The Effectiveness of Low-Pass Filters in English Pronunciation Training

著者（英）	Midori IBA
journal or publication title	Language and Culture : The Journal of the Institute for Language and Culture
volume	13
page range	1-16
year	2009-03-15
URL	http://doi.org/10.14990/00000478

The Effectiveness of Low-Pass Filters in English Pronunciation Training

Midori IBA

The purpose of this study is to investigate the effectiveness of applying low-pass filters to computer-assisted pronunciation training of English. Electric low-pass digital filtering of speech has been used in a considerable number of experiments to highlight the prosodic features of speech. Some attempts to apply low-pass filters to language learning have been made but the effectiveness has not yet been empirically proved. If the application is found to be effective to acquire the prosody of the target language, we might develop the training software using low-pass filtered sound. The experiment, using a pretest-posttest design, provided 13 native Japanese-speaking learners of English with 10 time training sessions focused on prosody using a real-time computerized pitch display. Multiple exemplars produced by native speakers of English provided training feedback. A group of seven learners were trained with low-pass filtered models while another six students were given non-filtered examples. Learners' recorded pre- and posttest productions were analyzed by computer software. Acoustic analyses were used to determine how accurately the suprasegmentals were produced. In comparison of the low-pass group with the non-filtered group, a regression ANOVA was used and it revealed a significant difference between the two. Questionnaire responses indicated a greater awareness of the various aspects of speech and increased confidence in producing the target language.

1. Introduction

1.1. Low-pass filters

By definition, a low-pass filter is a circuit offering easy passage to low-frequency signals and difficult passage to high-frequency signals. It eliminates certain frequency components of sounds. Such a filter was originally used to direct high frequencies to a tweeter speaker for music or speech. Eliminating the high frequencies of speech in a signal that sounds muffled. Segmental content of speech is no longer intelligible though the prosodic information remains. As it highlights prosodic features, the low-pass digital filtering of speech has been applied to various fields such as speech therapy, experiments with learning disability, and neuroscientific experiments.

For instance, low-pass filtered speech effectively separates children with learning problems from those who are normally achieving (Keith T & Farrer, 1981). Filtered word testing is one subtest of the SCAN-C auditory test battery (Keith, 2000). Hearing loss and processing problems should be evaluated as separate issues, although they may be closely related. Children with learning disabilities often show signs of auditory processing difficulties. According to Rosenkötter, high-frequency filtered music seems to be highly effective in the treatment of children with hearing problems (Rosenkötter, 1996). A psychophysical experiment was undertaken to investigate whether male and female listeners differed with respect to which frequencies were important in the perception of a male voice as a natural sound object (Hunter, Phang, Lee & Woodruff, 2005).

Children with hypersensitivity suffer from many stressful and disturbing symptoms; they may over-react to common noises, and be distressed by classroom sounds. Filtered sound training for those children has resulted in significant improvements in their hearing ability and behaviors. This suggests that masking segmental and semantic information by filters may affect auditory processing in the human brain. If we adopt filtered training to language teaching for normal-hearing learners, what would happen? This was the first question I had upon reading the papers mentioned above.

1.2. Application of low-pass filters to pronunciation training

Perception and production training with low-pass filters has already been adopted as a therapy for hearing-impaired people. It was originally created by Petar Guberina (Guberina, 1972, 1976), a Croatian psycholinguist who was working on problems of perception and production with hearing-impaired people as well as people with normal hearing. His work in this area is based on what he called verbo-tonal theory. This theory provides us with interesting ways of thinking about perception and learning in general. Guberina’s notion is that deafness can be thought of not so much as a condition caused by a physical defect but as a way of organizing the world which differs from strategies which have been learnt by people who are not deaf. Guberina’s work has subsequently been augmented and reframed through application of the thinking of Jack Derrida (1982), Pierre Boudieu (1991, 1995), Ann Freadman (1994), and Ania Lian (2003). Lian et al created MMExplore, a system designed to enable the exploration of authentic text in a variety of ways with emphasis on development skills. It enables the use of electronic low-pass digital filtering of speech to highlight intonation patterns.

