

Learning English in China: A Tablet-Based App Using the Voices of Native Speakers

Dhonam Pemba¹, Virginia Mann^{1,2*}, Trina Sarkar³, Kaveh Azartash¹

¹Kadho, Inc., Irvine, CA, USA

²Department of Cognitive Sciences, University of CA, Irvine, CA, USA

³Department of Computer Science, Stanford University, Stanford, CA, USA

Email: vmann@uci.edu

Received 5 May 2016; accepted 18 July 2016; published 25 July 2016

Abstract

This paper concerns the acquisition of English by two groups of five-to-six year old children living in China: 1) Those taught by the traditional classroom methods using non-native speakers of English as teachers and 2) Those taught by a new computer application designed by Kadho, Inc. to train children in the perception of English consonants vowels and suffixes produced by native speakers of English. The data consist of judgments of children's accent in mimicking English sentences, at the onset, 1, 3, 6 and 10 months into treatment. Children who received the computer app were judged to have significantly better accents from the first month of training onwards.

Keywords

English Language Learning, Computer-Based Training, Primary Education

1. Introduction: Age and Source Factors in Second Language Acquisition

China is a country that is rapidly adopting English language instruction into its school and universities. However the ever-increasing numbers of Chinese people being exposed to English as a foreign language are attaining varying degrees of success. Although many achieve a high command of written vocabulary and grammar, most are unable to achieve native-like fluency in speaking and listening. The accents of Chinese Mandarin speakers trying to speak English can be traced to two factors. Those individuals who do achieve native-like fluency all tend to share an early age of exposure; and an extended exposure to the speech of native speakers of English.

Research has shown that the earlier an individual begins learning a second language, the better their pronunciation becomes [1]-[3]. As the age of acquisition increases, the child's eventual chance of native-like fluency linearly decreases such that it becomes virtually impossible for adults to acquire native-like pronunciation in a foreign language [4]-[9]. That is because, although every child is born with the ability to learn and produce the sounds of any human language [10], unless he or she is engaged with stimuli from other languages, the first language (L1) develops and matures, but the potential for acquiring sound systems for other languages (L2 and beyond) withers with increasing age. As a child ages, this loss becomes more significant, ultimately leading to a

*Corresponding author.

pattern in adults where L1 interferes with the learning of L2 [11], challenging perception and production [12].

Age of acquisition aside, it is further the case that language learning abilities are compromised by a lack of early exposure to L2 native speakers. Munro and Mann [3], for example, linked the achievement of superior English accents to the age at which a Mandarin Chinese speaker had immigrated to the US and become immersed in the speech of native speakers of English. Thus age is one part of the equation, but the source of language engagement is another. One problem facing the Chinese educational system is that, although younger and younger children are exposed to English language learning, the teachers who are providing the instruction, while well-trained in the English language, are not native speakers of English. Non-native English speakers make up a large portion of English teachers abroad in non-native English speaking countries, and China is no exception.

2. Some Particular Difficulties of Chinese EFL

Chinese people who are learning English as a foreign language offer an excellent example of how native language, L1 significantly influence the pronunciation of a foreign language, L2. Their native Chinese language, which is Mandarin in the population we consider, causes multiple problems with their being able to speak English correctly and without a foreign accent. That is because the English and Mandarin Chinese sound systems are very different from each other. Their differences in stress, intonation, rhythm and syllable structure, not to mention sentence structure, can make attaining native-like fluency in English a difficult challenge for the Chinese Mandarin speaker [3] [12] [13].

Two aspects of the sound systems are noteworthy. One is the use of different consonants and vowels. Mandarin Chinese does not have any voiced stops, affricates or fricatives, which leads to consonant substitutions such as pronouncing “pill” instead of “bill” or “ket” instead of “get”. The English vowels also contrast with Mandarin; English has more vowels, and Mandarin Chinese speakers must learn new vowel distinctions if they are to master English. Because Mandarin Chinese has no lax vowels, it is particularly difficult for Chinese people to hear and produce contrasts like “bit” vs. “beet” or “cot” vs. “cat”.

