Perception and Analysis of Chinese Accented German Vowels

Hongwei Ding, Oliver Jokisch and Rüdiger Hoffmann

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

ABSTRACT
This report describes a new approach to obtain specific knowledge about the production of German vowels by Chinese speakers. Experiments have been conducted to acquire both perceptual and acoustic measures of the vowels. 11 Chinese speakers were selected as subjects to read 20 German words. Assessment was conducted by 10 German native speakers. We came up with the conclusion, that the Chinese speakers have less difficulty in pronouncing the long vowels. They are not accurate in producing the short vowels. With further acoustic investigation, it was further discovered that the Chinese speakers can not always aware of the duration contrast of long-short vowels, and they can not find the right position of the tongue and grasp the degree of lip-rounding in producing central vowels.

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 English spoken as a second language [1]. But there exist less acoustic experiments on the interference of Chinese with speaking German as the second language. This paper is devoted to describe the examination in this field, which gains increasing importance with the development of international speech communication.

The direct motivation of this experiment was aroused 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 speaker selection, besides Chinese, the pronunciation of English and German was also considered. The experiments were designed to find out which vowels are indicated by the native speakers with more foreign accents and acoustic measures were supplied to account for it. The investigation can at the same time provide implication for the German Speech Recognition of speakers with Chinese accent, it can further gain a better insight into the exact learning problems of pronunciation, and supply the second language acquisition with appropriate guidance.

2. Differences between German and Chinese vowel systems

The difference between German and Chinese vowel systems can lead to inaccurate pronunciation of vowels. Mandarin has six basic monophongs. They are slightly different in the phonological systems in the “openness” of the mouth, the level of tongue position and the ways lips are stretched or rounded. Figure 1 depicts F1xF2 of six basic vowels (represented with Sampa-C) uttered by a professional female speaker in isolation.

![Figure 1. Chart of basic Chinese vowels from a female speaker](image1)

![Figure 2. Chart of German vowels from a native speaker](image2)

On the other hand, German phonetic system consists of much more vowels. 9 vowels were selected to be presented in Figure 2. The formant values were taken from the literature [3], represented here in Sampa. All the long vowels are connected with solid lines, short vowels with dotted lines.

With a comparison of the two charts, one is easy to find, that German vowels are spread in the acoustic vowel space, but in the case of Chinese, the centralized position of the vowel triangle is absent. The first question popped up, whether the Chinese speakers have difficulty in producing the central vowels? Another consideration is, there is no long-short features in Chinese, whereas this is a prominent feature in German vowels. We wondered, does it pose a problem for Chinese learners when pronouncing German vowels?

3. Preparation of Speech Database

In order to provide answers to the questions, the first step was to develop speech material and select subjects.
3.1. Reading Material

Because it is not possible for the speaker to pronounce the vowels in isolation or in nonsense carrier words, on the other hand, it is also difficult for the naive native listeners to make any assessment. Therefore all the vowels were embedded in meaningful carrier words. Thus 20 simple words, which include all the German vowels were selected as the reading list. These carrier words are listed in Table 1 together with the embedded vowels:

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

3.2. Subjects

Chinese, who speak good standard Mandarin and at the same time have advanced German skills, were selected to take part in this test. The 10 female speakers are all candidates for the acoustic inventory speaker of our Chinese synthesis. Another male speaker was also included in the test for the possible analysis of sex differences. All the 11 Chinese subjects were students or staff at Dresden University. They have been living in Germany for many years. It is also uncommon experience that such advanced speakers speak the second language with foreign accents. In order to facilitate the comparison, a normal native female speaker was also taken in the experiment as a reference.

3.3. Recording of Database

Recordings were made in an anechoic chamber with a high-quality microphone with pop protection. Recorded tokens were directly digitalized at a sampling rate of 32 kHz with 16-bit resolution and stored onto the hard disk of the work station. A German and a Chinese expert monitored the recording 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 vowels. Instruction was stressed that they should judge the pronunciation in reference to German from the broadcasting, and the focus was the vowel. Fortunately, all the consonants were not difficult for these advanced
speakers, the vowel quality became vital to the perception. The stimuli for the experiment were 20 words from 12 speakers, altogether resulted in an amount of 240 words. Because these advanced speakers did not have the problem of intelligibility, it means one vowel or word should not be confused with another. The problem was that the produced vowel is not accurate as that of the native speaker, or in other words, they have Chinese accents. So the identification of the words was unnecessary, the intended words were presented in the 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. Scale 5 means it is a standard pronunciation without any accent. Scale 1 means it is so poor, it can be confused with another word.

4.2. Results

Responses from the listeners were analyzed and compared.

4.2.1. Results of the Speakers

It was clear to notice that these advanced speakers were still indicated by the native speakers as having foreign accents. Reference can be made in Figure 3.

Among them Speaker 4 (noted as 4-ref.) is a normal native speaker with MOS of 4.62. The MOS of Chinese speakers ranges from 3.27 to 4.27. The male speaker is one with most foreign accent. We can conclude that the advanced Chinese speakers can grasp the properties of vowels well, but there is still discrepancy between their pronunciation and standard pronunciation.

