

THE EXPERIMENTAL ANALYSIS OF THE EFFECTIVENESS AND
SUSTAINABILITY OF A CHINESE TUTORING PACKAGE

BY

Hang Wu

Submitted to the graduate degree program in Applied Behavior Analysis
and the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Master's of Arts.

L. Keith Miller
Chairperson

Committee members

Claudia L. Dozier

Glen White

Date defended: April 16, 2009

The Thesis Committee for Hang Wu certifies
that this is the approved Version of the following thesis:

THE EXPERIMENTAL ANALYSIS OF THE EFFECTIVENESS AND
SUSTAINABILITY OF A CHINESE TUTORING PACKAGE

Committee:

Chairperson L. Keith Miller
 Claudia L. Dozier
 Glen White

Date approved: June 16, 2009

Abstract

This experiment investigated two effects of training four native Chinese speakers to use a Chinese Tutoring Package to teach American students pronunciation and translation of Mandarin Chinese vocabulary. Tutors were trained in use of the Tutoring Package at staggered times to form a multiple baseline design. The effects of the training on teaching pronunciation were analyzed. After the tutors were trained, the average pronunciation test score for students increased from 45% to 90% and tutor's correct use of the tutoring package increased from 68% correct trials to 92%. The effects of tutor training were replicated for the teaching of translation. Results suggest that the training produced an increase in correct tutoring by the tutors, which in turn increased the students' performance. Furthermore, continued use of the package under naturalistic conditions, high social validity ratings, and extended follow up on two tutors during a second semester suggest that use of the package by tutors may be sustainable.

Introduction

Chinese is the sixth most frequently taught foreign language in institutions of higher education in the United States, with enrollment increasing by 20% from 1998 to 2002 (Welles, 2004). Learning Mandarin Chinese presents two difficulties to English speakers. One of the major factors contributing to this difficulty is the use of four different tones in Mandarin Chinese. Two words having the same phonemes can have entirely different meanings depending upon the tonality with which they are pronounced. Another difficulty is lack of correspondence between the non-phonetic Chinese characters and the component sounds of the word that it represents.

The Chinese government took a partial step toward solving these two problems by adopting the phonetic system called “Pinyin” in 1979 (Shibles, 1994). Pinyin uses Latin letters to represent the sounds of Chinese words. In addition it uses diacritical marks to represent the tonality of syllables within that word. For example, the Pinyin for Chinese characters “今天” is “jīn tiān” Where the Latin letters represent the sounds and the diacritical marks above the “i” and the “a” represent the tonality for those sounds. While Pinyin helps bridge the gap between Chinese and western systems of writing, mastering the pronunciation and translation of Chinese words remains a difficult challenge for American students (Yang, 2009).

Wu and Miller (2007) developed a tutoring protocol to help an American student learn to pronounce Chinese characters correctly. The tutoring package used flashcards, Pinyin, modeling, feedback and reinforcement to teach the translation and pronunciation of Chinese characters. The tutoring package included praise, error

correction, and practice. They used a multiple-baseline experiment across two groups of Chinese words to analyze the effect of the tutoring package on the student. The results showed that the tutoring package produced mastery pronunciation of targeted Mandarin Chinese

The generality of the Wu and Miller (2007) study is unknown due to several limitations. First, because the study was conducted with a single student it is unknown if the package will be effective with other students. Second, because the researchers implemented the package, it is unknown if other tutors, especially Chinese tutors, not having knowledge of behavior analysis can also use it. Especially, in Chinese culture, positive reinforcement is not encouraged in education (Cai, 2006). Third, it is unknown if other tutors will find the package practical enough that they will choose to use it under non-research conditions. Fourth, it is unknown if the tutoring packages can be applied in college-level Chinese classes. The current experiment was carried out to address these critical limitations.

The purpose of this study was to determine (a) if the Chinese Tutoring Package in the previous study remained effective when implemented by different tutors with different students and (b) if the tutors continued to use the package over time when they were not required to do so.