These attempts at using low-pass filers in language training have been made over the last few decades, but are still not common in the field of language teaching. The emphasis in foreign language teaching is on achieving communicative effectiveness. Many learner-centered communicative approaches aim at enabling learners to successfully communicate in the target language. Pronunciation is an obvious component of communication and serious pronunciation problems are known to hamper communication or put learners at a social and professional disadvantage (Munro & Derwing, 1995). Recent studies have shown that tailor-made training is effective in improving perceptive and productive skills (Akahane-Yamada et al. 1998, Moyer 1999, Hardison 2004). However, in reality, the time that is generally available for pronunciation training in traditional classroom instruction has remained relatively limited in Japan where the grammar-translation method played an important role for a long period of time in importing necessary knowledge and information for modernization. Although the old educational paradigms have shifted to communicative approaches, pronunciation training is still peripheral. Computer-assisted pronunciation training might improve the current situation.

Low-pass training, as mentioned above, already exists in language training but mainly for hearing-impaired people and is highly limited for normal-hearing learners. Regarding the effectiveness of low-pass filters, it has not been empirically proven. Low-pass filters are used for language training and speech therapy on the assumption that they are factually effective. If the system of perception of speech sounds of the hearing-impaired differs from that of normal learners as Guberina mentioned, we should be more prudent about adopting the filter training for the non-disabled group. The research reported in this paper is aimed at examining the effectiveness of using low-pass filters in English pronunciation training for normal-hearing learners. The details of these research questions follow below.

1.3. Purpose of the study

This study is intended to investigate the effectiveness of applying low-pass filters to computer-assisted pronunciation training of English. If the application is found to be effective to acquire the prosody of the target language, low-pass filtered sounds can be used for designing efficient pronunciation training programs.

My research questions in this study were as follows:

1. Is there any difference in speech production between the group of subjects

who attended 10 sessions of training with electric low-pass filtered sounds and the controlled group who did the same training without filtered sounds? Did the low-pass group become more accurate in production than the non-filtered group, or vise versa? Were both improved?

2. How did the participants feel after they finish the 10 training sessions?

2. Experiment

2.1. The materials

The pretest for this experiment was basically the same as I designed and used in my previous study (Iba, 2007). The total number of stimuli in the pretest was 40 and 10 of them were compared to the same stimuli in the posttest which included 80 stimuli. When I designed the pre/post tests, I made six groups (Group A to Group F) of sentences and words as follows.

Table 1: The stimuli of the pre/posttests

A	1	Thank you very much for everything.	D	1	Will you read it again?
	2	I bought them drinks.		2	There's a crack in the glass.
	3	What's she saying?		3	I think I'll take a bath.
	4	They are all afraid.		4	Let's keep in touch.
	5	What would you like to do?		5	That class is easy.
	6	Give it to him.		6	He is on vacation.
	7	This is better than that.		7	I saw a flash of lighting.
	8	I know it's true.		8	Is it true that he is ill?
	9	think / sink		9	think/ sink
	10	right/ light		10	clothe/ close
B	1	He went over the path.	E	1	They're leaving next week on a trip around the world.
	2	They thought about it.		2	What's the matter with you?
	3	We've fired them.		3	Why won't you believe me?
	4	Repeat the word.		4	Richard and Christine won the state lottery!
	5	I knew it was wrong.		5	Is that what you want to say?
	6	He didn't collect the papers.		6	Are you criticizing me?
	7	He didn't correct the papers.		7	What are you going to do tomorrow?
	8	Put all these things in the bag.		8	Don't worry; I'll do it for you?
	9	food / hood		9	year/ear
	10	cars / cards		10	woos/ooze
C	1	Look at the train.	F	1	What kind of person is willing to send his children to wars?
	2	Don't disturb them while they are praying.		2	What would you like to have for dinner?
	3	Humpty Dumpty sat on a wall		3	My breakfast is always bread and butter.
	4	4. Humpty Dumpty had a great fall		4	I just can't wait.

	5	All the king's horses and all the king's men		5	How many times do I have to tell you this?
	6	Couldn't put Humpty together again		6	It won't be long.
	7	What a wonderful life he lived!		7	I have to admit that I was a little drunk.
	8	How beautiful you are!		8	I was wondering if you could babysit tomorrow night.
	9	ban/ van		9	I thought you were a normal person.
	10	deaf/ death		10	I think I'm coming down with something.

Selection of the stimuli for testing followed these guidelines: 1) familiar vocabulary, 2) structural variety, 3) sustained phonation which may provide a visually obvious display of pitch contour, 4) relatively short sentences to facilitate easy production, and 5) sets of minimal pairs which include consonants that are difficult to produce for Japanese learners of English.

As for the pretest, the four groups of stimuli (A, B, B, D) are selected and set in the software which I used in the previous experiment. The order of the stimuli is A, B, B, D. The stimuli in the first two groups are displayed only in the text style on the computer display whereas the latter two groups are both in the text style and sounds (See Figure 1). The subjects were required to read and record the stimuli for the first two, and then for the latter two, they were asked to read the text, listen to the model voice, and record the stimuli.

