NEUROBIOLOGIÆ EXPERIMENTALIS

Short communication Acta Neurobiol Exp 2014, 74: 98–103

Auditory perception of temporal order: A comparison between tonal language speakers with and without non-tonal language experience

Yan Bao[1,2,4,5]*, Yuan Fang[1], Taoxi Yang[1], Lingyan Wang[1,3], Aneta Szymaszek[6,7], and Elzbieta Szelag[4,6,7]

[1]Department of Psychology, *Email: baoyan@pku.edu.cn, [2]Key Laboratory of Machine Perception (MoE), [3]College of Life Science, Peking University, Beijing, P.R. China; [4]Human Science Center, Ludwig-Maximilian University of Munich, München, Germany; [5]Parmenides Center for Art and Science, Pullach, Germany; [6]Laboratory of Neuropsychology, Nencki Institute of Experimental Biology, Warsaw, Poland; [7]University of Social Sciences and Humanities, Warsaw, Poland

It has been shown recently that temporal order perception is modulated by language environments. The present study focused on the specific question whether a secondary language experience influences temporal order perception by comparing the temporal order thresholds (TOTs) between Chinese subjects with and without a secondary non-tonal language (i.e., English) experience. Besides monaurally presented paired clicks, binaurally presented two different types of tone pairs were used in order to better capture a potential difference between tonal and non-tonal languages. The results showed a non-significant language effect on monaurally presented click TOTs, but a significant language effect for binaurally presented tone TOTs. Compared to click performance, Chinese subjects without English proficiency demonstrated a significantly lower TOT only for close frequency tone pairs, while Chinese subjects with English proficiency demonstrated lower TOTs for both close frequency and distant frequency tone pairs. These results confirm on the one hand a common and language independent temporal mechanism for perceiving the order of two monaurally presented stimuli, and indicate on the other hand specific mechanisms of neuronal plasticity for perceiving the order of frequency-related auditory stimuli for tonal language speakers with or without a secondary non-tonal language experience.

Key words: time perception, auditory processing, order threshold, language experience

One challenge for the human brain to understand speech is to extract the temporal order of speech sounds like different consonants and/or vowels, thus providing the operative basis for identifying phonemes, syllables, words and sentences (Pöppel 2009, Bao et al. 2013). The close link between temporal processing and language comprehension has been suggested on a theoretical level quite some time ago (Pöppel 1971, Martin 1972), and it is well established by a number of studies (e.g. Albert and Bear 1974, Tallal 2004, Szelag et al. 2011a). Evidence for such a link comes for instance from clinical studies in which parallel deficit in language and timing were reported. Deficits in speech perception at the phoneme level such as in patients with left-hemispheric lesions and aphasia (von Steinbüchel et al. 1999, Wittmann et al. 2004, Sidiropoulos et al. 2010), in children with specific language impairment (Tallal et al. 1998, Fitch and Tallal 2003), and in children and adults with dyslexia (Fink et al. 2006, Vandermosten et al. 2011, see also Szelag et al. 2010 for a review) are often associated with difficulties in processing rapidly changing auditory signals as indexed by an increased temporal order threshold (TOT) for discriminating the sequence of successively presented acoustic stimuli.

In addition to the evidence showing the connection between speech processing and temporal perception, it is also suggested that native language experience may influence auditory information processing at both behavioral and neuronal levels. Compared with English subjects, Chinese subjects who are embedded in the tonal language environment perform significantly bet-

Correspondence should be addressed to Y. Bao,
Email: baoyan@pku.edu.cn

Received 04 October 2013, accepted 28 November 2013

ter in frequency-modulated identification task (Luo et al. 2007). Furthermore, native Chinese speakers are found to be more sensitive to tone contours while native English speakers are more sensitive to early pitch differences in discriminating lexical tones (Kann et al. 2007). Brain stem responses also revealed that pitch-tracking accuracy of whole tones and pitch strength of 40-ms tonal sections were generally higher in subjects with the tonal language experience (Chinese and Tai) compared to subjects with non-tonal language experience (English) (Krishnan et al. 2010).

In a recent study (Szelag et al. 2011b) on cross-linguistic comparisons of TOTs for two monaurally presented clicks, no difference between Polish and German subjects was found across the life span from 20 to 69 years. This study seems to suggest that auditory perception of temporal order is independent of language experience. However, the absence of evidence does not prove evidence of absence; not having observed a potential language effect on temporal order perception in monaurally presented click tasks does not ensure that no such an effect can be observed in other temporal order tasks such as binaurally presented tone tasks. In addition, although German and Polish languages are different from each other, they both fall into the same category of non-tonal languages.

