Effects of Visual Feedback on Second Language Productive Phonology
By Jilani S. Warsi

Introduction

The acquisition of second language productive phonology is seldom ever completely successful with adult learners. Most researchers contend that adult language learners cannot achieve native-like phonology in their second language (L2), and attribute the failure, principally, to language transfer and age-dependent factors. For example, Scovel (1969, 1988) maintains that no adult ever achieves native-like pronunciation in a L2. Some researchers suggest that successful attainment of L2 phonology is extremely rare (Oyama, 1976; Flege & Fletcher, 1992; Flege, Munro, & MacKay, 1995; Young-Scholten 1995). However, with individualized practice, there is evidence that the learners’ performance is improved (Hill 1970, Neufeld 1977). These scholars argue that second language productive phonology is attainable regardless of the learner’s age and first language. They maintain that there are methods that can enhance the teaching of L2 pronunciation and that can help students acquire native or near native proficiency in pronunciation. The present research builds on this direction of instruction.

Hypothesis

This study proposes to test the hypothesis that adult learners practicing L2 sound, with the ability to see on a diagram articulatory movements (point and manner of articulation) and conscious modifications of their researcher-prompted output, will approximate closer the target sounds, with the result of more native-like production and a
more rapid progress. The idea behind this proposal is that it isn’t just practice of sounds that improves the productive phonology, but informed practice. Teachers may give feedback to the student, but my hypothesis is that instruction that can be seen and then the output-modified will work better. The hypothesis is based on the assumption that to acquire new speech sounds, L2 learners need visual instruction.

**Phonological Fossilization**

As Selinker (1972) points out, the most important fact concerning L2 phonology is the phenomenon of fossilization. He claims that “fossilizable linguistic phenomena are linguistic items, rules, and subsystems which speakers of a particular native language will tend to keep in their interlanguage relative to a particular target language, no matter what the age of the learner or amount of explanation or instruction he receives in the target language.”

Similar views are put forth by Tarone (1976), Nemser (1971), and Sridhar (1980), who have tried to explore the causes of fossilization in language learner’s interlanguage phonologies. There are two related questions here which have baffled scholars:

1. Is phonological fossilization inevitable for L2 learners?
2. What are the causes of such fossilization?

According to Scovel (1969), the answer to the first question is a resounding yes. He contends that adult language learners maintain a typical accent which is indicative of their first language (L1). Scovel has named this the ‘Joseph Conrad Phenomenon’ after the prominent British author who achieved native-like fluency in English syntax (his L2) but retained a Polish accent (his L1). Scovel is so confident of his theory that he promises to offer a free dinner to anyone who can show him someone who learned a L2
after puberty and who speaks that L2 with perfect native like pronunciation. No one has, hitherto, been able to produce such an individual to Scovel.

As mentioned previously, some researchers do not go along with this idea. Hill (1970) maintains that phonological fossilization is by no means inevitable. Neufeld (1977) argues that there are methods that can enhance the teaching of pronunciation of a L2 and that can help students acquire native or near native proficiency in pronunciation. However, the subjects of Hill and Neufeld have not been examined by L2 acquisition researchers to determine whether they really achieved native-like pronunciation in their respective second languages. It seems that the question of the inevitability of phonological fossilization in adults remains undecided.

The second question is complicated and requires serious attention. One possible explanation for the causes of phonological fossilization is the atrophy of the nerves and muscles necessary for articulation. This theory maintains that the nerves and muscles instrumental in articulating second language pronunciation patterns have atrophied so that native-like pronunciation is almost impossible. This notion, however, has not been proven empirically.