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

4.2.2. Results of the Vowels

Our next interest is the quality of the vowels uttered by the Chinese speakers. This is presented in Figure 4. It is clear to observe, 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/, /l/, /y/, /E/, /a/) and one long vowel (/2:/). This result can be inferred from the difference of the vowel charts. The short vowels and the umlaut /2:/ are absent in Chinese vowel system, therefore the Chinese subjects can not produce them accurately as their long counterparts, which are similar in their native language.

5. Analysis of Acoustic Properties

The goal of the final experiments was to discover the acoustic parameters that might correlate with the pronunciation. Dynamic formants and duration values are argued to be important in the production of vowels, which were examined in our experiment.

5.1. Duration

Because the duration of a certain vowel is influenced by the speaking rate, the structure of the word, and so on, the absolute duration can hardly be compared with that from standard sources. But we had a reference native person, she had the same utterances. If the speaking rate was not the same, the comparison between the relative duration of the long short pairs should be proportional.

A comparison was thus made with the seven minimal pairs in German. Mean duration values of all Chinese subjects were presented in Figure 5. For the convenience of comparison, the statistics were also presented from the native speaker in Figure 6.

![Figure 5. Duration Values of minimal pairs of Chinese speakers](image1)

![Figure 6. Duration Values of minimal pairs of the female native speaker](image2)

According to the published sources [3], the opposition long-short is a striking phonetic feature of 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 in most pairs of the native speaker in Figure 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 Figure 5 the first four minimal pairs can also display such a ratio, but the difference is not so large in the fifth pair. Moreover, in the last two pairs (/a/ and /a:/, /O/ and /o:/), the short vowels have even longer duration values than their long counterparts. This also partly accounts for the reason, the four
vowels /O/, /a/, /o:/, /a:/ are among the more accented utterances in the perception test. It demonstrates that the Chinese speakers are aware of distinguishing the long-short distinction of vowels in German, but are still not successful in the manipulation for all vowels.

5.2. Formants

As we have already known, the contribution of dynamic properties of vowels is much more significant to the perception. The quality of vowel sounds depends on the formant frequencies, especially the first and second formant (F1 and F2). Vowel charts and the formant patterns were thus compared respectively in the following. Assuming that the male speaker can have different formant frequency than the female ones, the male speaker was excluded in the formant comparison.

5.2.1. Comparison based on F1-F2 vowel charts

Because the vowel formants are influenced from the neighboring consonants at the beginning and ending parts, which is also mentioned by [5], we just averaged the values at the middle point of the monophongs from the Chinese speakers, which are plotted in Figure 7, and those of the native reference speaker are also presented in Figure 8.

![Figure 7. Chart of German vowels from Chinese speakers](image)

In Figure 7, we can observe, that the Chinese speakers have a similar pattern of vowel charts as that of an ideal native speaker (ref. Figure 2): F1 of all the lax vowels (/I/, /E/, /a/, /O/, /U/, /Y/, /ɔ/) are shorter than their corresponding tense vowels (/i/, /e:/, /a:/, /o:/, /u:/, /y:/, /ɔ:/); and all the short vowels connected with dotted lines are located in the triangle of the long vowels connected with solid lines. It is the same case for the native speaker, only that /a:/ is only one exception, but this is indicated as normal for many German native speakers [3].

5.2.2. Comparison based on the dynamic formant trajectory

In order to compare 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, that means they were calculated at 20%, 35%, 50%, 65%, and 80% (which are represented in Figure 9-10 as point 1-5 respectively) of the durations of each vowel. Because the absolute dynamic formant movements are quite different from different speakers, we just plot them separately of all speakers. We can not, however, present all the plots of every vowel here, only one pair of lax /E/ and tense /e:/ is illustrated as a representative example in Figure 9 and 10.

Figure 9: Formant pattern of /E/  
Figure 10: Formant pattern of /e:/

In these figures, the thin dotted lines illustrate the formant trajectory of ten Chinese speakers, the thick solid lines represent that of the German speaker. It is discovered, that the Chinese native speakers have generally higher F1 and F2 than the German native speakers. However, F1 is still comparable to that of the native speaker. As to F2, the Chinese speakers have much higher values of lax vowels (such as /E/) compared to the native speaker.

5.2.3. Results of the comparison

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

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

Some explanations can be provided for the differences:
1. The close 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/, /E/, /Y/, /E/, etc. can account for the more accent of these vowels.

As we know that vowel height is negatively correlated with F1 frequency; vowel frontness is correlated with F2. We can conclude that Chinese speakers should have a little backness in producing front and central lax vowels. Due to the fact that the Chinese speaker does not open his mouth wide enough or shape his lips as needed to produce the sounds correctly, and thus resulted in a foreign-accented German pronunciation of some vowels.

6. Conclusions

This investigation presents a new approach to understanding of the interference of Chinese in pronunciation of German vowels and the correlate of acoustic properties to account for that. Although the database employed was not very large, the statistics was carefully conducted, and the problem was explicitly elaborated. They are representative in the investigation in this field. It is clear, foreign-accented is not only resulted from monothongs, other factors such as diphthongs, consonants, and intonations also contribute to it. A larger body of speech database is needed to obtain more refined results.

References