Method

Setting

The experiment was conducted in a small room used for tutoring in a Midwest university. Tutoring took place outside the students' routine coursework and was

voluntary. Each session lasted for 30 minutes. A covert web camera with a microphone was used to record the participants' vocal trials and other relevant behaviors during each session. The researcher was not present in the setting during tutoring sessions. The researcher provided no feedback or other commentary on the tutor's or student's performance for any tutoring session.

Participants

There were ten participants in this study: six students and four tutors. The participating students were four males and two female aged from 18 to 25. These six students were native English speakers. They were all enrolled in a first-year Chinese language class. These students demonstrated mastery of writing Pinyin equivalents for approximate 200 Chinese characters. The participating tutors were two females and two males aged from 24 to 31. Each tutor was a native Chinese speaker enrolled as a student at the same university. The author was the researcher while she was a graduate student in applied behavior analysis.

Measurement

The researchers observed two student responses for each of 10 Chinese words presented to them during an end-of-session test for each session of the experiment. They observed the percentage of correct pronunciations and the percentage of correct translations for each of the 10 Chinese words. Two observers, the author and a second native speaker, independently replayed the recordings for each session and judged the correctness of the pronunciation and translation. The reliability observer was not informed of the purpose of the experiment until the scoring was finished.

Reliability data for student responses were collected for 100% of the sessions. Inter-observer agreement was calculated by dividing the number of agreements in each session by the number of agreements plus disagreements and multiplying by 100%. Reliability for student responses was 100% for all sessions.

The researchers observed two responses for each tutor. They observed the percent of trials during which the tutor correctly used the tutoring package to teach the student how to pronounce Chinese characters. A correct pronunciation trial was recorded when (1) the tutor held up a card with Chinese characters and asked the student to pronounce the word and (2A) praised the student if their pronunciation was correct or (2B) demonstrated the correct pronunciation and asked the student to pronounce the word again if their pronunciation was wrong.

The researchers also observed the percent of trials during which the tutors correctly used the package to teach the translation of the Chinese characters. A correct translation trial was recorded when (1) the tutor held up a card with Chinese characters and asked the student to translate the word and then (2A) praised the student if their translation was correct or (2B) demonstrated the correct translation and asked the student to translate the word again if their translation was wrong.

The same reliability observer was used for the tutoring behavior. Reliability data for number of correct tutoring trials were collected for 50% of the sessions and at least once in each condition. Inter-observer agreement was calculated by dividing the number of agreements in each session by the number of agreements plus disagreements and multiplying by 100%. Reliability for correct tutoring of

pronunciation ranged from 82% to 100%, and averaged 92%. Reliability for correct tutoring of translation ranged from 86% to 100%, and averaged 95%.

Social Validity

The researcher measured two aspects of the Tutoring Package's social validity. Students rated the extent that they "would like to participate in similar tutoring programs in the future." Tutors rated the extent that they "would like to use this same tutoring package again." Both groups rated their agreement on a seven point scale with the endpoints defined as "Extremely disagree" and "Extremely agree."

Procedure

The researcher arbitrarily selected Chinese characters for 800 words which the students had not learned before the experiment. These characters were written on flash cards for later presentation to the participant. The front of each 12cm X 8cm card presented a Chinese word represented by two Chinese characters. The other side of the card presented the Pinyin of each characters and the English translation of the word. At the beginning of each session, the researcher gave ten new flashcards to each tutor. Each tutor taught ten new words in each session. Each session lasted for 30 minutes, with no more than two sessions per day. Each tutor decided when to introduce new vocabulary during each session. Each tutor tutored one student from the beginning to the end of the experiment. Two tutors taught two additional students during a second semester for follow up of the tutor's correct tutoring. At the end of each session, the students were quizzed over the pronunciation and translation of the ten Chinese words. During each quiz, the tutor asked the student to pronounce and

translate all ten new vocabularies once without providing any error correction. After each student finished the vocabulary quiz of the 10 words, the student and the tutor were permitted to use the remainder of the 30 minute session to conduct free conversation in English.

