Chinese: 7%; Dutch: 17%).



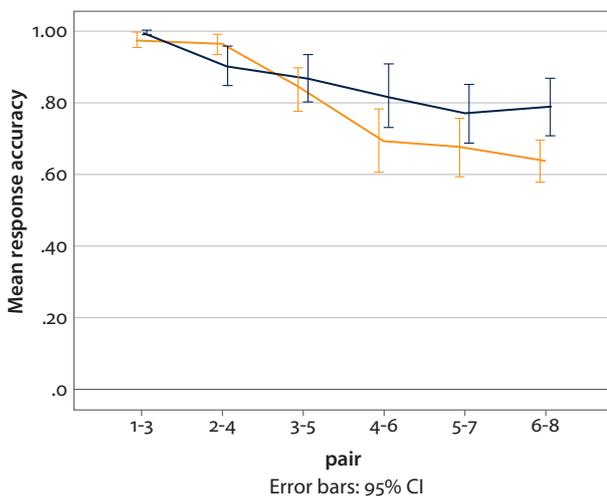
**Figure 6.** T1-T4 AX pair-wise discrimination curves for Chinese (solid line) and Dutch (dash line) listeners

In the Chinese group, the correct responses to pairs 1–3 and 2–4 were significantly higher than all other pairs ( $p < .01$ ). None of the other pairs were significantly different from each other. Note that CP predicts that high response accuracy indicates category separation of the target pairs, as is reflected in Chinese listeners' discrimination of the leftmost 2 pairs. Meanwhile, Chinese listeners performed poorly at distinguishing the rest of the pairs. From pair 4–6, the accuracy rate dropped to less than 20% and further decreased to less than 10% from pair 5–7 onwards. This was in accordance with the boundary location measured in identification tasks by means of a PROBIT analysis. The curve of the Dutch listeners resembled the Chinese group in this discrimination task, which was not unexpected given earlier studies showing that non-tone-language listeners performed better in perceiving the speech distinction in a discrimination task than an identification one (Hallé et al. 2004; Wu & Lin 2008). Dutch listeners appeared to be sensitive when distinguishing the subtle difference in tonal stimuli. Without tonal categories, they appeared to perform slightly better at discriminating the within-category pairs than Chinese listeners.

#### *AXB discrimination*

Similar to AX discrimination, An RM ANOVA was conducted with language group (Chinese and Dutch) as the between-subject factor and mean proportion of response accuracy (i.e. choosing “first” when  $A = X \neq B$ ) at each pair as the within-subject variable. The effect of pair was significant,  $F(5, 34) = 37.687$ ,  $p < .001$ , illustrating different responses across pairs. The interaction between language group and pair was also significant,  $F(5, 34) = 13.873$ ,  $p < .001$ , reflecting

the perceptual difference between Chinese and Dutch listeners along the tonal continua (Figure 7). Test of between-subjects effects showed a trend for the language group,  $F(1, 38) = 3.643, p = .064$ .



**Figure 7.** T1-T4 AXB pair-wise discrimination curves for Chinese (solid line) and Dutch (dash line) listeners

For both language conditions, the mean accuracy rate was above 60% across pairs. Chinese participants' performance was high in pair 1–3 (97.5%) and pair 2–4 (96.2%), and relatively low near the right edge (66.7% on average in pairs 4–6, 5–7, 6–8). Pair 3–5 stood in the middle of this steep drop and differed significantly from all the other steps. On the other hand, pair 1–3 was significantly different from all the other pairs ( $p < .001$ ) for Dutch listeners, while the differences between all other adjacent pairs did not yield significance.

Chinese and Dutch participants' pair-wise response accuracy scores between AX and AXB discrimination tasks were compared using a Multivariate ANOVA. The language group ( $F(6, 71) = 17.183, p < .001$ ), the tasks ( $F(6, 71) = 56.637, p < .001$ ) and the interaction of the two factors ( $F(6, 71) = 2.836, p = .016$ ) were all significant. Tests of between-subjects effects showed that pair 3 was the only pair that did not differ significantly between the two language groups ( $p = .822$ ), indicating similar perception across participants in the middle of the continua. In addition, pair 1 was the only pair that did not differ between the two discrimination tasks ( $p = .188$ ), showing relatively stable discrimination of the pair. Splitting the language groups, One-way ANOVAs reported that the response accuracy of Pairs 1–3 did not differ significantly across the two tasks for Chinese ( $F(1, 38) = 2.08, p = .158$ ) and for Dutch ( $F(1, 38) = 1.00, p = .324$ ) participants. The response accuracy of all other pairs differed significantly for Chinese ( $F_s(1, 38) > 12.61$ ,

