Analysis of phone confusion matrices in a manually annotated French-German learner corpus

Denis Jouvet, Anne Bonneau, Jürgen Trouvain, Frank Zimmerer, Yves Laprie, Bernd Möbius

Analysis of phone confusion matrices in a manually annotated French-German learner corpus

Denis Jouvet[1], *Anne Bonneau*[1], *Jürgen Trouvain*[2], *Frank Zimmerer*[2], *Yves Laprie*[1], *Bernd Möbius*[2]

[1] Speech Group, LORIA
Inria, Villers-lès-Nancy, F-54600, France
Université de Lorraine, LORIA, UMR 7503, Villers-lès-Nancy, F-54600, France
CNRS, LORIA, UMR 7503, Villers-lès-Nancy, F-54600, France

[2] Computational Linguistics & Phonetics, Saarland University, Saarbrücken, Germany
denis.jouvet@inria.fr

Abstract

This paper presents an analysis of the non-native and native pronunciations observed in a phonetically annotated bilingual French-German corpus. After a forced-choice automatic annotation a large part of the corpus was checked and corrected manually on the phone level which allows a detailed comparison of the realized sounds with the expected sounds. The analysis is reported in terms of phone confusion matrices for selected error-prone classes of sounds. It revealed that German learners of French have most problems with obstruents in word-final position whereas French learners of German show complex interferences with the vowel contrasts for length and quality. Finally, the correct pronunciation rate of the sounds, for several phonetic classes, is analyzed with respect to the learner's level, and compared to native pronunciations. One outcome is that different sound classes show different correct rates over the proficiency levels. For the German data the frequently occurring syllabic [=n] is a prime indicator of the proficiency level.

Index Terms: Language learning, speech corpus, pronunciation variants, non-native speech

1. Introduction

The first language (L1) influences the learning of a target language (L2) on all linguistic levels including lexicon, morphosyntax, pragmatics, sound structure and its phonetic implementation (e.g. [1,8]). Whereas many studies and corpora related to foreign language learning involve the English language, the languages considered in the IFCASL project [9] are French and German.

The speech corpus was designed, using a two-step process [10], to focus on some phenomena of interest for the French/German language pair, covering segmental and prosodic levels as well as spelling problems. A non-exhaustive list of selected difficulties can be found in Table 1 (note that SAMPA notation [11] is used throughout this paper). The IFCASL corpus was recorded by French and German foreign language learners of various proficiency levels (see Table 2). The learners have also recorded data in their native language, which thus provide native French and native German speech data.

After a forced alignment procedure based on the transcriptions of the sentences, the phone level annotation of a large part of the recorded data was manually checked and corrected. Thus, the result of the manual re-annotation contains for each segment the expected sound as well as the realized sound. Using these detailed annotations, a study was conducted to investigate substitutions, insertions and deletions at the phone level. The paper focuses on this aspect, and does neither investigate other phenomena such as prosody, e.g. lexical stress, nor re-syllabification due to *liaison*.

It is also worth mentioning that the manual re-annotations were carried out by a rather high number of student annotators (n=14 student assistants with a phonetics background) and the amount of manually annotated data differ between languages and proficiency levels. Agreement between annotators [12] is not investigated in this paper; however results presented in the paper about the analysis with respect to the learner level are quite consistent.

Table 1. *Example of some expected difficulties.*

Phenomenon	Example French speaking German	Example German speaking French
nasal vowels	-	'bon' [O~] => [a~]
Schwa-n combinations	'hatten' [=n] => [9n]	-
vowel length and/or quality	'Polen' [o:] => [O]	'copain' [O] => [o]
plosives	'Paar' [p_h] => [p]	'pourboire' [b] => [p]
"ich-"sound	'Licht' [C] => [S]	-
final devoicing	'Nerv' [f] => [v]	'rouge' [Z] => [S]

The paper is organized as follows. Section 2 provides information on the bilingual speech corpus. Section 3 presents and comments confusion matrices for the French and German languages, for selected classes of sounds. Finally section 4 analyses the correct pronunciation rate per class of sounds with respect to the proficiency level, and a conclusion ends the paper.