Before training on the Chinese Tutoring Package, tutors were told that the target of tutoring was to teach the students to pronounce and translate the ten new words correctly. Also, the tutors were permitted to teach students by using their own tutoring methods. During training, each tutor was individually introduced to and trained to use the Chinese Tutoring Package by the researcher. The average training time for each tutor was 44 minutes (ranged from 35 to 48 minutes). After training, the researcher instructed the tutors to choose their preferred method to teach the students in the following sessions. There was no mastery criterion for both tutors and students' responses during the tutoring sessions.

Intervention

The author developed the tutoring protocol (Wu & Miller, 2007) with definitions of the correct teaching behavior. The researcher trained the tutors to use the Chinese Tutoring Package. The researcher gave the tutors written instructions in Chinese and English, asked the tutors to read those instructions. If the tutors could not understand the instruction, the researcher would do the modeling. In addition, the researcher demonstrated the correct method with the tutor playing the role of student and then gave feedback when the tutor used the tutoring methods with the researcher playing the role of student. All tutors followed the protocol 100% correctly at the end

of training sessions. For students, the intervention was the implementation of the Chinese Tutoring Package by the tutors. For tutors, the intervention was the training of using the Chinese Tutoring Package. Refer to Appendix for the written instructions.

Design

This experiment used a multiple-baseline design across four pairs of students and tutors.

Baseline. Each of four students was taught by their tutor using the tutor's own tutoring method. This condition lasted two to six sessions depending on the tutor-student pair.

Post Tutor Training. Each of the four students was taught by their tutor after being trained in the Chinese Tutoring Package. This condition lasted five to six sessions depending on the tutor-student pair.

Follow up. Two new students were taught by two of the original four tutors. This condition lasted 9 or 11 sessions depending on the tutor-student pair. These students were not exposed to baseline.

Results

Figure 1. Percent of correct student pronunciations before and after tutors were trained to use the Chinese tutoring package.

