Equations for generating IRN stimuli

T2

```
F0 = 103 -  (8.45/0.25).*(0:1/fs:0.25)  -  (76.32/(0.25^2)).*(0:1/fs:0.25).^2  +  (297.91/(0.25^3)).*(0:1/fs:0.25).^3  - (185.34/(0.25^4)).*(0:1/fs:0.25).^4
```

T2’

```
X = (0: 1/fs: 0.25)';

T2 = [103 - (8.45/0.25).*x - (76.32/ (0.25^2)).*x. ^2 + (297.91/ (0.25^3)).*x. ^3 - (185.34/ (0.25^4)).*x. ^4]';

Y = [(T2 (round (length(x)* 0.25/max(x)))-T2 (1))/ 0.25].*[0:1/fs: 0.25] +T2 (1);

Temp = T2 (1: length(y));

F0 = y + y-temp;
```

T4

```
F0 = 130.8 + (8.45/0.25).*(0 : 1/fs : 0.25)  + (76.32/ (0.25^2)).*(0 : 1/fs : 0.25).^2  - (297.91/(0.25^3)).*(0 : 1/fs : 0.25).^3 + (185.34/ (0.25^4)).*(0 : 1/fs : 0.25).^4;
```

Audio of IRN stimuli and stimulus condition

T2 *irnpitchT2.mp3*

T2' *irnpitchT2'.mp3*

T4 *irnpitchT4.mp3*

NoisetopitchT2 *irnnoisetopitchT2.mp3*

Results

Table S1. Fz-T7/T8 (linked) peak latencies in response to stimuli as a function of component and group

	Stimulus					
	T2		T2'		T4	
Group	C	E	C	E	C	E
Component						
Na	134(1.1)	139(1.9)	142(2.1)	158(2.3)	129(1.5)	143(1.9)
Pb	204(1.5)	203(2.8)	223(2.3)	230(3.2)	194(2.1)	196(2.1)
Nb	268(1.1)	273(3.2)	264(1.9)	274(2.2)	264(1.5)	272(2.7)

Note: Values are expressed in milliseconds as mean and standard error in parentheses (±1 SE).

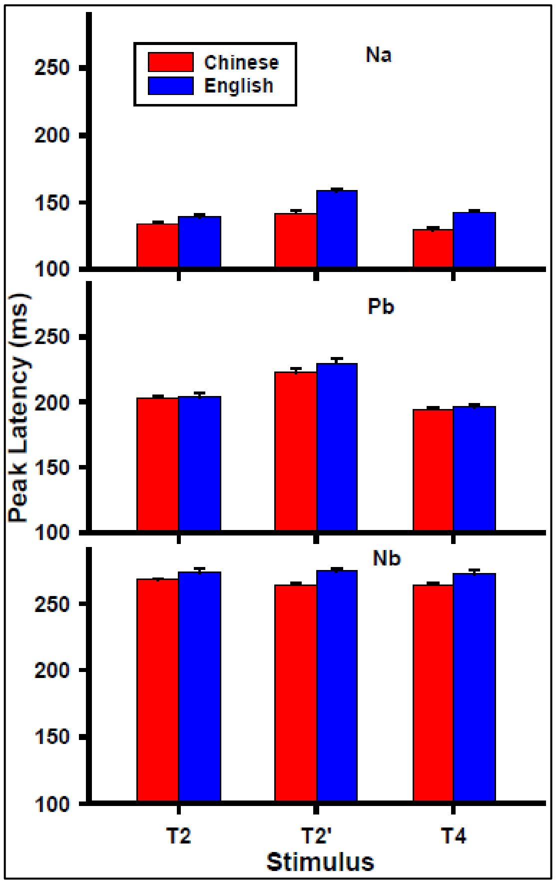


Figure S1a.