Regarding the posttest, the procedure is the same as the pretest but it includes more groups of sentences as follows.

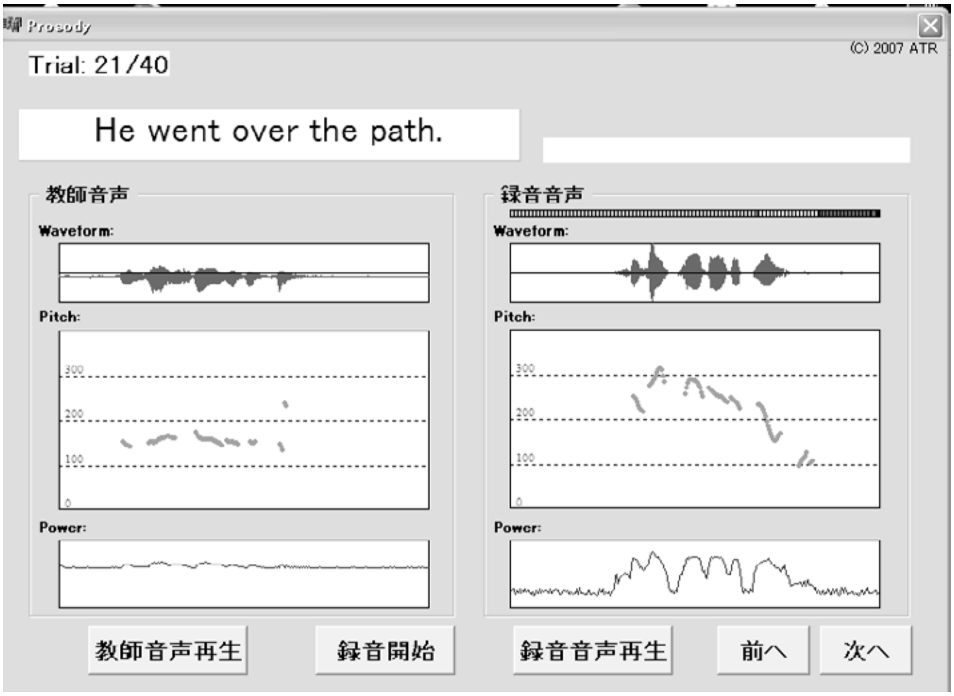
Table 2: The order and conditions of stimuli in the pre/posttest A~F are groups of the stimuli. (See Table 1) R for "Read," LR for "Listen and Repeat."

Pretest					Posttest								
Order	A→ B→ B→ C				Order	A→	B→	D→	E→	B→	C→	E→	F
Condition	R	R	LR	LR	Condition	R	R	R	R	LR	LR	LR	LR

I designed the six groups of stimuli in order to compare the results of this experiment with further research. In this experiment, I compared the following.

1) Pretest B under the condition of R (Read)
2) Pretest B under the condition of LR (Listen and Repeat)
3) Posttest B under the condition of R (Read)
4) Posttest B under the condition of LR (Listen and Repeat)

Figure 1. The interface of pre/posttest



2.2. Speakers

In the pre/posttests, a male speaker of standard British English recorded the test items. Between these two tests, subjects attended 10 sessions of pronunciation training. As there are 20 stimuli in each session, 200 hundred stimuli were recorded by two male speakers and two female speakers. All of them are professional recorders of standard American English.

2.3. Subjects

A total number of 13 native speakers of Japanese (10 female, 3 male) volunteered to participate in this study. All of them were undergraduate students at Konan University in Kobe. They belonged to different faculties of the university. None had spent more than two months in an English speaking country. Their ages ranged from 19 to 22. They reported normal hearing and vision. All of them were motivated to improve their production of English.

2.4. Procedure

A pretest-posttest design was used to measure the effects of one month's training (10 sessions of about 40 to 50 minutes each) using computerized visual displays of pitch contours and wave forms as feedback (See Figure 1). The same software was used for the pretest, 10 training sessions and the posttest. Its basis was provided by ATR (Advanced Telecommunications Research Institute International). Users can customize it by inputting the stimuli. This time 40 stimuli for the pretest, 200 stimuli for the training sessions and 80 stimuli for the posttest were put into the software.

The software was installed into 10 computers in a self-study room at the university. Subjects were asked to come to the room at any time during the training period. For the first time, they were asked to read instructions about using the software and took the pretest by computer. Their voices were automatically recorded and stocked in the server. During the training period, some of the subjects came daily, finished the training sessions and took the posttest relatively early. Some of them came to the room as regularly as twice a week and others came quite irregularly.

