Perception of French speakers' German vowels

Frank Zimmerer, Jürgen Trouvain

Phonetics, Saarland University, Saarbrücken (Germany)

`zimmerer|trouvain@coli.uni-saarland.de`

Abstract

Despite some similarities in their inventories, German and French exhibit marked differences in the vowel systems. Most notably, German uses both tenseness and length to differentiate vowels, whereas in French, vowel length is not distinctive. Therefore, interferences can be expected when French native speakers learn to speak German. Results of a vowel judgment experiment with vowels in minimal pairs produced by 56 French learners of German, indicate that these learners have indeed problems producing German vowels correctly. Advanced learners manage vowel productions better than beginners. Both groups show lengthening as well as shortening errors. Furthermore, rounded vowels seem to pose more severe problems in L2 acquisition than unrounded vowels. These results have important implications for language learning and teaching, particularly for individualized computer-assisted pronunciation training.

Index Terms: Second language acquisition, interference, L2, speech production, learner corpus, German, French

1. Introduction

Learning a foreign language after childhood is a hard task. Especially phonetic and phonological aspects of foreign languages are hard to acquire. One of the reasons for this is the fact that the native language (L1) has a strong impact on the way new sounds in the foreign language (L2) are perceived and produced, e.g. [1, 2, 3, 4, 5, 6]. This strong impact is often manifested in the production and perception of vowels, if the phonological systems of L1 and L2 are different with respect to their vowel inventories and the features with which the vowels are differentiated [3, 5, 6, 7]. For instance, Kingston [6] investigated the perception of German vowels by native English listeners and showed how perceptual training can improve perceptual problems in the acquisition of German vowels for English L1 listeners. Similarly, Italian learners of English have been found to have problems discriminating English vowels, such as /ɑ-ʌ/, /ɛ-æ/, or /i-ɪ/ [5], which are not contrastive in Italian but in English. In a study investigating both perception and production of German vowels, Darcy and Krüger [3] showed that children whose L1 was Turkish and who started to learn German as L2 had perceptual problems with German vowels but were able to produce them correctly.

The problems of acquiring new segments are not limited to perception, but they also affect production. Many speakers have an audible foreign accent when speaking in L2. Italian learners of English have been found to produce some vowels not correctly. In contrast, it has been found that Italian speakers often lengthen vowels in L2 speech (e.g., they produce something like *eataly* instead of *Italy*). Furthermore, they appear to not shorten these vowels [8].

Interference from L1 to L2 in production is the focus of this article, in which we ask to what extent French learners of German produce the L2 vowels 'correctly'. The situation of French speakers learning German is similar to the one for Italian speakers speaking English. The German vowel system is an interesting possibility to study L1-L2 interferences for several reasons. It shows 16 monophthongal vowels: /iː, ɪ, eː, ɛː, ɛ, aː, a, oː, ɔ, uː, ʊ, yː, ʏ, øː, œ, ə/ - the vocalization of /r/ as [ɐ] is not considered as an underlying monophthong here. Contrary to French, German uses both tenseness and length to distinguish these vowels from each other (e.g., [9]).

In this study, German serves as the target language for French native speakers who learn German as a foreign language. The French vowel system shows both similarities as well as crucial differences compared to the German system on the phonological and phonetic level [10]. The French vowel system consists of 11 oral monophthongs: /i, e, ɛ, a, y, ø, œ, u, o, ɔ, ə/, see e.g. [11]. Thus, concerning height, roundness and laxness, the two systems are quite similar, although small acoustic differences have been identified between the two languages (e.g. [12]). However, more importantly, French has no long and short vowel contrasts [11, 10]. Therefore, like Italian learners of English, native French speakers may have difficulties to acquire the German contrasts correctly.