Another physiological explanation comes from Lenneberg (1967) who suggests that after puberty, it is difficult to master the pronunciation of a L2 because a critical period in brain maturation has been passed and “...language development tends to freeze.” He calls this phenomenon “lateralization” - the completion of cerebral dominance. According to him, lateralization impedes the learning of the phonology of a L2 more than the learning of the syntax or vocabulary of a L2.
Contrary to Lenneberg, Flynn and Manuel (1991) argue that the effects of age-dependent variables on the language acquisition process, and the universal properties shared by language learners are not known clearly. They argue that lateralization does not increase by age, and it is hard to reconcile the fact that plasticity is the determining factor in language acquisition and that the brain becomes less functional with age. Discussing modularity and categorical perception, Flynn and Manuel (1991) note that perceiving and discriminating between speech sounds is a specialized behavior. They claim that adult L2 learners don’t lose their ability to perceive speech sounds, but they have difficulty with certain perceptual distinctions. To acquire new speech sounds, L2 learners need feedback which need not be auditory, they suggest. Finally, they point out that the critical period hypothesis is less convincing because it doesn’t account for successful second language learners. It should be noted, however, that they studied speech perception in a L2. Archibald (1992, 1993a; 1993b) suggests that speech perception and speech production are relatively independent skills and should be teased apart in trying to unravel the puzzle of phonological fossilization. As stated previously, this study is concerned with L2 productive phonology and, therefore, will focus on the articulatory aspects of phonology.

A somewhat different position has been taken by Krashen (1977) who opposes Lenneberg. He maintains that adolescents consciously construct abstract theories about the world during the course of their cognitive development. They tend to learn the L2 by abstracting grammar and pronunciation rules and applying them. It is obvious that this theory considers L2 acquisition the same as learning a L1. Krashen calls this process ‘creative construction’ and argues that the close of the critical period is related to Piaget’s
stage of formal operations. In another study, Krashen and Harshman (1972) reanalyzed Lenneberg’s data and came to a conclusion contradicting his finding. They argue that lateralization takes place long before the end of the critical period for language learning. However, Tarone (1978) does not agree with Krashen and Harshman and asks why formal operations should affect only the pronunciation and not the syntax or morphology. This indeed puts a question mark on the formal operation type of psychological explanation for phonological fossilization.

Another psychological explanation is related to the issue of language transfer. Theoreticians claim that transfer has its strongest effect on the pronunciation of a L2 (Broselow 1988). However, Neufeld (1977) reports on a study in which he used a new technique to enhance teaching second language pronunciation to adults. Instead of linking language transfer with L2 productive phonology, he says that adult learners tend to form inaccurate acoustic images of the target language sound patterns, thus attributing this to inappropriate learning situations. These acoustic images get set once they are formed. This leads to the fixation of the learner’s pronunciation patterns. He maintains that the learner’s inability to perceive and articulate a new sound could result from his or her psychological inability to alter the criteria used to categorize speech sounds. It is, however, not clear from his discussion why adults are affected by acoustic images and children are not.

A third type of explanation is very different from psychological habit formation and uses arguments related to affective factors to prove that interlanguage pronunciation is a sensitive indicator of adult learners’ lack of empathy with the native speakers and culture of the L2. Unlike children, who are more compatible to L2 culture, adults have
more rigid language ego boundaries. They may be inclined to establishing their cultural and ethnic identity and this they do by maintaining their stereotypical accent (Guiora et al. 1972)

According to Guiora et al. (1972), adults do not have the motivation to change their accent and to acquire native-like pronunciation. These researchers attempted to mitigate the empathy level of their subjects by administering increasing amounts of alcohol. They found that the learners’ pronunciation of the target language sounds improved to a certain point and then decreased as they drank increasing amounts of alcohol. However, a different explanation could be that subjects were under the influence of alcohol and had less difficulty in articulating the target language sounds because of muscle-relaxation.

That socio-emotional factors are powerful in determining degree of proficiency in pronunciation cannot be denied. It should be noted that these factors are hard to determine in an experimental setting. Nevertheless, the findings of Guiora et al. may have some feasible implications for the use of socio-emotional factors in enhancing the learning process.

At this point, we don’t have a clear understanding of what causes phonological fossilization. It is obvious that none of the above discussed explanations provides deep insights into this debatable phenomenon. There is persuasive evidence that supports the existence of different processes and constraints that cause phonological fossilization. It is not clear, however, whether it is influenced and determined by inadequate phonetic input, by lack of motivation to acquire the L2 sounds, by gradual deterioration of some
basic speech learning mechanisms, or by inability to keep the L1 and L2 phonological systems from interacting with one another (language transfer).

