Evaluating the "Critical Period" Hypothesis: Perceptual Learning of Mandarin Tones in American Adults and American Children at 6, 10 and 14 Years of Age

Yue Wang†‡ and Patricia K. Kuhl‡

† Department of Linguistics, Center for Cognitive Science, University at Buffalo, State University of New York, USA

‡ Center for Mind, Brain and Learning, Department of Speech and Hearing Sciences, University of Washington, USA

E-mail: yuewang@buffalo.edu, pkkuhl@u.washington.edu

ABSTRACT

In this study, we examined the perceptual learning of lexical tones in Mandarin Chinese by American young adults, and children from 6 to 14 years old, covering the age range surrounding the "critical period". The participants received a two-week computerized Mandarin tone training program designed to be child-friendly. The results showed that, for the trainees in each of the four age groups but not the controls, percent correct identification increased significantly from the pre-training test to the post-training test, indicating significant improvement after training across age groups. However, comparing the pre-puberty and the post-puberty groups, we did not find an abrupt decrease in the degree of improvement, as would have been predicted by the Critical Period Hypothesis. These results support the view that language learning is not a strictly timed developmental process with rigid cut-off periods.

1. INTRODUCTION

It is generally claimed that adults are inferior to children in the ability to acquire second language speech sounds. According to the Critical Period Hypothesis (CPH), language-learning ability is reduced at puberty, resulting from the loss of neurological plasticity of the brain [1]. On this account, puberty marks a closing of the window of opportunity for learning [2].

An alternative theory has been proposed that future learning may be limited by prior learning, which itself alters the brain, producing a kind of interference effect that impacts later learning [3]. That is, if language learning involves the creation of mental maps for speech, learning must "commit" neural structure in some way. Once committed, the learned structure may interfere with the processing of information that does not conform to the learned pattern. In light of this, experience itself, rather than time, is the critical variable [3].

Indeed, recent studies have indicated that cortical representations may be continuously shaped throughout life [4], and that language-related cortical responses may differ as a function of linguistic experience [5]. Empirical research has also revealed that, with sufficient experience and exposure, adult second language (L2) learners can authentically perceive or produce novel L2 sounds which have no native language (L1) phonemic counterparts [6]. More explicitly, auditory training studies have successfully trained adult listeners to better distinguish non-native sounds [7]. These findings suggest that adult perceptual mechanisms have more plasticity than was previously recognized.

Despite these findings on adult learning, little research has investigated second language speech learning in children, and there has been few direct tests comparing across the age range from childhood into adolescence, and into adulthood. This leaves unanswered the question of whether children exhibit the similar or different learning patterns from adults. With respect to the Critical Period Hypothesis, direct comparison between children at different ages and adults is particularly interesting. If puberty marks the closing of the window of opportunity for learning, one would expect a discontinuity in speech learning abilities for pre-puberty and post-puberty children (and adults).

Using the perceptual training procedure, we examined the learning of lexical tones in Mandarin Chinese by American adults, and children at 6, 10, and 14 years of age, covering the age range surrounding the "critical period". Mandarin phonemically distinguishes four tones differing primarily in pitch height and contour shape. Since this functional use of pitch is not characteristic of American speakers, it has presented great difficulty for learners of Chinese. However, perceptual training research has shown that adult American learners' ability to identify the four Mandarin tones could be significantly improved after training [8]. In the present study, we extend this training procedure to the investigation of Mandarin tone learning by children.

2. METHOD

The perceptual training program [8] was designed to be a high-variability procedure exposing learners to the full

ISBN 1-876346-48-5 © 2003 UAB

range of stimuli and speaker variability to facilitate the construction of novel phonetic categories [7].

Participants. Eighty-six children and adults in four age groups (6, 10, 14, and 19 years of age) participated in the study. They were all native speakers of American English with no known history of speech and hearing impairments, and no previous experience with Mandarin or other tone languages. The participants were randomly assigned to the trained and control groups.

Stimuli. The stimuli are real monosyllabic Mandarin words among which the four tones are equally distributed. They were provided by six native speakers of Mandarin Chinese (3 males, 3 females). A total of 684 stimuli were included: 144 in Pre/posttest, and 540 in Training (for details of the nature of the stimuli, see [8]).