The possible difficulties in acquiring the correct pronunciation of German vowels may lead to an audible foreign accent. This might not be problematic in many cases where context can be used to disambiguate the two words, however, in some situations, French speakers might be misperceived. If the vowels in words like *Höhle* 'cave' and *Hölle* 'hell' are not produced correctly, then a sentence like: *Geh zur X* 'Go to(the(CLITIC)-CASE) X' may have very different meanings. To examine the possible difficulties in the vowel production of French learners of German in more detail, we investigate the following research questions:

- To what extent do French speakers show problems producing German vowels correctly?

- Is there a difference between advanced learners and learners at the beginner level?

- Can we identify individual tendencies for different speakers?

- Overall, do French learners of German behave like Italian learners of English and only lengthen vowels, or do they also produce short vowels instead of long ones?

A forced-choice judgment experiment was carried out to answer these questions. This experiment aims at investigating to what extent the French speakers were successful to produce the vowels in such a way that German listeners can identify them correctly. In the upcoming section, this experiment is described in more detail, before the results are presented and discussed. The article ends with some implications for language

September 6–10, 2015, Dresden, Germany

learning strategies and discuss possibilities for individualized computer-assisted learning.

2. Vowel judgment experiment

The stimuli used in the vowel judgment experiment are based on productions of French learners of German from the IFCASL corpus [13, 14]. The part of the corpus we are interested in here consisted of 60 German sentences read aloud by French speakers. These sentences (i.e., the words in question) were used as the basis for the judgment experiment which will be described in the next sections.

2.1. Materials and methods

Overall, 56 (24 female, 32 male) French speakers produced the items that were used for the judgment experiment. Of these, 22 (12 female, 10 male) were advanced learners (ADV) - B2, C1 according to the European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), whereas 34 were beginners (BEG) - A1, A2, B2; 12 of them female, 22 male. The sentences were recorded in a quiet office with use of the JCorpusRecorder software [15] at LORIA in Nancy. The regional background of ADV and BEG were comparable.

In the corpus, 11 minimal pairs occurred in sentences where the vowel duration and/or quality was the distinctive feature of the crucial segment. An example is the sentence *In Berlin zahlt man wenig Miete* ('In Berlin, you pay a small rent'). We decided to use the excised words rather than the sentence context, to avoid any semantic influences. Note that for ease of understanding we refer to the first word in the minimal pair as a word with a 'long' vowel, and the second one as a word with a 'short' vowel, ignoring differences in tenseness:

- /iː/ vs. /ɪ/: *Miete* 'rent', *Mitte* 'mid'; *schief* 'aslope', *Schiff* 'ship';

- /eː/ vs. /ɛ/: *Beet* 'bed (gardening)', *Bett* 'bed'; *Weg* 'way', *weg* 'away';

- /ɛː/ vs. /ɛ/: *Täler* 'valleys', *Teller* 'plate';

- /aː/ vs. /a/: *Kahn* 'barge', *kann* 'can'; *Hase* 'rabbit' and *hasse* 'hate-(1ST PERSON SG.)'[1]

- /oː/ vs. /ɔ/: *Polen* 'Poland, Poles', *Pollen* 'pollen';

- /yː/ vs. /ʏ/: *Hüte* 'hats', *Hütte* 'hut'; *Wüste* 'desert', *wüsste* 'would know';

- /øː/ vs. /œ/: *Höhle* 'cave', *Hölle* 'hell';

The words were excised from the sentential context with PRAAT [16]. In order to avoid clicks, the items were excised at zero crossings. Of the 1232 possible items (56 speakers * 11 minimal pairs * 2 words), 1157 could be used for the experiment. We discarded 75 items because the word in question was not produced correctly (for instance, if the /a/ vowel in *hasse* was produced as [aɪ], the item was not used). Of the 748 words that were uttered by BEG, 58 items could not be used (7.8%), whereas for ADV 17 items of the 484 utterances had to be discarded (3.5%). After exclusion of mispronounced words, BEG contributed between 11 and 22 items to the experiment (M: 20.3), whereas individual contributions from advanced speakers ranged between 15 and 22 items (M: 21.2). Concerning individual words, the number of items that had been produced correctly

---

[1] Note that this is a near minimal pair where apart from the vowel difference, the fricative following the vowel is voiced in the former word, and voiceless in the latter.

ranged from all 56 speakers (e.g. *Polen* 'Poland, Poles' and *Pollen* 'pollen') to 45 (M: 52.6) usable instances (*Höhle* 'cave' where most of the excluded items had been produced with a back vowel). After the exclusion, there were 574 left that had a long vowel, and 583 items with a short vowel.