One question that is pertinent to these issues is whether adult learners can produce the L2 sounds just like native speakers of the target L2. The present study examined the production of English /l/ and /r/ by native speakers of Japanese since these are the most problematic L2 sounds for them to pronounce. Japanese speakers of English often identify English liquids /l/ and /r/ with Japanese liquid /r/, and, as a result, approximate and substitute the target L2 sounds with Japanese /r/. The process is called interlingual identification, and is triggered when the perceptually similar L2 and L1 sounds differ acoustically and auditorily. What is interesting is that this identification can extend from a perceptual level to a productive level (Lehiste, 1988; Flege, 1988).

**Research Question**

Can adult Japanese speakers of English improve their pronunciation of English /l/ and /r/ with informed practice and, thus, transcend the process of interlingual identification?

**General Methods**

**Subjects**

Four Japanese speakers of English, studying multi-skill English as a Second Language (ESL) courses at Salem State College, Massachusetts, were chosen for this study. They were drawn from the same ESL core class, which met Monday through Friday from 9 am to 12 pm. That is, they had 15 hours of intensive English a week and had the same instructor, who was a native speaker of English. In addition to taking the core class, they were taking a Test of English as a Foreign Language (TOEFL)
preparatory course in order to get a minimum score of 500 on the TOEFL test. Their age varied from 16 to 18. At the time of the study, they were living in campus housing.

All subjects had had 6 years of English in Japan before they came to the US. They started learning English in Japan at the age of 12. Their instructors were Japanese, and the use of English, which mainly consisted of reading and writing, was restricted to the classroom. All of them spoke Japanese in Japan to communicate with their parents, relatives, and friends. This was contrary to the situation in America where they used English for academic and communicative purposes. However, they frequently used Japanese among themselves and with their Student Counselors.

The subjects were divided into two groups: control and experimental. There were two subjects in each group.

Data Collection/Analysis Procedures

Word list

The following five words consisting of the target sounds /l/ and /r/ were given to the subjects:

1. root
2. star
3. late
4. chocolate
5. call

The reason for selecting these words was the fact that even though learners sometimes approximate the phones in isolation, they still have trouble pronouncing them in different word positions, mainly because of the sounds preceding and following them.
The target sounds /l/ and /r/ in the above words occur in both word-initial and word-final positions, which is effective in determining whether the subjects have mastered these sounds in different phonological environments. The data for word medial /r/ were collected, but not all data were used for analysis for purposes of this study and availability constraints of the judges.

**Pre-test**

Both the control group and the experimental group pronounced the words in the order in which they were presented, i.e. from 1 to 5. Their production was recorded on audiotapes.

**Training**

The subjects in the experimental group spent 2 hours practicing the target sounds twice a week for a period of 4 weeks. They were shown two diagrams, one for each sound, that showed the exact point of articulation of the target sounds. The investigator also explained the organs of speech and manner of articulation. Technical details such as liquid and retroflex were not discussed to preclude complexity and misunderstanding. The investigator reinforced the technique at every meeting so that the subjects could understand the speech mechanism and the role of the organs of speech in producing the sounds being studied. During the 2 hours of training, the investigator did not model the target sounds for them. As the subjects attempted to pronounce the target sounds, the investigator prompted them frequently, indicating how far and/or close they were getting from/to their target.

/l/

**Diagram**
In addition to the diagrams, they were given the following oral and written instructions for each sound:

/r/

Diagram

/l/

1. Place the tip of your tongue against your upper gum ridge.
2. As you make the sound, air flows out over the sides of your tongue.

/r/

1. Raise the tip of your tongue towards the upper gum ridge but do not touch it.
2. The tip of your tongue should not touch anything.
3. Press the sides of your tongue against your upper back teeth.

As mentioned before, the target sounds were not modeled by the investigator. Therefore, the subjects had to rely on visual instruction. They only had the list of words, a list of instructions, and the two diagrams to look at as they pronounced the sounds targeted in the study. The training lasted 4 weeks. The control subjects did not receive the special training. They were orally tested along with the test subjects.

Post-test

After the completion of the training, both the control group and the experimental group’s production of the target sounds was recorded on audiotapes. They pronounced the same five words that they had pronounced before the training. However, this time the words were presented in a mixed order to keep the subjects from memorizing the sequence in which the words were presented before the training.

