Devoicing of vowels in German, a comparison of Japanese and German speakers

Rei Yasuda\textsuperscript{1,2}, Frank Zimmerer\textsuperscript{1,3}

\textsuperscript{1} Institute of Phonetics, Goethe-University, Frankfurt, Germany
\textsuperscript{2} Graduate School of Language and Culture, Osaka University, Japan
\textsuperscript{3} Institute of Phonetics, Saarland University, Saarbrücken, Germany

zimmerer@em.uni-frankfurt.de, yasuda@em.uni-frankfurt.de

\textbf{Abstract}

In Tokyo Japanese, vowel devoicing is a common process, that leads to the reduction of high, unstressed vowels (\textit{\textless}i\textgreater{} and \textit{\textless}u\textgreater{}) between unvoiced consonants. This article investigates to what extent native Japanese speakers (L1) learning German as foreign language (L2) show a strong tendency to produce these vowels in the foreign language as devoiced, too. Furthermore, the question is addressed whether German native speakers also devoice vowels in the same context. To this end, a production study of German words with German (L1) and Japanese (L2) native speakers was carried out. Results of this production task show that Japanese speakers devoice vowels in German words quite regularly, whereas German speakers show this pattern only rarely. For the Japanese speakers, the reduction patterns are comparable to those of the native language. Thus, interference of Japanese (L1) patterns can be observed in German (L2) indicating that this process is deeply rooted in Japanese speakers’ phonetic/phonological knowledge and leads to interference when learning a foreign language, irrespective of the existence of the process in that language (L2).

\textbf{Index Terms}: speech production, devoiced vowels, Japanese, German, interference

1. Introduction

When adults learn a foreign language, they rarely reach perfection in that language. Their native (L1) phonological system and their phonetic knowledge (e.g. the way consonants or vowels are produced) of their native language are part of the reason why this is the case. This knowledge leads to the production of processes of the L1 in the foreign language (L2), creating interference in the target language, leading to an audible accent. Connected to this interference in production are also possible problems in perception (e.g. [1,2,3,4,5,6,12,13]). For instance, Dupoux and colleagues [5] found that the phonological system of French where accent (and word stress) is predictably final and arguably is not part of the lexical representation, leads to “stress deafness” in French speakers. On the other hand, the phonological system of Japanese makes Japanese listeners epenthesize a vowel that is physically not present in the speech signal in non-native words [12,13]. This result is also observable in production when speakers epenthesize a vocalic element into L2 consonant clusters that are banned in the phonotactics of L1 (e.g. [3] – although the exact phonological status of this epenthetic vowel is not clear).

One aspect of languages where differences occur are vowel systems. Languages can differ with respect to the overall number of vowels, but also in the way the vowel space is organized. If two languages have organized the vowel space differently, and if they also have a different number of vowels, perceiving and producing vowels in L2 is affected by interference of L1 very often (e.g. [4]). Additionally, not only the systems themselves can vary, but also phonological or phonetic processes that affect vowel pronunciations can be different across two languages.

In (Tokyo) Japanese, one such process that affects vowels and leads to a reduction is high vowel devoicing (e.g. [6,7,8]). In between voiceless consonants, speakers of Tokyo Japanese tend to devoice the high vowels \textit{[i]} and \textit{[u]} regularly (e.g. [5,6]) (subsequently, these vowels are referred to as \textit{\textless}i\textgreater{} and \textit{\textless}u\textgreater{}). In Tokyo Japanese, the high vowels in the system (i.e. \textit{\textless}i\textgreater{}, and\textit{\textless}u\textgreater{})- concerning the notation in this paper, a vowel between \textit{\textless}>\textgreater{} indicates orthographic conventions, \textit{\textless}>\textgreater{} concerns its underlying structure, \textit{[i]} denotes the actual realization of the vowel) are devoiced up to 90\% of the time, when they occur in the correct context (e.g. [6,7,8]). On the other hand, there is no phonological process of high vowel devoicing in German. Devoicing may occur, but it is not a regular process (e.g. [9]). The German vowel system consists of three high vowels (ignoring vowel length, and laxness), \textit{\textless}i\textgreater{}, \textit{\textless}y\textgreater{} and \textit{\textless}u\textgreater{}. Reduction to schwa may occur, but complete devoicing of vowels occurs rather seldom (e.g. 8, 10).