Unlike non-tonal languages such as English, German, Polish or Russian, the meaning of a Chinese word cannot be solely defined by phonemes without any lexical tone. For example, the Chinese syllable /fa/ has four distinct lexical meanings when spoken with different global pitch contours. It can mean 'becoming rich' when pronounced with a high level tone, the action of 'punishment' with a high rising tone, the 'law' with a low dipping tone, or the 'hair' with a high falling tone. Therefore, to extract the meaning of Chinese words, the pitch contour which features small changes in frequency range plays a crucial role. In contrast, non-tonal languages only have one single lexical meaning in one syllable regardless of some possible tone variations; thus, it is the pitch height, not the pitch contour, which is important in decoding semantic information. In other words, non-tonal languages are characterized by the large changes in frequency range, and the pitch contour plays barely any role in decoding lexical meanings.

Considering the significant difference between tonal and non-tonal languages, the present study aimed to address whether language experience influences temporal order processing by comparing the temporal order thresholds (TOTs) between tonal language speakers (Chinese) with and without a secondary non-tonal language (English) experience. In order to capture the possibly higher sensitivity of Chinese subjects to small changes in frequency range featured by pitch contour, we applied not only monaurally presented clicks as stimuli but also binaurally presented two different types of tone pairs as stimuli. One is "close frequency" tone pairs using 600 Hz and 1 200 Hz tones as stimuli; the other is "distant frequency" tone pairs using 400 Hz and 3 000 Hz tones. If a common temporal mechanism exists for click TOT as suggested by Szelag and coauthors (2011b), then no difference on click TOTs should be observed between the two subject groups. Since monaurally presented clicks might not be appropriate for capturing a potential language effect on temporal order perception as has been shown previously (Bao et al. 2013), the main focus of the present study is to possibly reconfirm the previous observations and to further examine whether the experience of a secondary non-tonal language (i.e., English) may improve the TOT performance on the "distant frequency" tone task. As a selective impact of tonal and non-tonal language environment on TOT performance of tones has already been demonstrated (Bao et al. 2013), i.e., Chinese subjects tend to show lower TOT for close frequency relative to distant frequency tones, while Polish subjects demonstrate the opposite, we expected in the present study that Chinese subjects without English proficiency would show a lower TOT only for close frequency and not for distant frequency tones as compared to click TOT, while Chinese subjects with English proficiency would show lower TOTs for both close frequency and for distant frequency tones.

Eighteen native Chinese speakers without proficiency of a non-tonal language (NC group) and 18 native Chinese speakers with proficiency of English (NCE group) participated in the present study. All subjects were college students. They speak standard Mandarin as well as one local Chinese dialect and have no experience of other languages except English. The subjects of the NCE Group had passed the formal College English Test 4 (CET, Band IV) at the National Ministry of Education in China. All participants were right-handed (Oldfield 1971) and had no history of neurological or psychiatric disorders, or any indication of cognitive impairment. None of the participants had

Table I

Detailed characteristics of participants

Group	*n* (Male/Female)	Age Range (Years)	Mean Age (SD)	Handedness (Left / Right)	Mosaic (SD)	Hearing Status (Normal /Not)
NC	18 (9/9)	20–29	25 (2.4)	Right	37.4 (7.3)	Normal
NCE	18 (9/9)	19–27	23 (1.8)	Right	37.8 (5.8)	Normal

(NC) native Chinese speakers without proficiency of a non-tonal language; (NCE) native Chinese speakers with proficiency of English; (*n*) number of participants; (SD) standard deviation

received a formal musical education, which possibly could increase the sensitivity of auditory perception of acoustic stimuli.

To ensure normal hearing for all participants, pure-tone audiometry screening was assessed (Audiometer GSI 17). Participants with hearing thresholds higher than 30 dB HL and differences of more than 20 dB HL between the two ears were excluded from the study (ANSI 2004). In all participants intellectual abilities were assessed (nonverbal Mosaic Test, see Tewes 1994) and matched scores were obtained for the two groups. Detailed descriptive data of the participants are listed in Table I.

The stimuli were paired sounds presented in rapid succession with varied inter-stimulus-intervals (ISI, i.e., the time distance between the offset of the first stimulus and the onset of the second stimulus). Three types of paired stimuli were used in the study: paired clicks, paired close frequency (600 Hz and 1 200 Hz) sinusoidal tones, and paired distant frequency (400 Hz and 3 000 Hz) sinusoidal tones. The clicks were 1 ms rectangular pulses presented in an alternating monaural stimulation mode, i.e., one click was presented to one ear, followed by another click to the other ear. Subjects were asked to indicate the sequence of the two clicks by pointing to one of the two response cards: "left-right" or "right-left". The tones were generated with the program Cool Edit 2000 (sampling rate 44 100 Hz, 16-bit), and each tone lasted 10 ms with 1-ms rise-and-fall time. The paired two tones in each trial were presented in a binaural stimulation mode, i.e., each tone was presented to both ears with a short gap in between. The subjects had to indicate the temporal order of the two tones by pointing to one of the two response cards: "low-high" or "high-low".