In the training sessions, a group of seven subjects were trained with low-pass filtered models while another control group of six subjects were given non-filtered examples. Before they participated in this experiment, they took a proficiency test and were divided into nearly homogeneous two groups. For the low-pass filtered (LP) group, each session had 10 stimuli and each stimulus was repeated 10 times. For the first five times, the stimulus was filtered and the second five times, non-filtered. For instance, an LP subject saw the text of stimuli on the display, listened to the ambiguous filtered sound and recorded her voice just as she listened. Thus the first five recorded voices of the LP subject sounded quite indistinct, like humming. Then the second five sounded normal because she was listening to a non-filtered voice. For the non-filtered (NF) group, the same stimuli were used as the LP group but they didn't listen to the filtered sounds at all.

Both groups took the same posttest. Their voices were saved in the computer server as WAV files. In this study, ten stimuli in the pretest and the posttest were analyzed by computer software called WaveSurfer.

Regarding the questionnaire, subjects were asked to complete a questionnaire consisting of the following questions after they took the posttest: 1) How do you feel after finishing the training program? 2) What have you noticed about your own pronunciation in English? 3) (LP subjects only) How did you feel when you listened to LP sounds?

3. Results

3.1. Duration of eight sentences

Table 3 shows the duration of eight sentences used in the study. This will be the base of the following analyses.

Table3. The duration of 8 sentences (ms)

Subject Number	Low-lass or Non-filtered	Pretest or Posttest	Read or Listen & Repeat	Duration of 8 sentences							
				Sen-1	Sen-2	Sen-3	Sen-4	Sen-5	Sen-6	Sen-7	Sen-8
Model				1685	1158	1068	1101	1118	1608	1761	1683
S01	LP	PRE	RD	1519	1239	937	885	1256	2325	2226	1976
S01	LP	PRE	LR	1668	1369	1026	1198	1181	2014	1968	1830
S01	LP	POST	RD	1510	1450	1133	1190	1753	1908	1944	1999
S01	LP	POST	LR	1680	1357	1056	1129	1360	1864	1780	1818
S02	LP	PRE	RD	1587	1320	1983	886	1419	1861	1993	2065
S02	LP	PRE	LR	1562	1108	1412	1107	1385	1941	1948	2104
S02	LP	POST	RD	1284	1112	1202	840	1050	1817	1573	1712
S02	LP	POST	LR	1192	1140	1083	1110	1115	1618	1760	1660
S03	LP	PRE	RD	1954	1430	1172	1230	1488	2058	2043	2204
S03	LP	PRE	LR	1686	1129	1100	1112	1447	1902	2175	2087
S03	LP	POST	RD	1545	1326	1083	1146	1402	1718	1710	1905
S03	LP	POST	LR	1739	1276	1071	1108	1452	1662	1817	1937
S04	LP	PRE	RD	2489	1913	2739	1289	1820	2804	2035	2954
S04	LP	PRE	LR	1846	1581	1457	1146	1948	2184	2021	2203
S04	LP	POST	RD	1744	1574	1607	1122	1609	2032	2051	2235
S04	LP	POST	LR	1717	1502	1298	1100	1512	1958	1866	1854
S05	LP	PRE	RD	1713	1388	1584	1078	1596	1864	2300	2220
S05	LP	PRE	LR	1722	1375	1319	1099	1705	1776	1947	2126
S05	LP	POST	RD	1588	1360	1610	930	1408	1855	2004	2143
S05	LP	POST	LR	1535	1128	1261	1196	1307	1655	1782	1846
S06	LP	PRE	RD	1734	1503	1549	1227	1536	2119	1956	2289
S06	LP	PRE	LR	1665	1480	1302	1142	1439	2091	1979	2018
S06	LP	POST	RD	1622	1443	1276	1194	1374	1724	1873	2050
S06	LP	POST	LR	1411	1399	1164	1100	1350	1623	1780	1701
S07	LP	PRE	RD	1297	1206	1112	1105	1236	2117	1791	1697
S07	LP	PRE	LR	1763	1117	1217	1133	1189	1940	1912	1694
S07	LP	POST	RD	1488	1367	1009	1290	1438	1970	2112	2095
S07	LP	POST	LR	1784	1357	1055	1056	1112	1651	1901	1800
S08	NF	PRE	RD	1492	1280	1218	902	1288	1784	1877	2364
S08	NF	PRE	LR	1716	1195	1236	1077	1215	1901	1910	1923
S08	NF	POST	RD	1423	1301	1036	1054	1231	1642	1567	1910
S08	NF	POST	LR	1303	1233	981	1001	1102	1613	1655	1724
S09	NF	PRE	RD	1424	1252	949	949	1135	1752	1528	2082
S09	NF	PRE	LR	1479	1059	950	1090	1313	1599	1889	1971
S09	NF	POST	RD	1413	1116	934	963	1162	1675	1803	1852

