ARCHIVES OF ACOUSTICS
32, 1, 89–100 (2007)

PERCEPTION AND ANALYSIS OF CHINESE ACCENTED GERMAN VOWELS

Hongwei DING, Oliver JOKISCH, Rüdiger HOFFMANN

Dresden University of Technology
Laboratory of Acoustics and Speech Communication
e-mail: {hongwei.ding;oliver.jokisch;ruediger.hoffmann}@ias.et.tu-dresden.de

(*received October 16, 2006; accepted December 16, 2006*)

This report describes an investigation to obtain specific knowledge about the production of German vowels by Chinese speakers. The experiments have been conducted to acquire both perceptual and acoustic measures of the vowels. 11 Chinese speakers have been selected as subjects to read 20 German words. The assessment was conducted by 10 German native speakers. We came to the conclusion that the Chinese speakers have little difficulty in pronouncing long vowels. However, they are not accurate in the short vowels production. In the further acoustic investigation, it was discovered that the Chinese speakers are not always aware of the duration contrast of long-short vowels, and they have some difficulty in finding the right position of the tongue and to grasp the degree of lip-rounding in central vowels production. Those deviations can contribute to the foreign accent of the Chinese German speakers.

Keywords: foreign accent, German vowels, vowel formants, Chinese.

1. Introduction

Phonetic characteristics of the native language are thought to interfere with the production of the second language. There have been many acoustic and perceptual studies examining the effects of Asian languages on the phonetic features of speaking English as the second language [1]. However, there are relatively few acoustic experiments on the interference of Chinese with speaking German as the second language. This paper describes the investigation in this field, which gains an increasing importance with the development of international speech communication.

The direct motivation of this experiment was to fulfill the task of providing a synthesized Chinese voice for our mutilingual speech synthesis [2], with the aim that this voice can also pronounce some English and German words. In the selection of the speakers, besides the Chinese origin, also the pronunciation quality of English and German was considered. The experiments were designed to find out which vowels were perceived by

the native speakers with more foreign accent, and acoustic measures were supplied to account for it. The investigation can at the same time provide an implication for the German Speech Recognition of speakers with Chinese accent; it can further provide a better insight into the exact learning problems of pronunciation, and supply the acquisition of the second language with an appropriate guidance.

2. Differences between German and Chinese vowel systems

The difference between German and Chinese vowel systems can lead to inaccurate pronunciation of German vowels spoken by Chinese speakers. Mandarin has six basic monothongs. They are slightly different in the phonological systems in the "openness" of the mouth, because the level of tongue position and the ways lips are spread or rounded. Figure 1 depicts F1 $\times$ F2 chart of six basic Chinese vowels (represented with SAMPA-C symbols for Cantonese) uttered in isolation by a professional female speaker. This chart can be regarded as representative, because the values are quite similar to those documented in other researches, for example in [7].

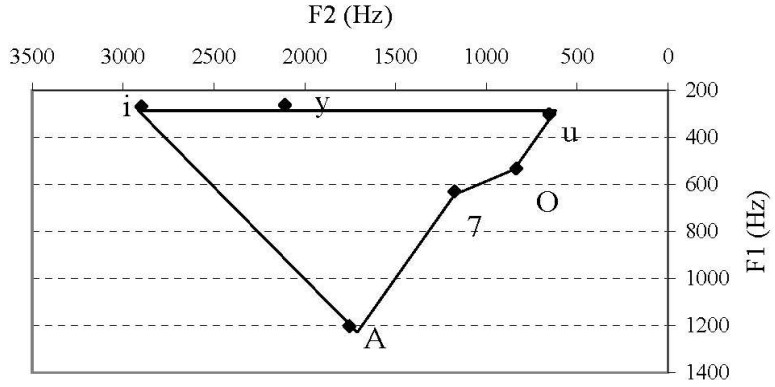

Fig. 1. The chart of basic Chinese vowels of a female speaker.