2. Bilingual learner corpus

2.1. Speech corpus

The bilingual speech corpus was recorded by French learners of German and by German learners of French in their native and second languages. Hence, the corpus is made up of four sub-corpora: two native language sub-corpora (French reading French, and German reading German) which are considered here as native speech, and two non-native sub-corpora (French reading German, and German reading French) which represent the main part of analysis. At the current state of the corpus there were more than ninety speakers who pronounced about 60 French sentences and 50 German sentences and a short text.

Each non-native sub-corpus consists of four sets of sentences, corresponding to different speaking conditions: (1) reading sentences (about 30 sentences); (2) repeating sentences (about 30 sentences), (3) some sentences eliciting focus, and, (4) reading of a short text ("the three little pigs"). The subjects were seated in a quiet room and read the sentences from the screen of a Windows laptop (or repeated them, depending on the speaking condition), with a headset microphone (AKG C520) connected to an Audiobox (M Audio Fast track). For details see [10].

2.2. Annotation of pronunciations at the phone level

For the phonetic annotation we used the machine-readable SAMPA symbols [11]. For example, /Z/ is the SAMPA symbol for the voiced post-alveolar fricative. Also, some sounds are represented with multiple characters, e.g. the nasal vowels [U~/] as in French 'bain' vs. [o~] in 'bon'.

For facilitating the annotation process, an automatic speech-text alignment was first carried out for all the sentences. This was achieved using hidden Markov models.

Then, annotators checked and corrected the annotations for a part of the corpus. The convention used at the phone level for corrections explicitly indicates the expected phone as well as the realized phone. For example, "s-z" indicates that a phoneme [s] was expected, but that the realized sound is actually a [z]. This notation also allows insertions to be handled, as for example "-@" for the insertion of a schwa. Deletions are treated in a similar way, for example "@-" indicates the deletion of a schwa, but in such a case, the segment length is reduced to a very short duration (less than 1 ms). This is just an annotation convention that allows to keep track of deleted phone segments in the annotation file, without impacting significantly on the adjacent phone segment boundaries.

In the annotations, voicing and devoicing phenomena are mentioned, if relevant, using suffixes "_0" and "_V". For example "b_0" indicates a devoiced [b] sound, and "t_V" indicates a voiced [t] sound (see explanations in Section 3.1).

Table 2 indicates the amount of annotated data that are later used for computing the reported statistics. The table reports the breakdown per language and per proficiency level based on the European reference frame for language learning CEFR [13]: A for A1 and A2 (beginners), B for B1 and B2 (intermediate level), and C for C1 and C2 (advanced learners).

Table 2. *IFCASL corpus data, as used for computing the statistics on native and non-native pronunciations.*

	Language level	Non-native			Native
		A	B	C	
French	No. of speakers	15	13	12	52
	No. of sentences	734	641	696	2344
German	No. of speakers	24	16	16	40
	No. of sentences	636	404	463	183

3. Confusion matrices

The detailed manual annotations at the phone level are used to compute confusion matrices that show how expected phones are realized by the learners. The amount of sentences used for computing these statistics are indicated in Table 2, for each language and each proficiency level as well as for native speech. In this section, most of the confusion matrices are computed using "all levels", i.e. speech of all learners, which means using more than 2000 manually annotated sentences for German learners speaking French, and more than 1400 manually annotated sentences for French learners speaking German.

In all the reported confusion matrices, as for example in Table 5, the lines correspond to the expected sounds of the considered class of sounds. The columns correspond to the realized sounds, possibly taking into account the voicing suffixes (that indicate voicing or devoicing modifications). The numbers indicate the percentage of occurrences of the expected sound (line) that are realized in a given form (column) according to the manual annotation. For easier reading, a "*" replaces a value lower than 1.0% and a dot '.' replaces a zero value. Only expected sounds occurring more than 50 times are reported. The second column reports the number of occurrences of the expected sound (in square brackets). In some cases we summarized the confusion matrices in tables focusing only on the correct pronunciation rates.

