A COMPUTER-ASSISTED VOWEL TRAINER FOR THE GERMAN LANGUAGE

Jeanin Jügler [1,2] and Frank Zimmerer [1]
Computational Linguistics and Phonetics, Saarland University
Institute for the German Language (IDS), Mannheim
juegler | zimmerer@coli.uni-saarland.de

Kurzfassung: Learners with different native language backgrounds often show problems in the acquisition of German vowels caused by interferences of the phonological vowel system of the native language with the complex German vowel system. German not only distinguishes between tense and lax vowels but also between short and long vowels. The aim of the vowel trainer is to help learners of German with different native language backgrounds to train and acquire the correct pronunciation of vowels with respect to quality and quantity. Using a game-like pronunciation training system, users can learn intuitively. This way, no provision of phonetic information about the production of German vowels, the German vowel system, and specific differences to the native language of the learner are necessary.

1 Introduction

The acquisition of the correct production and perception of vowels in a second language (L2) is very difficult for language learners [1, 2, 3, 4, 5, 6, 7]. In general, the German vowel system seems to be rather complex for many leaners since it consists of sixteen monophthongs and three diphthongs. In addition, German distinguishes between tense and lax vowels as well as short and long vowels. Furthermore, the rounded front vowels (/yː ʏ, ø, œ/) are typologically marked. Analyses of the IFCASL corpus [8] (read speech) showed that French learners of German have problems with the correct production of vowels with regard to length and quality [7, 9]. Zimmerer and Trouvain [7] investigated the perception of French speakers' production of German vowels by German native listeners. They performed an identification experiment using German minimal pairs. Results indicated that learners showed problems producing German vowels correctly. Although advanced learners' productions were identified more often correctly than productions of beginners, both groups showed lengthening as well as shortening errors. Interestingly, rounded vowels seemed to cause more difficulties than unrounded vowels. For example, in the sentence „Im Frühling fliegen Pollen durch die Luft." (In spring pollen hurtle through the air), some leaners produce the lax /ɔ/ in <Pollen> as a tense [oː]. In addition, orthographic induced interferences often cause French learners of German to produce <u> as [yː, ʏ]. In addition, Jouvet et al. [9] created phone confusion matrices allowing for a comparison of the manually corrected annotation of the produced sounds in the IFCASL corpus with the automatic alignment of the expected sounds. An analysis of these confusion matrices revealed that French learners of German showed complex interferences with vowel contrasts for length and quality. In general, learners with different native language backgrounds often show problems in the acquisition of German vowels caused by interferences of the phonological vowel system of the native language with the German vowel system.

Figure 1 - Welcome Screen of the *Hüpf* software.

2 Vowel Trainer

Within the scope of a software engineering project at the Department of Informatics at Saarland University, employees of the university can propose a software concept. The suggested vowel learning environment for learners of German was subsequently developed by seven informatics students within one semester[1].

The vowel trainer called *Hüpf* is a prototype which was not yet evaluated and tested by learners for its effectiveness in pronunciation training. The purpose of the software is to help the user to learn the correct pronunciation of vowels and difficult vowel contrasts in German, but could be extended to other languages and sets of vowels. The user receives automatic visual feedback from the software on how well the vowel was produced in terms of formant values and duration. The speaker also has the possibility to play back a reference recording produced by a German native speaker and the users own recorded utterance to compare the productions of the words.

2.1 Graphical User Interace

The name of the software *Hüpf* is based on the graphical user interface (GUI) of a frog that jumps to specific locations of water lilies in a pond. The training software was intendedly created as a game. This way, both adults and children would hopefully enjoy to use the software as it is both helpful and entertaining. Simultaneously, feedback can be presented in an intuitive way so users without any background in phonetics can use the system equally well as users with a background in phonetics. Each water lily represents mean values of F1/F2 coordinates which were fed into the system [10]. Figure 1 shows the welcome screen of the software. From this screen the user can start the training, look at the overall statistics to see which vowels have been trained so far and how well the pronunciation is, and change a few settings.

