SPOKEN L2 TEACHING
WITH CONTRASTIVE VISUAL AND AUDITORY FEEDBACK

Anne-Marie Öster

Department of Speech, Music and Hearing, KTH,
100 44 Stockholm, Sweden

ABSTRACT

Teaching strategies and positive results from training of both perception and production of spoken Swedish with 13 immigrants are reported. The learners participated in six training sessions lasting thirty minutes, twice a week. The training had a positive effect on the L2-speakers' perception and production of individual Swedish sounds, stress, intonation and rhythm. The positive results were obtained through auditory and visually contrastive feedback provided through a PC running the IBM SpeechViewer software. Skill building modules together with the Speech Pattering Module "Pitch and Loudness" was used, that displays the speech signal as graphical curves and diagrams. A split screen offered a comparison of the student's production with a correct model by the teacher. Pitch and loudness were displayed either separately or combined.

1. INTRODUCTION

Traditionally, teaching of Swedish as a second language has seldom focused on pronunciation training. However, nowadays the importance of an acceptable pronunciation of Swedish immigrants is stressed and the uses of efficient training methods are accentuated. Computer-based speech training with visually and auditory contrastive feedback has shown to provide a valuable resource in spoken L2 teaching.

The system used consists of a PC running the IBM SpeechViewer software. A microphone, an amplifier, and a speaker connected to a sound-card allow the user to input, store and analyse speech and then display it and play it back. The software contains several interactive programs and has shown to be successful in assisting hard-of-hearing children in achieving awareness and control over various speech attributes such as voicing, timing, pitch, and loudness as well as refined articulation and prosody [1], [2], [3].

Our results and experiences from visual speech training with hard-of-hearing children have shown that in order to be efficient the aid should provide a contrastive visual training, that is, the correct model of the teacher and the deviant production of the learner must be shown simultaneously to be compared with each other. Moreover, the visual pattern should be attractive, easily comprehensible and shown without delay.

Ongoing research in the phonetics of second language acquisition has pointed out the learners´ failure to perceive and produce difficult and new contrasts in the target language by ear alone [4], [5], [6]. The use of visual information as an alternative feedback in the training of L2-speakers vowel production has been used by Flege [7] among others. Some preliminary work based on the SpeechViewer has been applied to the pronunciation training of immigrant learners of Swedish by Öster with good result [8].

This paper reports preliminary results from a project together with the Unit for Languages and Educational Research and Development at KTH where this new L2-teaching strategy was used with adult international engineers learning Swedish.

2. METHOD

2.1. Speakers

The 13 speakers (4 females) studied technical Swedish at KTH. All of them except one had an academic degree in engineering from a university in their home country. Their age varied from 25 to 46 years. Six of them originated from Bosnia, the others from Cuba (2), Peru (2), Saudi Arabia (2) and Russia (1). On an average they participated in six training sessions lasting half an hour, twice a week.

2.2. Speech materials

Before training the speakers were recorded when they read a text of 17 sentences and 110 isolated words. This diagnostic speech material contained all Swedish vowels (long and short) and all Swedish consonants in initial, medial and final positions.

Two trained phoneticians transcribed the recorded words and made an assessment of each L2-speaker´s individual segmental deviations. The recorded sentences were also evaluated by the phoneticians according to the speakers´ accuracy in producing stress patterns, intonation, accent and rhythm. The segmental and suprasegmental deviations found in the two assessments formed the basis for each speaker's individual speech material for the pronunciation training that followed.

2.3. Visual feedback

By means of the visual and optional auditory feedback given in the module "Pitch and Loudness" efficient training of both perception and production of spoken Swedish was viable. It was easy to make the learners aware of in what way his/her production deviated and show him/her how to produce it correctly.

The visual feedback was very instructive and pedagogical because it was logical, motivating and shown in real-time. In

the following two figures the visual information that this module provides can be studied.

The program displays pitch and loudness individually or combined over time. The split screen offers an exact comparison of a learner's utterance with a model of the teacher.