$ps < .002$ ) and Dutch ( $F_s(1, 38) > 35.63$ ,  $ps < .001$ ) listeners. Participants across language backgrounds displayed better response accuracy in the AXB than in the AX task in the current study. This finding is intriguing since AXB tasks commonly exert a higher cognitive/memory load and should supposedly tap a more abstract level than AX tasks. The overall high performance in AXB task is likely to be task- and/or contrast-dependent. Task-wise, since the AX task was conducted prior to the AXB task, listeners may have had some, though limited, exposure to tones. Limited exposure to non-native tones can facilitate perception (Kuhl et al. 2003) and learning (Liu & Kager 2011), resulting in faster mapping in the AXB task. Contrast-wise, listeners have more referents to discriminate in the AXB than in the AX task. Tone is a salient property compared to the perception of segmental information (Lieberman et al. 1961; Liu & Kager 2015a). The additional referent per trial may greatly facilitate discrimination with salient targets. High homogeneity between the two discrimination tasks can be observed from Figures 6 and 7. Dutch participants outperformed Chinese participants in almost every pair except for pair 2–4, the one that fell between two tonal categories for Chinese listeners. The overall perceptual patterns pointed to a stronger CP effect in Chinese listeners.

### 3. The CP index analysis

CP predicts correlations between one's identification and discrimination, yet very few have investigated the interactions between the two domains. Contrast salience and quality of stimuli influence perception (van Hessen & Schouten 1999; Liu & Kager 2014, 2015b). Consonants appear to be more categorically perceived than vowels (Schouten & van Hessen 1992). Since tones are primarily realized on vowels and cannot be processed until vowel information is available (Gandour et al. 2000), CP for tones should be established from a more gradient perspective. In the current study, we adopt the CP index analysis (van Hessen & Schouten 1999) to measure the degree of CP of tones, and to capture the relationship between identification and discrimination.

$$\text{The degree of CP} = [r / (1 + 2 * |p(\text{obt}) - p(\text{pred})|)] * 100$$

By definition, the index value refers to how categorically the target is perceived, the value of which varying from 0 to 100. The numerator  $r$  represents the correlation coefficient between an identification function and a discrimination function: The higher the correlation, the larger the degree of CP. The denominator includes " $p(\text{obt}) - p(\text{pred})$ ", which represents the mean differences between the discrimination function ' $p(\text{obtained})$ ' and the identification function ' $p(\text{predicted})$ ' across data: The smaller the difference, the larger the degree of CP.

The discrimination function was obtained directly from the discrimination tasks. It is the averaged accuracy rates (reflecting discrimination ability) across pairs for each participant. In the current experiment,  $p(\text{obt})$  was the averaged value across six 2-step pairs along the continuum, from pair 1–3 to pair 6–8. The identification function  $p(\text{pred})$  is calculated from the identification differences within the relevant stimuli pairs through “Haskins prediction” formula (Pollack & Pisoni 1971; Cutting & Rosner 1976; Macmillan et al. 1977; Xu et al. 2006):

$$P(c) = 0.5[1 + (p_1 - p_2)^2]$$

In this formula,  $P(c)$  equals  $p(\text{pred})$ , and refers to the probability of a correct discrimination.  $P_1$  is the probability of assigning stimulus A to a particular category, while  $P_2$  is the probability of assigning stimulus B to the same category. Predicted probabilities can thus be obtained from identification data. The mean identification and discrimination index scores representing CP degree for each tone pair are listed in Table 3.

**Table 3.** Mean cross-task CP index score for T1-T4

No.	Task correlation		CP index	
	Identification	Discrimination	Chinese	Dutch
1	Forced choice	AX	41.144	35.974
2	Forced choice	AXB	44.382	20.212
3	Free choice	AX	58.641	52.409
4	Free choice	AXB	48.134	31.482
	T1-T4 Mean		48.075	35.019

Chinese and Dutch participants' CP index scores obtained from identification and discrimination task correlations were compared using a One-way ANOVA. Significant or marginal differences were observed in each of the comparisons between ( $F_{\text{No.1}}(1, 38) = 3.81, p = .058$ ;  $F_{\text{No.2}}(1, 38) = 5.34, p = .026$ ;  $F_{\text{No.3}}(1, 38) = 4.23, p = .047$ ;  $F_{\text{No.4}}(1, 38) = 4.89, p = .033$ ). The degrees of CP of Chinese participants appear to be higher than those of Dutch participants in all of the task correlations. Chinese listeners' performances in identification tasks agree with their behavior in discrimination tasks more than their Dutch peers. The results of the CP index analysis confirm the hypothesis that Chinese listeners perceive tones more categorically than Dutch listeners.