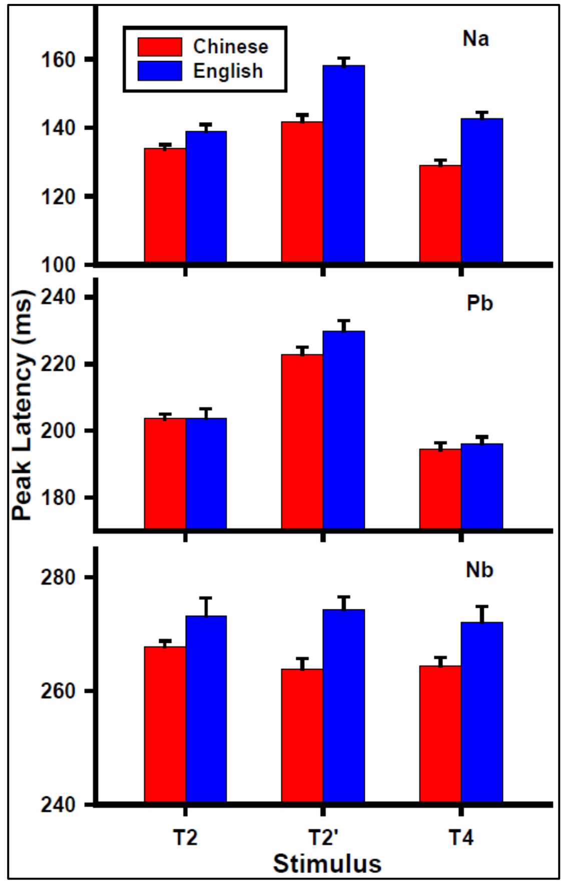
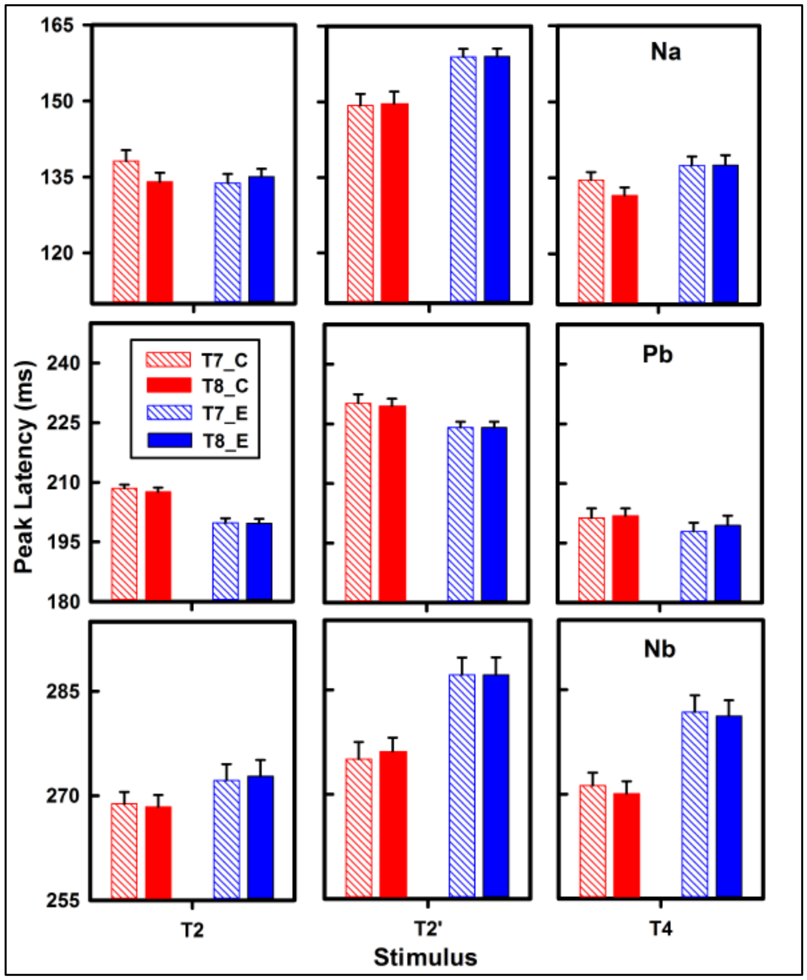