2.2. Participants and procedure

Two groups of participants judged the accuracy of vowels produced by French learners of German. The first group (*phoneticians*) consisted of three trained phoneticians at Saarland University. They participated without compensation in the experiment. As a second group, six undergraduate Phonetics students of the University of Frankfurt (*students*) took also part in a vowel judgment experiment. They participated for course credit. All judges had German as L1.

PRAAT was used to run the experiment [16]. During the experiment, each participant heard a word and was asked whether the word in question was produced with the long or the short vowel. At the top of the screen, the two words of the minimal pair in question were displayed. For instance, if the item that was played was one member of the *Höhle/Hölle* minimal pair, at the top of the screen, the question *'Höhle' or 'Hölle'?* occurred. Then, the participants judged whether the word was the one with the long or the short vowel by clicking with a computer mouse on a response box that appeared below the question. Participants had the option to replay each item up to three times. Responses of long/short as well as the confidence rating (for the phoneticians only) were recorded. The experiment was split into two parts, but participants could take breaks whenever they wanted, they could always start with replaying the last item and then continue. Experimental sessions lasted between one and two hours. The experiment was set up in a way that participants were forced to respond to each item. Note that there was no possibility to change the response once it was given. Participants were encouraged to click correctly on the buttons, and there was no time-constraint, to decrease the likelihood of unintended responses. Therefore, we analyzed all responses as 'correct'. All statistical analyses were carried out with JMP, Version 11 [17]. For the analyses, we coded each item as correctly produced, if the listeners judged the word as the intended item of the minimal pair. For the purpose of statistical analyses, this was coded with a '1' whereas cases in which the intended word was not judged correctly were coded as '0'.

3. Results

3.1. Overall judgment of vowel productions

A first LME analysis of the responses (was the vowel perceived as it should have been?) was performed to find out whether the phonetician group and the student group differed from each other. Responses (i.e. CORRECTNESS) were the dependent factor, PARTICIPANT and ITEM were entered as random factor, and L2LEVEL, VOWELLENGTH and GROUP were entered as fixed factors. Results showed that the two groups did not differ from each other significantly. Therefore, the responses of both groups were combined for all further analyses.

To find out how well the productions of French learners of German were perceived as the intended words, a mixed model was calculated with CORRECTNESS as dependent variable, PARTICIPANT and ITEM as random variables, and VOWELLENGTH (*long, short*) as fixed factor. Results indicate that long vowels were identified correctly 76.6% of the time, whereas short vowels were perceived correctly in 63.9%
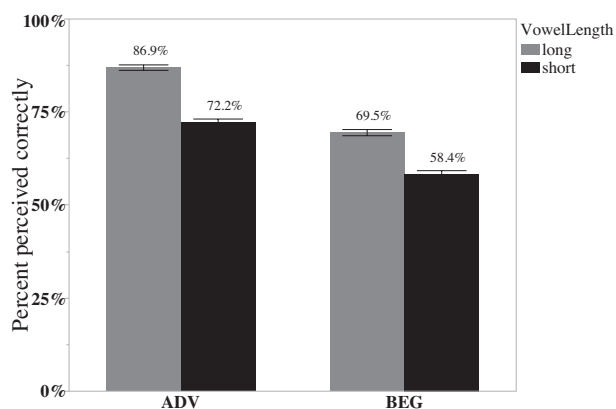
Figure 1: *Percent of correctly perceived vowels depending on the L2 level and vowel length.*

Figure 2: *Percent of correctly perceived vowels depending on the vowel and L2 level. The vertical line separates unrounded from rounded vowels. ADV are grey, BEG are black.*

of the cases. This difference was significant in the model (F(1,1155)=37.8, p<0.0001). This means that more than one third of the short vowels were perceived as long, and about 25% of the long vowels were perceived as short by German listeners.

