

The Combined Effect of Instruction and Monitor in Improving Pronunciation of Potential English Teachers

Qian Liu (Corresponding author)

School of Foreign Languages, Linyi University

Shuangling Road, Linyi Shandong 276005, China

Tel: 86-539-2158-360 E-mail: Liuqian@lyu.edu.cn

Zunwei Fu

School of Sciences, Linyi University, Shandong, China

Received: April 11, 2011 Accepted: April 29, 2011 doi:10.5539/elt.v4n3p164

The research is financed by Linyi University. No. HX10202 (Sponsoring information)

Abstract

The role that instruction and monitor can play in improvement of pronunciation has long been a focus of argument among linguists and researchers. It is assumed that their combination will result in positive effect. An empirical study is carried out and it is confirmed that their combination contributes greatly to the improvement of pronunciation accuracy and positive transfer of pronunciation knowledge. Monitor strategies are also discussed in this paper.

Keywords: Instruction, Monitor, Target sounds, Transfer

1. Introduction

Grammar-translation method was once very popular in TEFL (Teaching English as a Foreign Language) classes and is still widely used in many areas in China. When students trained in such a traditional program are finally encouraged to speak, they generally produce a kind of Chinglish accent because of the long-term neglect in pronunciation. With the rise of communicative pedagogy in recent decades, the interest of language teaching has been largely directed to the encouragement of fluency. Pronunciation, though being an integral part of the speaking process, has been pushed again to the sidelines of language teaching. Consequently, it is not surprising that many learners at higher language proficiency levels may have developed habitual, systematic pronunciation errors in spite of their reasonable fluency. Though intelligibility will finally be achieved with the assistance of immediate linguistic and situational context in actual communicative situations, the frequency and type of error will possibly obscure meaning and thus cause delay in the process of information conveyance and understanding.

Therefore, it is important for one to attain as high as possible a degree of accuracy in pronouncing a foreign language. This is especially true for English majors in normal universities, for their utterance will in future serve as a main source of input on the part of their students and inaccurate pronunciation seriously disadvantages them as potential English teachers. Accordingly, these learners are expected to pay special attention to their pronunciation not only for the goal of success, but also for survival in their future career.

However, fossilization has become apparent at this stage and pronunciation errors will persist without efforts and endeavours. Furthermore, Hammerly (1991) argues that while some permanent improvement is possible even after years of mispronouncing a SL (Second Language), correcting habitual distortions is far harder than helping the students form good pronunciation habits from the start. With reference to the possible measures that could be taken to change the so-called fossilized pronunciation, there exists much disagreement theoretically on the role that instruction and monitor can play. Nevertheless, few empirical studies have been carried out in this field. Therefore, it is the intent of this paper to do an empirical research to testify their combined effect in improving pronunciation of English majors in normal universities.

2. Literature review

2.1 The effect of instruction

A number of research studies have investigated the effect of instruction on the learning of pronunciation, but the results are inconclusive. Though some studies (Acton 1997, Hammerly 1991) report positive effect of explicit

teaching of pronunciation, some researchers posit that formal training does little to facilitate accurate pronunciation. For example, the role of instruction is also strongly refuted by Krashen (1985) in his Acquisition-Learning Hypothesis and Monitor Hypothesis, which are two important components of the Input Hypothesis.

In the Acquisition-Learning Hypothesis, acquisition is a subconscious process identical in all important ways to the process children utilize in acquiring their NL (Native Language), while learning is a conscious process that results in “knowing about” language. It is further claimed that the learners’ ability to produce utterances in another language comes from acquired competence, i.e. from subconscious knowledge. Conscious knowledge resulting from learning serves only as an editor, or monitor and monitor use is limited, for the conditions in which it can be applied are difficult to meet. However, after having reviewed some studies, Krashen & Terrell (1988) also acknowledge that adults, at least theoretically, can utilize learned knowledge a great deal.

2.2 The effect of monitor

According to Krashen (1988), conscious learning is only available as monitor, which can be used to alter the output of the acquired system so that accuracy can be improved. He further claims that monitor is fairly limited with respect to what sorts of “repairs” it can perform. While there is considerable individual variation, even the best adult monitor users confine most of monitoring to the simpler rules, for complex permutations require too much mental energy and processing time.