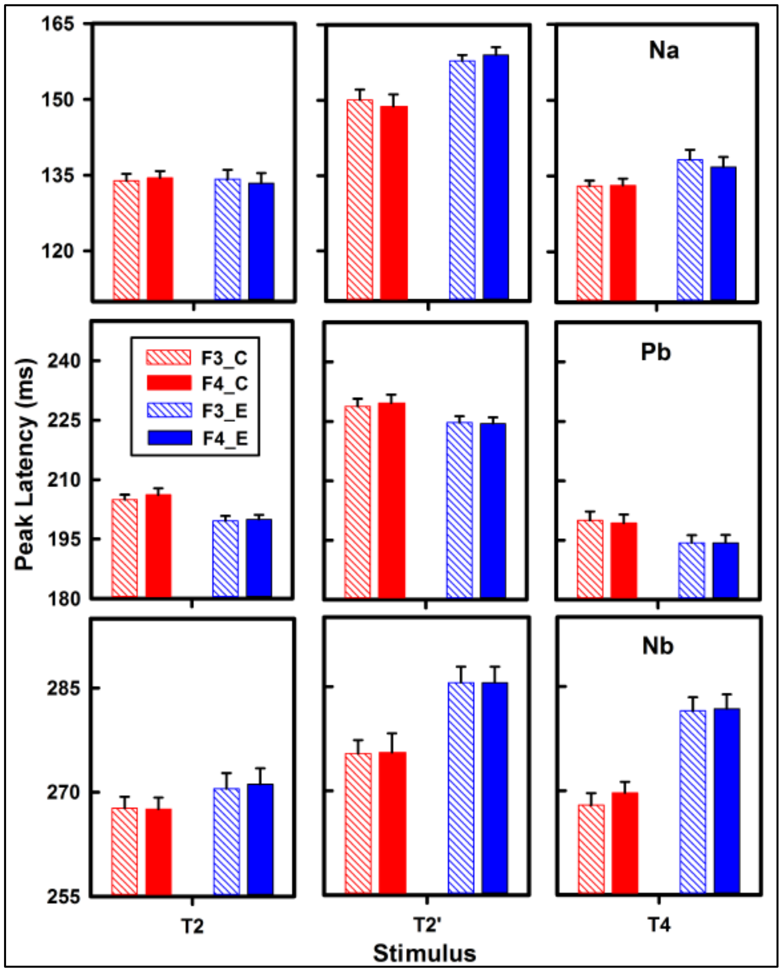
Figure S1b.

Mean peak latency of CPR components (Na, Pb, Nb) extracted from Fz as a function of stimulus (Fig. S1a: same Y-axis scale across stimuli; Fig. S1b: different Y-axis scales across stimuli). For both language groups, mean Fz peak latencies of Na, Pb, and Nb increase systematically across stimuli in temporal order of occurrence (S1a). As indexed by Na (S1b), T2', the pitch stimulus exhibiting an early peak of acceleration, elicits longer peak latencies than either T2 or T4 in both groups. By stimulus, response peak latencies for T2' and T4 are longer in English than Chinese. As indexed by Pb (S1b), both groups exhibit a longer latency evoked by T2' than either T2 or T4, i.e., longer latency to a pitch contour with an early acceleration peak (T2') relative to those with a late peak (T2, T4). As indexed by Nb (S1b), the latency in the English is longer than the Chinese group across stimuli. Error bars = ±1 SE.