3.2. BEG versus ADV and individual vowels

The next analysis was carried out to investigate whether ADV learners produced the vowels more successfully than the BEG group. Therefore, a mixed model was calculated with COR-RECTNESS as dependent variable, PARTICIPANT and ITEM as random variables, as well as VOWELLENGTH (*long,short*), L2LEVEL (*ADV, BEG*) and their interaction as fixed factors. The words with long vowels produced by BEG were perceived as long in 69.5% of the cases, their short vowels were perceived as short in 58.4%. The correct perception of ADV productions was 86.9% for long vowels, and 72.2% for the short ones (see Figure 1). Both VOWELLENGTH (F(1,1153)=39.6, p<0.0001), and L2LEVEL (F(1,1153)=57.7, p<0.0001) were significant factors, the interaction, however, was not.

In a next step, the model was changed to include the VOWEL as factor instead of VOWELLENGTH, all other factors (and the interaction) was kept identical. In this model, VOWEL (F(12,1131)=15.9, p<0.0001) and L2LEVEL (F(1,1131)=72.5, p<0.0001) were significant factors. The interaction was also not significant, but a trend could be identified (F(12,1131)=1.71, p=0.0588). As can be seen in Figure 2, ADV produced all vowels more successfully then BEG, but not all vowels showed the same differences.

Therefore, we separated *rounded* and *unrounded* vowels and added the factor ROUNDNESS into a model, together with VOWELLENGTH, L2LEVEL, and all their two-way interactions as fixed factors. Dependent variable and random factors were the same as before. Results indicate that all factors were significant: VOWELLENGTH (F(1,1150)=77.38, p<0.0001), L2LEVEL (F(1,1150)=64.8, p<0.0001), ROUNDNESS (F(1,1150)=47.71, p<0.0001); furthermore, the interaction of VOWELLENGTH and ROUNDNESS was significant (F(1,1149)=74.26 <0.0001). This was driven by the relatively poor performance for the short, rounded vowels (43%) which were perceived significantly worse than the other vowels (short unround: 75%, long round: 79%, long unround: 76%), which did not differ significantly from each other.
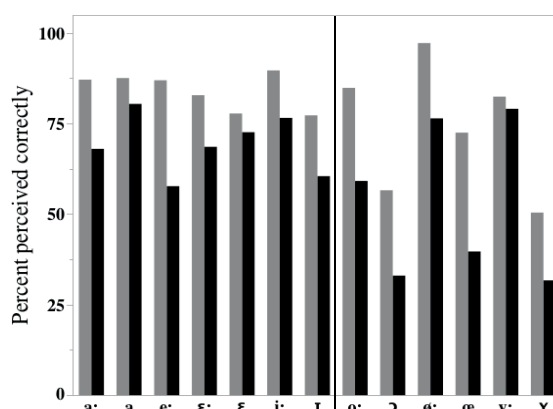
3.3. Learner-specific vowel productions

In a final analysis, we investigate whether we can identify patterns for individual speakers. We therefore entered into the mixed model CORRECTNESS as dependent variable, PARTICIPANT and ITEM as random variables, and SPEAKER, VOWELLENGTH and their interaction as fixed factors. Results indicate that all factors (including the interaction) were significant: SPEAKER, (F(55,1045)=2.98, p<0.0001), VOW-ELLENGTH (F(1,1045)=36.06, p<0.0001) SPEAKER x VOW-ELLENGTH (F(55,1045)=2.79, p<0.0001). Figure 3 depicts the different patterns for ADV, whereas Figure 4 shows how the productions of BEG were perceived. Some speakers (e.g. 30) were almost always perceived correctly (long: 95%, short: 98%). Some speakers show relatively low correct perception rates (e.g. 51, long: 52%, short: 56%). Many speakers follow the overall tendency that long vowels were perceived better than short ones (e.g. 09, 17). Some speakers show also the reverse tendency, where the short vowels are perceived better than the long ones (e.g. 01).