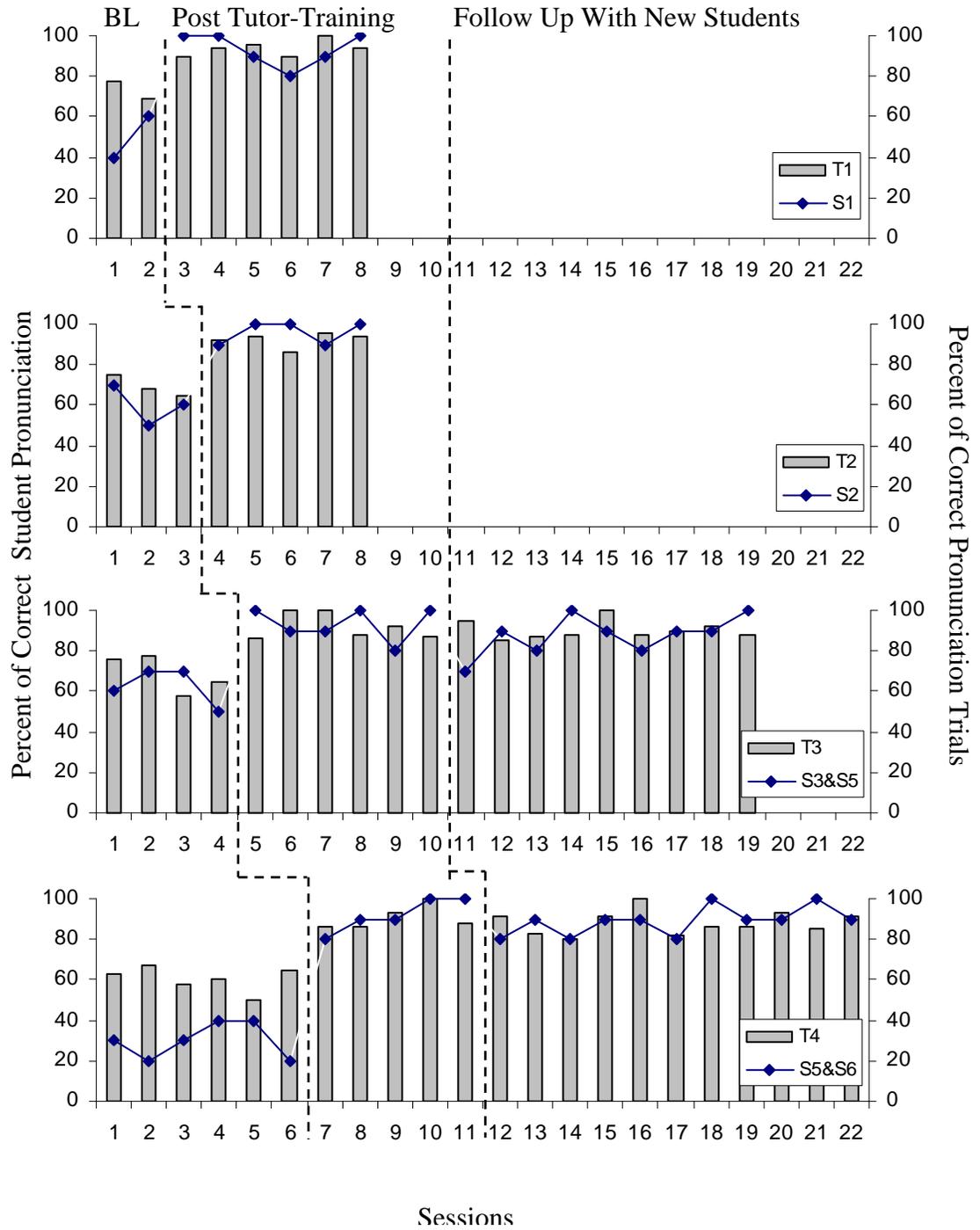


Figure 1 shows the effect of training the tutors in the use of the Chinese Tutoring package on both student and tutor performance with respect to pronunciation. The average pronunciation test score for students during baseline (i.e., before tutor training) was 45% (ranged from 10% to 70%) which increased to 90% (ranged from 70% to 100%) post tutor-training. The average percentage of correct pronunciation trials implemented by tutors during baseline was 68% (ranged from 58% to 78%) which increased to 92% (ranged from 80% to 100%). Thus training the tutors increased the students' pronunciation scores and the percentage of correct pronunciation trials implemented by tutors.

Figure 1 also shows the extent to which two tutors continued to use the tutoring package during follow up pronunciation trials with two new students. The students averaged pronunciation test scores of 88% (ranged from 88% to 89%) while the tutors averaged 89% correct trials (ranged from 80% to 100%). Thus both student and tutor pronunciation performance maintained at about the same level during the follow up period.

Figure 2. Percent of correct student translations before and after tutors were trained to use the Chinese tutoring package. The bars represent the percent of correct tutor translation trials. The line represents the percent of correct student translations.