2.2 Settings

Concerning the settings), the user can change the volume, choose a different language (currently, only German and English are available. This setting only refers to the language of the GUI, not the language that should be trained.), and redo the calibration. Calibration is automatically done before starting the very first learning session and consists of recordings of /a

[1]A very special thanks to Christian Becker, Elizabeth Pich, Florin Foss, Kevin Müller, Lukas Stemmler, Ngoufack Yol Daniel Hilaire, Sven Ziegler for developing the software.

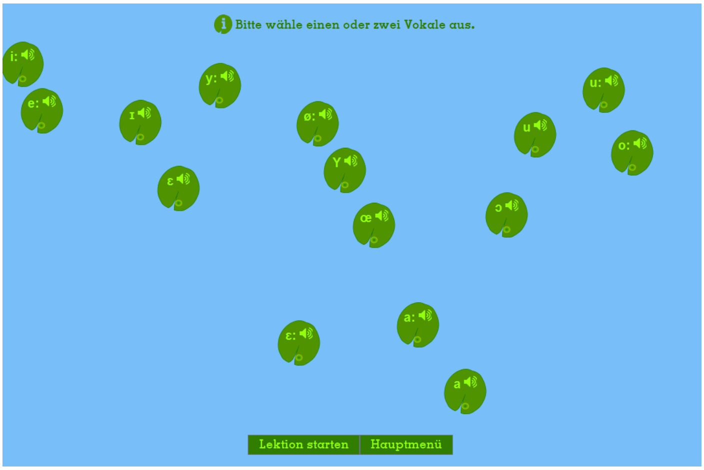

Figure 2 - Vowel selection screen of the *Hüpf* software.

i u/. These vowels were chosen since they exist in many languages and define a large vowel space. The determined formant values are used in order to normalize for the reference values [10] which are saved in the system.

2.3 Training and Feedback

After the calibration, the user is directed to the training screen where up to two vowels can be chosen and trained. The user is not restricted in the choice of the vowel contrast. Each vowel is displayed as a water lily and the position of each water lily is based on F1/F2 reference values. Users have the possibility to listen to random examples for each vowel by either a male or female German native speaker which might be useful if the learner is not familiar with the displayed phonetic symbols.

By clicking on 'Lektion starten' (start session), the user enters the training session (see Figure 3) which consists of two training units per vowel so far. However, the number of training units can be changed. Again, the location of the vowels or water lilies depends on their F1/F2 values. The user can listen to a recording of the word by a male or female German native speaker and can also record the word to receive feedback by clicking on the microphone symbol. If the frog 'lands' on the correct water lily, the production of the vowel was correct. If, however, the frog 'lands' on a different water lily or in the water, the production was incorrect (see Figure 4). The position of the frog represents the normalized F1/F2 values for the vowel of the recorded word. Currently, vowel recognition is not very advanced and no automatic speech recognition is integrated. The vowel of each word is detected with the help of the intensity contour of the complete word. It is assumed that the vowel of the stressed syllable shows the highest intensity. Based on the intensity peak of the vowel, vowel boundaries are estimated.

Regarding vowel duration, two bars are displayed around each water lily. The black bar illustrates the reference length that was fed into the system and the red bar illustrates the the length of the produced vowel.

Figure 3 - Training screen of the *Hüpf* software for the vowel pair /o: ɔ/.

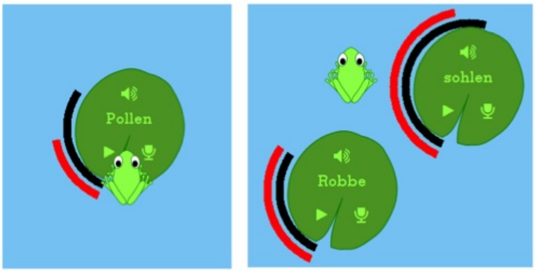

Figure 4 - Training screen of the *Hüpf* software. Left: The frog landed on the correct water lily, the production of the vowel formants was correct. But the duration was too short. Right: The frog landed in the pond, the production of the vowel formants was incorrect. But the duration is correct.