4. Discussion and conclusions

The results reported here can shed some light on the research questions that were posed in the Introduction section. First, it has been shown that French speakers do in fact have some problems to produce German vowels correctly. This problem affects both long and short vowels, although to a somewhat different degree. Overall, the words that had been produced by French learners of German were correctly perceived in about 70% of the time. In the other cases, German listeners misidentified these members of a minimal pair. Since vowel length is not distinctive in French, such problems were expected. Overall, French learners had less problems with long vowels than with short ones. Words with long vowels were correctly identified by German listeners in 77% of the cases, the words with short vowels were identified in about 64% of the learners' utterances.

Interestingly, French speakers seem to have more problems with the correct pronunciation of short rounded vowels than with long rounded vowels or with unrounded ones. The number of items is rather limited, however, and therefore, far-reaching interpretations should not be made. Yet, one possible explanation lies in the acoustic and articulatory differences between the languages, where, for instance, French /y/ seems to be realized
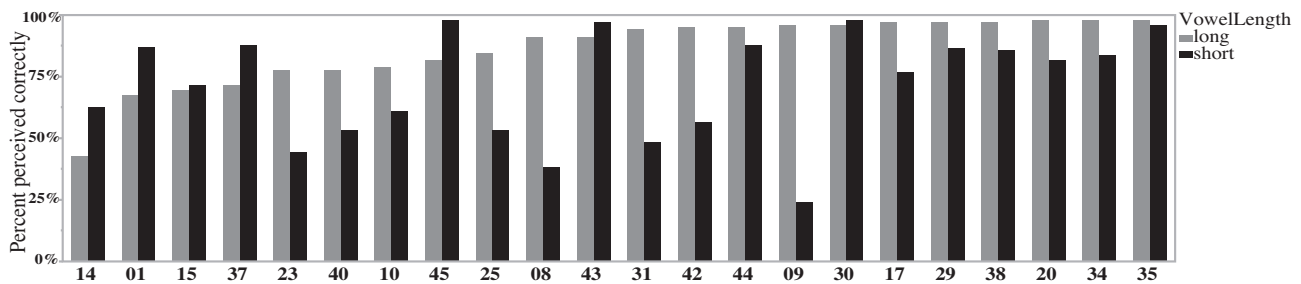
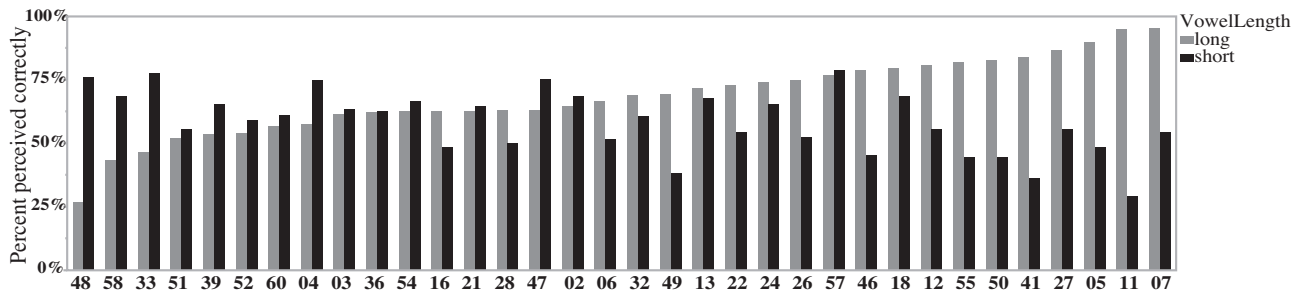Figure 3: *Percent of correctly perceived vowels depending on the speaker and vowel length for the ADV group.*

Figure 4: *Percent of correctly perceived vowels depending on the speaker and vowel length for the BEG group.*

more front than the German /yː/ [18]. This could lead to a perception of a tense vowel interpreted as /yː/ by German listeners if French speakers also produce [ʏ] more front.

