The impact of accent familiarity on the perception of difficult sound contrasts for German learners of English

Katrin Wolfswinkler[1], Eva Reinisch[1]

[1]Institute of Phonetics and Speech Processing, LMU Munich

k.wolfswinkler@yahoo.de, evarei@phonetik.uni-muenchen.de

Abstract

Second language (L2) learners usually retain a foreign accent even after years of training. The present study aimed at finding reasons for why accents are so persistent. We hypothesized that familiarity with an accent due to frequent exposure leads to adaptation which in turn allows listeners to understand the accent better, reducing the need for improvement. If this was the case, L2 learners should be better at understanding words spoken in a familiar than unfamiliar accent. To test this account, English minimal pairs containing two difficult sound contrasts for German learners (/ɛ/ vs. /æ/; voiceless vs. voiced word-final stops) were presented to native German listeners who had to identify the intended word. The tokens were produced by native speakers of English who served as a control, fellow Germans whose accent was highly familiar, Italians whose accent was somewhat familiar, and Finnish whose accent was unfamiliar. Results showed that words spoken by native English speakers or fellow Germans were recognized best, but words with the unfamiliar Finnish accent were recognized better than Italian accented words. A closer analysis of the acoustic differences that our speakers produced suggests that the acoustic cues available in the speech signal determined how well the words were identified.

Index Terms: Second language acquisition, Speech perception, Accent, Intelligibility, Familiarity, Foreign accent, Accent Familiarity

1. Introduction

Learners of a second language (L2) mostly retain an accent in L2 production even if they have long lasting experience with the L2. An obstacle in overcoming an accent appears to be that L2 learners perceive L2 sounds through a 'grid' of their first language (L1) sound inventory [1],[2]. That is, the sounds of the L2 tend to get assimilated to the perceptually closest L1 sound, affecting the correct perception – and consequently also the production – of L2 sounds that do not exist in the L1. Critically, if two different L2 sounds get assimilated to one single L1 category, L2 sound contrasts tend to get neutralized in perception as well as in production.

Although it has been shown that over time L2 learners get better at perceiving and producing difficult L2 sound contrasts, the ultimate attainment appears variable. One reason for this may be the quality of the input [3]. If the L2 is learned in the learners' native L1 environment, the main input is likely to come from other non-native speakers. This may include fellow learners at school and sometimes even the teachers. Therefore L2 learners usually have ample experience with listening to their own accent and hence are highly familiar with this accent. The purpose of the present study was to further investigate the role of non-native input and specifically, the role that accent familiarity plays in second language word recognition.

There are two lines of previous research that speak to this issue. First, native speakers of the target language have been shown to rapidly adapt to foreign accents. Even after brief exposure to an L2-speaker, native listeners of that L2 become better in understanding this learner's productions [4],[5].

Second, L2 learners have been shown to better understand an L2 when spoken by other learners of the same L1 background than by speakers of other L1 backgrounds. Sometimes L2 learners are even better at understanding L2 speakers with the same native language than native speakers of the target language. This has been called the '*interlanguage speech intelligibility benefit*' [6]. It is explained by the shared phonetic and phonological knowledge of the L1. Since L1 and L2 interact at segmental, phonotactic and prosodic levels this shared knowledge helps learners understand the accent.

The present study builds on these findings and asks whether, in addition to the shared L1 phonetic inventory, familiarity with an accent could affect L2 learners' performance in recognizing L2 words.

To test this issue, the present study examined the recognition of English words from minimal pairs where the contrasting sounds have been shown to be difficult to distinguish for native German learners of English [7],[8]. These were the vowels /ɛ - /æ/ and word-final voiced vs. voiceless stops. To investigate the influence of accent familiarity, the English minimal pairs were spoken by talkers of four different native languages whose accents differed in how frequently Germans were likely exposed to these accents. A native English accent (US) served as a control. German was the accent shared with the listeners hence frequently heard and highly familiar. An Italian accent was supposed to be somewhat familiar and Finnish was likely an unfamiliar accent.

Native German listeners were then exposed to these productions and had to identify which word of the minimal pair has been produced. We hypothesized that if familiarity with an accent affects how well this accent is understood, then listeners should be better able at understanding words spoken in a familiar than unfamiliar accent and rate this accent as less severe.

Specifically we tested four hypotheses:

H1: English words produced by native English speakers will be understood best since the words of the minimal pairs will be well differentiated.

H2: English words produced by native German speakers will be understood as well as the native English speakers' productions or second best since German is the listeners' own and hence most familiar accent. Listeners will be used to the specific cues that Germans tend to use for differentiating the words of the minimal pairs.