Though the role of monitor in improving pronunciation is strongly refuted by Krashen, many researchers (Celce-Murcia et al. 1996, Morley 1991, Derwing and Munro 2005) still argue for its positive effect on pronunciation accuracy. Kenworthy (1987), for example, claims that the possibility of change or adjustments in pronunciation will be blocked unless learners develop the ability to monitor their own speech and make this a habit. In spite of the importance of this issue and the intense debates involved, there are few empirical studies carried out to verify the different views illustrated above, except one done by Yule et al. (as cited in Celce-Murcia et al., 1996), noting that accuracy seems to have a more solid basis in terms of the learners’ monitor skills.

2.3 Hypotheses of this research

In spite of the disagreement on the effect of instruction and monitor, most researchers argue for their respective value in improving pronunciation. Nevertheless, it is our opinion that they are interrelated to each other in pronunciation learning and their combination can greatly facilitate the improvement of pronunciation, for effective instruction will inevitably lead to reflection and conscious use of monitor. Unfortunately, few empirical studies have been carried out to testify their combined effect on pronunciation. Therefore, it is the intent of this research to bring the following two hypotheses into scrutiny.

The first one is that the combination of instruction, resulting from the teachers’ efforts and monitor, which is a kind of learning strategy adopted by the students, will greatly promote the improvement of pronunciation accuracy. In addition, it is assumed that once the learners have mastered the specific sounds contained in some particular words on which they have been instructed, they are capable of pronouncing these sounds accurately even if they appear in uninstructed words. Therefore, the second hypothesis is that the combination of instruction and monitor will result in positive transfer of pronunciation knowledge.

3. The empirical study

3.1 Pre-test: data obtained in vocabulary reading from both groups

Subjects. The subjects are juniors majoring in English education in a local university. One class was randomly assigned to be experimental group and another control group. There were 7 male students and 23 female students in the experimental group. The control group consisted of 6 males and 24 females. All of the 60 subjects from both groups attended the pre-test.

Materials. Pronunciation includes various aspects such as sounds, weak forms, stress, accent, rhythm, intonation etc. Kenworthy (1987) maintains that the learners cannot be expected to do everything at the same time in pronunciation. Obviously, it is also impossible for the adult students learning to correct their pronunciation to pay attention to all the aspects simultaneously. Therefore, according to the insights gained from teaching practices, 10 sounds (including segments and combinations of segments) often mispronounced were chosen as target sounds of this research. As can be seen from table 1, altogether 10 vocabulary items were presented in the pre-test, each containing an underlined target sound.

Insert Table 1 here.

Recording. Prior to the recording, the subjects were instructed to read the list of 10 words printed on a piece of paper aloud. Then they were tested individually in a quiet classroom in the presence of the researchers and an assistant.

Their performances were recorded for later analysis. Each subject was allowed to proceed with their reading after they reported their identification number.

Scoring. Prior to the beginning of the research, two highly qualified teachers, Ms. X and Ms. Y, were invited to participate in the study. Ms. X was asked to rate the vocabulary data for accuracy. The ratings were carried out “blind”, for she was neither an instructor of these subjects nor was she present at the recording. She was informed to only listen to the pronunciation of target sounds and to ignore other mistakes, such as possible deviations in word stress etc. That is to say, the items were correct if the recording showed no replacement or modification of the target sounds. The subjects would receive one point for each correctly pronounced item (self-corrections would also be taken into account if occurred) and thus a maximum of ten could be achieved by the most accurate learners.

Results. The mean score of the pre-test (Test 1) was computed for each group. As a result, the experimental group obtained a mean value of 2.7333 and the control group, 3.1333. Then the data were submitted to independent samples test, a subprogram of SPSS, to determine whether there was a significant difference between these two groups. A p level of .05 was applied in all the analyses reported in this research. The result indicated no significant difference, with sig. (2-tailed) = .348, $p > .05$ (see table 5).

3.2 Intervention for the experimental group

Both of the two groups in this research attended the same regular classes every week, whereas the participants in the experimental group attended an extra course of about two hours the day after the pre-test, during which they received systematic instruction on pronunciation rules and monitor strategies.

Teacher and instructional materials. Ms. Y was responsible for the instruction in this research. The nature of the research was explained to her and we negotiated to determine the content and procedures of the instruction. The instructional materials were just the ten vocabulary items utilized in the pre-test. The researchers were also present at the instruction to ensure that all was going smoothly. It was observed that an overwhelming majority of the learners appeared to enjoy the two-hour course during which time no break was taken.

Procedures of pronunciation instruction. The 10 vocabulary items were presented to the subjects one by one. The common mispronunciation of each word, caused generally by the specific target sound contained in it, was pointed out and the pronunciation rules of the target sound were introduced to the learners. Then they were trained to discriminate between the acceptable and unacceptable pronunciation of each word and to imitate the correct one after the instructor. Once a satisfactory result was obtained from the chorus, several learners would be asked to produce the particular word in public to ensure instruction had been properly received individually. Feedback and necessary correction were provided in the end.