T7/T8. As indexed by Na over T7/T8 electrode sites (Fig. S2), the omnibus ANOVA yielded a group x hemisphere x stimulus interaction ($F_{2,52} = 4.23$, $P = 0.0198$). Simple effects of group and hemisphere revealed that peak latency of T2' was longer than T2 or T4 across hemispheres for both Chinese (C+LH, T2 vs T2': $t_{52} = -4.54$, $P = 0.0004$; C+LH, T2'vs T4: $t = 6.01$, $P < 0.0001$; C+RH, T2 vs T2': $t = -6.34$, $P < 0.0001$; C+RH, T2' vs T4: $t = 7.38$, $P < 0.0001$) and English (E+LH, T2 vs T2': $t = -10.22$; E+LH, T2'vs T4: $t = 8.74$; E+RH, T2 vs T2': $t = -9.75$; E+RH, T2'vs T4: $t = 8.76$, $P < 0.0001$). Across hemispheres, T2' evoked a longer latency in English than Chinese (LH+T2', C vs E: $t = -3.67$, $P = 0.0034$; RH+T2', C vs E: $t = -3.67$, $P = 0.0047$). These findings on group and stimulus for Na are consistent with those obtained from the Fz-T7/T8 (linked) site. As indexed by Pb, the ANOVA yielded main effects of group ($F_{1,26} = 14.79$, $P = 0.0007$) and stimulus ($F_{2,52} = 164.36$, $P < 0.0001$). The findings on stimulus for Pb agree with those obtained from the Fz-T7/T8 (linked) site (T2' > T2, T4). The group effect, however, shows that Chinese has a longer latency than English (C > E), contrary to that obtained from the Fz-T7/T8 (linked) site. As indexed by Nb, the ANOVA similarly yielded main effects of group ($F_{1,26} = 20.58$, $P < 0.0001$) and stimulus ($F_{2,52} = 14.53$, $P < 0.0001$). The findings on stimulus (T2' > T2, T4) converge with those obtained across components over both Fz-T7/T8 (linked) and T7/T8 electrode sites. In the



case of Nb, the group effect agrees with that of the Fz-T7/T8 (linked) (E > C), but is opposite to that obtained from Pb (C > E). The group effect (C > E), pooled across stimuli and hemisphere for T7/T8, is difficult to interpret. As indexed by Pb, we speculate that Chinese listeners may be using a longer integration window to more accurately capture the dynamic, perceptually-salient portions of the pitch contours. This view is consistent with our earlier hypothesis that Pb indexes the more dynamic portions of the pitch contours (Krishnan *et al.*, 2014). However, it is not clear why this integration time would be extended to a nonnative pitch contour such as T2'.



F3/F4. As indexed by Na over F3/F4 electrode sites (Fig. S3), the omnibus ANOVA yielded a group x stimulus interaction ($F_{2,52}$ = 4.81, P = 0.0121). Simple effects of group and revealed that peak latency of T2' was longer than T2 or T4 for both Chinese (T2 vs T2': t_{52} = -7.13; T2' vs T4: t = 7.65, P < 0.0001) and English (T2 vs T2': t = -11.51; T2' vs T4: t = 9.79, P < 0.0001). English latency was longer than Chinese in response to T2' (t = -3.84, P < 0.0003). Across hemispheres, T2' evoked a longer latency in English than Chinese (LH+T2', C vs E: t = -3.67, P = 0.0034; RH+T2', C vs E: t = -3.67, P = 0.0047). These findings on group and stimulus for Na are consistent with those obtained from the Fz-T7/T8 (linked) site. As indexed by Pb, the ANOVA yielded main effects of group ($F_{1,26}$ = 10.89, P = 0.0028) and stimulus ($F_{2,52}$ = 187.69, P < 0.0001). The Pb findings on peak latency for both group (C > E: t_{26} = 3.30, P = 0.0028) and stimulus (T2' > T2: t_{52} = -14.80; T2' > T4 t = -

14.80, $P < 0.0001$) are consistent with those obtained from the Fz-T7/T8 (linked) site. As indexed by Nb, the ANOVA showed a marginally significant group x stimulus interaction ($F_{2,52} = 3.32$, $P = 0.0438$). By group, Chinese peak latency evoked by T2' was longer than T2 ($t_{52} = -2.85$, $P = 0.0188$); English peak latency, on the other hand, was longer for T4 as compared to T2 ($t = -3.96$, $P = 0.0007$). Although the group trends (C > E) are similar to T7/T8 with respect to Pb latency the weak interaction would suggest that observed differences in peak latency between language groups is more robust over the temporal electrode sites.