On the other hand, the German phonetic system consists of much more vowels. 9 selected vowels are presented in Fig. 2. The formant values were taken from the literature [3], and the German vowels are represented here by the SAMPA code. All the long vowels are connected by solid lines and the short vowels by dotted lines.

From the comparison of the two charts it is obvious that the German vowels are spread in the acoustic vowel space, whilst in the case of Chinese vowels, the centralized position of the vowel triangle is absent. The first question that arose was whether the Chinese speakers have difficulty in producing the central vowels. Another consideration is that there is no long-short feature in Chinese, whereas this is a prominent feature of the German vowels. We wondered whether it poses a problem for Chinese learners in the pronouncing of German vowels?

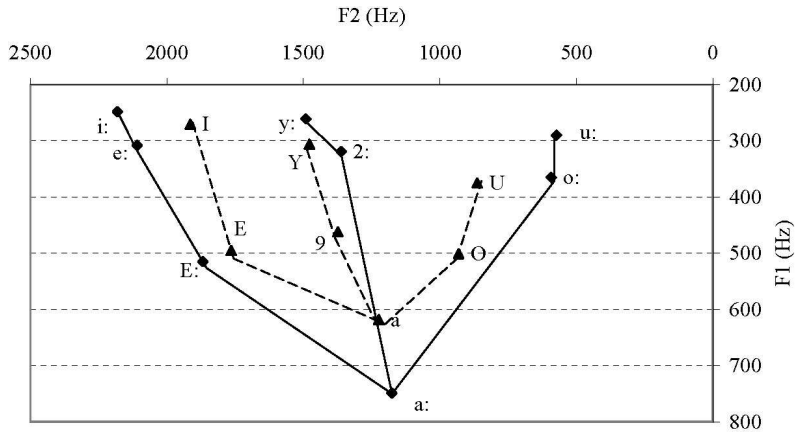

Fig. 2. The chart of German vowels of a native speaker.

3. Preparation of the Speech Database

In order to provide answers to the mentioned above questions, the first step was to develop the speech material and select test subjects.

3.1. Reading material

It is quite difficult for non-professional speakers to pronounce the vowels in isolation or to read nonsense carrier words. On the other hand, it is also difficult for naive native listeners to make any assessment by listening to isolated vowels or nonsense carrier words. Therefore, the idea to read the vowels in isolation or in nonsense carrier words was rejected. The vowels were thus embedded in meaningful carrier words. 20 simple words were selected as the reading list. These carrier words are listed in Table 1 together with the embedded target vowels:

Table 1. Table of German vowels and words.

Vowels	I	E	A	O	U	Y	9
Words	Sitz	Gesetz	Satz	Trotz	Schutz	hübsch	plötzlich
Vowels	i:	e:	E:	a:	o:	u:	y:
Words	Lied	Beet	spät	Tat	rot	Blut	süß
Vowels	2:	aI	aU	OY	@	6	
Words	blöd	Eis	Haus	Kreuz	bitte	besser	

3.2. Subjects

Subjects chosen for the investigation were 10 Chinese female students and teaching staff at the Dresden University. There were two reasons that only female speakers were included.

Firstly, female speakers were observed to have higher proficiency in German than male speakers recorded in our database. In this research, only advanced speakers were included. The aim was rather to pinpoint the foreign accent than to detect wrong pronunciation. Thus, we had the preference in choosing female speakers. Only one male speaker was included for the general assessment. Secondly, limiting the set of speakers to female voices only was advantageous to the comparison of different acoustic parameters.

All the 10 speakers had already received high education in China; they came to Germany at the age of 20–30 to continue their master or doctoral study. All of them had completed more than a six-month intensive German course and have already lived in Germany for several years. They tried to speak German as accurately as possible to accomplish their study and work in Germany.

It is also an uncommon experience that such advanced speakers still have a foreign accent in speaking the second language. In order to facilitate the comparison of some acoustic parameters, a native female speaker was also included in the experiment as a reference.

3.3. Recording of database

The recordings were carried out in an anechoic chamber using a high-quality microphone with pop protection. The recorded tokens were digitalized directly at a sampling rate of 32 kHz with a 16-bit resolution and stored on the hard disk of the work station. A German expert and a Chinese monitored the recordings and corrected the speakers' reading errors that might have occurred.