H3: English words produced by Italian speakers will be well understood, possibly as well as the Germans, since listeners are likely familiar with the accent.

H4: English words produced by Finnish speakers whose accent is largely unfamiliar will be understood significantly worse than the other accents (as familiarity cannot aid perception).

2. Method

2.1. Participants

Nineteen native German listeners (8 male) who were students at LMU in Munich participated. They were between 19 and 30 years old, were not enrolled in any Language and Literature class, and had not spent more than six months in an English speaking country. All had studied English for an average of nine years at school. According to self-report questionnaires administered after the experiment, all participants considered themselves as medium proficient in speaking and understanding spoken English.

2.2. Material

Twenty-two minimal word pairs were selected that contained either the vowel contrast /ɛ– æ/ (e.g., men-man, pen-pan, bed-bad,...; 11 pairs) or a voicing contrast in word-final stops (e.g., bag-back, wide-white, robe-rope,...; 11 pairs). Both contrasts have been shown to be difficult to distinguish for German learners of English in perception and production [7],[8].

All words were recorded by speakers of four different native languages (American English, German, Italian, and Finnish). For the recordings the words were presented at the end of semantically neutral carrier sentences (e.g., 'The next word is...') and speakers were asked to read the sentences at a comfortable pace. The list of words was randomized such that words of a minimal pair would not appear in direct succession. Native speakers of English and German had been recorded for another experiment and tokens for the present experiment were selected from this larger corpus [9]. Three native speakers of each, Italian and Finnish were recorded specifically for the present study. Two speakers per language were selected for the experiment. All speakers were female and between 18 and 30 years of age.

To assess how the speakers of the different accents produced the critical words with respect to the critical sound contrasts, acoustic measures were taken using Praat [10]. For the words with the vowel as target sounds the length as well as the first two formants of the vowel were measured. For the voicing contrast in word-final stops we measured duration of the preceding vowel and aspiration duration.

A summary of the measures is given in Table 1. The acoustic analyses showed that for the vowels, Italian and German learners produced smaller contrasts between the words of the minimal pairs than the native English or Finnish speakers, although they did not entirely neutralize the contrasts. The Finnish speakers differentiated the vowels clearly but did not produce the same vowel qualities as the native English speakers. For the stops, the German speakers produced large differences, close to the native English speakers. Italians and Finns produced smaller differences and their productions were more variable than the native English speakers' productions.

2.3. Design and Procedure

From the set of recorded words, eight minimal pairs with the vowel contrast and six pairs with word-final stops were selected for the perception experiment. The selection was based on the quality of the recordings and consistency of acoustic characteristics within the two speakers of a given accent.

For each speaker, four carrier sentences were selected and randomly combined with all words produced by this speaker. Carriers were selected such that they did not contain any hesitations and did not contain any of the critical sounds that listeners might use as an anchor for identifying the critical sound in the target word.

Listeners were seated in a sound-proof booth in front of a notebook computer wearing high-quality headphones. On each trial, listeners were presented visually the two words of a minimal pair. After 800ms preview the sentence was presented at a comfortable listening level. The listeners' task was to decide by button press (0 and 1 key on the computer keyboard with sides matched to the visual layout) which of the two words had been said. After logging their response the next trial started automatically. Each participant heard all words from all eight speakers once, resulting in a total of 224 unique trials (28 words x 8 speakers). Words and speakers were intermixed and presented in a separate random order for each participant.

In the same session after the identification test, participants were given a short accent rating task to assess the perceived strength of the speakers' accents. They were presented the carrier sentences without target words and asked to rate on a scale from 1 to 5 how strong they considered this speakers' accent. Endpoints and labels of the scale ('nearly no accent' – 'very strong accent') were counterbalanced across participants. The four sentences of the eight speakers were repeated three times in fully random order.

Table 1: *Means and Standard Deviations (in brackets) of the acoustical measures. Durations in milliseconds.*

	Vowels				Stops			
	/æ/		/ɛ/		voiced		unvoiced	
	F2-F1	**duration**	**F2-F1**	**duration**	**aspiration duration**	**vowel duration**	**aspiration duration**	**vowel duration**
English	1067 (422)	230 (60)	1266 (256)	160 (20)	30 (20)	230 (50)	80 (20)	150 (30)
German	1103 (176)	200 (60)	1198 (158)	170 (60)	200 (10)	220 (70)	50 (20)	160 (40)
Italian	975 (232)	180 (30)	1038 (276)	180 (50)	70 (20)	260 (60)	90 (40)	220 (70)
Finish	719 (119)	160 (40)	1380 (242)	140 (30)	50 (20)	230 (70)	60 (30)	200 (80)