#### 4. General discussion

This paper discusses the degree of tonal CP effect of tone- and non-tone-language listeners. Similar to previous studies, identification and discrimination tasks were used to illustrate participants' perception. Different from these studies, we adopted CP index analysis, originally developed for consonant and vowel perception (van Hessen & Schouten 1999), to the perception of tones. The analysis allows for a more nuanced expression of the degree of CP observed. CP, by definition, means that perception is categorical rather than continuous. The degree of CP may yet remain gradient, possibly influenced by listeners' native language experience (Francis et al. 2008), stimuli acoustic salience (Huang & Johnson 2010; Liu & Kager 2014), and types of tasks (Wu & Lin 2008; Chen et al. 2015, 2016). In the current study, we have demonstrated that Dutch listeners are not tone-deaf, but perceive tones less categorically than their Chinese peers. The gradient nature of tonal processing as a function of the above mentioned factors matches the perceptual development of linguistic pitch (Quam & Swingley 2010), which has been argued to present in virtually all languages lexically, grammatically or on a phrasal level (Myers 1987).

CP predicts better cross-boundary perception and worse within-category perception in native but not in non-native listeners. The results from this study are in line with this prediction and support the view that lexical tones are perceived (more) categorically by Chinese (tone-language) listeners, and (more) psycho-acoustically by Dutch (non-tone-language) listeners. Results of the identification tasks illustrate a high degree of tonal CP among Chinese listeners, whereas Dutch listeners seem to lack tonal categories. In the discrimination tasks, Dutch listeners perform no worse than their Chinese peers in response accuracy. On the one hand, Chinese listeners' perception may be hindered by their native tonal categories when discriminating small acoustic differences within the same category. On the other hand, the lack of categories in Dutch listeners may assist them to perceive the subtle differences more accurately. Dutch listeners outperform their Chinese peers when discriminating the tonal differences in some tasks, even though they do not identify tones in the same manner as Chinese listeners. This is in line with a psycho-acoustic perceptual pattern.

The CP index analysis provides a meaningful interpretation of identification and discrimination tasks, which helps to address the interaction between different modes of tonal perception. As mentioned, the claim of CP solely based on one type of task should not be considered adequate (Abramson 1979; Francis et al. 2003). The convergence between identification and discrimination is crucial for the interpretation of a CP effect. The correlation across the two abilities marks the strength of CP. In the current study, the analysis reveals consistency between

the observed probabilities from the discrimination data and the predicted probabilities from the identification data among Chinese listeners, indicating a strong degree of CP. On the other hand, Dutch participants' identification and discrimination data present less consistency, resulting in a weaker degree of CP than their Chinese peers. Previous studies comparing the CP index for segments demonstrate that the degree of CP is stronger for consonants than for vowels (Schouten & van Hessen 1992). We hypothesize that the degree of CP for tones should be similar to that for vowels for native tone listeners and leave this cross-domain comparison open for future research.

The degree of tonal CP may be stronger in non-tone-language listeners who have more pitch variations in their native languages. Comparing the current outcomes with previous studies, the differences in identification results between Dutch and Chinese participants resemble those between English and Chinese, and between French and Taiwanese listeners (Hallé et al. 2004; Wu & Lin 2008). In all scenarios, tone language listeners outperform their non-tone language peers in tone identification, and demonstrate categorical perception. Non-tone language listeners identify step 1 (T1) without difficulty, but fail to show robust and consistent identification in other steps along the tonal continua. In discrimination tasks, the perceptual pattern of Dutch listeners resembles that of English (Wu & Lin 2008) but not French listeners (Hallé et al. 2004). The tone discrimination peak of French listeners is relatively smooth, whereas English listeners discriminate tones as accurately as Chinese listeners. This may relate to listeners' L1 influence. From a language-general perspective, Dutch and English can be seen as "rich" in speech prosody with abundant stress and intonation patterns, whereas languages like French or Korean are relatively "poor" in prosody (Jun & Fougeron 2000), such as being "deaf" in stress perception. In particular, French does not have lexical stress, and French listeners have been shown to have difficulty encoding contrastive stress in their phonological representation (Dupoux et al. 1997; Dupoux et al. 2008; Peperkamp et al. 2010). From a language specific perspective, prosodic contrasts resembling lexical tones exist in Dutch. At the lexical level, Dutch has minimal pairs marked by stress placement (i.e. /'ka:nɔn/ 'canon' vs. /ka:'nɔn/ 'cannon'). The Dutch nuclear pitch accent exhibits phonetic properties that are similar to Mandarin T2 and T4 (Chen et al. 2014). It has been found that English speakers misperceive Mandarin T4 as T1 far more likely in phrase-final position than elsewhere (Broselow et al. 1987). The authors argue that English listeners may hear the Mandarin T4 in final position disregard the fall, analyzing it as the surface reflex of a low tone associated with the intonation contour of the sentence, and identify what remains when this sentence contour is stripped away as a high tone. Dutch and English have similar declarative tonal contour (H\*L%). Although each pair of tones was presented individually with 500ms interval in the current study, the low