The second challenge is subtler, but just as profound. Mandarin Chinese, like many Asian languages, allows many consonants in syllable-initial position but allows only the consonant “n” at the end of syllables. Thus final consonants in syllables like “tub” “had” and “bag” are difficult for Mandarin Chinese speakers even though they are familiar with “b”, “d” and “g” at the beginnings of syllables [13]. Thus Mandarin speakers make two common errors: 1) leaving off the final consonant, and 2) adding a vowel to the final consonant making the syllable into 2 syllables instead of one (e.g., “ba-gu” instead of “bag”). There are also consonant clusters within English but not within a syllable in Mandarin Chinese. This results in Mandarin Chinese speakers inserting neutral vowels between sequences of consonants, such as saying “suturawu” for “straw” or “besutu” for “best”.

3. Mobile Technology

The development and widespread adoption of mobile technology has brought new potentials for language learning [14] [15]. The user-friendly capabilities of mobile technologies have found uses in vocabulary [16], grammar [17], reading practice [18] [19], and listening and speaking [20] [21]. Mobile language learning provides the capability and convenience of achieving learning goals anytime and anywhere.

Currently, although mobile-based language learning technologies are being used to teach some aspects of language, the potential for using native language stimuli in the “sensitive period” of childhood has yet to be leveraged. Although software has been used to teach pronunciation by focusing on mouth positions for speech therapy, it is agreed that specialized teaching in phonetics typically produces poor outcomes because improvements in pronunciation best occur through exposure rather than training [22] [23]. In the natural world, language is not acquired by teaching rules to students and having them memorized. Instead, developing the sounds of a second language seems to require extensive interactive experience and not just passive exposure [24] within early childhood. When used effectively, mobile technologies can prove an important tool for achieving language fluency through early exposure to native speakers in a truly immersive, interactive environment.

4. The Design of the Kadho English App

“Kadho English” is a new app whose software provides realistic, contextualized and interactive video of spoken native English scaffolded by spoken Mandarin Chinese. More detail about the construction of the exercises is

available in [25]. Kadho English engages participants with meaningful audio-visual English stimuli by having them respond often to these stimuli through perception-led button presses and spoken imitations. Task-based speaking and listening activities include context-based matching, discrimination and speaking activities that have the potential to make learning a more realistic, rewarding and fun experience [20] [26] [27]. Though they are encouraged to produce what they hear, only their perceptual responses are evaluated by the app.

The Kadho English curriculum can be broken down into 4 stages reflecting 4 main content areas: Vowels, Consonants, Consonant Clusters, and Inflectional affixes, with each stage having its own set of subtasks involving increasingly difficult contrasts. Each stage has its focus on a content area, but all share the following four properties: 1) a goal of teaching sound perception and not orthographic perception. 2) use of words as well as sounds 3) the introduction of basic vocabulary and phrases and 4) the encouragement of production and repetition as well as the forced-choice assessment of perception. Four learning activities were designed for use with each stage in the curriculum to accomplish the goals of Kadho English. These are summarized in **Table 1**.

5. Experimental Method

To measure and compare possible improvements in the pronunciation quality of short English sentences after traditional training versus training with the Kadho application, two groups of Chinese students were studied before (onset), during (1, 3 and 6 months from onset) and 4 months after (*i.e.* 10 months from onset) EFL training began. The experimental group worked with the Kadho application during individual sessions while the control group received instruction in the form of traditional English as a foreign language taught by non-native English teachers using a commercially available written guide [28]. The experiment consisted of one measurable outcome: Degree of Accent [3]. This is the perceived accent in English sentences produced by the children and judged by a panel of 10 native English speakers using a 10 point scale from 0, strong accent to 10, no accent.

5.1. Participants

The 1022 subjects were 5- and 6-year-old children attending kindergarten and primary schools in Zhengzhou, China. They served with the permission of their parents who also supplied the tablets that children used. Children were divided into two different groups, treated and control. There were 510 treated participants who received the Kadho English app on their tablets; 362 males and 148 females whose mean age was 5 years 208 days (s.d. 210 days). There were 512 untreated, control participants who attended traditional English classes; 354 males and 158 females whose mean age was 5 years, 208 days (s.d. 219 days). Participants were from schools with no native English-speaking teachers and no participant had received private English lessons from native English speakers. At the end of the study 993 students remained in the study; 29 dropped out due to unforeseen reasons.

5.2. Training Materials

Audiovisual speech stimuli for training were produced by two different native speaking English speakers for use in the 4 stage series of sequential activities described in Section 4 above. A more complete description of the Kadho English materials and the computer program are available in Pemba *et al.* [25]. Children gave perceptual responses on a tablet and were given feedback as to the accuracy of those responses. At each stage, they were encouraged to use the computer to record their own productions but were not given feedback as to the accuracy

Table 1. Learning activities design.