By convention, the four Mandarin tones are named Tone 1, Tone 2, Tone 3, and Tone 4, visually represented by tonal diacritics. However, in the present study, the abstract association of tones with numbers was replaced by animals. So the children (as well as adults) were told that they would hear Bird tone, Frog tone, Cow tone, and Dog tone; and these animals were also visually presented to them. In addition, background pictures, music and sound effects were incorporated to make the program interactive and child-friendly. Both testing and training were created using E-prime, an experiment generation software developed by Psychology Software Tools, Inc.

Procedure. The training program consisted of four phases: introduction, pretest, training, and post-test. The testing and training for the children were conducted in a quiet computer classroom at their school, and for the adults this took place in a testing room in the Center for Mind, Brain, and Learning at the University of Washington. The procedure for the children and adults was identical.

Before the pretest, all participants received a 15-minute introduction of the four tones, which allowed them to hear examples of the tones, and be familiar with the tasks they were to do in the test.

Both the trainees and controls took the 10-minute pretest, in which they were seated wearing headphones at a computer terminal with a monitor and keyboard. The stimuli were presented in four blocks, with each tone being the target tone for one block. The order of tone block presentation was counter-balanced across listeners. For each block, listeners performed a go/no go task. With the picture of the target animal (tone) appearing on the screen, they were instructed to press the "yes" button when hearing the target tone, while pressing the "no" button for the other tones. They were encouraged to press as fast as they could, but the inter-stimulus interval was set to 3 seconds.

After the pretest, the trainees received the two-week training program, consisting of six sessions of 40 minutes each. In the meantime, the children in the control groups were given computer games for the same amount of time. The training program was designed such that both the tone presentation and tasks allowed for a systematic increase in difficulty of tone contrasts. The four tones were presented pairwise from the least confusing pair to the most confusing one [8]. Three tone pairs were trained in each session, thus it took a training set of two successive sessions to complete a cycle of all six tone pairs. During the first two sessions, the trainees performed an ABX task to decide which of the two precursor tones was the same as the third tone by pressing the corresponding button (labeled as the pictures of the animals). During the next two sessions, the trainees' task was two-alternative forced-choice identification. The last two sessions required the same identification task, except that the speaker voice was mixed including all the native talkers. The children took a short break after each tone pair training, and they were rewarded (stickers, pencils, Chinese postcards, bookmarks, etc.) after completing each session.

Immediately after the training program, both the trained and the control listeners took the post-test, which was identical to the pretest.

3. RESULTS

Correct identification scores for the trainees and the controls across the four age groups at pretest and post-test are displayed in Figure 1. As shown in the upper panel of the figure, the trainees in each of the four age groups showed an improvement in their identification scores from the pretest (6 yr: 55%, 10 yr: 67%, 14 yr: 74%, and adults: 86% correct identification) to the post-test (61%, 76%, 82%, and 95% correct identification, respectively). The overall increases in tone identification accuracy for the four age groups are: 6%, 9%, 8%, and 9%, respectively.

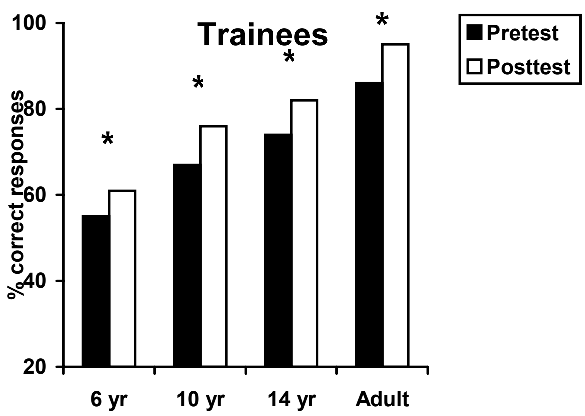
In contrast, as shown in the lower panel, although the control listeners started at approximately the same level as the trainees in the pretest (6 yr: 56%, 10 yr: 63%, 14 year: 73%, and adults: 88% correct identification), they exhibited little improvement in the post-test (55%, 66%, 73%, and 89% correct identification, respectively).

In addition, across trainees and controls and across tests, there appears to be a progressive increase in the identification scores as a function of age, with the younger participants having the lower scores than the older ones.