In figure 1 the voiced and voiceless contrast is clearly indicated by different colors in the loudness application of the module "Pitch and Loudness". Voiced sounds are red and voiceless sounds are green. In the upper portion the teacher's correct production of the word pairs /puss/ (kiss) and /buss/ (bus) is shown and in the lower portion a Bosnian speaker's production is shown. It can be seen that the Bosnian speaker has difficulties to produce the phonetic contrast between the initial unvoiced aspirated plosive and the voiced plosive in Swedish.
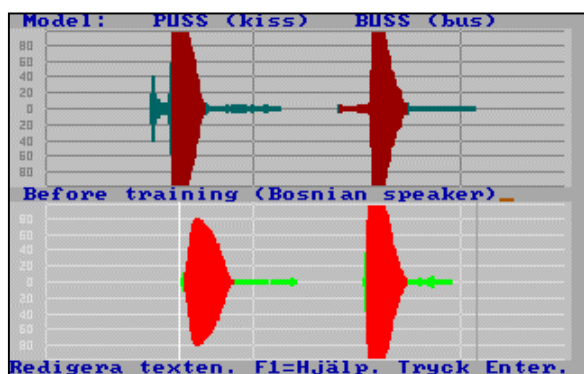

Figure 1: The upper portion shows the teacher's model of correct produced voicing contrast between the initial plosives /p/ and /b/. The lower portion shows a deviant production of a Bosnian speaker.
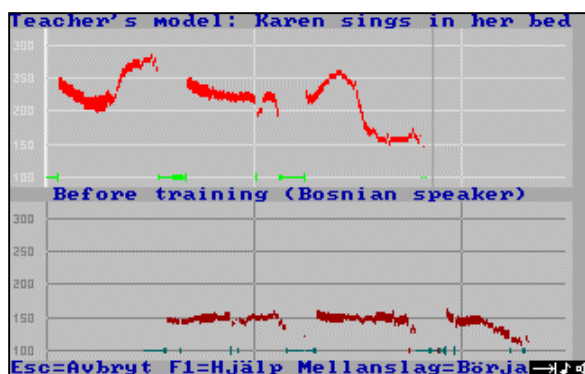

Figure 2: In the upper portion, the teacher's correct intonation pattern of the phrase "Karin sjunger i sängen" (Karen sings in her bed) is shown. In the lower portion, a Bosnian speaker's intonation pattern is shown before training.

Figure 2 shows a training session of the intonation pattern in the Swedish sentence "Karin sjunger i sängen" (Karen is singing in the bed) in the application where pitch and loudness are combined over time in the module "Pitch and Loudness". Stressed syllables are visible on the teacher's screen through pitch modulation, longer duration and higher intensity (thickness of the line). On the lower screen a Bosnian speaker's quite monotonous production is shown. The contrastiveness of the visual patterns gave a learner possibilities to discriminate between various distinctive features that underlie phonological contrasts in Swedish.

3. TRAINING STRATEGY

While the specific structure of spoken L2-teaching may vary somewhat from teacher to teacher there are certain general steps to be followed to obtain an efficient training. Following steps are recommended:

3.1. Diagnosis of individual deviations and selection of an appropriate speech training material

The first stage involves an assessment of the deviations that should be trained. All assessments are time-consuming and laborious. However, to prevent speech training from becoming a series of meaningless "articulatory gymnastic" sessions, the teacher must be prepared to spend some time carrying out an appropriate diagnosis prior to training.

As an example of the result of a general diagnosis segmental deviations by a group of six Bosnian speakers are shown. Deviantly produced Swedish vowel-quality is shown in vowel diagrams, fig 3, and deviantly pronounced consonants are inserted in a table that shows applicable positions fig 4. A summary of general prosodic deviations are also given below.

Vowels: To insert deviations in vowel diagrams could be a good pedagogical aid to describe the role of the tongue in vowel production.

Swedish language has 9 phonologically long and 9 short vowels. Quantity is an important contrast in Swedish, which is realised as a difference in duration and, for some pairs, also with a difference in vowel quality. Furthermore, the following consonant has opposite quantity, i.e.; long stressed vowels are followed by short consonants and vice versa.