accuracy among Dutch listeners may be due to the interference of Dutch intonation when hearing a tone stimulus by itself, which may arguably represent a phrase final scenario. In brief, listeners' performance may be related to their native categories as well as the cross-language mappings including prosodic knowledge of both the native and target languages (Francis et al. 2008; Liu 2014). Future studies should focus on how different L1 and pitch exposure influences tone perception by testing the same speech stimuli across non-tone-language groups differing in the degree of linguistic functional use of pitch (Wu et al. 2012).

The tone perception patterns have constantly been reported among tone and non-tone language speaking adult listeners (Stagray & Downs 1993). From a developmental angle, it is perhaps more important to investigate and understand how infants' perception develops. The majority of infant experiments target their discrimination ability as reflected by looking time patterns. When habituated in one stimulus, infants' attention shifts or looking time increases upon hearing a novel stimulus indicate discrimination. For tone language learning infants, lexical tones are typically discriminated across ages and language backgrounds. Sensitivity to native and even non-native lexical tonal contrasts (e.g. Cantonese high-rising versus mid-level tones) is maintained in tone language learning Cantonese infants of 4 months, Yorùbá infants of 6 months, and Mandarin Chinese infants from 4 to 9 months (Harrison 2000; Mattock & Burnham 2006; Yeung et al. 2013). Non-tone language learning infants, on the other hand, show a sensitivity decline from 4 to 9 months during language-specific perceptual attunement. Reduced sensitivity to lexical tones in Thai, Cantonese, and Mandarin Chinese has been found in English and Dutch infants at 9 months after birth (Mattock & Burnham 2006; Mattock et al. 2008; Yeung et al. 2013; Liu & Kager 2013, 2014).

Comparing the perceptual patterns between infant and adult listeners across the life span, non-tone-language adult listeners' performance is linked with that of non-tone-learning infants for insights on the nature of tonal perception. In a recent study, Dutch infants' tonal sensitivity is examined using a visual habituation paradigm, revealing their tonal discrimination ability. Specifically, a strong discrimination is found for Dutch infants across ages from 5 to 18 months when discriminating the T1-T4 contrast in Mandarin Chinese, and a U-shaped perceptual pattern for a tonal contrast contracted from T1-T4, with a perceptual decline in tonal sensitivity at 8–9 months, and a perceptual rebound at 17–18 months (Liu & Kager 2014). Importantly, the nature of tone perception, especially at the perceptual rebound stage, remains unclear for non-tone language learning infants. Since typical paradigms target infant discrimination abilities, how infants identify and categorize lexical tones remains unanswered. It is hypothesized that the non-tone language learning infants' perceptual patterns at the perceptual rebound stage resemble that of non-tone-language adult listeners, acoustically instead of

linguistically. In the current experiment, Dutch adults have been shown to be sensitive to the tonal contrasts with even lower acoustic salience especially in the AXB discrimination task. This perceptual pattern is in line with Dutch infants' perception at the tonal perceptual rebound stage (around 17–18 months) after perceptual reorganization (Werker & Tees 1984). At 18 months after birth, infants have established native sound categories and use them to guide word learning (Dietrich et al. 2007). Non-tone-learning infants' ability to use tonal information in word learning deteriorates in the second year of life (Singh et al. 2013; Hay et al. 2015; Liu & Kager, under review). No matter whether such loss is due to a natural decay of linguistic function without tonal exposure from the ambient input, or the deterioration of the ability to abstract and create a tonal proto-category without sufficient input, non-tone-learning infants' psycho-acoustic tonal perceptual pattern is argued to have stabilized by the second year of life, resembling that of non-tone-language adult listeners tested in the current study (Liu & Kager 2014, 2017a). Future research should investigate the perceptual development of pitch in non-tone-language listeners across ages (Liu & Kager 2017b), and the different perception and strategies among non-tone-language listeners (Liu & Kager 2016b).