Instruction on monitor strategies. Once the words were properly produced by the subjects, they were suggested to monitor, with the rules learned, whether they had made similar mistakes when pronouncing other words containing the same target sounds. With reference to the monitor strategies, they were mainly advised to identify the possibly mispronounced words from their reading, for relatively ample time might be guaranteed for monitor to take place in this preliminary stage of correcting their habitual pronunciation errors. For example, having learned that [w] is often substituted for /v/ in their utterance, they might pay attention to the pronunciation of “vine” to make sure that it should not be mispronounced as [waɪn] when they see the word in print. Once the unacceptable pronunciations of particular words are identified and corrected during their reading, less energy might be needed for them to incorporate the accurate patterns into their output. In addition to the monitor strategies mentioned above, many other strategies such as “self-rehearsal technique”, “post hoc monitoring” and so on introduced by several researchers (Morley, 1991; Kenworthy, 1987) were also recommended for use in their regular studies. The implementation of these strategies will be discussed in detail in implication section.

3.3 Post-test

3.3.1 Data obtained in vocabulary and sentence reading from both groups

Subjects and materials. All the participants from both groups attended the recording of vocabulary reading (Test 2) and sentence reading (Test 3) in post-test. In these two reading tasks, altogether twenty words (including a letter) containing the target sounds were tested; half of them were still presented in isolation, whereas the other ten were incorporated into 4 specifically prepared sentences designed by the researchers (see table 2). The inclusion of sentence reading was to elicit a speech sample intermediate in degree of monitoring between vocabulary reading and spontaneous speech.

Recording and scoring. The recording, which was held two weeks after the instruction, was similar in process to that of T1. The subjects were required to read the vocabulary items before they proceeded to read the sentences. Ms. X was informed to apply the same criteria as in T1 for the scoring of these two tasks.

Results. First, the mean values of T2 and T3 were computed for the control group. An examination of the results revealed that its mean of T2 was 3.2333 and that of T3 was 3.1667. Then they were compared respectively with the mean of T1. Paired samples test signified that the difference between each pair was not significant (see table 3). Therefore, it was clear that the participants in the control group showed little improvement in pronunciation accuracy in both T2 and T3.

The same procedures were repeated for the experimental group. The results showed that its mean of T2 was 8.1677 and that of T3 was 7.533. As can be seen from table 4, the learners in this group had achieved considerable improvement in both T2 and T3, with that of T2 the most significant.

Insert Table 2, 3 and 4 here.

In addition, an independent samples test was carried out to ascertain the possible differences in T2 and T3 between these two groups and the results were shown in table 5, where their difference in T1 was also included. Though the performances of these two groups were judged to be similar in T1, the differences between them proved to be significant in both T2 and T3.

Insert Table 5 here.

3.3.2 Data obtained in spontaneous speech from the experimental group

Subjects. Only the subjects in the experimental group attended the recording of spontaneous speech, whereas the control group was dropped in this stage in view of time limit and practical need.

Materials. This task (Test 4) was designed to ascertain the subjects' performances in such a situation where, compared with in vocabulary and sentence reading, there was the least time for monitor to occur. However, the disadvantage of this task was that the appearance times of the target sounds in the utterances of a particular participant could not be predicted. For example, in the speech sample of an individual, the target sound /æ/ might appear in different words many times, whereas other target sounds such as /wɪn/ and /si:/ might not be presented at all. Therefore, in order to elicit a speech sample containing as many target sounds as possible, each participant was expected to produce longer utterances in the relatively limited time. Consequently, the subjects were required to present for 3 to 5 minutes on one of the five topics (see Appendix One) that had been dealt with in their composition.

Recording. The subjects in the experimental group attended the collection of spontaneous speech data one day after their recording of T2 and T3. Each one was given a list of five topics after they came into the recording room and was required to choose one from them. Then the recording began, with no time spared for them to prepare for their speech.

Scoring. Prior to the scoring, several procedures were taken so as to mitigate the burden of the rater. First, the subjects were required to copy, listening to the recording of their own, exactly what they had uttered onto a piece of paper, including their self-corrections if ever occurred. Then they were requested to underline all the words bearing the target sounds. Finally, their labour was proofed and necessary revision was implemented by the researchers. Consequently, the rater could evaluate the target sounds contained in the underlined words with the help of the written materials, and the scoring method was the same as in the previous tests. In view of the great possibility that a specific target sound might appear several times in a speech sample, its pronunciation could be judged to be accurate only when it was properly produced in all the words containing it. For example, the segment /θ/ was supposed to have been mastered by a subject if it was accurately pronounced in all the words containing this sound, otherwise the subject would score 0 instead of 1 for this target sound. However, if the subject perceived his or her pronunciation error of this particular sound and made prompt self-correction, s/he would not lose score for the error that had been properly corrected.