High and mid-high long Swedish vowels are diphtongised in open syllables. Noteworthy is also the number of Swedish front vowels and the fact that some of them are rounded. As a consequence, it is often difficult for L2-speakers to have a complete productive and perceptual mastery of front vowels.
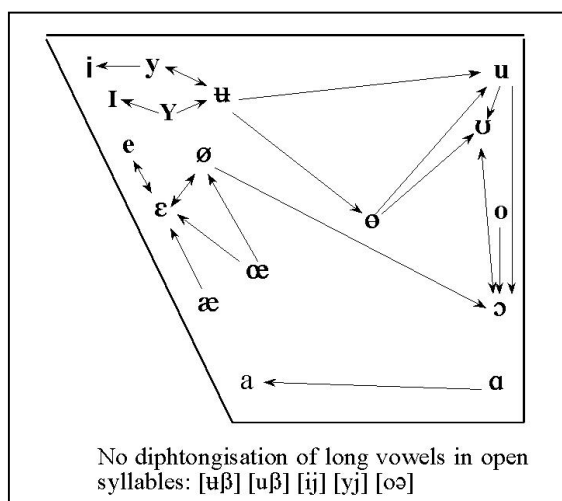
No diphtongisation of long vowels in open syllables: [ʉβ] [uβ] [ij] [yj] [oɔ]

Figure 3: Deviantly produced Swedish vowel-quality by six Bosnian speakers. Lines and arrows indicate typical confusions.

Consonants: Deviantly produced consonants are inserted in a table that shows all possible positions. Some of the Swedish consonants cause production and perception difficulties for L2-speakers, especially the voicing contrast between plosives, aspirated voiceless plosives, the great number of fricatives, the nasal velar sound and (not in assessment table) the fact that dentals are retroflexed when they are followed by /r/.

IPA	initial	medial	final
p	b	b	b
t	d	d	d
k	g	g	g
b			p
d	t	t	t
g		k	k Ø
f			
v			
s			
ʃ	ç ʃj sj	çj	
ç	ʃ		
j			
h	ç χ		
m			
n			
ŋ		ŋg	ŋg
l		ɫ	ɫ
r			

Figure 4: Deviantly produced Swedish consonants by six Bosnian speakers.

Prosody: Prosodic deviations are evaluated and summarised when listening to the recorded text. To sum up, the most common prosodic deviations made by the L2-speakers in this study were:

- difficulties in perceiving and producing accent II

- difficulty in producing a long consonant after a short vowel in stressed syllables

- no quantity differences between phonologically long and short vowels.

3.2. Instructions that aim at awareness, correct realisation, and understanding

It is important to make the learner aware of in what way his/her production deviates and show him/her how to correct it. This was done through a contrastive instruction; that is, the split screen showed simultaneously the correct model of the teacher and the deviant production of the learner. Instruction was based on this possibility of comparison in the module "Pitch and Loudness".

The most instructive speech material has been shown to be minimal word pairs that contain the contrast that the learner produces deviantly.

3.3. Further training to establish automaticity and transfer to untrained material

When studying the following figures it is obvious that this contrastive training with both visual and auditory feedback had a certain effect on the L2-speakers´ production. The figures show positive results after some training of the same material as shown in figure 1 and 2 (actually exaggerated).
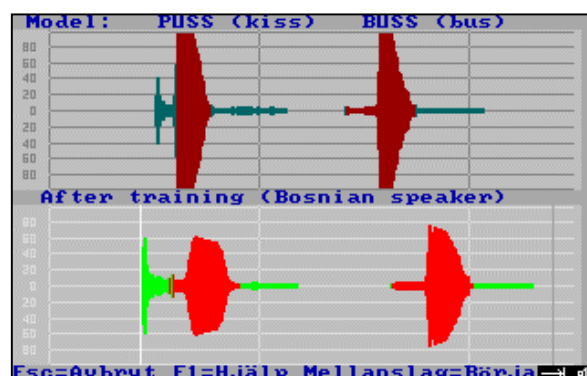

Figure 5: The upper portion shows the teacher's model of correct produced voicing contrast between the initial plosives /p/ and /b/. The lower portion shows a correct production by a Bosnian speaker after some training (cf. Figure 1).