S09	NF	POST	LR	1520	1160	1056	1080	1129	1955	1779	1903
S10	NF	PRE	RD	1764	1470	1359	1237	1724	2122	2609	2835
S10	NF	PRE	LR	1684	1126	1395	1101	1520	2035	2002	2002
S10	NF	POST	RD	1186	1003	1202	1179	1401	1910	1872	1813
S10	NF	POST	LR	1364	1185	1076	1174	1345	1679	1798	1770
S11	NF	PRE	RD	1475	1170	1209	865	1477	2375	2782	3126
S11	NF	PRE	LR	1379	1200	1057	916	1419	1831	1889	1930
S11	NF	POST	RD	1426	1178	1073	902	1206	1684	1910	1181
S11	NF	POST	LR	1368	1106	952	887	1293	1851	1774	1730
S12	NF	PRE	RD	1544	1169	1108	1062	1154	1811	1980	2056
S12	NF	PRE	LR	1681	1230	1154	1024	1307	1929	2056	1783
S12	NF	POST	RD	1579	1244	1164	1099	1259	1813	1786	1937
S12	NF	POST	LR	1737	1359	1134	990	1264	2082	2146	1700
S13	NF	PRE	RD	1605	1326	1407	964	1379	2187	2467	2700
S13	NF	PRE	LR	1570	1162	1110	1152	1130	2138	2232	2142
S13	NF	POST	RD	1874	1453	1209	1370	1370	2134	2145	2687
S13	NF	POST	LR	1652	1362	1192	1164	1164	2082	2157	2004

3.2. Analyses

Analysis No.1

This experiment classifies subjects into two groups, and lets them experience the LP program and the NF program. The purpose of this experiment is to see the difference in subjects' achievements between the LP and the NF group, and whether there are differences in reading conditions (RD or LR). Normally, we use two-way or multiple-way ANOVA for this kind of analysis, but since the population of each group is different, I have performed "regression ANOVA" , in which I regress "Duration variable" on the following binary variables.

I have created the Duration variable by integrating the observations of duration from Sentence 1 to Sentence 8 and the number of observations is 416. See Table 4.

Table 4. Regression analysis (LP vs. NF)

Multiple correlation coefficient: R	0.833104
R-square: R2	0.694063
Corrected R2	0.688049
Standard error	232.4233
number observed	416

The first binary variable is the NF. The elements of the NF are 1 for the subjects who experienced the NF program and 0 for subjects who did the LP program.

As for the second binary variable, Sentence 2, its elements are 1 for the subjects who read the sentence 2 and 0 for the other subjects. Sentence 3 and the other binary variables have similar properties to Sentence 2: for example, the elements of Sentence 8 are 1 for those who read sentence 8 and 0 for the others.

Table 5. Analysis of variance (LP vs. NF)

	Degrees of freedom	Sum of squares	Distribution	Observed variance ratio	Significant F
Regression	8	49879245.16	6234905.6	115.4172418	1.032E-99
Residual	407	21986373.6	54020.574		
Total	415	71865618.76			

Empirical results: The benchmark of these eight binary variables is the subjects who read sentence 1 and did the LP program. Their average duration is the value of the intercept of estimated equation. The value of intercept and standard error are 1623.56 and 33.91, respectively. Therefore, the confidence interval at 95% level for them ranges from 1556 to 1690.

Table 6. Results of each sentence

	Parameter	Standard error	T-value	P-value	Lower bound 95%	Upper bound 95%
Intercept	1623.564217	33.91405784	47.87289757	2.7602E-169	1556.895652	1690.232782
NF	-71.8058036	22.85870584	-3.1412891	0.001804944	-116.7416563	-26.86995086
sen2	-296.423077	45.58195238	-6.50307987	2.30993E-10	-386.0284952	-206.8176586
sen3	-360.076923	45.58195238	-7.89955024	2.63801E-14	-449.6823414	-270.4715048
sen4	-506.769231	45.58195238	-11.1177605	2.96058E-25	-596.3746491	-417.1638125
sen5	-227.461538	45.58195238	-4.99016665	8.95728E-07	-317.0669568	-137.8561202
sen6	325.6346154	45.58195238	7.143937422	4.21422E-12	236.0291971	415.2400337
sen7	595.8846154	45.58195238	13.07281905	7.49191E-33	506.2791971	685.4900337
sen8	435.4807692	45.58195238	9.553798082	1.1976E-19	345.8753509	525.0861875