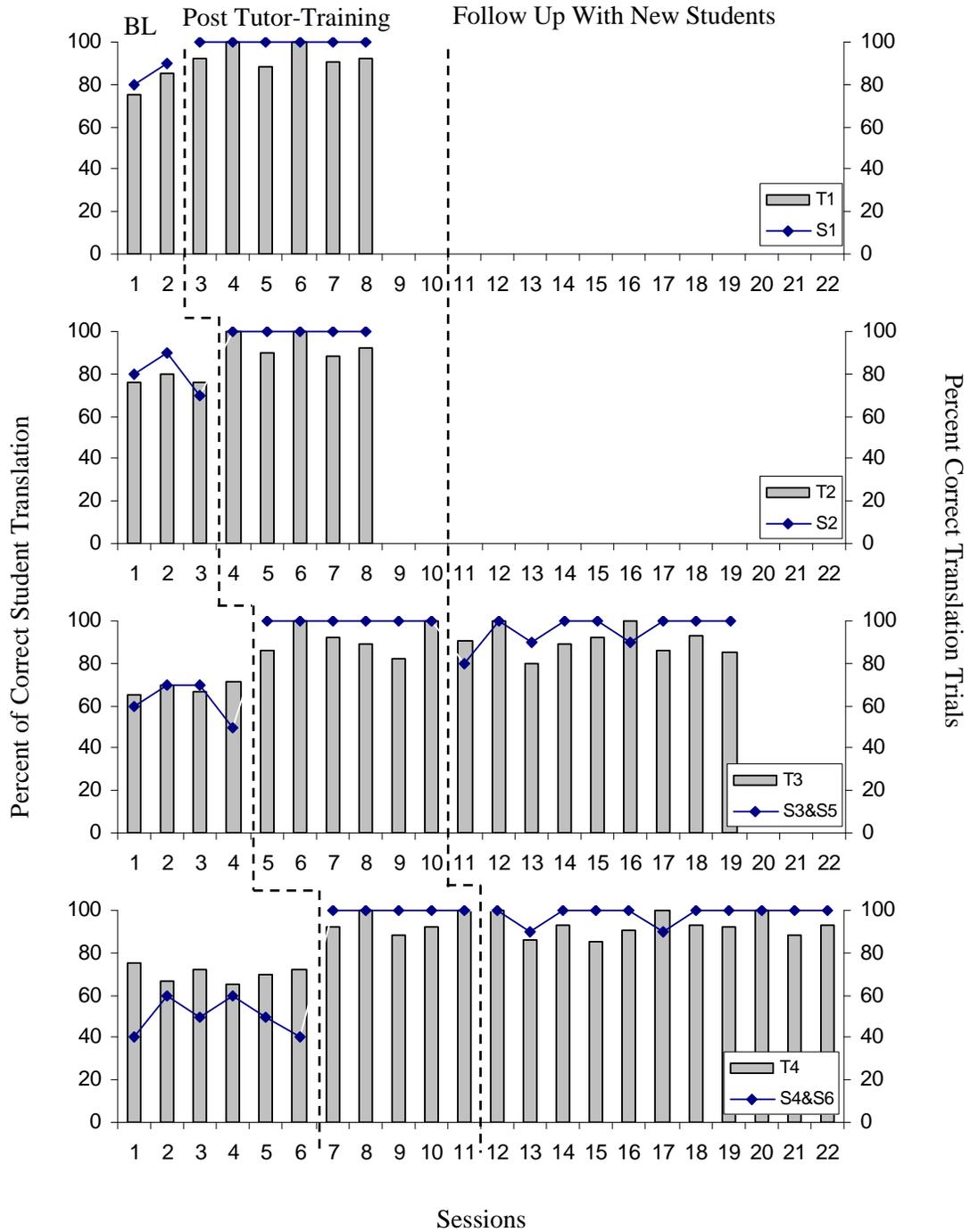


Figure 2 shows the effect of training the tutors in the use of the Chinese Tutoring package on both student and tutor performance with respect to translation. The average translation test score for students during baseline was 65% (ranged from 40% to 90%) which increased to 100% post tutor-training. The average percentage of correct translation tutoring trials during baseline was 71% (ranged from 65% to 85%) which increased to 93% (ranged from 88% to 100%). Thus training the tutors increased the students' translation scores and the percentage of correct translation tutoring trials implemented by tutors.

Figure 2 also shows the extent to which two tutors continued to use the tutoring package during follow up translation trials with two new students. The students averaged translation test scores of 97% (ranged from 96% to 98%) while the tutors averaged 91% correct trials (ranged from 80% to 100%). Thus both student and tutor performance maintained at about the same level during the follow up period.

Table 1. Average Number of Trials during Baseline and Post-Tutor training

Participant	Baseline	Post-tutor-training
Tutor 1	33	45
Tutor 2	21	35
Tutor 3	30	42
Tutor 3	12	38

Table 1 shows that the average number of trials of each tutor increased after tutoring training.

Table 2. Average Minutes of Tutoring During Baseline and Post-Tutor Training

Participant	Baseline	Post-tutor-training
Tutor 1	30	25
Tutor 2	30	27
Tutor 3	30	24
Tutor 3	30	28

Table 2.shows that the average time of trials of each tutor decreased after tutoring training.

Discussion

The experiment investigated two effects of training tutors to use the Chinese Tutoring Package. The first was the effect of tutor training on students' acquisition of the pronunciation and translation of Chinese words. The performance of the initial four students was higher after the tutors began using the Chinese Tutoring Package than before they used it. The fact that each student's performance increased at different times and only after the tutors had been trained to use the tutoring package suggests that other factors were not responsible for the increase. Therefore it is reasonable to conclude that training tutors to use the tutoring package was responsible for the increase.