3. Results

For the identification task, a generalized linear-mixed effects model was fitted with response (the correct/intended word coded as 1, the incorrect as 0) as the dichotomous dependent variable for which a logistic linking function was used. Fixed factors were Accent (English, German, Italian and Finnish), sound Contrast (/ɛ- æ/, voiced vs. unvoiced word-final stops) and their interaction. For the factor Accent the Level 'German' was mapped onto the intercept. The factor Contrast was added to the model since the acoustic measures of the productions had shown differences in how well speakers of the different accents had differentiated the words of the minimal pairs. The level 'Vowels' was mapped onto the intercept. The model was fitted with a full random-effects structure [11]. Figure 1 shows the results that were confirmed by the statistical analyses which are reported in Table 1.

Table 2: *Results of the mixed-effects model.*

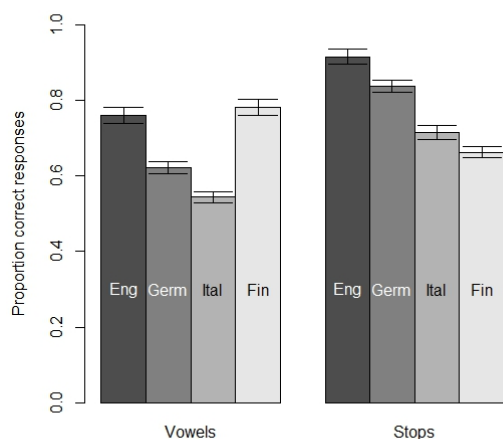|	Fixed effect	*b*	*z*	*p*
Vowels	Intercept (German)	0.62	2.77	¡.01
	Accent (English)	0.72	2.79	¡.01
	Accent (Italian)	-0.42	-1.41	.16
	Accent (Finnish)	0.72	2.98	¡.005
Stops	Contrast	1.25	3.57	¡.001
	Contrast:Accent (English)	0.06	0.16	.88
	Contrast:Accent (Italian)	-0.41	-0.87	.38
	Contrast:Accent (Finnish)	-1.89	-5.05	¡.001

Figure 1: Proportion of correct word identifications dependent on accent for each sound contrast. Error bars indicate one standard error.

Results showed that for the vowels listeners were able to identify the correct word above chance if spoken with a German accent (see Table 2, Intercept). However, they were significantly better at identifying words when spoken by native En-

glish speakers or speakers with the Finnish accent. Words produced by speakers with the Italian accent were identified significantly worse.

For the stops listeners were overall better than for the vowels when produced by the Germans (see Table 1, Contrast). The same was the case for the productions of the English as well as the Italian speakers as evident in Figure 1. However, the stops of the native English speakers were identified slightly better compared to the German accented words while the stops produced by the Italian speakers where identified slightly worse but neither difference was significant. Just the difference to the Finnish speakers was significant as their consonant contrast was discriminated worst of all.
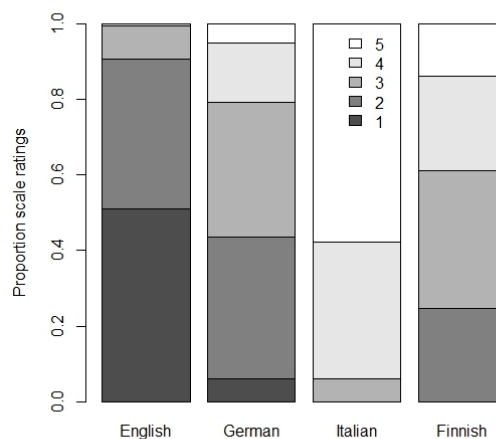

Figure 2: Results of the accent rating with the scale adjusted such that 1 = 'nearly no accent' and 5 = 'very strong accent'.

The results of the accent ratings which are presented in Figure 2 confirm the overall tendencies of the identification task (though note that for accent ratings the carrier sentences were presented without the target words). The native English speakers were rated the best suggesting that they were mostly recognized as native speakers. In the identification task their productions were understood best. German accented speakers were rated to have the next weakest accents which again matched the overall results from word identification. The accent ratings for the Finnish speakers was rated slightly stronger than the German accent, and the Italian speakers were rated as having the stronges accent.