The overall results were analyzed using a three-way ANOVA of Test (Pretest, Post-test), Group (Trainee, Control), and Age (6 yr, 10 yr, 14 yr, adult), with Test as the repeated measure. There was a significant main effect of Test [$F(1, 73)=22.1$, $p<.0001$], Age [$F(3, 71)=34.6$, $p<.0001$], and Test x Group interaction [$F(3, 71)=9.4$, $p<.003$]. However, there was no main effect of Group [$F(1, 73)=2.2$, $p>.144$]; nor was there a Test x Age interaction [$F(3, 71)=0.4$, $p>.726$], or a Group x Age interaction [$F(3, 71)=0.4$, $p>.728$], or a Test x Group x Age interaction [$F(3, 71)=0.5$, $p>.654$].

To further examine the main effect of Test (pretest, post-test), a paired-samples T-test was calculated separately for the trainees and controls for each age group, with Test as the factor. For the trainees, there was a significant improvement from the pre- to the post-test for each age group: 6 yr: [t(17)=2.4, p<.028], 10 yr: [t(6)=3.6, p<.011], 14 yr: [t(13)=3.2, p<.007], and adult: [t(6)=3.3, p<.016]. However, none of the four control groups showed any significant increase: 6 yr: [t(11)=1.0, p>.330], 10 yr: [t(6)=0.9, p>.366], 14 yr: [t(9)=0.21, p>.837], and adult: [t(6)=0.66, p>.534].

Regarding the main effect of Age, *Post hoc* (Tukey-HSD) comparison showed that the score for the 6 year olds was significantly lower than the other three groups, and the adult score was significantly higher than the other three groups. However, the scores for the 10 and 14 year olds were comparable, although they were significantly higher than that for the 6 year olds, and significantly lower than that for the adults.



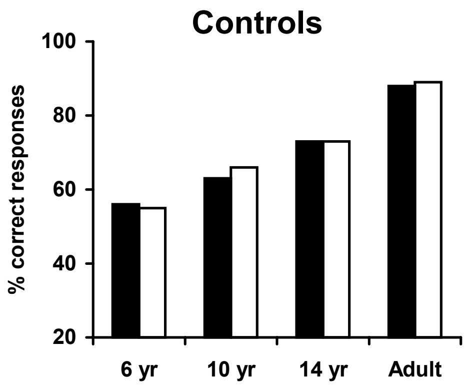


Figure 1: Mean percent correct identification of the four Mandarin tones for the trainees (upper panel) and controls (lower panel) of the four age groups. [*] indicates that the difference is significant at p<.05.

Overall, these results show that, for the trainees in each of the four age groups but not the controls, percent correct identification increased significantly from the pre-training to the post-training test, indicating significant improvement after training across age groups. Comparing the performance of the four age groups, the degree of improvement as the result of training was comparable (6 yr: 6%, 10 yr: 9%, 14 yr: 8%, adult: 9% increase), although the identification scores were progressively higher with the increase of age.

4. DISCUSSION AND CONCLUSIONS

The results showed significant improvement in the identification of Mandarin Chinese lexical tones after the 2-week perceptual training for American children at 6, 10, and 14 years old, as well as American adults. This study was an initial attempt to apply the auditory training procedure to the investigation of second language speech learning by school-age children. Our results for children are consistent with the findings for adults, showing significant effect of training [7, 8].

One major question asked in this study was whether children exhibit the similar or different learning patterns from adults, and whether there was any difference in learning patterns between the pre-puberty and post-puberty children. According to the Critical Period Hypothesis, we would expect that children before the critical period perform markedly better than the older children and adults. However, comparing the pre-puberty and post-puberty groups, we did not find an abrupt decrease in the degree of improvement, as the increase was comparable across the four age groups. That is, our result does not seem to support the CPH claim of rigid cut-off periods for language learning. Instead, it seems to suggest that given the same amount of exposure, the degree of learning would be the same regardless of age.