3.1. Annotations of obstruents

In the corpus, a series of sentences has been devoted to the [voice] feature. From now on, we will use quotes ("") when the terms *voiced* or *voiceless* are related to phonemic categories distinguished by the [voice] feature, and no quote when these terms are related to the articulatory phenomenon (vocal fold vibration). There are two major differences between German and French systems with respect to the [voice] feature. The first one is phonological and concerns final devoicing in German: in this language, the opposition between "voiced" and "voiceless" obstruents (fricatives and stops) is neutralized in final position in favor of the realization of "voiceless" categories, whereas in French this feature is kept distinctive in final position. This difference between both systems is known to be a source of error for German speakers, who tend to produce "voiceless" obstruents in final position when speaking French instead of the expected "voiced" consonants.

The second difference between French and German is related to the phonetic implementation of the [voice] feature for stop consonants. To be short, the presence vs. absence of voicing due to vocal fold vibration is an important cue (not the only one) in the distinction of French "voiced" vs. "voiceless" stops, whereas the absence vs. presence of aspiration is an important cue for the same distinction in German. Voicing during closure is not mandatory for German "voiced" stops, French "voiceless" stops are not aspirated. Hence, German speakers might realize the closure of French "voiced" stops without glottal buzz, whereas French speakers tend to realize German "voiceless" stops without aspiration.

Both phenomena, the absence of (expected) periodicity during stop closure, and the absence of (expected) periodicity during the production of an obstruent in final position, have been indicated at the phonetic level by a "_0" code added at the end of the expected segment. The code "_V" indicates the presence of voicing during "voiceless" consonants as well as voicing for "voiced" German plosives.

3.2. German learners speaking French

Regarding German speakers of French (as L2), Table 3 shows the correct pronunciation rate and the percentage of devoiced sounds for French fricative consonants, and Table 4 for the stop consonants, when the fricative or the stop consonant occurs in any position (left-hand side) or in word-final position (right-hand side). The results confirm the influence of L1 (German) on L2 (French), since about 20% of the "voiced" consonants have been incorrectly devoiced by German speakers. Note that we can also observe, in Table 9, that some French "voiced" consonants pronounced by French native speakers were considered as voiceless by annotators. This is probably partly due to (1) assimilation processes and (2) aerodynamics (that also works to explain differences between places of articulation for German speakers).

Table 3 shows the correct pronunciation of the **fricatives** of the German learners. As expected there is no problem at all for the voiceless fricatives but there is indeed a problem for the voiced fricatives which can be mainly explained with the final devoicing in word-final position (right part of Table 3). It is interesting to see that final devoicing does not happen all the time but between 18% and 49%. In addition the three voiced fricatives behave differently, with [z] being much more often annotated as voiced than [v]. As can be seen in Table 4 the **plosives** follow a similar pattern for the final devoicing.

Table 3. *Correctness rates for fricatives in word-final position (right) and in all positions (left) for German learners of French (all levels); _0 indicates devoicing.*

all positions				word-final position			
	nb. occ.	**corr.**	**_0**		**nb. occ.**	**corr.**	**_0**
v	[1255]	87	11	v	[206]	44	49
z	[1146]	79	20	z	[905]	80	18
Z	[498]	69	30	Z	[249]	53	45
f	[595]	99		f	[<50]	---	
s	[2215]	98		s	[602]	99	
S	[536]	99		S	[73]	99	

Table 4. *Correctness rates for plosives (closure phase) in word-final position (right) and in all positions (left) for German learners of French (all levels); _0 indicates devoicing.*

all positions				word-final position			
	[c]	**corr**	**_0**		**[c]**	**corr**	**_0**
b	[1519]	88	11	b	[96]	72	28
d	[1799]	84	15	d	[168]	76	23
g	[584]	69	30	g	[65]	57	43
p	[1569]	99		p	[<50]	---	
t	[1945]	93		t	[383]	90	
k	[1749]	97		k	[390]	97	

Indeed, voicing is due to the vibration of vocal folds during the production of "voiced" stops and fricatives, and disappears when the intra-oral pressure becomes too high with respect to the subglottal pressure. The results are thus coherent with this explanation since, as we can observe, consonants whose place of articulation is closer to the glottis such as /g, Z/ (hence for which the intra-oral pressure tends to be higher than that of other points of articulation) received the higher number of "_0" codes.