The second was the effect of the training on tutors' correct use of the tutoring package for both translation and pronunciation. The performance of the four tutors was higher after the training than before training. The fact that each tutor's

performance increased at different times and only after training suggests that other factors were not responsible for the increase. Therefore it seems reasonable to conclude that the training program produced the increases in the tutors' performance.

The effect of tutors using the tutoring package on the students' pronunciation was similar to that found in the Wu and Miller (2007) study and extends the generality of that finding. The effectiveness of the tutor training extends the generality of the Wu and Miller study by confirming that the tutoring package can be used by non-behavior analysts and can produce improved student performance. Specifically, tutors learned to praise their student immediately after a correct response and to initiate a correction procedure immediately after an incorrect response. These behavioral methods are very important for Chinese teachers to learn because positive reinforcement is not applied often by Chinese teachers (Cai, 2006). Table 2 showed that the Chinese native speakers used less time in tutoring while Table 1 showed they have more tutoring trials during post tutoring sessions.

The present study also extends the generality of the Wu and Miller (2007) study by giving the tutors, after they have been trained, the choice to use the Chinese Tutoring Package or their own methods. They chose to use the tutoring package at a high-level of integrity after training without any further encouragement or feedback from the researcher after tutoring sessions. This finding is further strengthened by the tutors' continued correct use of the tutoring package during the Follow Up condition in the next semester. This suggests that tutors will choose to continue using the package during non-research conditions. This observation is consistent with the

speculation that the improved performance of the students may be a natural reinforcer for native Chinese tutors' use of the tutoring package.

The current experiment was conducted under sustainability testing conditions (Miller, Welch, Altus & Zwicker, 2006). These are conditions in which all researcher services and influences are minimized in order to simulate the naturalistic conditions that prevail in a normal, non-research situation. Several previous experiments have been conducted under these conditions (e.g., Altus, Welsh & Miller, 1991; Welsh, Miller & Altus, 1994; DeWein & Miller, 2008). They found that continued use of an intervention by service staff under simulated non-research conditions predicted continued use during follow up observations as long as 25 years later (Miller, 2006).

Further evidence that tutors may choose to continue using the tutoring package comes from the results of measuring the social validity ratings of students and teachers. Students gave a rating of 7.00 (out of 7) for the statement that they "would like to participate in similar tutoring programs in the future." Tutors gave a rating of 6.25 for the statement that they "would like to use this same tutoring package again." These measures suggest that the tutoring package was socially valid for both students and tutors and that the tutors would continue to use the package in non-research conditions.

Although the present study clearly extended the generality of the previous study by Wu & Miller (2007) by including multiple students and tutors, there remain limitations. One limitation of this study was that all tutors were native Chinese speakers. Therefore we do not know whether the tutoring package can be used by

non-native speakers. This may not be important because in the Department where this study was conducted all TA's for the prior ten years were native Chinese speakers.

This may be true nationally as well.

Another limitation is that the role of the error correction and praise cannot be specified. The Wu and Miller (2007) study found that simple practice did not lead to improvements in the absence of the tutoring package. Table 1 showed in the present study tutors increased the number of trials that they implemented from 24 per session during baseline to 40 post-training. It is possible that a combination of improved tutor implementation of error correction, praise and increased number of trials may be necessary to produce improved student performance. Future research should be directed at separating these factors.

Future research might also investigate the possible use of the tutoring package to teach Chinese pronunciation and translation to an entire class by teachers. If the package can be used in conjunction with actual Chinese language classes, the sustainability of the package in language departments could be investigated. In addition, this package can be modified to teach other languages in future studies.