Since any studies have shown that regular processes in L1 can lead to interference when learning a foreign language, such as vowel epenthesis in production and perception (e.g. [3], [12], [13]). The process of vowel devoicing is an excellent way to test the extent to which such interference might occur when native Japanese speakers learn a L2 where the process of high vowel devoicing has a different phonological (or phonetic) status. A production study is aimed to test this interference and the amount of devoicing produced by German speakers for the same words.

2. Methods

2.1. Participants

A production study was carried out to compare devoicing patterns for Japanese and German speakers. To that end, 10 Japanese (9 females and 1 male; 7 from the Tokyo area, two from the Kyushu area and one from the Kinki area, all dialects show vowel devoicing – since there is no evidence that there is a gender difference concerning vowel devoicing (e.g. [8]) we did not balance for male and female speakers) and 10 German (7 females and 3 males) native speakers volunteered to participate in the experiment. All the participants were recruited in Frankfurt (Germany). All of the Japanese participants had studied German (M=2.5 years, ranging from 0.5-3.5, SD=0.9) and lived in Germany (M=19.1 months, ranging from 1-42, SD=17.3) at that time. Their average age was 23.1 years old (ranging from 21-29, SD=2.9). Average age of the German participants was 29.5 years (ranging from 21-43, SD=6.3). None of the German speakers had learned

\textbf{Devoicing of vowels in German, a comparison of Japanese and German speakers}

\textbf{Introduction}

\textbf{Participants}

\textbf{Methods}

\textbf{Results}

\textbf{Discussion}

\textbf{Conclusion}
Japanese, none of them had stayed in Japan, nor had they heard about vowel devoicing in Japanese before.

2.2. Stimuli

For the experiment, we chose 126 trisyllabic words of German. The target vowels /i, u, y/ were located in different contexts and syllable conditions (initial, middle, final). All of the vowels occurred in the structure C1 V C2, that is, in the vowel (V) was always surrounded by two voiceless consonants (C1 and C2). C1 could be any of the following consonants (or clusters): ː/h, k, ks, t, s, ŋ, x, ts/, whereas C2 consonants were one of these: ː/f, k, p, t, s, ŋ, x, ts/.

The distribution of vowels was not symmetric, reflecting on the one hand the distribution of trisyllabic words in German with the respective vowels in the stimuli, on the other hand, they reflected the Japanese vowel system. In the system, <i> vowels are also part of Japanese, whereas <u> is produced as unround vowel in Japanese ([ui]) and <y> is not a phoneme of Japanese. In all, 88 stimuli included /i/ (e.g. Lexikon – ‘lexicon’), 20 /u/ (e.g. Milchzucker ‘milk sugar’) and 18 /y/ (e.g. Hypothese ‘hypothesis’).

The stimuli were presented in two conditions. First, the sentence condition, where they were spoken in the carrier sentence “Ich mag das Wort X nicht” (‘I don’t like the word X’ – X denotes the place of the respective word) or “Nicht Heidi, sondern Peter mag das Wort X nicht” (‘Not Heidi, but Peter doesn’t like the word X’). Note that in both cases, the main accent had to be placed on NICHT ‘not’, which was indicated by a ‘‘” on the word to avoid differences in accent and ensure for rather unstressed words of interest in the production). This condition was produced at normal speaking rate. The second condition was the word condition, where participants had to repeat each word three times in a row at a faster speaking rate.

Definitions of these speaking rates were as follows: participants were asked to read a ‘comfortable’ speed in the sentence condition, and at ‘the fastest rate without making an excessive amount of errors’ in the word condition. Furthermore, in the latter condition, there was additional time pressure (see procedure section below). The materials were recorded with 16 bit quantization and a 44.1 kHz sampling rate using an Audio Technica condenser microphone, pre-amplified and subsequently connected to a Mac Mini computer with PRAAT [11].

2.3. Procedure

Stimuli were presented as PowerPoint file on a computer screen. In the sentence condition, the task was self paced, that is, participants clicked for the next sentence (slide) as fast or as slow as they wished, after having finished reading the sentence.

In the word condition, participants had a fixed amount of time in which they were to produce the three repetitions of the word, after which a warning tone signaled the next word, and the slides changed automatically. Japanese participants had 4.5 seconds for the three renderings of the word, whereas German participants had to finish the repetition of each item within 3 seconds. The warning tone was played to participants via headphones. The different intervals were chosen deliberately. In a pre-study, we established that the task was harder for Japanese speakers and that they needed more time for their productions. An experimental recording session lasted about 17 minutes for the German participants and about 25 minutes for Japanese participants. Between the two conditions, they could take a break as long as they wished.