It should be noted that the annotations use two segments for each stop consonant: one for the closure part, and another one for the release part. To make the display more readable, Table 4 reports only the correctness rates and the percentage of devoiced sounds as annotated on the closure part. A detailed analysis shows that the closure and the release part of the French plosives in word final position have a rather similar behavior with respect to voicing/devoicing annotation. The main difference between the closure and release part concerns deletions, a few percent more deletions are observed for the release of unvoiced plosives, than for the deletion of the closure part.

French **oral vowels** were corrected in the annotations only if a noticeable error was observed. For example, differences between mid-open and mid-close vowels were, in general, not corrected. Hence confusion matrices on French vowels are not relevant.

The **nasal vowels** were better matched than expected by the German learners. Although the results are slightly worse than the oral vowels, we find a general performance larger than 90% which can be considered as less problematic for the learners.

3.3. French learners speaking German

For the **fricatives** of French learners we concentrate on the word-final coda position, because it shows the same tendencies as the overall fricative productions. As can be seen in Table 5 the *"ich-"sound* [C] clearly distinguishes from the other voiceless fricatives with a rather low correct rate of 56%. As expected [S] was most often used as substitute for [C]. In contrast the *ach-sound* [x] which does not exist in French either reached a fairly high correct rate of 94%.

Table 5. *Confusion matrix on fricatives in word-final coda position for French learners speaking German (all levels).*

	f	v	s	z	S	Z	C	x
f [150]	**85**	13	.	.	.	.	.	.
s [568]	*	.	**87**	8	.	.	.	.
C [281]	.	.	.	.	20	8	**56**	1
x [134]	.	.	*	.	.	.	*	**94**

The behavior of the closure and release parts or **plosives** in word-final position (Table 6) is not so similar to each other, contrary to the French data. Quite a few deletions are observed on the release part.

Table 6. *Confusion matrix on plosives (closure and release parts) in word-final position for French learners speaking German (all levels).*

	t	t_	d	d_	k	k_	g	g_	Del
t [1460]	**92**	.	2	.	*	.	*	.	2
t_ [1815]	.	**75**	*	2	.	*	.	*	19
k [190]	1	.	.	.	**81**	.	18	.	.
k_ [207]	.	*	.	.	.	**74**	.	17	8

Regarding the syllabic consonants (Table 7) the French learners do not delete schwa as the expected pattern for German would predict. Only 32% follow the expected German pattern whereas 52% keep the schwa. Note however that a missing schwa deletion does not lead to a wrong pronunciation, rather to a tendency to hyperarticulate - in contrast to the realization

Table 8. *Confusion matrix on vowels for French learners speaking German (all levels).*