Results. As had been expected, each target sound was produced by a different number of subjects in T4, and "N" is used to represent these different numbers of subjects. Consequently, the performances of N on each sound in all the four tests were summarized in table 6, where "Percentage of N" is used to indicate how many of N have properly pronounced a specific sound in a particular test. As might be seen from table 6, the learners in the experimental group still achieved great progress in T4, though the improvement was not as significant as in T2 and T3.

Insert Table 6 here.

3.4 Discussion and implication

The first hypothesis of this thesis predicts that the combination of instruction and monitor is able to promote pronunciation accuracy significantly. As a result, it was confirmed by the great progress in the post-test scores achieved by the subjects in the experimental group. In addition, it could be seen from the post-test materials that

improvement was not limited to the vocabulary that had been dealt with in the instruction. Having learned the pronunciation rules of target sounds, the participants in the experimental group were capable, with the help of monitor strategies, of identifying other words containing these sounds and thus improving their pronunciation accuracy of these words. Hence, positive transfer in pronunciation knowledge occurred owing to the combined effect of instruction and monitor, and thus the second hypothesis was also confirmed.

As might be seen from the previous discussion, capability in monitor strategies serves as an important factor in the improvement of pronunciation. Fortunately, it can be strengthened through strategy training, which assumes that the strategies are teachable and conscious attention to learning strategies is beneficial (Murphy, 1991, Dłaska and Krekeler, 2008). Therefore, a variety of practices are introduced as follows so as to help the learners enhance their monitor use.

Once the learners succeed in their imitation of target sounds, they can rehearse and stabilize their modified pronunciation patterns in various activities. Reading, as introduced earlier, might serve as an effective way of practice. In addition, “covert rehearsal”, an activity advised by Dickinson (1988), also serves for this purpose. In such an activity, the learners can practice privately before they actually present their speech in public. For instance, they may think about the pronunciation accuracy of their utterances with the help of pronunciation rules learned, and they may also compare them with their memory of native-speaker models.

Compared with “covert rehearsal”, both “post hoc monitoring” proposed by Acton (1997) and “action replay” suggested by Kenworthy (1987) represent a completely opposite process. The commonality of these two techniques is that they require the learners to suppress the urge to monitor their pronunciation when they speak. Then the speakers, familiar with the content and organization of their speech, are assumed to have more mental spaces to concentrate on their pronunciation problems when they scan their initial production.

Kenworthy (1987) recommends that the learners can record their speech onto a tape while they are involved in relevant activities. Therefore, they are capable of evaluating their own performances according to the immediate feedback from the tape. Furthermore, the tape also provides a record of progress for the learners so that they may know what they have accomplished and what they still have to do in this extended pronunciation learning process.

In addition to the methods suggested above, peer-monitoring can also be exploited by the learners as an effective activity, which may promote both their perceptive and productive accuracy. A usual practice of peer-monitoring is to have students work in groups. According to Celce-Murcia et al. (1996), the learners should be divided into groups of three or four instead of pairs, because there might be disagreement in pair-work about whether the speaker produced the utterance inaccurately or the listener heard it incorrectly. Bigger groups are expected to work better because the consensus of group members might contribute to the settlement of disagreement.

In short, improvement can be rather limited if the learners depend exclusively on their teachers without frequently monitoring their own pronunciation, for sufficient time is not guaranteed for the teachers to focus on their pronunciation problems individually. In addition, owing to the rapid development of information technology, multimedia software might also be used to assist pronunciation learning (Marks, 2005). Therefore, the learners must take responsibility for their own pronunciation and learn to monitor their own speech inside and outside the classroom.

4. Conclusion

The results of this small-scale study suggest that even the adult learners can exhibit, with the help of systematic instruction and the application of monitor strategies, great potential for change in pronunciation in a relatively short period of time. However, it is not certain whether such effects persist over time and carry over to everyday use of their TL. Furthermore, it is seen in this research that pronunciation of individual sounds may not improve as radically in spontaneous speech as in vocabulary and sentence-reading. Nevertheless, it is expected that most of the participants will develop into optimal monitor users with constant practices and finally succeed in producing target-like pronunciation in normal communication. Apparently, this is only a preliminary research carried out in this field and it appears to be an impossible task to carry out a prolonged investigation due to time limit. Therefore, a considerable amount of further research will be needed to evaluate the efficacy of this study and the ultimate effect of instruction and monitor.