4. Perception and assessment

4.1. Procedure

Ten native listeners were asked to judge the quality of the recorded vowels. They were native, not professional listeners without phonetic background. They could only tell whether the vowel sounded German or foreign, but they were not able to point out the possible reason for the accent.

In the listening test, the instruction was focused on that they should judge the pronunciation in reference to the German known from the broadcasting, and they should only evaluate the quality of the vowel in the word. Fortunately, all the speakers had no problem in pronouncing the consonants. Thus, the quality of the target vowel is cru-

cial to the assessment of the perception. The stimuli for the experiment consisted of 20 words spoken by 12 speakers and they resulted altogether in an amount of 240 words.

The native German listeners could identify correctly what the Chinese speakers read, but they found it sometimes not accurate enough. A perception test to identify the words was thus unnecessary. So we adopted a listening test to obtain the general impression of the pronunciation.

The words were presented in a test paper. The listeners were instructed to give an assessment of the intended words in a five-scale assessment like that of speech synthesis in a mean opinion scores (MOS) test. The score 5 indicates a standard pronunciation without any foreign accent. The score 1 indicates poor pronunciation which could cause a misunderstanding of the word.

4.2. Results

Listeners' responses were analyzed and compared. The results are presented in Fig. 3.

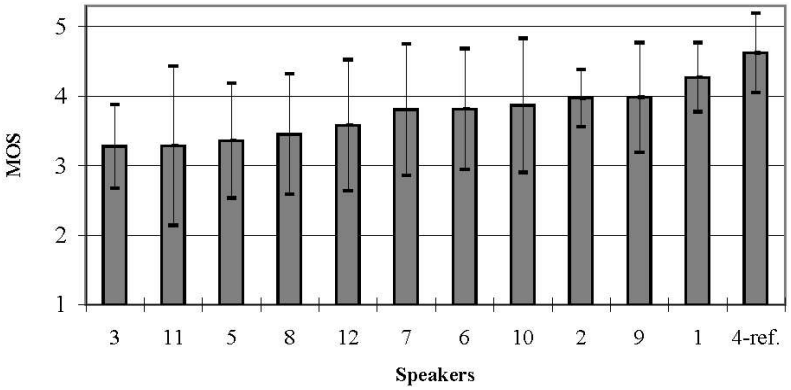

Fig. 3. The speakers foreign accent assessment in producing German vowels.

4.2.1. Results of the Speakers

The result shows that those advanced speakers were still indicated by the native speakers as having foreign accent. The reference can be made in Fig. 3.

Among the Speakers, 4-th one (noted as 4-ref.) is a native speaker with the MOS of 4.62. The MOSes for the Chinese speakers range from 3.27 to 4.27. The male speaker was judged to have the most foreign accent. We can conclude that the advanced Chinese speakers can grasp the properties of vowels well, but there is still as discrepancy between their pronunciation and the standard pronunciation of German vowels.

The evaluation demonstrated that the performance of Chinese speakers was considered across the listeners. Some of the subsequent experiment results are presented as the group average for particular vowels.

4.2.2. Results for Vowels

Our next interest is the assessment of vowels uttered by the Chinese speakers. The results are presented in Fig. 4.

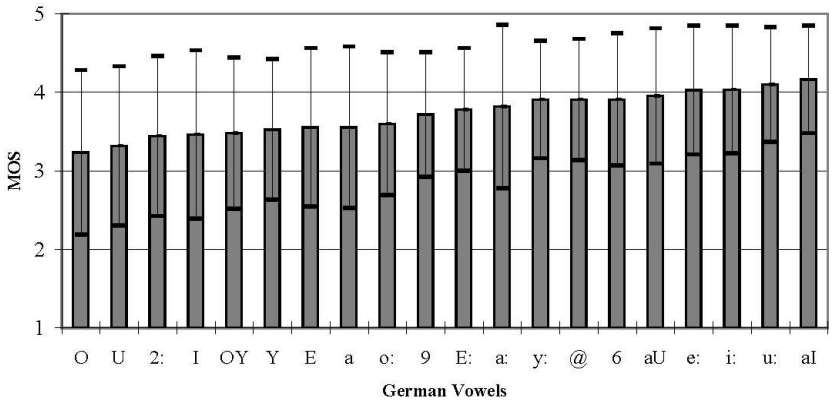

Fig. 4. The assessment of the quality of German vowels uttered by Chinese subjects.