References

- Altus, D. E., Welsh, T. M., & Miller, L. K. (1991). A technology for program maintenance: Programming key researcher behaviors in a student housing cooperative. *Journal of Applied Behavior Analysis, 24*, 667-675.
- Cai, J. (2006). U.S. and Chinese teachers' cultural values of representations in mathematics education. *Mathematics Education in Different Cultural Traditions-A Comparative Study of East Asia and the West, 9*, 465-481.
- DeWein, M. & L. Keith Miller (2008). The Effect of a Teacher Report on the Sustainability of an Intervention to Facilitate Engagement by a Child with Developmental Disabilities. *Education and Treatment of Children. 31*, 333-350.
- Shibles, W. A. (1994) *Chinese romanization systems: IPA transliteration*. Philadelphia, PA, USA: University of Pennsylvania Dept. of Asian and Middle Eastern Studies; 52.
- Miller, L. K. (2006). The design and analysis of a worksharing program that survives. Paper presented at the annual meeting of the Association for Behavior Analysis, Atlanta.
- Miller, L. K., Welch, T., Altus, D. E., & Zwicker, T. (2006). Creating Sustainable Behavior Analytic Interventions. *Current Repertoire, 25*, 3.
- Welles, E. (2004). Foreign language enrollments in United States institutions of higher education, Fall 2002 [Electronic version]. *ADFL Bulletin, 35*, 1-24.

- Welsh, T. M., Miller, L. K., & Altus, D. E. (1994). Programming for survival: A meeting system that survives 8 years later. *Journal of Applied Behavior Analysis, 27*, 423-433.
- Wu, H. & Miller, L. K. (2007) A tutoring package to teach pronunciation of Mandarin Chinese characters. *Journal of Applied Behavior Analysis, 40*, 583–586.
- Yang, J. (2009) Chinese borrowings in English. *World Englishes, 28*, 90-106.

Appendix: Tutoring Protocol – English version

Step 1: Teach the cards one by one

Ask the student to pronounce and translate the word on each card

- A. If the student can pronounce and translate the word, ask the student to make a sentence using the word on the card, and praise the student, then put the card on the table (face up). Move on to the next card.
- B. If the student can't pronounce and translate the word correctly, demonstrate the correct pronunciation and/or translation of the word. Ask the student to pronounce and/or translate the word. You can show them the pinyin and translation on the back of the card. Repeat this step until the student can pronounce the word AND translate it correctly. Perform Step 1, part B one more time once the student pronounces and translates it correctly. However, if the student still cannot pronounce the word correctly after three attempts, ask the student to practice pronouncing the word three times, and ask the student to make a sentence, then put the card on the table, and move on to the next card.
- C. Arrange the cards in the following order:

1	2
3	4
5	6
7	8
9	10

Step 2: Review cards

- A. Point at the first card, and ask the student to pronounce and translate the words on the card. If the student can do both tasks correctly, praise them and move on to the second card
- B. If the student cannot pronounce and/or translate the word correctly, then demonstrate how to pronounce and/or translate it and ask the student to pronounce and translate it. If the student still can not pronounce and translate the word correctly (after three attempts), ask the student to pronounce the word three times, and move on.
- C. Repeat part A to review the words on the second card. If the student can pronounce and translate the word correctly, praise them and then ask them to pronounce and translate the word on the first card AND on the second card. Move on if the student can pronounce and translate both cards correctly. If not, review only the card(s) with trouble as in Step 2, part B, and move on.
- Review the other cards by using the same methods as above. For example, once the student can pronounce and translate the third card correctly, praise them and ask them to pronounce and translate the first, second, and third card. Move on if the student can pronounce and translate all the three reviewed cards correctly.
- For another example, once the student can pronounce and translate the fourth card correctly, praise the student and ask the student to pronounce and translate the words on the first, second, third and fourth cards. Move on if the student can pronounce and translate all the four reviewed cards correctly. If the student, for

example, cannot get the second and the third card correct, review only the second and the third card as in Step 2, part B, and move on.

During this step, you have the choice to randomly pick up any reviewed cards to ask the student to pronounce and translate the word on it at any time. If the student makes mistakes again, give feedback and teach the student again as in Step 2, part B.

Step 3: Ask the student to pronounce and translate the words on all the cards for the last time

Collect all the cards and shuffle them, and put the cards randomly into the following arrangement:

Any card	Any card

Ask the student to pronounce and translate each card once, and finish the tutoring session.

Step 4: Write down the words which student did not pronounce or translate correctly in step 3.

Step 5: Use any remaining time to have a free conversation with the student if there is time left.