References

Acton, W. (1997). Seven suggestions of highly successful pronunciation teaching. *The Language Teacher* 21 (2), 21-25.

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Derwing, T & Munro, M. (2005). Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39 (3), 379-397. doi:10.2307/3588486, <http://dx.doi.org/10.2307/3588486>

Dickinson, L. (1988). Learner training. *ELT Documents*, 131, 45-53.

Dlaska, A. & Krekeler, C. (2008). Self-assessment of pronunciation. *System*, 36 (4), 506-516. doi:10.1016/j.system.2008.03.003, <http://dx.doi.org/10.1016/j.system.2008.03.003>

Hammerly, H. (1991). *Fluency and accuracy*. Clevedon, UK: Multilingual Matters.

Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.

Krashen, S. (1985). *The input hypothesis: issues and implications*. London: Longman.

Krashen, S. (1988). *Second language acquisition and second language learning*. Oxford: Oxford University Press.

Marks J. (2005) Review: English pronunciation in use. *ELT Journal*, 59 (2): 177-179. doi:10.1093/eltj/cci032, <http://dx.doi.org/10.1093/eltj/cci032>

Murphy, J. (1991). Oral communication in TESOL: integrating speaking, listening, and pronunciation. *TESOL Quarterly*, 25 (1), 51-75. doi:10.2307/3587028, <http://dx.doi.org/10.2307/3587028>

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25 (3), 481-520. doi:10.2307/3586981, <http://dx.doi.org/10.2307/3586981>

Table 1. Ten vocabulary items in the pre-test

Word	Acceptable pronunciation	Unacceptable pronunciation	Word	Acceptable pronunciation	Unacceptable pronunciation
very	/vɛri/	[vɛri]	road	/rəʊd/	[roud]
thing	/θɪŋ/	[sɪŋ]	window	/wɪndəʊ/	[wendəʊ]
had	/hæd/	[hed]	sun	/sʌn/	[sang]
pleasure	/pleɪʒə/	[plɛrə]	long	/lɒŋ/	[lang]
see	/si:/	[seɪ]	down	/daʊn/	[dang]

Table 2. Words tested in vocabulary and sentence reading

Words tested in vocabulary reading		Words tested in sentence reading	
five	row	live	tomorrow
faith	wind	think	windy
mass	run	bad	fun
measure	song	treasure	strong
C	town	sea	downstairs
Sentences			
(1)	It's windy tomorrow.		
(2)	I think there are lots of treasures in the sea.		
(3)	He feels a strong desire to play downstairs.		
(4)	It's not bad to live in this city, where life is full of fun.		

Table 3. Results of paired samples test for the control group

	Pairs	Std. Deviation	t	df	Sig.(2-tailed)
1	T2 vs. T1	.3051	1.795	29	.083
2	T3 vs. T1	.4901	.372	29	.712

Table 4. Results of paired samples test for the experimental group

	Pairs	Std. Deviation	t	df	Sig.(2-tailed)
1	T2 vs. T1	2.0625	14.429	29	.000
2	T3 vs. T1	2.0410	12.882	29	.000

Table 5. Differences between the two groups in T1, T2 and T3

Experimental vs. Control		t	df	Sig.(2-tailed)
T1	vs. T1	-.947	58	.348
T2	vs. T2	13.309	58	.000
T3	vs. T3	10.747	58	.000

Table 6. Performances of N in T1, T2, T3, and T4

Target sound		N	Percentage of N			
			T1	T2	T3	T4
1	/v/	30	.63	1.00	.87	.77
2	/θ/	30	.57	1.00	.80	.70
3	/æ/	30	.03	.93	.77	.33
4	/ʒ/	15	.40	.67	.53	.47
5	/si:/	11	.36	.91	.82	.73
6	/rəʊ/	7	.57	.86	.71	.71
7	/wɪn/	3	.00	.67	.67	.33
8	/ʌn/	20	.05	.90	.80	.65
9	/vŋ/	11	.09	.82	.82	.64
10	/aʊn/	7	.00	.57	.43	.29

Appendix One: Five Topics in Spontaneous Speech

- 1 On Happiness.
- 2 Sexual Discrimination in China.
- 3 A Typical Chinese Festival.
- 4 The Necessity of Environmental Protection.
- 5 Elaborate on anything which you are interested in.