It can be observed that some of the diphthongs (/aI/, /aU/), most of the long vowels (/u:/, /i:/, /e:/, /y:/, /a:/, /E:/), and two schwas (/@/, /6/) have better scores than the short vowels (/O/, /U/, /I/, /Y/, /E/, /a/) and one long vowel (/2:/). This can be inferred from the differences in the vowel charts. The short vowels and the umlaut /2:/ are absent in the Chinese vowel system and therefore the Chinese subjects can not produce them as accurately as their long counterparts which are similar to vowels existing in their native language.

5. Analysis of the vowels' acoustic properties

The goal of the final experiments was to discover acoustic parameters that might correlate for the vocalic pronunciation in both languages. The dynamic formants and duration values are argued to be important in the perception of vowels examined in our experiment.

5.1. Duration

Because the duration of a vowel is influenced by the speaking rate, and the word structure as well, the absolute duration values can hardly be compared with those for standard pronunciation. However the reference was a female native speaker which produced the same utterances as Chinese subjects. Even the speaking rate was different, we could expect that the relative duration values of the long-short pairs should be proportional.

Thus, a comparison was made with the seven minimal pairs in German. The mean duration values obtained for all the Chinese subjects are presented in Fig. 5.

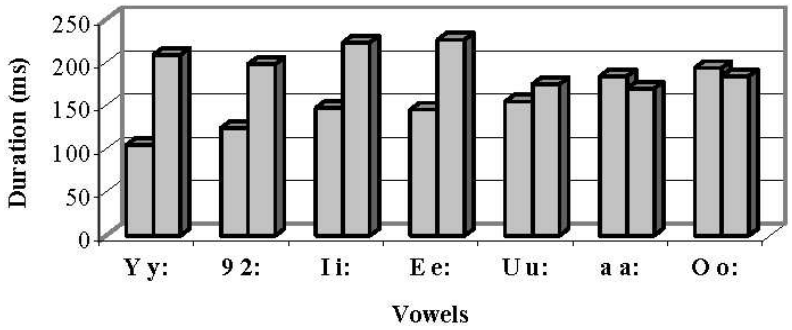

Fig. 5. Duration values of minimal pairs measured for Chinese speakers.

For the convenience of comparison, the statistics is also presented for the native speaker in Fig. 6.

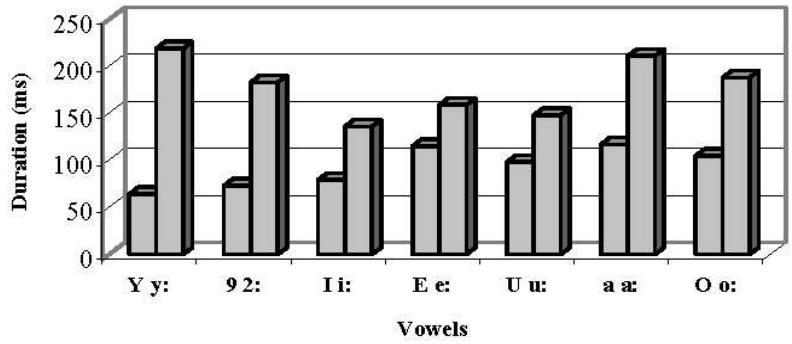

Fig. 6. Duration values of minimal pairs measured for a female native speaker.