## Acknowledgements

We dearly thank Annemarie Kerkhoff, Desiree Capel, Elise de Bree, Maartje De Klerk, Theo Veenker, and Xiaoli Dong for their kind help in the experiment; Bert Schouten, Hugo Quene, Jurriaan Witteman, and Pierre Hallé for their assistance in statistics; Bert Schouten, Jingwei Zhang, Tianlin Wang, and Tomas Lentz for their valuable feedback. We sincerely thank *Language and Linguistics* anonymous reviewers for their valuable comments on this paper.

## References

- Abramson, Arthur S. 1979. The noncategorical perception of tone categories in Thai. *Frontiers of Speech Communication*, ed. by B. Lindblom & S. Öhman, 127–134. London: Academic Press.
- Best, Catherine T., & Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20.3: 305–330.
- Boersma, Paul, & David Weenink. 2012. *Praat: Doing Phonetics by Computer* [Computer program] version 5.5. Available at <http://www.praat.org/> (accessed 1 May 2016).
- Broselow, Ellen, Richard R. Hurtig, & Catherine Ringen. 1987. The perception of second language prosody. *Interlanguage Phonology: The Acquisition of A Second Language Sound System*, ed. by Georgette Ioup & Steven H. Weinberger, 350–361. Cambridge: Newbury House.
- Brown-Schmidt, Sarah, & Enriqueta Canseco-Gonzalez. 2004. Who do you love, your mother or your horse? An event-related brain potential analysis of tone processing in Mandarin Chinese. *Journal of Psycholinguist Research* 33.2: 103–135.  
doi: 10.1023/B:JOPR.0000017223.98667.10

- Chen, Ao, Aoju Chen, René Kager, & Patrick Wong. 2014. Rises and Falls in Dutch and Mandarin Chinese. Paper presented at The 4th International Symposium on Tonal Aspects of Languages, May 13–16, 2014. Nijmegen: Radboud University Nijmegen & the Max Planck Institute for Psycholinguistics.
- Chen, Ao, Liqian Liu, & René Kager. 2015. Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences* 48: 62–69. doi: 10.1016/j.langsci.2014.12.002
- Chen, Ao, Liqian Liu, & René Kager. 2016. Cross-domain correlation in pitch perception, the influence of native language. *Language, Cognition and Neurosciences* 31.6: 751–760. doi: 10.1080/23273798.2016.1156715
- Content, Alain, & Noémi Perwez. 2011. Categorical perception of tones in Vietnamese. Paper presented at the 16th International Congress of Phonetic Sciences, August 6–10, 2007. Saarbrücken: Saarland University.
- Cutting, James E., & Burton S. Rosner. 1976. Discrimination functions predicted from categories of speech and music. *Perception & Psychophysics* 20.1: 87–88. doi: 10.3758/BF03198712
- Dietrich, Christiane, Daniel Swingle, & Janet F. Werker. 2007. Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences* 104.41: 16027–16031. doi: 10.1073/pnas.0705270104
- Dupoux, Emmanuel, Christophe Pallier, Núria Sebastián-Gallés, & Jacques Mehler. 1997. A destressing “deafness” in French? *Journal of Memory and Language* 36: 406–421. doi: 10.1006/jmla.1996.2500
- Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete, & Sharon Peperkamp. 2008. Persistent stress ‘deafness’: the case of French learners of Spanish. *Cognition* 106.2: 682–706. doi: 10.1016/j.cognition.2007.04.001
- Francis, Alexander L., Valter Ciocca, & Brenda Kei Chit Ng. 2003. On the (non)categorical perception of lexical tones. *Perception & Psychophysics* 65.7: 1029–1044. doi: 10.3758/BF03194832
- Francis, Alexander L., Valter Ciocca, Lian Ma, & Kimberly Fenn. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics* 36.2: 268–294. doi: 10.1016/j.wocn.2007.06.005
- Gandour, Jackson T. 1978. The perception of tone. *Tone: A Linguistic Survey*, ed. by Victoria A. Fromkin, 41–76. New York: Academic Press. doi: 10.1016/B978-0-12-267350-4.50007-8
- Gandour, Jackson T., Donald Wong, & Gary Hutchins. 1998. Pitch processing in the human brain is influenced by language experience. *NeuroReport* 9.9: 2115–2119. doi: 10.1097/00001756-199806220-00038
- Gandour, Jackson T., Donald Wong, Li Hsieh, Bret Weinzapfel, Diana Van Lancker, & Gary D. Hutchins. 2000. A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience* 12.1: 207–222. doi: 10.1162/089892900561841
- Gandour, Jackson T., Mario Dzemidzic, Donald Wong, Mark Lowe, Yunxia Tong, Li Hsieh, & Joseph Lurito. 2001. Differential neural circuitry underlies the perception of Chinese tone and intonation: an fMRI study. *NeuroImage* 13.6: 533. doi: 10.1016/S1053-8119(01)91876-2
- Hallé, Pierre A., Yueh-Chin Chang, & Catherine T. Best. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32.3: 395–421. doi: 10.1016/S0095-4470(03)00016-0
- Harnad, Steven. 1987. *Categorical Perception: The Groundwork of Cognition*. New York: Cambridge University Press.
- Harrison, Phil. 2000. Acquiring the phonology of lexical tone in infancy. *Lingua* 110.8: 581–616. doi: 10.1016/S0024-3841(00)00003-6