Activity type	Learning objective	Description
Video Immersion	Exposure to English sounds being used through real vocabulary and engaging sentences	Immersive and entertaining video that uses words and sounds in a child-engaging manner
Animation Lesson	Teach sound and word association	Targeted lesson to introduce key sounds and associated words
Video Game	Test and train sound and word identification, discrimination	Interactive task to identify whether sounds are the same or are in present in a word
Speech Imitation	Practice sound production	Video-based task for a child to repeat words said in the video

of those productions owing to present technology's known difficulty with the recognition and evaluation of children's voices.

5.3. Assessment Materials

The aim of assessment was to evaluate the overall degree of perceived foreign accent in English sentences produced by the participants. We used ratings of elicited sentence repetitions, following previous research on foreign accent assessment [3]. For eliciting the children's productions, two female native-English speakers each produced eight short English sentences that were integrated into a mobile app. All sentences were statements of 5 - 8 words in length, and each contained from 2 - 3 different words that were drawn from the materials used in perceptual training. (e.g., "The *bear* threw the *acorn* in the basket." "The *flower* and the *book* are on the *table*.")

Another mobile app was created to record each participant's elicitations, with Chinese language instructions telling the participant to repeat each English sentence that was seen and heard on the screen. A total of 4 elicitations of each sentence were taken per participant. All elicitations were subsequently rated by a panel of 10 university students who were native speakers of American English having a mean age of 19 years. These raters were blind as to whether samples were produced by control or treated participants. All had passed a pure-tone audiometry screening prior to participating, none had special training in speech or language. Each rater used a scale ranging from 1 to 10 to rate each sentence, with sentences heard at a 2 second interval.

5.4. Procedure

The treated group used the Kadho English application on their tablets for 30 minutes a day, 2 times a week at the time their untreated, control peers were learning in traditional classrooms taught by their classroom teachers who were not native speakers of English. The Kadho English software required children to complete a minimum of 4 learning activities each day that covered new material in the stages of focus; the remaining time was spent with previously taught material in a stage of the child's choice. Inflectional affixes were taught throughout the program, vowel perception was taught from weeks 1 to 12, and consonant perception during week 13 to 42. Both groups were allowed to bring tablets to school, but only the treated group used the Kadho English app. Treatment lasted for six months and measurements of perceived accent were taken at treatment onset (baseline), and at 1 month, 3 months, 6 months (completion of treatment) and 10 months (*i.e.* 4 months post-treatment).

Speech samples of the desired sentence production were captured for each participant at each time of assessment with an app that had the participant hold down a button and repeat the English sentence he/she had just heard. Participants in the treated group had already been familiarized with this task since it is included in the training even though the productions in training were not judged for quality. Participants in the control group were familiarized with the task for 5 minutes before the elicitations made at each time of testing.

6. Results

The data consist of evaluations of perceived accent among the Kadho English treated subjects and the untreated control subjects attending traditional classrooms. A split-plot ANOVA showed a significant interaction of time of measurement (*i.e.* assessment month) and group $F(3.88, 4056) = 1209.51, p < 0.001$, that explains 28.9% of the variability in degree of accent. The results are displayed in **Table 2** where it may be seen that English accents improved to a greater extent in the treated group than in the control group. One-way repeated measure ANOVAs showed that the effect of time of measurement is significant within both groups, for the control group $F(3.95, 2040) = 57.587, p < 0.001, \eta^2 = 0.101$, and for the treated group $F(3.74, 2016) = 2698.112, p < 0.001, \eta^2 = 0.842$. However, while the Kadho treated group shows no advantage at pretest, they show an advantage after a single month of treatment, $t(1014) = -8.904, p < 0.001$, and as time passes the final effect size of receiving the Kadho app is very large $r = 0.991$, Cohen's $d = 15.39$. As seen in **Table 2**, there is effectively no overlap between the final English accents of the treated and untreated children.

7. Discussion

This study examined the success of two different means of teaching spoken English on the English accents of 5- to 6-year-old children. One is an individually delivered tablet-based program, Kadho English, that engages children in activities that use real-time video of speech produced by native speakers of English. The other is the

Table 2. Mean, standard errors and confidence intervals for Treated and Untreated groups.