The NF parameter means the overall difference between the duration of the LP and the NF subjects. This is −71.8 with t-statistics equal to −3.14 (P-value is 0.018). Therefore, I conclude that the NF subjects differ from LP subjects in terms of duration for all sentences, and this fact is statistically significant at 1.8% level.

Analysis No.2

The rest of the experiments are the sub-analysis of Analysis No.1. In this analysis I perform regression ANOVA to see the effect of Pre/Post effect within LP subjects. I am interested in the parameter of the binary variable, Post, the element of which is 1 for the post-subjects in the LP group, and 0 for pre-subjects. The number of observations is 224. The empirical result is that the parameter of POST is −145.15 with t-statistics and P-value 6.87 and 0.00, which means the duration of post-subjects is shorter than that of pre subjects.1 This result holds for all the readers from sentence 1 to 8.

Table 7. Regression analysis (Analysis 2)

Multiple correlation coefficient: R	0.829797985
R-square: R2	0.688564696
Corrected R2	0.676976406
Standard error	222.7721788
number observed	224

Table 8. Analysis of variance (Analysis 2)

	Degrees of freedom	Sum of squares	Distribution	Observed variance ratio	Significant F
Regression	8	23590507	2948813.375	59.41900604	2.491E-50
Residual	215	10669900.39	49627.44367		
Total	223	34260407.39			

Table 9. Results of each sentence (Analysis 2)

	Parameter	Standard error	T-value	P-value	Lower bound 95%	Upper bound 95%
Intercept	1717.004464	44.65377689	38.45149468	2.55065E-98	1628.989328	1805.0196
POST	-145.151786	29.76918459	-4.87590734	2.10407E-06	-203.8285429	-86.47502849
sen2	-289.107143	59.53836918	-4.85581226	2.3057E-06	-406.4606573	-171.7536284
sen3	-329.535714	59.53836918	-5.53484616	9.0201E-08	-446.8892287	-212.1821998
sen4	-532	59.53836918	-8.93541438	1.86209E-16	-649.3535144	-414.6464856
sen5	-219.892857	59.53836918	-3.69329661	0.00028075	-337.2463716	-102.5393427
sen6	285.9642857	59.53836918	4.803025169	2.92838E-06	168.6107713	403.3178002
sen7	512.8571429	59.53836918	8.613893022	1.52564E-15	395.5036284	630.2106573
sen8	363.5	59.53836918	6.105306628	4.73453E-09	246.1464856	480.8535144

Analysis No.3

This analysis is similar to Analysis No.2. My purpose is to see the difference between the Pre/Post subjects within the NF group. The number of observation is 192. The empirical result is that the parameter of Post is −121.18 with t-statistics and P-value 3.80 and 0.00, which also means the duration of post-subjects is shorter than that of pre subjects.2 This result holds for all the readers from sentence 1 to 8.

3.3. Pitch comparison

The fundamental frequencies of the sound wave are said to be closely related to the pitch of the voice. The pitch of a recorded sound cannot be literally measured

Table 10. Regression analysis (Analysis 3)

Multiple correlation coefficient: R	0.873323104
R-square: R2	0.762693244
Corrected R2	0.752319178
Standard error	219.2569704
Number observed	192

Table 11. Analysis of variance (Analysis 3)