- Hay, Jessica F., Katharine Graf Estes, Tianlin Wang, & Jenny R. Saffran. 2015. From flexibility to constraint: the contrastive use of lexical tone in early word learning. *Child Development* 86.1: 10–22. doi: 10.1111/cdev.12269
- Hsieh, Li, Jack Gandour, Donald Wong, & Gary D. Hutchins. 2000. A PET study of the perception of Chinese tones. *NeuroImage* 11.5: 313. doi: 10.1016/S1053-8119(00)91245-X
- Huang, Tsan, & Keith Johnson. 2010. Language specificity in speech perception: perception of Mandarin tones by native and nonnative listeners. *Phonetica* 67.4: 243–267. doi: 10.1159/000327392
- Jun, Sun-Ah, & Cécile Fougeron. 2000. A phonological model of French intonation. *Intonation: Analysis, Modeling and Technology*, ed. by Antonis Botinis, 209–242. Dordrecht: Kluwer. doi: 10.1007/978-94-011-4317-2\_10
- Kaan, Edith, Christopher M. Barkley, Mingzhen Bao, & Ratree Wayland. 2008. Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: an event-related potentials training study. *BMC Neuroscience* 9: 53. doi: 10.1186/1471-2202-9-53
- Klein, Denise, Robert J. Zatorre, Brenda Milner, & Viviane Zhao. 2001. A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage* 13.4: 646–653. doi: 10.1006/nimg.2000.0738
- Kuhl, Patricia K., Feng-Ming Tsao, & Huei-Mei Liu. 2003. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences* 100.15: 9096–9101. doi: 10.1073/pnas.1532872100
- Lee, Chao-Yang. 2007. Does horse activate mother? Processing lexical tone in form priming. *Language and Speech* 50.1: 101–123. doi: 10.1177/00238309070500010501
- Lee, Lisa, & Howard C. Nusbaum. 1993. Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception and Psychophysics* 53.2: 157–165. doi: 10.3758/BF03211726
- Lieberman, A. M., Katherine S. Harris, Jo Ann Kinney, & H. Lane. 1961. The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology* 61.5: 379–388. doi: 10.1037/h0049038
- Liu, Liqun. 2014. *The Effects of Bilingualism on Infant Language Development: The Acquisition of Sounds and Words*. The Netherlands: LOT.
- Liu, Liqun, & René Kager. 2011. Is perceptual reorganization affected by statistical learning? Dutch infants' sensitivity to lexical tones. *Proceedings of the 35th Annual Boston University Conference on Language Development*, ed. by Nick Danis, Kate Mesh & Hyunsuk Sung, 404–413. Somerville: Cascadilla Press.
- Liu, Liqun, & René Kager. 2013. How bilingualism alters non-tone-learning infants' tone perception in the first year of life. *Proceedings of the 37th Annual Boston University Conference on Language Development*, ed. by Sarah Baiz, Nora Goldman & Rachel Hawkes, 231–240. Somerville: Cascadilla Press.
- Liu, Liqun, & René Kager. 2014. Perception of tones by infants learning a non-tone language. *Cognition* 133.2: 385–394. doi: 10.1016/j.cognition.2014.06.004
- Liu, Liqun, & René Kager. 2015a. Bilingual exposure influences infant VOT perception. *Infant Behavior and Development* 38: 27–36. doi: 10.1016/j.infbeh.2014.12.004
- Liu, Liqun, & René Kager. 2015b. Understanding phonological acquisition through phonetic perception: the influence of exposure and acoustic salience. *Phonological Studies* 18: 51–58.
- Liu, Liqun, & René Kager. 2016a. Perception of a native vowel contrast by Dutch monolingual and bilingual infants: a bilingual perceptual lead. *International Journal of Bilingualism* 20.3: 335–345. doi: 10.1177/1367006914566082