Group	Time of measurement	Mean	Std. Error	Lower and upper value of mean in the 95% confidence interval	
Treated	Pre treatment	2.408	.028	2.353	2.463
	After one month	3.119	.031	3.058	3.180
	After three months	4.663	.037	4.591	4.736
	After six months	5.801	.037	5.728	5.873
	Post treatment	7.027	.036	6.956	7.098
Untreated	Pre-treatment	2.556	.028	2.501	2.611
	After one month	2.730	.031	2.670	2.791
	After three months	2.734	.037	2.662	2.806
	After six months	2.977	.037	2.905	3.049
	Post treatment	3.159	.036	3.089	3.230

traditional method of a classroom curriculum taught by native speakers of Chinese who, though well versed in the English language, are not native speakers of English.

The results demonstrate that the tablet-based technology offers a strong and significant improvement in children's spoken English. The children who trained with Kadho were able to improve their accents 4.6 points on a 10 point scale where children in traditional classrooms only improved by only 0.6 points. This finding is made all the more impressive by the fact that the treated children trained for only 30 minutes per session, two times a week analogous to the two thirty-minute classroom lessons the controls received from their teachers. After what would amount to a little over six hours of application-based training—1 month of lessons—the treated children had made gains that were four times that of the controls. By the end of 10 months the gain was seven times greater and the effect size approaches 1.0. In addition, the children improved on accent measures even though the application encouraged speaking but did not evaluate it, *per se*. Though perceptual judgment was the focus of evaluation within the app, the children using Kadho English spoke English to the app and appear to have experienced a more targeted and natural immersion in native English through their interactive engagement with the video content.

There are certain improvements and further experiments to do. The level of training effects in the domains that were trained—vowel perception, consonant perception and affix production—should be measured and we will be providing those measurements in our forthcoming study [29]. We should show inter-rating agreement. We should try the app among children at older ages. We should see if there are benefits of using the app within the context of traditional classes instead of as part of an either-or design. We should sample written English as well as spoken.

Technology is becoming more widely used for language learning and pronunciation training in foreign languages. However, very little is known about the actual effectiveness of this new type of training, during treatment, post-training and over extended times. This study represents one of the first attempts to show training effects in a large cohort of children studied over an extensive amount of time (*i.e.* ten months). The results indicate that video-based, real-time instruction can indeed provide much-needed native English speech stimuli within the “sensitive period” of language acquisition.

Acknowledgements

The authors wish to thank the actresses who provided our stimuli, the raters who judged accentedness and the staff of Kadho who worked on the materials. We are indebted to the principals, teachers and students of local kindergarten and primary schools in Zhengzhou, Xiamen and Shenzhen. We would also like to thank Zheng Ming Kai of SIAS University, who graciously recruited the subjects in our experiment and oversaw testing.