E\R	[c]	i:	e:	E:	a:	o:	u:	y:	2:	e	i	o	I	E	a	O	U	Y	9	2	6	@	u	y
i:	[938]	**85**	*	.	.	*	.	.	.	.	4	.	9	.	.	.	.	.	.	.	.	*	.	.
e:	[560]	4	**83**	2	.	.	.	.	.	*	*	.	4	*	*	.	.	.	*	.	*	*	.	.
E:	[75]	.	24	**68**	.	.	.	.	.	.	.	.	1	.	1	.	.	.	.	.	1	.	.	.
a:	[514]	.	*	.	**83**	.	.	.	.	.	.	.	.	.	16	*	.	.	.	.	.	.	.	.
o:	[347]	.	.	.	.	**82**	1	.	*	.	.	9	.	.	.	5	.	.	*	*	.	.	*	.
u:	[395]	.	.	.	.	*	**70**	*	.	.	.	1	.	.	.	*	18	6	.	.	.	*	3	*
y:	[140]	.	.	.	.	.	4	**83**	4	.	.	.	.	.	.	*	1	2	1	.	.	.	.	2
2:	[128]	.	.	.	.	3	.	.	**89**	.	.	2	.	.	.	3	.	.	2	*	.	.	.	.
o	[79]	.	.	.	.	14	.	.	.	.	.	**77**	.	.	.	6	.	.	.	.	.	.	1	.
I	[2049]	5	*	.	.	.	.	*	.	.	5	.	**85**	.	.	.	.	*	*	.	.	*	.	.
E	[106]	3	20	27	2	.	.	.	.	.	.	.	12	**2**	4	2	.	.	6	.	4	10	.	.
a	[1377]	.	.	*	3	.	.	.	.	.	*	*	*	.	**91**	*	.	.	*	.	1	*	.	.
O	[270]	.	.	.	.	6	.	.	.	.	.	9	.	.	*	**72**	4	.	2	*	1	.	1	.
U	[260]	.	.	.	.	*	3	.	.	.	.	*	.	.	.	*	**80**	2	.	.	.	.	*	*
Y	[78]	.	.	.	.	.	.	14	.	.	.	.	6	.	.	.	3	**63**	.	.	.	.	.	10
9	[84]	.	.	.	.	1	.	.	10	.	.	.	1	.	.	5	.	12	**57**	14	.	.	.	.

Table 9. *Confusion matrix on French fricatives, in word-final position, with respect to proficiency level.*

A level

		f	v	v_0	s	z	z_0	S	Z	Z_0	Z_V
v	[57]	11	**46**	44	.	.	.	.	.	.	.
s	[225]	.	.	.	**99**	*	.	.	.	.	.
z	[292]	.	.	.	*	**76**	20	.	.	.	.
Z	[88]	.	.	.	.	.	.	1	**45**	52	1

B level

		f	f_V	v	v_0	s	s_V	z	z_0	S	Z	Z_0
v	[69]	6	.	**39**	55	.	.	.	.	.	.	.
s	[182]	.	.	.	.	**98**	2	.	.	.	.	.
z	[302]	.	.	.	.	.	.	**79**	19	.	.	.
Z	[76]	.	.	.	.	.	.	.	.	1	**50**	47

C level

		f	v	v_0	s	z	z_0	S	Z	Z_0
v	[80]	6	**48**	46	.	.	.	.	.	.
s	[195]	.	.	.	**99**	.	.	.	.	.
z	[311]	.	.	.	*	**85**	14	.	.	.
Z	[85]	.	.	.	.	.	.	.	**64**	36

Native

		f	f_V	v	v_0	s	s_V	z	z_0	S	Z	Z_0
f	[54]	**94**	4	2	.	.	.	.	.	.	.	.
v	[237]	.	.	**76**	23	.	.	.	.	.	.	.
s	[645]	.	.	.	.	**96**	2	.	.	.	.	.
z	[1066]	.	.	.	.	.	.	**97**	3	.	*	.
S	[91]	.	.	.	.	.	.	.	.	**100**	.	.
Z	[287]	.	.	.	.	.	.	.	.	.	**79**	21

Table 10. *Confusion matrix on German fricatives, in word final position, with respect to proficiency level.*

A level

		f	v	s	z	S	Z	C	x
f	[65]	**82**	14	.	.	.	.	.	.
s	[240]	.	.	**80**	12	.	.	.	.
C	[114]	.	.	.	.	19	12	**43**	3
x	[57]	.	.	2	.	.	.	2	**95**

B level

		f	v	s	z	S	Z	C	x
s	[143]	2	.	**86**	8	.	.	.	.
C	[76]	.	.	.	.	22	11	**50**	1

C level

		f	v	s	z	S	Z	C
f	[54]	**87**	13	.	.	.	.	.
s	[185]	.	.	**96**	3	.	.	.
C	[91]	.	.	.	.	18	1	**78**

Native

		s	z
s	[92]	**98**	2

of [E] instead of [@] which happened in 7% of the cases, and [9] for [@] in 3% of the cases.