The finding that French speakers shorten and lengthen vowels is somewhat different from the behavior of Italian learners of English, who have been found to only lengthen short vowels, but do not shorten long ones [8] (citing [19]). The data from French learners of German suggests that this learner group both lengthens and shortens vowels, with a tendency towards the behavior of Italian speakers, though. Further studies need to be carried out to investigate the reason for both this tendency as well as the difference between the speakers of the two languages. The finding that French speakers do not necessarily produce vowel contrasts which are part of their native language vowel system more successful than vowels which are not, may be connected. For instance, French has both /ø/ in words like *deux* 'two' and /œ/ in words like *neuf* 'nine'. However, the production seems to be not more successful than /iː/ and /ɪ/, where there is no contrast in French. One possible caveat of this study is the difference of context (both sentential as well as segmental) of the vowels. Studies have shown that context is a crucial factor (e.g. [7, 12]). Also, the acoustic differences of comparable phonemes could be important here (e.g. [18, 12]).

Furthermore, we have shown that increased proficiency in the L2 (as measured by the L2-Level, based on the CERF scale), leads to an increase in correct pronunciations. The patterns concerning the correct production of long and short vowels remained the same, however, for the two groups. In the results, speaker performance ranged between 53% and 97% (the former being BEG, the latter being ADV). The perception of the vowels of the French learners with the lowest number of correct words was at about chance level for the German speakers. However, as can be seen in Figure 3, there are ADV speakers that show problems with at least one vowel length (e.g., 09 or 23).

The results reported here can be seen as indicating that training of German vowels is a crucial task for French learners,

especially for BEG. However, there were also quite a number of ADV speakers who show some problems in the production of vowels. Therefore, this kind of training needs to be maintained at higher proficiency levels, depending on the ability of the individual learner. Training the pronunciation of vowels is not only important to reduce foreign accent, but it is crucial also for intelligibility. Considered from a broader psycholinguistic aspect there is evidence that depending on whether native speakers are familiar with mispronunciations from L2 speakers, they can hinder word recognition (e.g. [8]).

One possibility to tailor language training individually is to use computer-assisted language learning platforms. Depending on the level and knowledge of the learner, the program can select exercises for L2 training. If successful vowel production and perception is to be addressed, such platforms need to train both vowel quality as well as vowel quantity. If the software was to focus on the improvement of learners' productions, for instance, easy to understand feedback needs to be given to language learners (e.g. [20]).

Finally, this article raises further questions to be addressed. The relation between perception and production needs to be studied further. For instance, Strange and Bohn have shown that patterns of perceptual similarity patterns are not easily predicted by acoustic similarity [7, 12]. Therefore, it is crucial to know whether the problems in production can be traced back to perceptual problems on the learner side. Furthermore, the duration and formant cues have to be investigated in more depth: for instance, results of Strange and colleagues suggest that both cues may be used to a different extent in second language acquisition [7], or that in absence of spectral counterparts in an L1, listeners may pay more attention to duration cues [21].

5. Acknowledgements

6. References

[1] C. T. Best, "A direct realist view of cross-language speech perception," in *Cross-language studies of speech perception: A historical review*, W. Strange, Ed. York: Timonium, 1995, pp. 171–206.

[2] C. T. Best and M. D. Tyler, "Nonnative and second-language speech perception," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, M. J. Munro and O.-S. Bohn, Eds. Amsterdam: John Benjamins, 2007, pp. 13–34.

[3] I. Darcy and F. Krüger, "Vowel perception and production in Turkish children acquiring L2 German," *Journal of Phonetics*, vol. 40, pp. 568–581, 2012.