- Liu, Liquan, & René Kager. 2016b. Is mommy talking to daddy or to me? Exploring parental estimates of child language exposure using the multilingual infant questionnaire. *International Journal of Multilingualism* 1–12. doi: 10.1080/14790718.2016.1216120
- Liu, Liquan, & René Kager. 2017a. Perception of tones by bilingual infants learning non-tone languages. *Bilingualism: Language and Cognition* 20.3: 561–575. doi: 10.1017/S1366728916000183
- Liu, Liquan, & René Kager. 2017b. Enhanced music sensitivity in 9-month-old bilingual infants. *Cognitive Processing* 18.1: 55–56. doi: 10.1007/s10339-016-0780-7
- Liu, Liquan, & René Kager. Under review. The associative word learning of tones by monolingual and bilingual infants learning non-tone languages. *Frontiers in Psychology*.
- Luo, Hao, Jing-Tian Ni, Zhi-Hao Li, Xiao-Ou Li, Da-Ren Zhang, Fan-Gang Zeng, & Lin Chen. 2006. Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences* 103.51: 19558–19563. doi: 10.1073/pnas.0607065104
- Macmillan, Neil A., Howard L. Kaplan, & C. Douglas Creelman. 1977. The Psychophysics of categorical perception. *Psychological Review* 84.5: 452–471. doi: 10.1037/0033-295X.84.5.452
- Mattock, Karen, & Denis Burnham. 2006. Chinese and English infants' tone perception: evidence for perceptual reorganization. *Infancy* 10.3: 241–265. doi: 10.1207/s15327078in1003\_3
- Mattock, Karen, Monika Molnar, Linda Polka, & Denis Burnham. 2008. The developmental course of lexical tone perception in the first year of life. *Cognition* 106.3: 1367–1381. doi: 10.1016/j.cognition.2007.07.002
- Myers, Scott P. 1987. *Tone and the Structure of Words in Shona*. Amherst: University of Massachusetts Amherst dissertation.
- Peng, Gang, Hong-Ying Zheng, Tao Gong, Ruo-Xiao Yang, Jiang-Ping Kong, & William S.-Y. Wang. 2010. The influence of tone language experience on categorical perception of pitch contours. *Journal of Phonetics* 38.4: 616–624. doi: 10.1016/j.wocn.2010.09.003
- Peperkamp, Sharon, Inga Vendelin, & Emmanuel Dupoux. 2010. Perception of predictable stress: a cross-linguistic investigation. *Journal of Phonetics* 38.3: 422–430. doi: 10.1016/j.wocn.2010.04.001
- Pollack, Irwin, & David Pisoni. 1971. On the comparison between identification and discrimination tests in speech perception. *Psychonomic Science* 24.6: 299–300. doi: 10.3758/BF03329012
- Quam, Carolyn, & Daniel Swingley. 2010. Phonological knowledge guides two-year-olds' and adults' interpretation of salient pitch contours in word learning. *Journal of Memory and Language* 62.2: 135–150. doi: 10.1016/j.jml.2009.09.003
- Repp, Bruno H., & Hwei-Bing Lin. 1990. Integration of segmental and tonal information in speech perception: a cross-linguistic study. *The Journal of the Acoustical Society of America* 87.1: 46. doi: 10.1121/1.2028239
- Schouten, M. E. H., & A. J. van Hessen. 1992. Modeling phoneme perception. I: Categorical perception. *Journal of the Acoustical Society of America* 92.4: 1841–1855. doi: 10.1121/1.403841
- Singh, Leher, Tam Jun Hui, Calista Chan, & Roberta Michnick Golinkoff. 2013. Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Developmental Science* 17.1: 94–109.
- Soto-Faraco, Salvador, Núria Sebastián-Gallés, & Anne Cutler. 2001. Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language* 45.3: 412–432. doi: 10.1006/jmla.2000.2783