2.4 Summary and Progress

After the training is completed, users will receive a summary of their duration and pronunciation (F1/F2) accuracy as well as information about how many exercises they have completed for each vowel (see Figure 5). For an overview of all vowels, the button *Statistik* (statistics) on the home screen can be clicked. The program logs the learners' progress and illustrates it by means of the size of a flower. It takes into account how good the pronunciation was, how accurate the duration matched the reference value, and how often the vowel was practiced. The more the users practiced and the better the formant values and duration matched the reference values, the bigger the flower gets (see Figure 6).

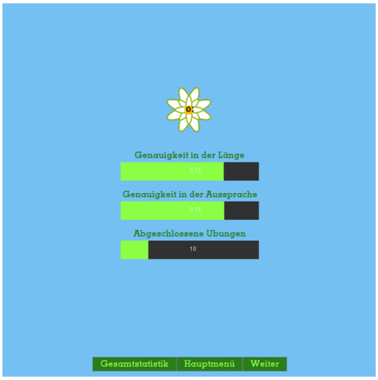

Figure 5 - Individual result screen of the *Hüpf* software for the vowel /o:/ after training. Results are displayed for accuracy in length, accuracy in pronunciation (F1/F2), and number of trained items.

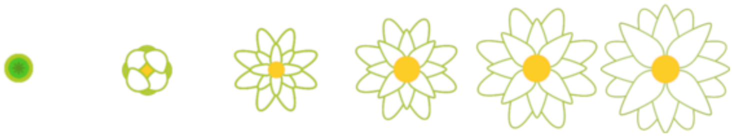

Figure 6 - Flower states in the overall statistics screen. The better the pronunciation of the vowel, the bigger the flower.

3 Outlook

Besides fixing some bugs in the software it would be desirable to evaluate and test the software with learners of German. It is important to know whether learners find the software intuitive and entertaining, and whether they will be able to improve their pronunciation of German vowels with regard to formant values and duration. Using a game-like pronunciation training system, users can learn intuitively. As a consequence, no provision of phonetic information about the production of German vowels (and in this case formants), the German vowel system, and differences to the native language of the learner are necessary. It is not argued that these information are not helpful in general but they might overwhelm a learner without any phonetic background. It would also be desirable to incorporate a forced alignment process for a more stable vowel recognition. Furthermore, this would give us the possibility to include longer words as well as phrases or sentences in the training condition. Also, as was mentioned previously, the vowel trainer can be extended to other languages. If the evaluation and testing are successful, including other target languages in the trainer could be a next step.

References

[1] DARCY, I., AND KRÜGER, F. Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics 40* (2012), 568–581.

[2] FLEGE, J. E., AND HILLENBRAND, J. M. Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America 76* (1984), 708–721.

[3] FLEGE, J. E., AND MACKAY, I. R. A. Perceiving vowels in a second language. *Studies in Second Language Acquisition 26* (2004), 1–34.

[4] IPDS. The Kiel Corpus of Read Speech, 1994. Volume 1, CD-ROM #1.

[5] JOUVET, D., BONNEAU, A., TROUVAIN, J., ZIMMERER, F., LAPRIE, Y., AND MÖBIUS, B. Analysis of confusion matrices in a manually annotated French-German learner corpus. In *Proc. Workshop on Speech and Language Technology in Education (SLaTE), Leipzig, Germany* (2015), pp. 107–112.

[6] KINGSTON, J. Learning foreign vowels. *Language and Speech 46* (2003), 295–349.

[7] LEUSSEN, J.-W. V., AND ESCUDERO, P. Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology 6* (2015).

[8] STRANGE, W., BOHN, O.-S., TRENT, S. A., AND NISHI, K. Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America 115* (2004), 1791–1807.

[9] TROUVAIN, J., BONNEAU, A., COLOTTE, V., FAUTH, C., FOHR, D., JOUVET, D., JÜGLER, J., LAPRIE, Y., MELLA, O., MÖBIUS, B., , AND ZIMMERER, F. The IFCASL corpus of French and German non-native and native read speech. In *Proc. Language Resources and Evaluation Conference (LREC), Portoroz, Slovenia* (2016), pp. 1333–1338.

[10] ZIMMERER, F., AND TROUVAIN, J. Perception of French speakers' German vowels. In *Proc. Interspeech, Dresden, Germany* (2015), pp. 1720–1724.