In all measurement, the stimuli were presented *via* headphone (SONY MDR-CD 480) at a comfortable listening level which is well above threshold. The inter-stimulus intervals between the two acoustic stimuli were controlled by a maximum-likelihood based algorithm – YAAP procedure (Treutwein 1997). According to the subjects' previous responses, the ISI of the present trial was set at the current best estimate of the threshold corresponding to 75% correct responses based on a logistic psychometric function. The main measurement for each type of stimuli was preceded by a practice session in which participants reported the temporal order of the two acoustic stimuli presented with a constant, relatively long ISI of 160 ms. The practice was continued until a criterion of 11 correct responses in a series of consecutive 12 presentations was reached.

The sequencing ability was indexed by the temporal order threshold (TOT) which was defined as the minimum time interval required for correctly identifying the temporal order of two successively presented stimuli. To obtain individual values of TOT, a logistic psychometric function was fitted to the subject's data, using MATLAB toolbox psignifit version 2.5.41 (see http://bootstrapsoftware.org/psignifit/), a software package which implements the maximum-likelihood method described by Wichmann and Hill (2001). This procedure estimates an inter-stimulus interval (ISI) corresponding to 75% correct order discrimination (Strasburger 2001). The click TOT was calculated by the estimated ISI plus the duration (1 ms) of the first click, and the tone TOT by the estimated ISI plus the duration (10 ms) of the first tone, thus both using SOA (stimulus onset asynchrony, the time interval between the onset of the first stimulus and the onset of the second stimulus) as TOT index.

A two-way mixed ANOVA was performed for the values of TOT with Stimulus Type (clicks, close frequency tones, and distant frequency tones) as a withinsubjects variable and Subject Group (NC group, NCE group) as a between-subjects variable. The results showed a significant main effect of stimulus type and a non-significant main effect of subject group, $F_{2,68}$=21.531, P<0.001, η_p^2=0.388, and $F_{1,34}$=1.563, P>0.05, η_p^2=0.044. More importantly, the stimulus type significantly interacted with subject group, $F_{2,68}$=3.467, P<0.05, η_p^2=0.093. Further analysis of this interaction using *t*-tests with Bonferroni corrections revealed very interesting TOT patterns: For the NC group, a significantly lower TOT was observed for discriminating the temporal order of the two close frequency tones relative to both the clicks (31 ms vs. 56 ms, P<0.01) and the two distant frequency tones (31 ms vs. 54 ms, P<0.01), and no TOT difference between the clicks and the distant frequency tones was observed (56 ms vs. 54 ms, P>0.05). However, for the NCE group, although a similar high TOT was observed for the clicks as compared to the NC group (63 ms vs. 56 ms, P>0.05), both the close frequency tones and the distant frequency tones resulted significantly lower TOTs relative to their click TOT (21 ms vs. 63 ms, P<0.001; 34 ms vs. 63 ms, P<0.01). No TOT difference between the two types of tones was observed (21 ms vs. 34 ms, P>0.05) (see Fig. 1).

The present study measured temporal order thresholds with both monaurally presented clicks and binaurally presented two types of tones in both NC and NCE subject groups. The results of click TOT confirmed previous observations (Szelag et al. 2011, Bao et al. 2013), i.e., no TOT difference between NC and NCE groups was observed. Thus, the click TOT results suggest a general phenomenon in temporal order processing, i.e., a common and language independent mechanism underlying milliseconds timing which governs monaurally presented click ordering.

With the substitution of frequency-related tones for clicks and the more natural mode of stimulus presentation, namely, binaural instead of monaural presentation of the stimuli, a more sensitive measurement was involved to capture a language effect on temporal order perception. As the TOT results of the two different types of tones indicated, the NC group showed a significantly lower TOT for discriminating the temporal order of two close frequency tones, while the NCE group demonstrated significantly lower TOTs for both the close frequency and the distant frequency tones. These results are consistent with our predictions based on the major difference between tonal and non-tonal languages. Since tonal language speakers like Chinese were experienced with pitch contour detection in their natural language environment, they were more sensitive to small changes in frequency range, thus a lower TOT for close frequency tones relative to distant frequency tones was observed for the NC group. As non-tonal languages are characterized by large changes in frequency range, being exposed to such a language like English apparently increases the sensitivity to large changes in frequencies, thus improving the NCE subjects' temporal order processing of two distant frequency tones.