A panel of 10 native English speakers tested the subjects’ production of the target sounds /l/ and /r/ by listening to the audiotape and ranking them on a scale of 1 to 5, with 1 being non-native and 5 being native-like. The judges consisted of 9 undergraduate students who were taking an introductory course in linguistics at Boston University and a professor of linguistics who was teaching the course. The data, including the pretest and the posttest productions, were randomized again to preclude a response bias pattern. The purpose of the study was not explained to the panel of judges in order to minimize any
conscious effort on their part to be biased in determining how non-native or native-like the subjects’ production was. They also had no prior knowledge of which word was pretest and which word was posttest.

The judges tallied oral production on a scale of 1-5. The errors which allowed minimal intelligibility but fell short of native-like production were weighted, with close to native form receiving 5 points, different but intelligible receiving 4, between non-native and native-like receiving 3, somewhat unacceptable receiving 2, and completely unacceptable receiving 1.

Scores given by the judges were computed. These scores were obtained by averaging over responses obtained for each subject. An overall mean was computed for laterals /l/ and retroflexes /r/ spoken by all 4 subjects. It appears that the judges were not consistent among themselves in determining exactly which particular production of the targeted sound was far from or close to native-like production.

Results

The first part of this section will discuss the pretest and the posttest productions of the target sounds that came from the subjects in the experimental group (S1 and S2). After that, the pretest and the posttest productions obtained from the subjects in the control group (S3 and S4) will be discussed in the second part.

<table>
<thead>
<tr>
<th>S1</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Root</td>
<td>Star</td>
</tr>
<tr>
<td>1 Non-native</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>
As shown in table 1, S1’s production of the target sound /l/ and /r/ slightly improved. Compared to retroflexes, laterals spoken by S1 received a higher rating by the judges. With the exception of /l/ in ‘chocolate’, this finding suggests that S1 produced highly intelligible laterals. It should be noted, though, that S1 produced more intelligible laterals in ‘late’ and ‘call’, but the mean ratings for the production of /l/ in chocolate (1.8) remained the same after the training. In ‘late’ and ‘call’, /l/ occurs in word-initial and word-final position. On the other hand, it occurs in ‘chocolate’ in word-medial position after a syllable break. One possible explanation for S1’s failure to produce native-like /l/ in ‘chocolate’ is that syllabification plays an important role in second language speech processing. Target language syllable structure, especially those which are not permitted in the native language, seem to exert some influence on the pronunciation of second language learners (for a detailed study see Mehler, Domergues, Frauenfelder, and Segni, 1981).

Retroflexes spoken by S1 received a low rating. S1’s production of /r/ in ‘star’ deteriorated from 3.4 to 3.0. This can be considered a case of backsliding. However, the mean rating for the retroflex in ‘root’ improved from 1.4 to 1.9. On aggregate, S1’s production of the targeted sound improved from 2.54 to 2.94, which is not remarkable but given the amount and period of training, it can be considered satisfactory.

### TABLE 2
Pretest and posttest productions of the subject in the experimental group (S2)

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Root</td>
<td>Star</td>
</tr>
<tr>
<td>S2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As evident in Table 2, S2 showed an overall improvement in producing both laterals and retroflexes. The table shows the average and aggregate ratings of the five targeted words. The retroflex /r/ in root and ‘star’ improved significantly from 2.3 to 2.7 and from 2.1 to 3.6 respectively. Laterals produced by S2 received an even higher rating by the panel. /l/ in ‘late’ progressed from 2.6 to 3.9, whereas the lateral in ‘chocolate’ advanced from 1.4 to 2.6. In comparison to S1’s production of this problematic word, the difference in improvement may be attributed to individual variation. The mean ratings for /l/ in ‘call’ decreased slightly from 4.2 to 4.1, which is neither improvement nor deterioration. In the aggregate, S2’s production of the target sounds improved significantly from 2.52 to 3.38.

As shown in Table 3, the subject in the control group, S3, who had not received the special training, did not improve. Barring the exception of the production of /l/ by this speaker in ‘chocolate’ (2.8 -3.1), which is not a noticeable improvement, laterals and
retroflexes received a rather low rating. Retroflexes in ‘root’ and ‘star’ relapsed from 2.3 to 2.1 and from 3.7 to 3.1 respectively. In a similar manner, laterals in ‘late’ and ‘call’ regressed from 3.9 to 3.7 and from 3.7 to 2.9 respectively. It is difficult to determine why S3’s production of laterals and retroflexes deteriorated on aggregate from 3.28 to 2.98, but the backsliding might have been caused by a gamut of factors that cannot be quantified such as personality and attitudinal.