- Stagray, James R., & David Downs. 1993. Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics* 21.1: 143–163.
- Sun, Kuo-Chan, & Tsan Huang. 2012. A cross-linguistic study of Taiwanese tone perception by Taiwanese and English listeners. *Journal of East Asian Linguistics* 21.3: 305–327. doi: 10.1007/s10831-012-9092-9
- van Hessen, A. J., & M. E. H. Schouten. 1999. Categorical perception as a function of stimulus quality. *Phonetica* 56: 56–72. doi: 10.1159/000028441
- Van Lancker, Diana, & Victoria A. Fromkin. 1973. Hemispheric specialization for pitch and “tone”: evidence from Thai. *Journal of Phonetics* 1: 101–109.
- Veenker, Theo J. G. 1998. *FEP, a tool for designing and running computerized experiments* [Computer software]. Utrecht: Utrecht Institute of Linguistics OTS, Utrecht University.
- Wang, William S.-Y. 1973. The Chinese language. *Scientific American* 228: 50–60. doi: 10.1038/scientificamericano273-50
- Wang, William S.-Y. 1976. Language Change. *Annals of the New York Academy of Sciences* 280.1: 61–72. doi: 10.1111/j.1749-6632.1976.tb25472.x
- Wang, Yue, Allard Jongman, & Joan A. Sereno. 2001. Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language* 78.3: 332–348. doi: 10.1006/brln.2001.2474
- Wang, Yue, Dawn M. Behne, Allard Jongman, & Joan A. Sereno. 2004. The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics* 25.3: 449–466.
- Werker, Janet F., & Richard C. Tees. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7.1: 49–63. doi: 10.1016/S0163-6383(84)80022-3
- Whalen, D. H., & Yi Xu. 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49.1: 25–47. doi: 10.1159/000261901
- Wu, Xianghua, & Hua Lin. 2008. Perception of Mandarin tones by Mandarin and English listeners. *Journal of Chinese Language and Computing* 18.4: 175–187.
- Wu, Xianghua, Jung-Yueh Tu, & Yue Wang. 2012. Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics* 33.3: 623–641. doi: 10.1017/S0142716411000506
- Xu, Yisheng, Jackson T. Gandour, & Alexander L. Francis. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America* 120.2: 1063–1074. doi: 10.1121/1.2213572
- Yeung, H. Henny, Ke Heng Chen, & Janet F. Werker. 2013. When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language* 68.2: 123–139. doi: 10.1016/j.jml.2012.09.004

## Further readings

- Henderson, A. I., & S. Nelms. 1980. Relative salience of intonation fall and pause as cues to the perceptual segmentation of speech in an unfamiliar language. *Journal of Psycholinguistic Research* 9.2: 147–159. doi: 10.1007/BF01067468
- Liu, Liquan, & René Kager. In press. Dutch infants’ lexical tone perception under distributional learning. *Journal of Child Experimental Psychology*.

**Appendix. F0 value in 4 interpolate of the stimuli along /ta1/-/ta4/ continua**

In these tables, /ta1/ = stimulus 1 while /ta4/ = stimulus 8. The rest numbers correspond to the stimuli created along the continua. The four letters (A, B, C and D) represent four continua created from different original /ta1/ - /ta4/ sounds produced by the same speaker.

## Continuum A

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
A1	288.300	289.000	289.000	290.000
A2	295.614	291.857	275.700	271.484
A3	302.928	294.714	262.425	252.970
A4	310.242	297.571	249.140	234.456
A5	317.556	300.428	235.855	215.942
A6	324.870	303.285	222.570	197.428
A7	332.184	306.142	209.285	178.914
A8	339.500	309.000	196.000	160.400

## Continuum B

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
B1	275.500	277.633	279.766	281.900
B2	283.582	282.443	267.202	264.592
B3	291.664	287.253	255.635	247.281
B4	299.746	292.063	244.068	229.970
B5	307.828	296.873	232.501	212.659
B6	315.910	301.683	221.934	195.348
B7	323.992	306.493	210.367	178.037
B8	332.077	311.300	198.800	160.726

## Continuum C

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
C1	278.795	276.670	274.525	272.390
C2	293.044	284.788	259.178	255.264
C3	307.293	292.906	243.829	238.136
C4	321.542	301.024	228.480	221.008
C5	335.791	309.142	213.131	203.880
C6	350.040	317.260	197.782	186.752
C7	364.289	325.378	182.433	169.624
C8	378.537	333.494	167.084	152.496

## Continuum D

	Starting Point	Interpolate 1	Interpolate 2	Ending Point
D1	272.502	272.826	273.150	273.474
D2	287.777	278.903	259.092	256.802
D3	303.052	284.980	245.032	240.133
D4	318.327	291.057	230.972	223.464
D5	333.602	297.134	216.912	206.795
D6	348.877	303.211	202.852	190.126
D7	364.152	309.288	188.792	173.457
D8	379.430	315.366	174.732	156.788

*Authors' addresses*

Liquan Liu  
 School of Social Sciences and Psychology  
 Western Sydney University  
 56 Second Avenue, DEMC Babylab  
 Sydney, New South Wales  
 Australia

[l.liu@westernsydney.edu.au](mailto:l.liu@westernsydney.edu.au)

**Publication history**

Date received: 10 November 2015

Date accepted: 2 May 2016