

---

# THE EFFECT OF RECORDING SETTING IN ACCENTEDNESS PERCEPTION<sup>1</sup>

Ksenia Gnevshcheva: *College of Arts and Social Sciences, Australian National University* <[ksenia.gnevshcheva@anu.edu.au](mailto:ksenia.gnevshcheva@anu.edu.au)>

---

## Abstract

This study investigates the effect of recording setting on accentedness perception. It specifically tests Piller's (2002) claims that certain settings are conducive to 'passing' for a native speaker. A number of non-native speakers of English were recorded in four different settings communicating on different topics (family, university, speaking with friends, and in short service encounters). Short clips from these recordings were subsequently played in two experiments to native speaker listeners who rated them on an accentedness scale. The results suggest that listeners rate clips recorded in the services and, under some conditions, family settings as less accented. The results of this perception study are discussed in light of a production study of sociolinguistic variation in non-native speakers (Gnevshcheva 2015b).

## 1. Introduction

An accent is a '... cumulative auditory effect of those features of pronunciation which identify where a person is from, regionally or socially' (Vishnevskaya 2008: 235). Every speaker has an accent; however, lay people often believe that they do not have an accent, and only people who speak differently from them do. A difference might come in many forms, such as when someone meets a speaker whose first language (L1) is different from their own, they

may hear differences in the person's pronunciation in the second language (L2) which are a result of the speaker's L1, cumulatively perceived as a 'second language accent'. When they meet a speaker of a shared L1, they will probably regard the speaker as a native speaker of a language without any second language accent; however, if the speaker is a native speaker of a different variety of L1, they will hear a 'foreign accent', and if they are a native speaker of the same variety of L1, they will most likely believe that the person does not have an accent in that language.

In accentedness perception tasks listeners are presented with utterances which they are asked to rate on an accentedness scale (e.g., from 'No foreign accent' to 'Strong foreign accent'). Presumably, listeners somehow assess the amount of deviation from their 'native speaker ideal' present in the utterance and assign it a numerical representation. This understanding is supported by Munro & Derwing (1995), who found that the majority of listeners in their experiment exhibited a significant correlation between their perceived accentedness scores and quantitative measures of intonation, segmental, and grammatical errors, and Munro (1993), who found that the speakers with more target-like F1 and F2 measurements in their vowels were rated to be less accented.

However, accentedness perception is not just affected by the linguistic content of the audio stimuli. Ratings are highly variable and are known to be influenced by a number of speaker-independent factors. Kraut & Wulff (2013) found an effect of familiarity with foreign accents such that listeners who often interacted with non-native speakers (NNSs) perceived a weaker foreign accent. The same study found that female speakers at lower and intermediate proficiency levels were rated to be significantly more accented compared to males of the same proficiencies. Word frequency was also found to be a significant predictor as words of higher frequency were perceived to be significantly less accented (Levi, Winters, & Pisoni 2007). This finding is compatible with usage-based accounts of perception (e.g., Johnson 1997) which may predict that less frequent words would be perceived to be more novel and, potentially, more accented compared to less novel and more frequent words. Listener-dependent factors have been found to affect not only accentedness ratings but comprehension of foreign accented speech. For example, teenagers were found to understand more of the foreign accented speech than younger children, and although not tested statistically, adults were informally assessed to understand more of the foreign accented speech than teenagers (Munro, Derwing, & Holtby 2011).

Rubin & Smith (1990) explored the way listeners' stereotypes and expectations as to the relationship between speaker accent, ethnicity, and certain topics (the former two also explored in, for example, Rubin 1992) affect their perception. The same Chinese-accented speakers delivered a 'science' and a 'humanities' lecture. Although no significant effect of topic on perceived accentedness was found, they were perceived to be significantly more 'oriental' in the humanities subject. One methodological criticism that could potentially account for this finding is that the humanities lecture was about an Indian classic tale, the Mahabharata, a clearly 'oriental' subject, and the effect of a 'culturally neutral' topic remains unknown. The lexical frequency effect found by (Levi et al. 2007) potentially predicts a higher accentedness rating for the 'science' lecture which might contain more lower frequency vocabulary, such as 'growing scarcity of helium supplies' (Rubin & Smith 1990: 342). The word frequencies of the vocabulary in these particular texts are not known, but it has been shown that a 'science' text contains more technical vocabulary than a humanities text (Chung & Nation 2003).

Piller (2002) claimed that certain settings (namely, communication with friends and short service encounters, such as buying coffee in a coffee shop) are conducive to 'passing' for a native speaker (NS), that is being perceived to be a NS. She explained such variation in passing as being a result of speakers' identity negotiation. The identity construction account of variation posits that linguistic variation reflects speakers' changes in projected identity as it is being negotiated in communication (e.g., Eckert 2000). Piller (2002) argued that L2 speakers may be less focused on projecting their L2 identity in communication with close friends who know them well or with strangers in short service encounters when identity is presumably less of an issue. This study was based on self-reports and the variation in passing was not confirmed experimentally. Piller's (2002) claims were specifically tested in a production study exploring within-speaker variation among settings (Gnevshcheva 2015b).