Discussion

Effects of acoustic properties of stimuli in early cortical pitch processing

From human psychophysical experiments, the just-noticeable difference for change in frequency of sweep-tone stimuli are smaller for rising than falling tones (Shore *et al.*, 1987) and detection thresholds are lower for rising than falling tones (Nabelek, 1978; Cullen & Collins, 1979). Discrimination of falling sweeps require longer duration and/or higher sweep rates than rising sweeps (Schouten, 1985). Detection of changes in slope of linear F_0 ramps show greatest sensitivity when one ramp is rising and the other is falling (Klatt, 1973). Using multidimensional scaling, the underlying perceptual dimension related to pitch direction in the stimulus space separates primarily rising from non-rising pitch contours (Gandour, 1983). From electrophysiological experiments at the level of the auditory brainstem, amplitude of the frequency following response is smaller for falling than rising tonal sweeps (Krishnan & Parkinson, 2000); pitch strength is greater for Mandarin rising than non-rising tones (Krishnan *et al.*, 2004). Similar selectivity to rising tonal sweeps is reported for cochlear microphonics (Shore & Cullen, 1984), eighth nerve compound action potentials (Shore & Nuttall, 1985), and responses of the ventral cochlear nucleus units (Shore *et al.*, 1987).

Fig. S4 Grand average waveforms (left) and their corresponding spectra (right) of the CPR components for the two language groups (Chinese; English) recorded at electrode sites F3 (dashed) and F4 (solid) for each of the three stimuli (T2, T2’, T4). The waveforms reveal that regardless of language group or stimulus, pitch-related components at frontal sites essentially overlap between F3 and F4 with no discernible difference in magnitude (left) and show essentially identical spectrotemporal plots (right). There is no evidence of a hemispheric preference. The zero on the x-axis denotes the time of onset of the pitch-eliciting segment of the three stimuli.