2.4. Data analysis

For the sentence condition, we expected 2520 stimuli (126 stimuli * 10 persons * 2 groups = 2520 in total). The word condition was set out to create three times that amount, that is 7560 stimuli (126 stimuli * 3 times * 10 persons * 2 groups = 7560 in total). Figure 1 shows the devoicing rates for the different speakers.

The analysis was conducted using visual scaleable spectrograms and oscillogram displays as well as auditory information with help of PRAAT [11] manually by two trained phoneticians. Only when they agreed, the item was treated as devoiced. Each item was judged in a binary decision either to have the vowel devoiced or not. Only cases of complete devoicing were labeled as devoiced. Sometimes, vowels seemed to be deleted completely. However, a distinction between devoiced and deleted was not always possible. Therefore, both cases were treated as devoiced (i.e. reduced). There could have been more elaborate judging systems, however, since the binary choice (completely reduced vs. unreduced) is rather conservative, this strategy was chosen.

3. Results

In this section, we present the results of the production experiment. First, we report the analysis for the words that were produced in the sentence condition. After that, the analysis for the word condition is presented, before we compare the results of the two conditions.

3.1. Sentence condition

Table 1 presents the number of devoiced vowels and the devoicing rate in the sentence condition. First, the data is analyzed for Japanese participants. Overall, the devoicing rate was 14%. Concerning the vowel /i/, 16% of the vowels were devoiced, the vowel /u/ was devoiced in 5% of the cases, and the vowel /y/ was reduced 12% of the time.

Table 1. Devoiced vowels in sentence condition (normal speed). “JAP” stands for Japanese, “DEU” for German.

<table>
<thead>
<tr>
<th>Speaker group</th>
<th>Vowel</th>
<th>Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAP</td>
<td>/i/ (n=88)</td>
<td>14.2</td>
</tr>
<tr>
<td></td>
<td>/u/ (n=20)</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>/y/ (n=18)</td>
<td>1.7</td>
</tr>
<tr>
<td>S.D. Rate</td>
<td></td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>DEU</td>
<td>/i/</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>/u/</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>/y/</td>
<td>0.1</td>
</tr>
<tr>
<td>S.D. Rate</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
</tr>
</tbody>
</table>

3.2. Word condition

Table 2 presents the number of devoiced vowels and the devoicing rate in the word condition. First, we report the analysis for the words that were produced in the sentence condition. After that, the analysis for the word condition is presented, before we compare the results of the two conditions.

Table 2. Devoiced vowels in word condition (normal speed). “JAP” stands for Japanese, “DEU” for German.
The devoicing rate for the /i/ vowel was higher than for the other vowels. A one way ANOVA was calculated and showed a significant vowel effect (F(2,9)=8.09, p<0.01) concerning the devoicing rate. A multiple comparison (holm) showed that the devoicing rate of /i/ was significantly higher than the devoicing rate for /u/, but there was no significant difference between /i/ and /y/, nor between /y/ and /u/.

For German participants, we observed only 0.1% devoicing rate in words with the vowel /i/, no devoicing was found in words with /u/ or /y/.

Since German speakers did produce devoiced vowels so rarely in this condition, that is, performed at ceiling (almost 100% of the vowels were voiced) no further statistical model was calculated, neither for the German group alone nor to compare the two language groups.

3.2. Word condition

Next, the word condition was analyzed. Table 2 presents the number of devoiced vowels and the devoicing rate in this condition. For Japanese participants, the overall devoicing rate was 18%. The devoicing rates of the respective vowels were 19% for /i/ words, 7% for /u/ items and for the vowel /y/, target vowels were devoiced in 15% of the cases. Similar to the sentence condition, the devoicing rate of vowel /i/ was the highest. A one way ANOVA (devoicing rate as dependent variable, with vowels (three levels: /i/, /u/, /y/) as random variables) showed a significant vowel effect (F(2,9)=18.77, p<0.001). A multiple comparison (holm) showed that the devoicing rate of vowel /i/ and /y/ were significantly higher than vowel /u/ (p<0.01). No significant difference was found for the vowels /i/ and /y/.

For German participants, we observed only 1% devoicing rate in vowel /i/, and the devoicing was not observed at all for the vowels /u/ and /y/. Again no further analysis was performed.