[4] J. E. Flege, M. J. Munro, and R. A. Fox, "Auditory and categorical effects on cross-language vowel perception," *Journal of the Acoustical Society of America*, vol. 95, no. 6, pp. 3623–3641, 1994.

[5] J. E. Flege and I. R. A. MacKay, "Perceiving vowels in a second language," *Studies in Second Language Acquisition*, vol. 26, pp. 1–34, 2004.

[6] J. Kingston, "Learning foreign vowels," *Language and Speech*, vol. 46, no. 2-3, pp. 295–349, 2003.

[7] W. Strange, O.-S. Bohn, S. A. Trent, and K. Nishi, "Acoustic and perceptual similarity of North German and American English vowels," *Journal of the Acoustical Society of America*, vol. 115, pp. 1791–1807, 2004.

[8] A. Weber, A. M. D. Betta, and J. M. McQueen, "Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners," *Journal of Phonetics*, vol. 46, pp. 34–51, 2014.

[9] R. Wiese, *The Phonology of German*. Oxford: Oxford University Press, 1996.

[10] P. Delattre, "Comparing the vocalic features of English, German, Spanish and French," *IRAL - International Review of Applied Linguistics in Language Teaching*, vol. 2, pp. 71–97, 1964.

[11] E. Pustka, *Einführung in die Phonetik und Phonologie des Französischen*. Berlin: Erich Schmidt Verlag, 2011.

[12] W. Strange, A. Weber, E. S. Levy, V. Shafiro, M. Hisagi, and K. Nishi, "Acoustic variability within and across German, French, and American English vowels: phonetic context effects." *The Journal of the Acoustical Society of America*, vol. 122, pp. 1111–1129, 2007.

[13] J. Trouvain, Y. Laprie, B. Möbius, A. Andreeva, A. Bonneau, V. Colotte, C. Fauth, D. Fohr, D. Jouvet, O. Mella, J. Jügler, and F. Zimmerer, "Designing a bilingual speech corpus for French and German language learners," in *Proc. Corpus et Outils en Linguistique, Langues et Parole: Statuts, Usages et Mésuages*, Strasbourg, 2013, pp. 32–34.

[14] C. Fauth, A. Bonneau, F. Zimmerer, J. Trouvain, B. Andreeva, V. Colotte, D. Fohr, D. Jouvet, J. Jügler, Y. Laprie, O. Mella, and B. Möbius, "Designing a bilingual speech corpus for French and German language learners: a two-step process," in *Proc. Conference on Language Resources and Evaluation (LREC) 2014*. Reykjavik, Iceland: European Language Resources Association (ELRA), 2014, pp. 1477–1482.

[15] V. Colotte, "JCorpusRecorder," Nancy, 2013.

[16] P. Boersma and D. Weenink, *PRAAT - Doing phonetics by computer*, 2013.

[17] SAS, *JMP*. Cary (NC): SAS Institute, 2014.

[18] W. Strange, A. Weber, E. Levy, V. Shafiro, and K. Nishi, "Within- and across-language acoustic variability of vowels spoken in different phonetic and prosodic contexts: American English, North German, and Parisian French," in *Proc. 144th Meeting of the Acoustical Society of America*. Acoustical Society of America, 2002, p. 2384.

[19] A. Duguid, "Italian speakers," in *Learner English: A Teacher's Guide to Interference and Other Problems*, M. Swan and B. Smith, Eds. Cambridge University Press, 2008, pp. 73–89.

[20] P. Carroll, J. Trouvain, and F. Zimmerer, "A visual feedback tool for German vowel production," in *Proc. 26. Konferenz Elektronische Sprachsignalverarbeitung 2015 (ESSV)*. Eichstädt, Germany, 2015.

[21] O.-S. Bohn, "Cross-language speech perception in adult: First language transfer doesn't tell it all," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, W. Strange, Ed. Timonium (MD): York Press, 1995, pp. 279–304.