Table 7. *Confusion matrix on syllabic consonants for French learners speaking German (all levels).*

		=n	'9 n'	'@ n'	'E n'
=n	[683]	32	3	52	7

The **vowels** provide a rather homogenous picture (see Table 8). As expected the long tense vowels [i:, e:, a., o:, u:, y: 2:] were often substituted with their short lax counterparts. An exception represents [y:] which was also replaced with [u:] and [2:] probably due to misinterpretation of the spelling-pronunciation rules. An even larger exception is [E:] which is only marginally substituted with its short counterpart [E] but tremendously with [e:]. This mismatch gives support to a wrongly applied or even unknown distinction between [E:] and [e:] which should be treated carefully in pronunciation programs. This finding is also mirrored in the annotations of the short [E]: only in 2% it was correct, with [e:] and [E:] as the main substitutions. The deviations of most short vowels show a broad diversity in vowel qualities.

Regarding the **diphthongs** in German (Table 11) the French learners are doing well except for the au-vowel which often was interpreted according a French spelling-pronunciation correspondence that led to [O, o, o:].

Table 11. *Confusion matrix on diphthongs for French learners speaking German (all levels).*

		aI	aU	O	a	o	o:
aI	[1036]	**96**	.	.	1	*	.
aU	[462]	.	**84**	10	2	1	1

4. Impact of proficiency level

This section focuses on the relation between the proficiency level and the mispronunciation errors that are observed on the data.

4.1. Example of confusion matrices

Table 9 and Table 10 display the evolution of learner's mispronunciation with respect to the level proficiency. For comparison, results are also provided using annotation of native pronunciations.

In the matrices, the bold term corresponds to the correct pronunciation of the expected sounds (here fricatives). The number of mispronunciation errors gets smaller as the learner's level increases.

4.2. Correct pronunciation vs. proficiency level

Table 12 shows the correct pronunciation rates for various classes of sounds, on French data, and Table 13 reports similar results for German data. Fricatives and plosives are based on the voicing features (in French). For reasons of completeness, nasal consonants, liquids and glides were also listed in the table, but they are not subject of discussion here.

Table 12. *Correct pronunciation rate of sounds (average over phones occurring more than 50 times in annotated data), for French*.

Class of sounds	**A**	**B**	**C**	**Native**
Fricatives *All positions*	86.6%	87.4%	91.2%	96.1%
Fricatives *Word-final pos.*	66.6%	66.6%	73.9%	90.4%
Plosives *All positions*	85.5%	90.7%	89.8%	97.9%
Plosives *Word-final pos.*	82.4%	86.3%	89.4%	96.6%
Nasal consonants	98.1%	99.4%	99.9%	99.8%
Liquids and glides	96.1%	98.5%	98.5%	99.1%
Nasal vowels	90.7%	94.3%	95.8%	99.7%

It can be observed for the French data that the main evolution of correct pronunciation rate took place for fricatives and plosives, and to a lesser extent, to nasal vowels.

Table 13. *Correct pronunciation of sounds (average over phones occurring more than 50 times in annotated data), for German*.

Class of sounds	**A**	**B**	**C**	**Native**
Fricatives	68.0%	74.7%	84.2%	84.8%
Plosives	82.8%	82.3%	83.0%	84.4%
Nasal consonants	82.1%	82.7%	85.0%	80.7%
Liquids and glides	80.4%	85.6%	88.6%	94.7%
Syllabic consonants	7.7%	28.7%	64.2%	83.1%
Tense long vowels	78.5%	77.8%	85.7%	85.1%
Short lax vowels	80.9%	78.0%	85.5%	95.9%

As expected, there is a general increase in the correctness rate for the various sound classes from beginners up to native speakers in both languages. For German (Table 13), this trend is best visible with the syllabic consonants starting at 7.7% for the beginners (A) with a rapid improvement to the learners at the intermediate level (B) up to the advanced learners (C) who not yet reach the degree of schwa deletion as the native speakers. Thus this represents an excellent example for the various proficiency levels when mastering the foreign language (here German) that could be used with the phone confusion matrices. Here, we see an optimal point to teach learners to reach faster the next level of proficiency.