3.3. Comparison of Conditions

On first sight, the devoicing rate for Japanese participants and for German participants seems higher in the word condition (fast speed) compared to the sentence condition (normal speed). This is true for every speaker as can be seen in Figure 1. What can be seen in this figure as well is that the German speakers with the highest devoicing rate show less devoicing than the Japanese speaker with the smallest devoicing rate (DEU_01 & DEU_02 vs. JAP_05). Figures 2 and 3 illustrate the deletion rates for the two language groups individually. The figures show the low amount of reductions for German speakers and the difference compared to native Japanese speakers.

![Figure 1](image1.png)

**Table 2. Devoiced vowels in word condition (fast speed). “JAP” stands for Japanese, “DEU” for German.**

<table>
<thead>
<tr>
<th>Speaker group</th>
<th>Vowel</th>
<th>/i/ (n=88*3)</th>
<th>/u/ (n=20*3)</th>
<th>/y/ (n=18*3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAP</td>
<td>Avg. no of dev.</td>
<td>47.9</td>
<td>4.4</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>9.0</td>
<td>3.1</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>Rate (%)</td>
<td>19</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>DEU</td>
<td>Devoiced vowel</td>
<td>4.4</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>4.0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Rate (%)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

In order to investigate the effect of different devoicing rates further, we performed a mixed model, with participant as random factor, and the factors language group (Japanese or German), condition (sentence or word) as fixed factor, and additionally their interaction. The model showed that the main effects language group (F(1,18)=90.04, p<0.0001) and condition (F(1,18)=25.75, p<0.0001) were significant as well as a significant interaction between language group and condition (F(1,18)=7.00, p<0.05). Post hoc tests indicated that the difference between the conditions was not significant for German speakers (p>0.05), but there was a significant difference for Japanese speakers in the two conditions (p<0.0001).
4. Discussion

The results indicate that Japanese speakers produce devoiced vowels in German rather regularly. In some words (e.g. Architekt – ‘architect’), the devoicing rate is 100%. The results thus are comparable to the findings in Japanese, where, depending on the segmental context, devoiced vowels occur (e.g. {[6,7,8]}). Furthermore, a vowel effect emerged, /ɪ/ and /y/ were reduced more often than /u/. This finding can be connected to the Japanese vowel system. As in Japanese, the vowel with the highest devoicing rate is /ɪ/. In the system, there is no /y/. Subsequently, participants produced /y/ items almost exclusively with /ɪ/ when it was not devoiced. Thus, it is not surprising that the devoicing rates of the two vowels are comparable. Concerning /u/, there is a vowel /u/ in Japanese. However, this vowel is rather /ɯ/. And many renderings of German /u/ words were produced as [u], some also as [u].

At the same time, German speakers produce devoiced vowels only rarely. There was only one case in all the sentence conditions, and even in the fast word repetition condition, only /ɪ/ was devoiced, but very rarely. This is a finding that finds support also in independent analyses of German speech. Vowels in German are rarely devoiced or deleted, and even reduction to schwa is not that common (e.g. [10]).

At the same time, German speakers produce devoiced vowels only rarely. There was only one case in all the sentence conditions, and even in the fast word repetition condition, only /ɪ/ was devoiced, but very rarely. This is a finding that finds support also in independent analyses of German speech. Vowels in German are rarely devoiced or deleted, and even reduction to schwa is not that common (e.g. [10]).

Note that the dichotomous decision of devoicing (or deleting) vs. voicing may not be sufficient to capture the complete process of devoicing. However, this is a rather conservative measure in that only cases of clear reduction are labeled as such. And the presence of (some) voiced parts of vowels is a clear cue for voicing. Furthermore, cases of partial devoicing occur for Japanese and German speakers. Thus, we opted for the choice of a dichotomous labeling.

5. Conclusions

This study was set out to investigate the extent to which Japanese speakers produce devoiced vowels when they speak German. This is a case of interference of L1 when learning L2 (in this study: German). The fact that the process of high vowel devoicing is produced very regularly, despite of the lack of regularity of this process in German indicates that it is deeply rooted in Japanese phonetics/phonology. German speakers on the other hand rarely devoice vowels, as a consequence, German listeners do not hear this process very often and are not familiar with this process.

For language learning, this is an interesting finding. Until now, the devoicing (or refraining thereof) is not part of L2 learning or teaching. It also seems that Japanese learners are not aware of the process. Thus, including this process in L2 teaching, by raising awareness of the process and by practicing the correct production of German vowels, could be a way of reducing the degree of foreign accent for Japanese L2 learners.

An interesting question is, whether devoiced vowels can lead to problems in perception by German listeners, since they only rarely find themselves exhibited to devoiced vowels.

6. Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG – SPP 1234) and the German Academic Exchange Service (DAAD).

7. References