According to the published data [8] and [3], the opposition long-short is a striking phonetic feature for the German vowel system. The average lengths of long (tense) – short (lax) vowels has a duration ratio of about 2:1. This proportion can be observed for most pairs of the native speaker, which is presented in Fig. 6. Because of the slow speaking rate of the single words, some of the pairs, such as /Y/ and /y:/, /9/ and /2/, have a larger contrast than 2:1. On the other hand, in Fig. 5, the first four minimal pairs can also display such a ratio, but the difference is not so large in the fifth pair. Moreover, for the last two pairs (/a/ and /a:/, /O/ and /o:/), the short vowels have even longer duration values than their long counterparts. This partly accounts also for the reason that the four vowels /O/, /a/, /o:/, /a:/ belong to the more accented utterances in the perception test. It demonstrates that the Chinese speakers are aware of the importance of the long-short distinction of vowels in German, but they are not successful in the duration control of the vowels.

5.2. Formants

As we have already known, the contribution of the dynamic properties of vowels is much more significant to their perception. The quality of vocalic sounds depends on their formant frequencies, especially on the first and second formant (F1 and F2). The vowel charts and the formant patterns were therefore compared respectively in the following manner. It is well known, that the formant frequencies for a male speaker can have considerably different values from those of a female speaker. For this reason the male speaker was excluded in the formant pattern comparison.

5.2.1. Comparison based on the F1-F2 vowel charts

Because the vowel formants are influenced by the neighboring consonants at the beginning and ending of the vocalic parts (see for example [5]), we just averaged the values at the middle point of the monothongs spoken by Chinese speakers. The measurement results are plotted in Fig. 7.

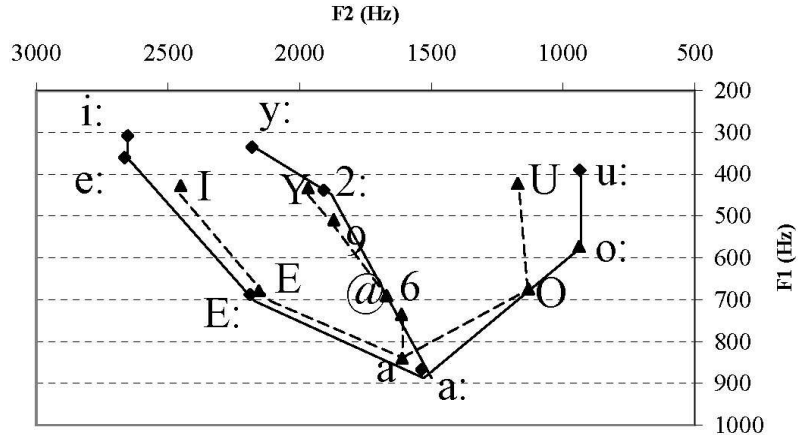

Fig. 7. The F1-F2 chart of German vowels obtained from Chinese speakers.

When the vowel patterns of the proficient in German Chinese speakers (Fig. 7) are compared with those of the ideal native speakers (Fig. 2), we find that the Chinese speakers have a similar pattern of the vowel charts: F1 of all the lax vowels (/I/, /E/, /a/, /O/, /U/, /Y/, /9/) are shorter than their corresponding tense vowels (/i/, /e:/, /a:/, /o:/, /u:/, /y:/, /2:/); all the short vowels connected by dotted lines are located on the triangle of the long vowels connected by solid lines. The difference is that lax vowels produced by Chinese speakers are not centralized as those produced by the German native speaker in Fig. 2.

When the vowel pattern (Fig. 8) of the only native speaker in our database is compared with that of the standard vowel space, they are not identical as well. This is a reasonable result because of differences in human vocal tract dimensions, although some