Although not much research has been done examining the process of L2 learning in children, the studies on adult L2 performance indexed by "age of arrival" (AOA, marking the onset of L2 exposure) in the second language country can indeed lend some support to the current findings [9]. These studies evaluated the CPH on the assumption that it could be demonstrated by showing a significant departure from linearity in L2 performance as a function of AOA. However, no such discontinuity was found for learners whose AOA was before the critical period and those whose AOA was after the critical period. Instead, experience-related factors, such as the amount of language use, did influence L2 performance. Consistently, studies investigating the cortical effects of L2 learning have also revealed that experience could be more of a determinant for L2 performance than age [5].

Thus, our study and these other adult age-related studies seem to support the view that L2 speech learning is probably not a strictly timed developmental process with rigid cut-off periods, as claimed by the CPH, but is rather a function of experience and exposure [3].

ISBN 1-876346-48-5 © 2003 UAB

However, one of the results that is bewildering is that the initial Mandarin tone identification score increased with age, contrary to the finding of a gradual decrease in L2 performance with the increase of age [9]. While we cannot rule out the factor of cognitive maturity that younger children were less able to perform the task well, we could also consider it in terms of linguistic influence. The general thesis with respect to experience is that successfully established L1 phonetic categories may impede L2 learning [3, 6]. It could be the case that tone, as a suprasegmental property with phonemic functions, is so novel for the native speakers of English, that there is no pre-existing comparable L1 categories. That is, the mental map for tone might be open for adults as well as children. Younger children's poorer initial performance might be attributed to developmental factors. It has been reported that younger children are less sensitive to category boundaries than older children and adults for novel foreign sounds, because their categories are more flexible [10].

The difference in performance between younger and older children, and adults could also be a matter of timing. Research examining L2 acquisition in the natural setting has shown that while adults and older children could learn an L2 better initially, younger children could out-perform in a longer term [11]. This raises the question regarding the current study as to what the pattern will be for the children and adults with further training; or as experience is further gained. Since language development is a dynamic process that is presumably continuing throughout life, more evidence is needed examining the various aspects of learning, various stages, and styles, with the consideration of maturational, developmental, linguistic, and cognitive factors.

Taken together, the current study does not support the Critical Period Hypothesis with the claim that the ability of learning a second language diminishes at puberty. The results of the Mandarin tone training suggest that both adults and post-puberty children still have the capacity to learn. At least at the initial stage of L2 learning, the degree of improvements could be comparable across age, given the same amount of exposure and experience.

REFERENCES

[1] E. Lenneberg, *Biological Foundations of Language*, New York: Wiley, 1967.

[2] T. Scovel, *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*, Cambridge: Newbury House Publishers.

[3] P. K. Kuhl, "A new view of language acquisition," *Proceedings of the National Academy of Sciences*, 97, 11850-11857, 2000.

[4] M., Van Turennout, T., Ellmore, and A., Martin, "Long-lasting cortical plasticity in the object naming system," *Nature Neuroscience*, 3, 1329-1334, 2000.

[5] D. Perani, E. Paulesu, N. Sebastian Galles, E. Dupoux, S. Dehaene, V. Bettinardi, S. F. Cappa, F. Fazio, and J. Mehler, "The bilingual brain: Proficiency and age of acquisition of the second language," Brain, 121, 1841-1852, 1998.

[6] J. E. Flege, "Second language speech learning: theory, findings, and problems," in *Speech Perception and Linguistic Experience*, edited by W. Strange, Baltimore: York, 1995.

[7] J. S. Logan, S. E. Lively, and D. B. Pisoni, "Training Japanese listeners to identify English /r/ and /l/: A first report," *Journal of the Acoustical Society of America*, 89, 874-886, 1991.

[8] Y. Wang, M. M. Spence, A., Jongman, and J. A. Sereno, "Training American listeners to perceive Mandarin tones," *Journal of the Acoustical Society of America*, 106, 3649-3658, 1999.

[9] J. E. Flege, G. H. Yeni-Komshian, and S. Liu, "Age constraints on second-language acquisition," *Journal of Memory and Language*, 41, 78-104, 1999.

[10] A. C. Walley and J. E. Flege, "Effect of lexical status on children's and adults' perception of native and non-native vowels," *Journal of Phonetics*, 27, 307-332.

[11] C. E. Snow and M. Hoefnagel-Hohle, "The critical period for language acquisition: Evidence from second language learning," *Child Development*, 49, 1114-1128, 1978.

ISBN 1-876346-48-5