**TABLE 4**

<table>
<thead>
<tr>
<th>S4</th>
<th>Pretest productions of the control group subject (S4)</th>
<th>Posttest productions of the control group subject (S4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Word</td>
<td>Root</td>
</tr>
<tr>
<td></td>
<td>1 Non-native</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5 Native-like</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Aggregate</td>
<td>1.8</td>
<td></td>
</tr>
</tbody>
</table>

Compared to S3, S4 in the control group was slightly better. While his production of laterals and retroflexes did not show any significant improvement, his aggregate ratings did not deteriorate. The production of the /r/ phoneme in ‘root’ fell in the same range (1.2 - 1.3), whereas /r/ in ‘star’ produced by this subject improved slightly from 1.5 to 1.9. Laterals produced by S4 were rated differently for different words. For example, the mean ratings for /l/ in ‘late’ regressed from 3.1 to 2.4, whereas the average ratings for /l/ in ‘chocolate’ increased from 1.1 to 1.6. Finally, the /l/ phoneme produced by this subject received similar mean ratings (2.1 - 2.2). On aggregate, S4’s production of the target sounds edged up slightly from 1.8 to 1.88, which does not draw serious attention.

**Significance**
As hypothesized earlier, the findings of this study show that the subjects in the experimental group, as compared to the subjects in the control group, improved their performance with informed practice and with the ability to see diagrams showing the point of articulation. The statistics also indicate that the subjects in the experimental group approximated closer English phonetic forms with the result of both more native-like production and a more rapid progress. On the other hand, the subjects in the control group did not show any significant improvement despite the fact that they were in the same class as were the subjects in the experimental group and that they were taught by the same English speaking instructor. It is, therefore, reasonable to assume that the special training had an impact on the productive phonology of the experimental subjects and that it accelerated, to a considerable extent, the process of acquiring problematic L2 sounds.

It would be inappropriate to claim, however, that native-like productive phonology can only be achieved with the special training the subjects of this study received. As mentioned earlier, even though it is extremely rare that adult language learners succeed in acquiring native-like phonology and manage to avoid producing their L2 phonetic forms without a detectable foreign accent, some learners may succeed in overcoming their problems in producing difficult L2 sounds without any phonetic training. The performance of the subjects in the control group may improve with more exposure to the target culture and individualized practice, but Flege et al., (1995b) consider it an exception.

From an acquisition point of view, the fact that visual training supplemented with articulatory instructions accelerated the learning process is quite significant. It is
needless to say that additional research is needed to determine whether all L2 sounds (vowels and consonants) can be mastered by learners of different language backgrounds, as well as to determine exactly what kind of phonetic training is needed to facilitate successful training. As mentioned previously, speech perception and speech production are relatively different skills (Archibald, 1993). It is, therefore, important that they should be kept separate by not providing teacher modeling, especially in teaching L2 productive phonology. The fact that the subjects were able to improve their pronunciation of English /l/ and /r/ without receiving external modeling of the target sounds is extremely important from the learning perspective.

The questions that arise from the present study are: Why visual training has a constructive effect on the learners’ productive phonology and why adult learners need that special training in order to approximate L2 phonetic forms? In other words, What is it that (might) decline with age that requires the provision of phonetic training in order for native-like phonology to be attained? Is the system attained comparable to that of a native speaker? Did the effects of the training last? Are all learners subject to such intervention equally successful? And finally, how is it that some adults seem to reach near-native competence without such intervention? These questions are of theoretical and practical interest and should be pursued further. Second language learners use information-processing strategies or problem-solving procedures, which make adult language learning very different from child language acquisition. Although the input processing strategy may not work sometimes, “the insight that acquisition involves input-processing strategies of some kind is important and should be pursued” (White, 1991).
Our understanding of second language speech learning is limited because of the complexity of phonological processes, individual variation, attitudinal factors, etc. Future research must be conducted to answer the questions raised in this study and to unravel the mystery of phonological fossilization.