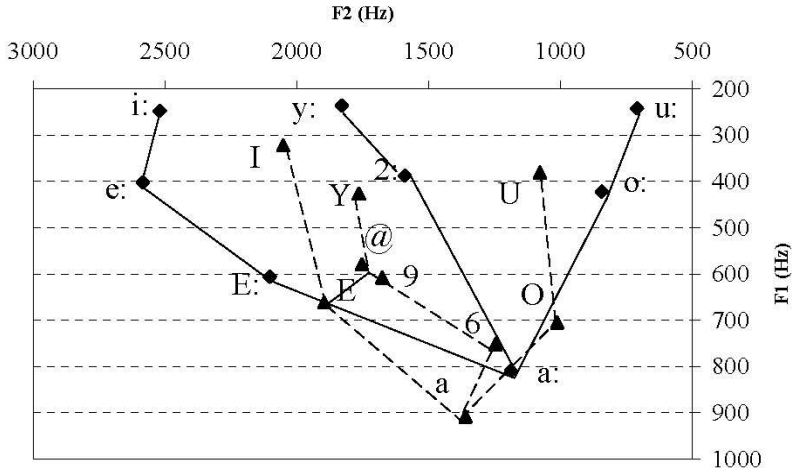

Fig. 8. The chart of German vowels spoken by a native female speaker.

common characteristics of these patterns are present. Except of the tense vowel /a:/, the lax vowels are located inside the corresponding tense vowels (this exception is indicated as normal for many German native speakers in [3]). Furthermore, the central vowels acquire a more central position in the reference vowel space, than those spoken by the Chinese speakers.

5.2.2. Comparison based on the dynamic formant trajectory

In order to compare the formant trajectories of the same vowel with different durations, the values were calculated at 5 equidistant points between 20% and 80% of the vowel duration, i.e. they were calculated at 20%, 35%, 50%, 65%, and 80% (which is represented in Figs. 9 and 10 as points 1–5, respectively) of the durations of each vowel. Because the absolute dynamic formant movements are quite different for individual speakers, we plotted the formant diagrams for each speaker separately. We can not, however, present all the plots of every vowel here, only one pair of the lax /E/ and tense /e:/ is illustrated as a representative example in Figs. 9 and 10.

In these figures, the thin dotted lines illustrate the formant trajectory of the ten Chinese speakers; the thick solid lines represent that of the German speaker. It is evident that the Chinese native speakers have generally a higher F1 and F2 than the German native speakers. However, F1 is still comparable to that of the native speaker.

Similar values of the German native speakers have also be found in other researches. For example, it is reported in [9] that the averaged formant frequencies for /e/ are 307 Hz (F1), 2080 Hz (F2), and 451 Hz (F1), and 1663 Hz (F2) for /E/. These values were measured during the steady-state portion of the citation syllable, which are comparable to our measures of the isolated words.

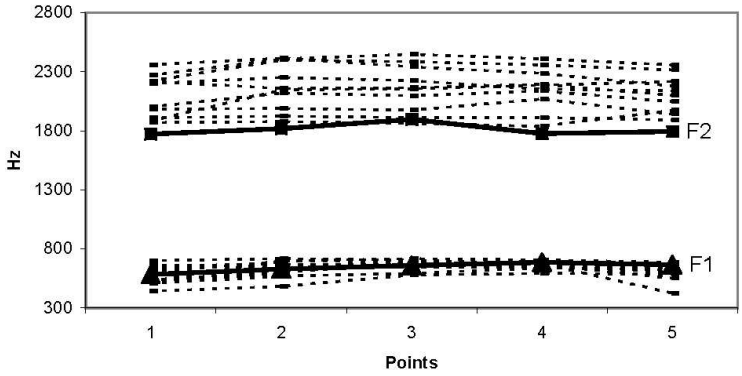

Fig. 9. Formant pattern of /E/.

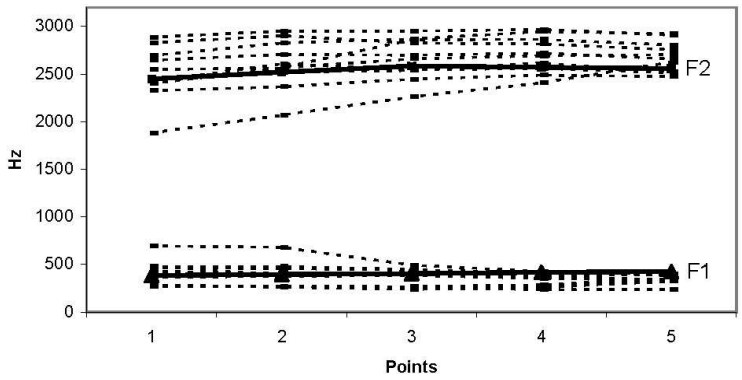

Fig. 10. Formant pattern of /e:/.