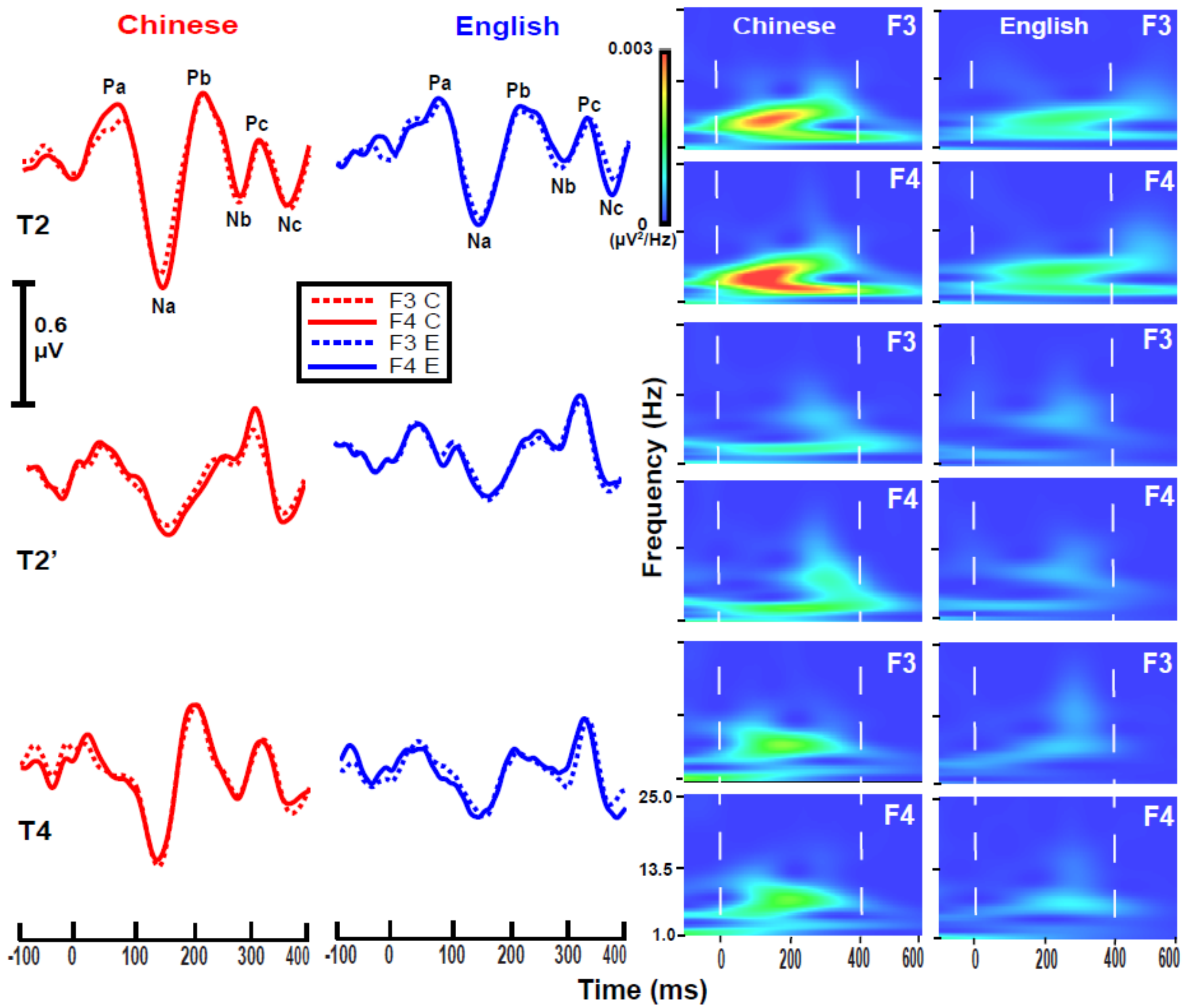
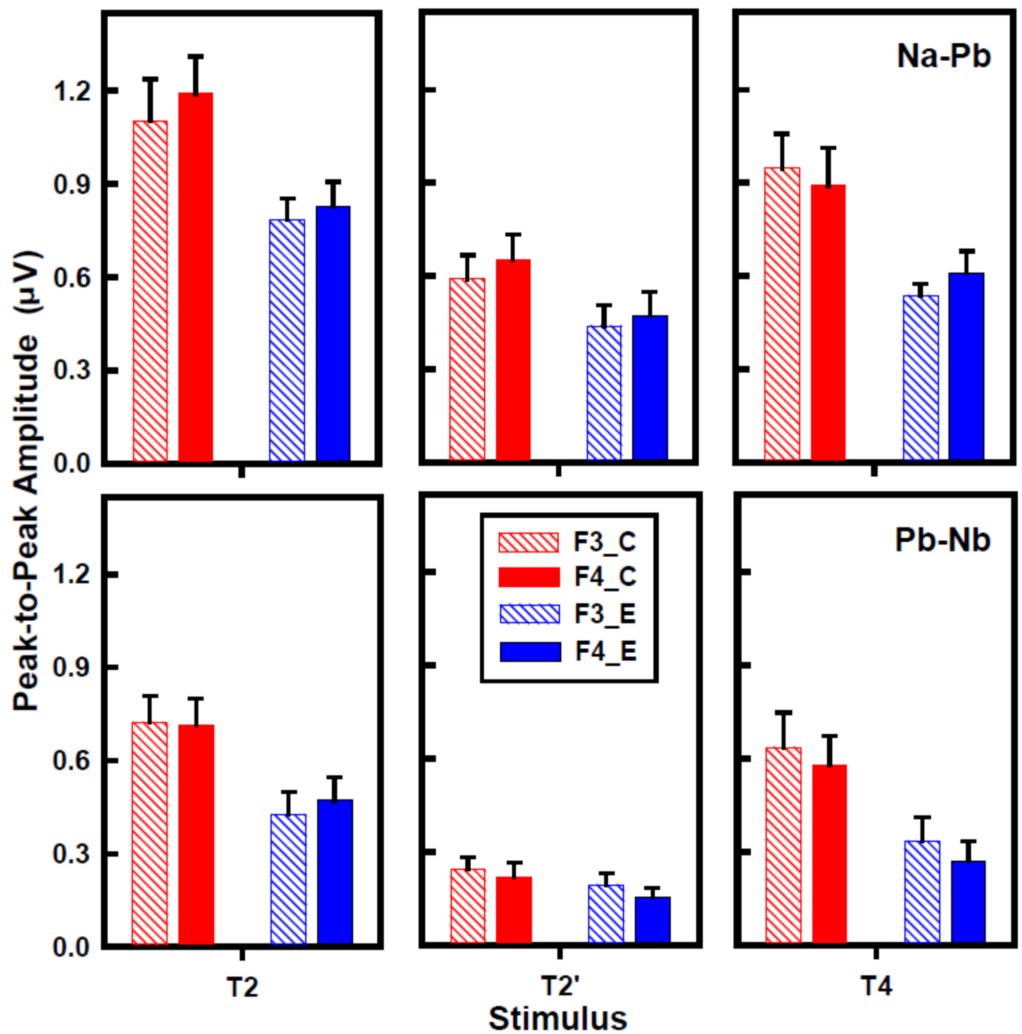