References

- [1] Flege, J. (1999) Age of Learning and Second-Language Speech. In: Birdsong, D., Ed., *Second Language Acquisition and the Critical Period-Hypothesis*. L. Erlbaum, Hillsdale.
- [2] Flege, J. and MacKay, I. (2004) Perceiving Vowels in a Second Language. *Studies in Second Language Acquisition*, **26**, 1-34. <http://dx.doi.org/10.1017/S0272263104261010>
- [3] Munro, M. and Mann, V.A. (2005) Age of Immersion as a Predictor of Foreign Accent. *Applied Psycholinguistics*, **26**, 311-341. <http://dx.doi.org/10.1017/S0142716405050198>
- [4] Krashen, L. and Scarcella, R. (1979) Age, Rate and Eventual Attainment in Second Language Acquisition. *TESOL Quarterly*, **13**, 573-582. <http://dx.doi.org/10.2307/3586451>
- [5] Long, M. (1990) Maturational Constraints on Language Development. *Studies in Second Language Acquisition*, **12**, 251-285. <http://dx.doi.org/10.1017/S0272263100009165>
- [6] Patkowski, M. (1980) The Sensitive Period for the Acquisition of Learning in a Second Language. *Language and Learning*, **30**, 449-472. <http://dx.doi.org/10.1111/j.1467-1770.1980.tb00328.x>
- [7] Scovel, T. (1969) Foreign Accents, Language Acquisition and Cerebral Dominance. *Language Learning*, **19**, 245-254. <http://dx.doi.org/10.1111/j.1467-1770.1969.tb00466.x>
- [8] Birdsong, D. (1992) Ultimate Attainment in a Foreign Language. *Language*, **68**, 706-755. <http://dx.doi.org/10.1353/lan.1992.0035>
- [9] Lenneberg, E.H. (1967) *Biological Foundations of Language*. Wiley, N.Y.
- [10] Polka, L. and Werker, J.F. (1994) Developmental Changes in Perception of Non-Native Vowel Contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 421-435. <http://dx.doi.org/10.1037/0096-1523.20.2.421>
- [11] Birdsong, D. and Molis, M. (2001) On the Evidence for Maturational Effects in Second Language Acquisition. *Journal of Memory and Language*, **44**, 235-249. <http://dx.doi.org/10.1006/jmla.2000.2750>
- [12] Defense Language Institute (1974) A Contrastive Study of English and Mandarin Chinese. ERIC Number: ED105774.
- [13] Crowther, C.S. and Mann, V.A. (1992) Native Language Factors Affecting Use of Vocalic Cues to Final Consonant Voicing in English. *Journal Acoustical Society of America*, **92**, 711-722. <http://dx.doi.org/10.1121/1.403996>
- [14] Liu, T.Y., Tan, T.H. and Chu, Y.L. (2009) Outdoor Natural Science Learning with an RFID-Supported Immersive Ubiquitous Learning Environment. *Educational Technology & Society*, **12**, 161-175.
- [15] Danan, M. (2004) Captioning and Subtitling: Undervalued Language Learning Strategies. *Meta*, **49**, 67-77. <http://dx.doi.org/10.7202/009021ar>
- [16] Chen, C.M. and Chung, C.-J. (2008) Personalized Mobile English Vocabulary Learning System Based on Item Response Theory and Learning Memory Cycle. *Computers & Education*, **51**, 624-635. <http://dx.doi.org/10.1016/j.compedu.2007.06.011>
- [17] Morita, M. (2003) The Mobile-Based Learning (MBL) in Japan. *Proceedings of the First Conference on Creating, Connecting and Collaborating through Computing*, Kyoto, 128-129. <http://dx.doi.org/10.1109/c5.2003.1222348>
- [18] Chang, K.-E., Lan, Y.-J. and Chang, K.-E. (2010) Mobile-Device-Supported Strategy for Chinese Reading Comprehension. *Innovations in Education and Teaching International*, **46**, 69-84. <http://dx.doi.org/10.1080/14703290903525853>
- [19] Wu, W.F., Zhang, H., Chen, J., Chen, J.Y. and Lin, C.Y. (2011) Development and Evaluation of a Computerized Mandarin Speech Test System in China. *Computers in Biology and Medicine*, **41**, 131-138. <http://dx.doi.org/10.1016/j.compbiomed.2011.01.002>
- [20] Liu, T.Y. (2009) A Context-Aware Ubiquitous Learning Environment for Language Listening and Speaking. *Journal of Computer Assisted Learning*, **25**, 515-527. <http://dx.doi.org/10.1111/j.1365-2729.2009.00329.x>
- [21] Edirisingha, P., Rizzi, C., Nie, M. and Rothwell, L. (2007) Podcasting to Provide Teaching and Learning Support for an Undergraduate Module on English Language and Communication. *Turkish Online Journal of Distance Education*, **8**, 87-107.
- [22] Krashen, S.D. (1982) *Principles in Second Language Acquisition*. Pergammon Press, N.Y.
- [23] Morley, M. (1994) *Pronunciation Pedagogy and Theory: New Views, New Directions*. Teachers of English to Speakers of Other Languages, Indiana.
- [24] Kuhl, P.K., Tsao, F.-M. and Liu, H.-M. (2003) Foreign-Language Experience in Infancy: Effects of Short-Term Exposure and Social Interaction on Phonetic Learning. *Proceedings of the National Academy of Sciences (USA)*, **100**, 9096-9101. <http://dx.doi.org/10.1073/pnas.1532872100>

- [25] Pemba, D., Mann, V., Azertash, K., *et al.* (In Preparation) Kadho: New Technology for Teaching Spoken English to native Speakers of Mandarin Chinese.
- [26] Purushotma, R. (2005) Commentary: You're Not Studying, You're Just... *Learning and Technology*, **9**, 80-96.
- [27] Wachowicz, K.A. and Scott, B. (1999) Software That Listens: It's Not a Question of Whether, It's a Question of How. *CALICO Journal*, **16**, 253-276.
- [28] English Language Learning Materials for Ages 3-11. Cambridge University Press, Cambridge, U.K.
<http://education.cambridge.org/us/learning-stage/3-11>



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing a 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>