5.2.3. Results of the comparison

Although the formant frequencies differ from person to person, they can hardly explain any difference in vowel properties [4]. However, if the positions of the cardinal vowels are employed as reference points, the properties of the other vowels can almost be inferred. From the comparison of the vowel charts and the formant plots, several differences can be listed:

1. The F1 frequency differences between /U/ and /u:/, /O/ and /o:/ are much smaller for Chinese speakers than those of the native speaker, whereas the quantitative differences of F2 are comparable.
2. All the lax vowels (except of /a/) of the native speaker are much more centralized in the vowel chart than those of the Chinese speakers.
3. The umlauts and front lax vowels of the native speaker lie fast in the central part of the vowel triangle, but the vowels of the Chinese speakers are located a little bit in the front part of the vowel triangle. This is also represented by the formant plots. These lax vowels have much greater F2 values than those of the native speaker.

Similar results have been reported also in researches on the foreign accent of Chinese English speakers. For example, in [6] it was found that "the experienced in English Mandarin subjects did not produce a significant B2–B1 (front/back) difference between /e/ and /{/ (English vowels)".

Some explanations of the differences can be provided:

1. The close spectral distance of F1 of /O/ and /U/ to their tense counterparts can account for the inaccurate pronunciation of these two vowels.
2. The relative front position (larger F2) of the central vowels /I/, /2:/, /Y/, /E/, etc. can account for the stronger accent of these vowels.

As we know, the vowel height is negatively correlated with the F1 frequency, while the vowel frontness is correlated with F2. We can conclude that Chinese speakers should have a little more backness in producing front and central lax vowels. In the second language learning, they have problems to pronounce the lax vowels accurately, Chinese speakers should pay attention not only to the durational differences in the short-long pairs, but they should also make efforts to differentiate the spectral distance between the tense-lax pairs, for example pay more attention to the lip-rounding, so as to imitate the native speakers.

6. Conclusions

This investigation provide some statistical results to the understanding of the influence of Chinese on the production of German vowels and to the acoustic properties of foreign accent. Although the database employed was not very large, the statistics was carefully conducted, and the problem was explicitly elaborated. They are representative for the investigation in this field. However, it is clear that the foreign accent results not only from monothongs, but other phonetic units such as diphthongs, consonants, and intonations also contribute to its perception. A larger body of a speech database is needed to obtain more refined results.

References

[1] FLEGE J., *Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training*, J. Acoust. Soc. Am., **86**, 1684–1697 (1989).

[2] JOKISCH O., DING H., *et al*., *Learning syllable duration and intonation of Mandarin Chinese*, ICSLP, Denver, 1777–1780 (2002).

[3] RAMERS K. H., *Vokalquantität und -qualität im Deutschen*, Niemeyer, Tübingen 1998.

[4] LADEFOGED P., *A course in phonetics*, Harcourt Brace Jovanovich College Publishers, America 1993.

[5] HESS W. J., *Speech synthesis – a solved problem?*, Signal Processing VI: Theories and Applications, Vandewalle J. *et al*. [Ed.], 37–46, 1992.

[6] FLEGE J. E., BOHN O., JANG S., *Effects of experience on non-native speakers' production and perception of English vowels*, Journal of Phonetics, **25**, 437–470 (1997).

[7] HOWIE J., *Acoustical studies of Mandarin vowels and tones*, Cambridge University Press, Cambridge 1976.

[8] MOEBIUS B., VON SANTEN J., *Modeling segmental duration in German text-to-speech synthesis*, ICSLP, 2395–2398, 1996.

[9] STRANGE W., BOHN O., *Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies*, J. Acoust. Soc. Am., **104**, 1, 488–504 (1998).