When the learner has acquired the target production, significant amounts of additional training must be invested to get him or

her to maintain correct production and to transfer skills to untrained situations. This is the most important element in a speech training program but the most difficult to carry out. The target production must be repeated and practiced in a variety of contexts. The ultimate goal is a system expansion that is accomplished when the learner's best production becomes his or her most common production.
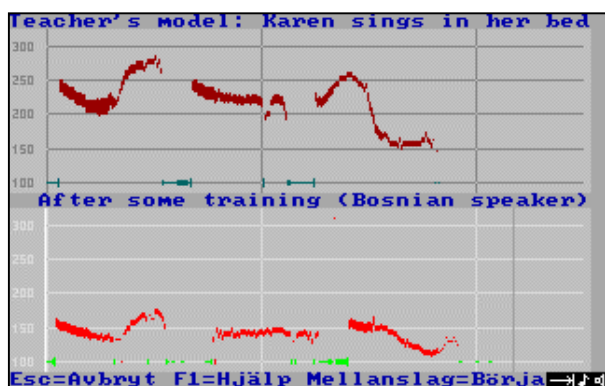

Figure 6: In the upper portion the teacher's correct intonation pattern of the phrase "Karin sjunger i sängen" (Karen sings in her bed) is shown. In the lower portion, a Bosnian speaker's improvement after some training is shown (cf. Figure 2).

Skill building modules were also used which provide a gamelike strategy to strengthen ability in refined articulation. Phonemes produced by the learner were matched against models by comparing target spectra. To avoid the problem of interspeaker variability the L2-speaker´s own "best production" was stored as a target. This strategy has also been used with the CISTA-Aid training [9].

4. FINDINGS IN THIS STUDY

The L2-speakers´ability to produce and discriminate between Swedish speech sounds, stress patterns, intonation, word accent and rhythm improved in a nativelike manner after training using this audiovisual tool. The auditory and visually contrastive feedback made them understand in what way their speech deviated from the Swedish norm.

Some of the Swedish phonetic contrasts were more easily learned than others. The L2-speakers´ competence to perceive and produce difficult and new phonetic contrasts were possibly depending on their mother tongue

A general opinion of the L2-speakers was that they became more aware of the Swedish pronunciation and spoke a bit more Swedish after training. However, most of them wanted that the training period should be extended for better long-term results.

Age of first exposure, motivation and first language may affect the L2-speakers´ ability to perceive and produce new speech contrasts in the target language. These factors will be investigated in a following study.

5. REFERENCES

1. Öster, A-M. 1989. "Applications and experiences of computer-based speech training", *STL-QPSR 1, 59-62.*

2. Öster, A-M. 1995. "Teaching speech skills to deaf children by computer-based speech training", *Proceedings of 18th International Congress on Education of the Deaf, Tel-Aviv, Israel.*

3. Öster, A-M. 1996. "Clinical applications of computer-based speech training for children with hearing impairment". *Proceedings of ICSPL 96, Philadelphia.*

4. Jamieson, D. G. 1995. "Techniques for training difficult non-native speech contrasts", *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm.*

5. McAllister, R. 1995. "Perceptual foreign accent and L2 production", *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm.*

6. Flege, J. E. 1998. "Second-language learning: The role of subject and phonetic variables", *Proceedings of Speech Technology in Language Learning (STiLL 98), Marholmen, Sweden.*

7. Flege, J. E. 1989. "Using visual information to train foreign language vowel production", *Language Learning 38, 365-407.*

8. Öster, A-M. 1997. "Auditory and visual feedback in spoken L2 teaching", *Reports from the Department of Phonetics, Umeå University, PHONUM 4.*

9. Youdelman, K. 1994. "Computer applications in teaching speech to deaf children", *Proceedings of Second International Symposium on Speech and Hearing Sciences, Japan.*