Fig. S5 Mean peak-to-peak amplitude of CPR components (Na-Pb, top row; Pb-Nb, bottom row) extracted from F3/F4 in the frontal lobe as a function of language group (Chinese; English), stimulus (T2, T2’, T4), and hemisphere (left, diagonal; right, solid). No laterality effects are observed for either group as reflected by Na-Pb and Pb-Nb. Error bars = ±1 SE. C, Chinese; E, English.



References

Cullen, J.K. & Collins, M.J. (1979) Rate effects in the detection of short duration tonal glides. *J. Acoust. Soc. Am.*, **65**, S38.

Gandour, J.T. (1983) Tone perception in Far Eastern languages. *J. Phonetics*, **11**, 149-175.

Klatt, D. (1973) Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *J. Acoust. Soc. Am.*, **53**, 8-16.

Krishnan, A., Gandour, J.T., Ananthakrishnan, S. & Vijayaraghavan, V. (2014) Cortical pitch response components index stimulus onset/offset and dynamic features of pitch contours. *Neuropsychologia*, **59**, 1-12.

Krishnan, A. & Parkinson, J. (2000) Human frequency-following response: representation of tonal sweeps. *Audiol. Neurootol.*, **5**, 312-321.

Krishnan, A., Xu, Y., Gandour, J.T. & Cariani, P.A. (2004) Human frequency-following response: representation of pitch contours in Chinese tones. *Hear. Res.*, **189**, 1-12.

Nabelek, I.V. (1978) Temporal summation of constant and gliding tones at masked auditory threshold. *J. Acoust. Soc. Am.*, **64**, 751-763.

Schouten, M.E. (1985) Identification and discrimination of sweep tones. *Percept. Psychophys.*, **37**, 369-376.

Shore, S.E., Clopton, B.M. & Au, Y.N. (1987) Unit responses in ventral cochlear nucleus reflect cochlear coding of rapid frequency sweeps. *J. Acoust. Soc. Am.*, **82**, 471-478.

Shore, S.E. & Cullen, J.K., Jr. (1984) Cochlear microphonic responses of the peripheral auditory system to frequency-varying signals. *Am. J. Otolaryngology*, **5**, 34-42.

Shore, S.E. & Nuttall, A.L. (1985) High-synchrony cochlear compound action potentials evoked by rising frequency-swept tone bursts. *J. Acoust. Soc. Am.*, **78**, 1286-1295.