4. Discussion

The present study was based on the observation that learners of a second language are frequently exposed not only to native speakers of the target language but also to fellow learners with either the same native language or different native languages. We addressed whether familiarity with the accent of fellow learners would help recognizing accented words in the L2 thereby assessing the boundaries of the interlanguage intelligibility benefit (ISI; [6]). The ISI suggested that shared phonetic knowledge allows learners to understand fellow learners with the same L1 (almost) as well as native speakers of the tar-

get language. We aimed at replicating this finding and extending it to other familiar vs. unfamiliar accents. Results from a word identification task with four different accents showed that German learners of English were best at identifying words from English minimal pairs when either spoken by native speakers of English or fellow German learners. This suggests that for listeners of medium proficiency native speech is well intelligible and that indeed shared L1 phonetics and frequent exposure to the accent facilitate word recognition (as was the case for the German accent). In other words, our results replicate the ISI and hypotheses 1 and 2 were confirmed with regard to the accents that were understood best. However, German accented words were identified worse than the native English tokens. Interestingly, our predictions about the other two accents were not borne out (i.e., pertaining to H3 and H4). That is, even though Germans could be considered familiar with Italian accent (e.g., Italy is a popular holiday destination and many Italians visit Munich) the Italian speakers' productions were understood less accurately than the other speakers' productions, even the Finnish speakers' productions whose accent was entirely unfamiliar.

However, our results can be explained by closer inspection of the acoustic differences that the speakers produced for the minimal pairs. For the vowels, Finnish speakers produced a large difference between the words of the minimal pairs. Therefore, even though the quality of the vowels differed from the native English speakers' productions, the intended words could be well identified – even better than the Germans' productions. The Italians in contrast, produced smaller differences between the vowels and were hence understood less accurately. For the stops, the picture looked somewhat different. Here, the Italian speakers produced larger differences than the Finnish speakers and words were hence identified better (details about the acoustic measures are reported in Table 1).

From a theoretical perspective this precedence of acoustic differences over accent familiarity may not be surprising, however, a hint of a role for familiarity can be seen when comparing the accent ratings. Although the Finnish speakers tended to produce the vowel contrast better than the German speakers, their overall accent (regardless of the any target words) was considered stronger. An additional experiment with better match of the speakers in overall accent ratings and the magnitude of differences they produce will shed further light on the issue of familiarity. With regard to the interlanguage intelligibility benefit we can conclude that L2 learners are best at understanding clearly produced words as found in native speech. In addition, the shared L1 phonetic inventory of fellow learners with the same L1 background, as well as the frequent exposure to the 'own' accent at school may make German listeners proficient at understanding the German accent. All these factors may contribute to the problem that for well understood speech there is little need for improvement which in turn could block learners from losing their accent even after years of learning.

5. Acknowledgements

6. Bibliography

[1] J. E. Flege, "Second language speech learning: Theory, findings, and problems," *Speech perception and linguistic experience: Issues in cross-language research*, pp. 233–277, 1995.

[2] J. E. Flege, G. H. Yeni-Komshian, and S. Liu, "Age constraints on second-language acquisition," *Journal of memory and language*, vol. 41, no. 1, pp. 78–104, 1999.

[3] T. Piske, I. R. MacKay, and J. E. Flege, "Factors affecting degree of foreign accent in an l2: A review," *Journal of phonetics*, vol. 29, no. 2, pp. 191–215, 2001.

[4] A. R. Bradlow and T. Bent, "Perceptual adaptation to non-native speech," *Cognition*, vol. 106, no. 2, pp. 707–729, 2008.

[5] M. J. Witteman, A. Weber, and J. M. McQueen, "Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation," *Attention, Perception, & Psychophysics*, vol. 75, no. 3, pp. 537–556, 2013.

[6] T. Bent and A. R. Bradlow, "The interlanguage speech intelligibility benefit," *The Journal of the Acoustical Society of America*, vol. 114, no. 3, pp. 1600–1610, 2003.

[7] M. Broersma and A. Cutler, "Competition dynamics of second-language listening," *The Quarterly Journal of Experimental Psychology*, vol. 64, no. 1, pp. 74–95, 2011.

[8] B. L. Smith, R. Hayes-Harb, M. Bruss, and A. Harker, "Production and perception of voicing and devoicing in similar german and english word pairs by native speakers of german," *Journal of Phonetics*, vol. 37, no. 3, pp. 257–275, 2009.

[9] N. Eger and E. Reinisch, "Is foreign-accented speech easier to understand if it is produced in one's own voice?" how Words Emerge and Dissolve: Evidence from Speech Production, Speech Perception, Acquisition and Disorders. [Online]. Available: http://www.phonetik.uni-muenchen.de/institut/veranstaltungen/how-words-emerge-and-dissolve

[10] P. Boersma *et al.*, "Praat, a system for doing phonetics by computer," *Glot international*, vol. 5, no. 9/10, pp. 341–345, 2002.

[11] D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily, "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *Journal of memory and language*, vol. 68, no. 3, pp. 255–278, 2013.