The language impact on tone TOT performance in the two subject groups can also be interpreted with reference to two different strategies used for temporal order perception (Bao et al. 2013). As already known, both analytic and holistic strategies can be used for decoding the sequence of two acoustic stimuli. When the analytic strategy is used, a subject has to identify the singular acoustic events in their identity and on that basis experiencing directly the temporal order of the two events. For example, in the TOT measurement with tones, the two successively presented tones were of different frequencies: one was a low frequency tone, and the other was a high frequency tone. Thus, the tone order judgment involves not only a temporal order task, but also a frequency discrimination task. Therefore, the subjects may identify the frequency of

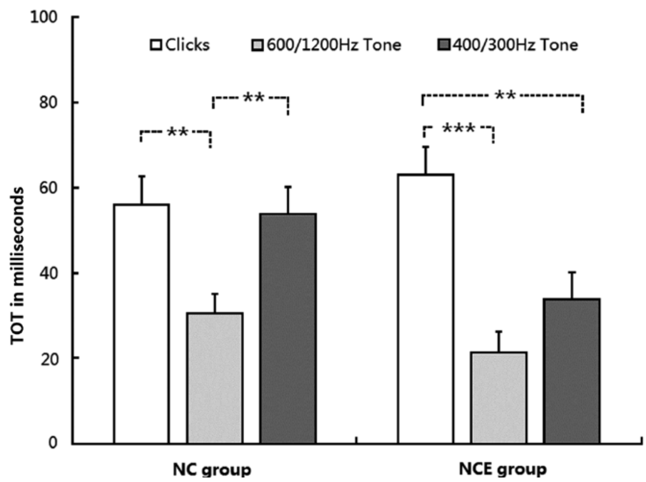

Fig. 1. Temporal order thresholds (TOTs) measured for clicks and two types of tone pairs (600/1 200 Hz and 400/3 000 Hz) in two subject groups. The "NC group" represents native Chinese speakers without proficiency of a non-tonal language, and the "NCE group" represents native Chinese speakers with proficiency of English (**P<0.01, ***P<0.001).

each tone immediately following the presentation of the individual stimulus, and on that basis report the temporal order of the two tones. When a holistic strategy is used, the two successively presented frequency-related tones are integrated into one unitary percept with either an up or a down patterning. On the basis of such global patterning, subjects reconstruct secondarily the temporal order of the two stimuli, leaving the identification of each individual tone unnecessary (Brechmann and Scheich 2005, Szymaszek et al. 2009). Compared to the analytic strategy, the holistic strategy usually leads to a lower temporal order threshold.

For native Chinese subjects who were experienced in detecting pitch contours, they might use a holistic strategy for perceiving the temporal order of two close frequency tones and an analytic strategy for the two distant frequency tones, thus showing lower TOT for close frequency tones relative to distant frequency tones. In contrast, for native Chinese subjects who were experienced also in a non-tonal language, it could be that being exposed to a non-tonal language their analytic strategy was shifted to a holistic one for decoding the temporal order of two distant frequency tones. The holistic processing mode being shaped by tonal or non-tonal language environment seems to be consistent with previous studies using single neuron recording technique, which suggest that neurons in the primary auditory cortex are selective to directions of frequency modulated sweeps (e.g., Tian and Rauschecker 2004). Possibly, such neurons are differently stimulated when being exposed to a tonal or a non-tonal language. The flexible holistic processing tuned by a tonal or non-tonal language experience suggests neural plasticity of the auditory brain as suggested previously (Merzenich and Sameshima 1993).

In conclusion, the main results of the present study indicated a clear impact of language experience on temporal order perception: While native Chinese speakers without English proficiency were good at discriminating the temporal order of two close frequency tones, native Chinese speakers with English proficiency were good at both the close frequency and the distant frequency tone order discrimination. Combining the similar performance for click order perception in both subject groups, a selective modification of language experience on temporal order perception is suggested in the present study.

This research was supported by grants from the National Natural Science Foundation of China (Projects 31371018, 91120004 and J1103602) and Polish Ministry of Science and Higher Education (No. 1082/P01/2006/31 and No. 507/1/N-DFG/2009).

Albert ML, Bear D (1974) Time to understand – a case study of word deafness with reference to the role of time in auditory comprehension. Brain 97: 373–384.