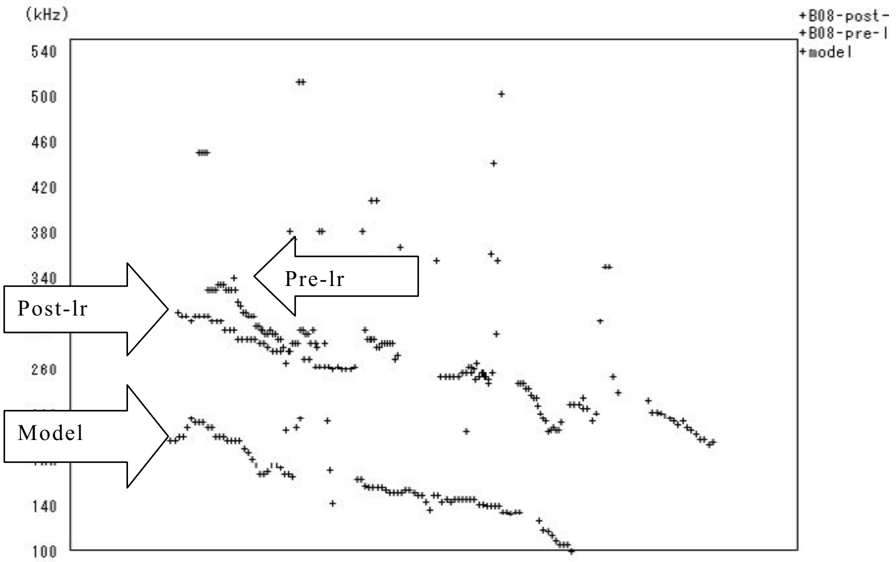
	Degrees of freedom	Sum of squares	Distribution	Observed variance ratio	Significant F
Regression	8	28274680.42	3534335.05	73.5192216	4.23202E-53
Residual	183	8797472.286	48073.6191		
Total	191	37072152.7			

Table 12. Results of each sentence (Analysis 3)

	Parameter	Standard error	T-value	P-value	Lower bound 95%	Upper bound 95%
Intercept	1587.505208	47.47052657	33.44191276	7.07142E-80	1493.845274	1681.165143
POST	-120.177083	31.64701771	-3.79742206	0.000198756	-182.6170395	-57.73712713
sen2	-304.958333	63.29403543	-4.81812119	3.03137E-06	-429.8382457	-180.0784209
sen3	-395.708333	63.29403543	-6.25190558	2.78306E-09	-520.5882457	-270.8284209
sen4	-477.333333	63.29403543	-7.54152157	2.09471E-12	-602.2132457	-352.4534209
sen5	-236.291667	63.29403543	-3.7332375	0.00025224	-361.1715791	-111.4117543
sen6	371.9166667	63.29403543	5.876014448	1.94917E-08	247.0367543	496.7965791
sen7	692.75	63.29403543	10.94494916	9.06855E-22	567.8700876	817.6299124
sen8	519.4583333	63.29403543	8.207066113	3.93802E-14	394.5784209	644.3382457

but the fundamental frequencies can. Figure 2 shows the fundamental frequencies of the sentence “Put all these things in the bag” pronounced by a model speaker and an LP subject. The arrow ‘Pre-lr’ is an abbreviation of ‘pretest-listen-and-read’, and the arrow ‘Post-lr’ means ‘posttest-listen-and-read.’ They are the names of the sound files. Each subject had two files (Pre-lr and Post-lr) for each sentence and they were compared with the equivalent model sound as shown in Figure 2. As there were seven subjects and each subject had eight sentences in the LP group, a total number of 56 figures (7 x 8) were visually examined. Regarding the NF group, 48 figures (6 x 8) were compared in the same way as the LP group. In comparison with the pitch of the pre/post files, there was a strong tendency of the pitch contours of ‘Post-lr’ to approximate to the model’s. This tendency can be observed both in the LP group and the NF group, but more obviously in the LP group (52 out of 56 figures: approximately 92.86%) than the NF group (29 out of 48 figures: approximately 60.42%).

Figure 2. An illustration of the pitch comparison of an LP subject. These are the pitch contours of "Put all these things in the bag."



3.4. Questionnaire responses: Subjects' feedback

At the conclusion of the training program, all subjects were given a questionnaire to complete and return to me anonymously. The purpose of giving the questionnaire to subjects is to assess their perceived value of speech technology such as this in foreign language instruction. The responses shown below are listed according to frequency of occurrence on the returned questionnaires.

1) How do you feel after finishing the training program?

- I feel more confident about my pronunciation.
- I was dismayed to see the difference between my pitch contours and the model's but gradually I came to approximate the model.
- I am not sure whether my pronunciation of English has improved but I feel as though I have achieved something.

2) What have you noticed about your own pronunciation in English?

- My intonation was rather monotonous.
- I have noticed how my pronunciation in English differs from that of native speakers.

3) (LP subjects only) How did you feel when you were listening to the LP sounds?

- I found it tiring at first. The LP sounds felt uncomfortable because I couldn't understand what was being said. This was frustrating. However, as I attended the training sessions, I became more used to them. Now I am not uncomfortable with the sounds at all.

- They sounded as if someone was talking under water.
- I recognized the rhythm and intonation of the speech although I didn't understand what was said.
- After repeatedly listening to the LP sounds followed by the NF sounds, I felt the NF sounds became clear and easy to understand.