5. Conclusions

One of the main outcomes of this analysis of phone confusion matrices is how much and how different French learners have trouble with the German vowel system. This might be surprising at first glance, because both German with 16 monophthongs and French with 11 monophthongs have a rather large vowel system (e.g., [14], [15]). It was expected that the long-short contrast lead to larger interferences. This general trend is valid, however, a more differentiating view helps with a better targeted support of the learners, e.g. when selecting individualized exercises in computer-assisted language learning. From a phonetical point of view, the contrasts between [a:] and [a], and [E:] and [E] rely almost purely on length whereas all other "length contrasts" include also contrasts of vowel quality. However, the [a]-vowels were comparably well mastered in contrast to the [E]-vowels with extremely low degrees of correctness. In a perception test with native speakers with a sub-set of the corpus data containing minimal-pair words [16] the problem emerged as well, however not in such an extreme way. In [16] we found that rounded vowels, particularly the [o:]-[O]-contrast, represent the main source of trouble. An advantage of the differences regarding the magnitudes of the learners' problems in both studies is that the different methods make different problematic areas visible, so that they complement each other. Concerning the problem mastering the three-way contrast between [e:] - [E:] - [E] in German, (as can be found in words like *"stehlen/Stelen"* – *"stählen"* – *"Stellen/Ställen"*) is rather dramatic. The difference between the German and the French vowel system is the correct use and production of vowel length and vowel quality in the correct combination, which is most crucial in this contrast. Because French speakers usually (as with a/a:) only have to focus on one aspect, their problems in correctly acquiring the vowel contrasts are likely most visible in this contrast. Thus, they can be traced back to the different set-up of the vowel systems of the two languages, but phonological systems have to be analyzed as a whole.

For German learners of French the vowels do not cause bigger problems (including nasal vowels) – in contrast to some of the investigated consonants. An interesting pattern is the problem of German speakers concerning the incorrect application of the final devoicing rule in the French productions. This result indicates that it is not only the set-up of phonological systems, including the use of (phonological) features of sounds that can lead to interferences, but also the overall patterns (or phonological rules) occurring in the L1.

6. Acknowledgements

This work has been supported by an ANR/DFG Grant "IFCASL" to the Speech Group LORIA CNRS UMR 7503 – Nancy France, and to the Phonetics Group, Saarland University – Saarbrücken Germany, 2013 – 2016.

7. References

[1] J. E. Flege and R. Davidian, "Transfer and developmental processes in adult foreign language production", *Journal of Applied Psycholinguistic Research* 5, pp. 323-347, 1984.

[2] J. E. Flege, "Second Language Speech Learning: Theory, Findings and Problems", in: *Strange, W. (ed). Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research,* pp. 233-272, 1995.

[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] 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, pp. 171–206, 1995.

[5] 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, pp. 13–34, 2007.

[6] 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.

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

[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] IFCASL project: http://www.ifcasl.org/

[10] 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 *LREC'2014, 9th Language Resources and Evaluation Conference, Reykjavik Iceland*, 2014.

[11] Sampa: http://www.phon.ucl.ac.uk/home/sampa/

[12] O. Mella, D. Fohr, and A. Bonneau, "Inter-annotator agreement for a speech corpus pronounced by French and German language learners", *Slate 2015, to appear.*

[13] Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR), http://www.coe.int/t/dg4/linguistic/cadre1_en.asp.

[14] R. Wiese, "*The Phonology of German*". Oxford et al.: Oxford University Press, 1996.

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

[16] F. Zimmerer and J. Trouvain. Perception of French speakers' German vowels. *Interspeech 2015, to appear*.