ANSI (2004) ANSI S3.6-2004. American National Standard Specification for Audiometers. American National Standards Institute, New York, NY.

Bao Y, Szymaszek A, Wang X, Oron A, Pöppel E, Szelag E (2013) Temporal order perception of auditory stimuli is selectively modified by tonal and non-tonal language environments. Cognition 129: 579–585.

Brechmann A, Scheich H (2005) Hemispheric shifts of sounds representation in auditory cortex with conceptual learning. Cereb Cortex 15: 578–587.

Fink M, Churan J, Wittmann M (2006) Temporal processing and context dependency of phoneme discrimination in patients with aphasia. Brain Lang 98: 1–11.

Fitch RH, Tallal P (2003) Neural mechanisms of language-based learning impairments: insights from human populations and animal models. Behav Cogn Neurosci Rev 2: 155–178.

Kann E, Wayland R, Bao M, Barkley CM (2007) Effects of native language and training on lexical tone perception: An event-related potential study. Brain Res 1148: 113–122.

Krishnan A, Gandour JT, Bidelman GM (2010) The effects of tone language experience on pitch processing in the brainstem. J Neurolinguistics 23: 81–95.

Luo H, Boemio A, Gordon M, Poeppel D (2007) The perception of FM sweeps by Chinese and English listeners. Hear Res 224: 75–83.

Martin JG (1972) Rhythmic (hierarchical) versus serial structure in speech and other behavior. Psychol Rev 79: 487–509.

Merzenich MM, Sameshima K (1993) Cortical plasticity and memory. Curr Opin Neurobiol 3: 187–196.

Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia 9: 97–113.

Pöppel E (1971) Oscillations as possible basis for time perception. Stud Gen (Berl) 24: 85–107.

Pöppel E (2009) Pre-semantically defined temporal windows for cognitive processing. Philos Trans R Soc Lond B Biol Sci 364: 1887–1896.

Sidiropoulos K, Ackermann H, Wannke M, Hertrich I (2010) Temporal processing capabilities in repetition conduction aphasia. Brain Cogn 73: 194–202.

Strasburger H (2001) Invariance of the psychometric function for character recognition across the visual field. Percept Psychophys 63: 1356–1376.

Szelag E, Dreszer J, Lewandowska M, Medygral J, Osinski G (2010) Time and cognition from the aging brain perspective: individual differences. In: Personality from Biological, Cognitive and Social Perspectives (Maruszewski T, Eysenck MW, Jakowska M, Eds). Eliot Werner Publications Inc., New York, NY. p. 87–114.

Szelag E, Skarżyński H, Senderski A, Lewandowska M (2011a) hearing loss and auditory processing disorders: Clinical and experimental perspectives. In: Culture and Neural Frames of Cognition and Communication (Han S, Pöppel E, Ed.). Springer-Verlag, Berlin, DE. p. 153–168.

Szelag E, Szymaszek A, Aksamit- Ramotowska A, Fink M, Ulbrich P, Wittmann M, Pöppel E (2011b) Temporal processing as a base for language universals: Cross-linguistic comparisons on sequencing abilities with some implications for language therapy. Restor Neurol Neurosci 29: 35–45.

Szymaszek A, Sereda M, Pöppel E, Szelag E (2009) Individual differences in the perception of temporal order: the effect of age and cognition. Cogn Neuropsychol 26: 135–147.

Tallal P (2004) Improving language and literacy is a matter of time. Nat Rev 5: 721–728.

Tallal P, Merzenich MM, Miller S, Jenkins W (1998) Language learning impairments: integrating basic science, technology, and remediation. Exp Brain Res 123: 210–219.

Tewes U (1994) Wechsler Adult Intelligence Scale (revision 1991) (in German). Handbuch und Testanweisung. Bern, CH.

Tian B, Rauschecker JP (2004) Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. J Neurophysiol 92: 2993–3013.

Treutwein B (1997) YAAP: Yet another adaptive procedure. Spat Vis 11: 129–134.

Vandermosten M, Boets B, Luts H, Poelmans H, Wouters J, Ghesquière P (2011) Impairments in speech and non-speech sound categorization in children with dyslexia are driven by temporal processing difficulties. Res Dev Disabil 32: 593–603.

von Steinbüchel N, Wittmann M, Szelag E (1999) Temporal constraints of perceiving, generating, and integrating information: Clinical indications. Restor Neurol Neurosci 14: 167–182.

Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63: 1293–1313.

Wittmann M, Burtscher A, Fries W, von Steinbüchel N (2004) Effects of lesion size and location on temporal-order judgment in brain-injured patients. Neuroreport 15: 2401–2405.