4. Discussion and Concluding remarks

The results of this experiment revealed significant differences between the LP group and the NF group. Throughout the low-pass applied training sessions, the LP subjects might have become more sensitive to prosodic features than the NF subjects. While more data from both of the groups are needed to fully validate the robustness of the hypothesis, it would be reasonable to conclude that computer-assisted pronunciation training using low-pass filters is more effective to train accurate pronunciation than training without digital filters.

The following are my research questions and their answers in this study.

1) Is there any difference in speech production between the group of subjects who attended 10 sessions of training with electric low-pass filtered sounds and the controlled group who did the same training without filtered sounds?

Yes, there is. As for the duration of the eight sentences in the posttest, the LP group was significantly different from the NP group and they approximated more to the model sounds. By comparing pitch contours of the pre/posttests, the LP group showed closer approximation of the model sounds.

Did the low-pass group become more accurate in production than the non-filtered group, or vise versa? Were both improved?

It depends on the definition of accuracy, but as for the approximation to the model sounds, the LP group became more accurate in the production of prosodic features than the NP group although both groups improved their production skills.

2) How did the participants feel after finishing the 10 training sessions?

See 3.4.Questionnaire responses.

There are some remaining issues. As the number of subjects was limited in this study, as mentioned earlier in this chapter, a follow-up experiment is needed with a larger pool of subjects. Subjective evaluation should be added to the experiment because the results of quantitative analysis and those of subjective evaluation are often different. Many attempts have been made to evaluate the improvement of pronunciation skills using technology. The precision of computer speech analysis is becoming ever more accurate. Yet, the human raters' sense of "Englishness" may differ from that of machines.

Notes

1) Without binary variables for sentences (i.e., binary variables Sentence 1 through to Sentence 8), I recognize this result is robust. The parameter is −145.15 with t-statistics 2.81, P-value 0.005, and n=416.
2) Without binary variables for sentences (i.e., binary variables Sentence 1 through to Sentence 8), this result is almost robust. The parameter is −120.18 with t-statistics 1.90 and P-value 0.059.

References

Akahane-Yamada, Reiko, Erik McDermott, Takahiro Adachi, Hideki Kawahara, and John S. Pruitt (1998). Computer-based second language production training by using spectrographic presentation and HMM-based speech recognition scores. *Proceedings of the 5th International Conference on Spoken Language Processing, Sydney, Australia, 1998* [CD-Rom] Paper 0429.
Bourdieu, P. (1991). Language and symbolic power. (Translated by J. B. Thompson), Cambridge, MA: Harvard University Press.
Bourdieu, P. (1995). The logic of practice. Palo Alto: Stanford University Press.
Derrida, J. (1982). Différance. In Derrida, J. (Ed.), Margins of philosophy (pp.1-27). Chicago: The University of Chicago Press.
Freadman, A. (1994). Models of genre for language teaching. Paper delivered as The Sonia Marks Memorial Lecture. Sydney: University of Sydney.
Guberina, P. (1976). Structuration et dépassement des structures perceptives et psycholinguistiques dans la méthodologie SGAV. In Actes du 3e Colloque international SGAV pour l'enseignement des langues (pp. 41-58). Paris: Didier.
Guberina, P. (1972). Restricted bands of frequencies in auditory rehabilitation of deaf. Zagreb: Institute of Phonetics Faculty of Arts.
Hunter, M.D., Phang, S., Lee, K., & Woodruff, P.W.R. (2005). Gender-specific sensitivity to low frequencies in male speech. Neuroscience Letters, Volume 375, Issue 3, 148-150.
Iba, M. (2007). The influence of model sounds on the speech production of Japanese learners of English. The Journal of the Institute for Language and Culture, 12, 45-66.
Keith, R.W. (2000). SCAN-C: Test for Auditory Processing Disorders in Children-Revised. San Antonio: Psychological Corporation.
Keith, R.W., & Farrer, S. (1981). Filtered word testing in the assessment of children with central auditory disorders. Ear Hear, 12, 267-269.
Lian, A.B. (2003). Beyond illusions and facts: Toward a methodology of dialogue and dialogue-

enhancing environments. Paper presented at The International Conference on Computers and Philosophy, ANU, Australia and Rice University, Texas, USA. Retrieved from http://www.anialian.com/On_Time.html

Moyer, Alene (1999). Ultimate attainment in L2 phonology. The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition* 21: 81–108.

Munro, Murray J. and Tracey M. Derwing (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45: 73-97.