Gnevshcheva (2015b) investigated non-native English speakers' (NNESs) production of several monophthongal vowels when speaking about their family, university life, and in short service encounters in New Zealand. German and Korean L1 speakers were recorded speaking English in interviews with a NNES of a different L1 about their family at home (family setting) and university life on the university campus (university setting); they also self-recorded themselves in short service encounters (e.g., buying a coffee in a coffee-shop) speaking predominantly to native speakers of New Zealand English (NZE) (see Section 2.2 for a summary of the recording procedure).

It was hypothesized that the speakers would use more L1-like vowels when speaking about their family to a NNES (family setting), that they would use more L2-like vowels when speaking about the university life in New Zealand to a NNES (university setting), and that their production of vowels would be most L2-like when speaking to native English speakers (NESs) in the services setting. Both groups of speakers were found to produce more L2-like vowels when participating in short service encounters compared to the other two settings, and the Korean group additionally produced more L1-like vowels when speaking about their family compared to when speaking about the university with the same NNES interviewer. In particular, German L1 speakers produced a lower KIT vowel (more NZE-like) in the services setting compared to the university one, and Korean L1 speakers produced a lower KIT and TRAP (less L1-like) in the services setting and a backer GOOSE (more L1-like) in the family setting compared to the university one. Despite an effect of speaker sex found in previous studies of the use of ethnically-marked variants in style-shifting (e.g., Sharma 2011), no effect of speaker sex was found in this production study.

Variation in perceived accentedness of a speaker may be affected by variation in both production and perception, and a production study by itself will not offer a clear picture of variation in passing as Purnell (2010) and Munro & Derwing (2015) have noted that acoustic measurements of production do not always correlate with perception ratings. This study explores the potential for between-setting variation in accentedness perception and aims to complement the Gnevsheva (2015b) production study. It particularly tests Piller's (2002) predictions that certain settings produce a different effect on perception of NNSs. It also considers the effect of variables which have been explored in previous studies: speaker sex and word frequency, among others. To address this matter, native English-speaking listeners were presented with clips from the recordings in different settings produced by the L2 speakers in Gnevsheva (2015b) in two accentedness rating experiments. I elaborate on the speakers and the two experimental conditions in the next section.

## 2. Method

### 2.1 *Speakers*

Nine German, nine Korean, and six English L1 speakers (2 from each of New Zealand, the USA, and England) were recruited for the study (24

speakers in total, evenly split by sex). The speakers' average age was 25 (age range = 21–34). All were studying towards or had completed a higher academic degree (8 Bachelor's, 4 Master's, and 12 PhD) and were affiliated with the same university in New Zealand. The NNESSs' age of acquisition of English was ten or higher. They all had demonstrated a level of English high enough to be accepted into an English-medium program at a New Zealand university; however, six participants in each L2 English group (3 males and 3 females) were informally categorized by the author to be of higher English proficiency, and three were categorized as lower English proficiency. The higher proficiency speakers exhibited less L1 interference, more fluency, and fewer grammatical mistakes while the lower proficiency speakers sounded clearly foreign-accented and made more grammatical mistakes. The speakers were recruited to represent a variety of proficiencies in order to create a range of accentedness in the perception experiments.

## 2.2 Recording procedure

The twenty-four speakers were recorded in four different settings with a head-mounted Opus 55.18 MKII beyerdynamic microphone and an H4n Zoom audio-recorder which allowed for speaker mobility (see Gnevsheva 2015a) for full details on recording procedure). Here setting is understood quite broadly and includes differences in topic, context, and interlocutor. I interviewed all speakers on family-related topics at their home (hereafter referred to as the *family* setting) and about their research and studies on the university campus (the *university* setting). In the family setting the speakers were asked about their family and childhood, and in the university setting the speakers were asked to describe their PhD or term paper research depending on their level of education. The order of the two types of interview was counterbalanced. After the first interview, the speakers were instructed to record themselves in natural communication with friends (the *friends* setting) and in a minimum of four short service encounters in a public space, such as ordering a drink at a coffee shop (the *services* setting). A total of about 1 hour of recorded speech was collected per speaker. This data now forms the 'Accent of Non-Native English' (ANNE) corpus (Gnevsheva 2015a).

In a post-recordings interview the speakers reported that they believed that more than 90% of their interlocutors in the services setting were NSs of New Zealand English, but there was much more variation in the friends setting as the speakers reported more than 50% of their interlocutors being L2 speakers. The friends setting varied most on topic as well with a wide range of topics

covered as the speakers were not constrained the way they were in the family and university semi-structured interviews which followed the same order and format of questions. Thus, the results pertaining to the friends setting should be interpreted keeping this variation in mind.

### 2.3 *Stimuli*

The audio stimuli in Experiments 1 and 2, reported below, were the same short clips extracted from the recordings of the twenty-four speakers in the four different settings (family, friends, services, and university). Where possible, the clips were extracted after the initial 5 minutes of recording, by which point the speaker would feel more comfortable with the recording procedure. Most clips were of a minimum of 25 words, but because service encounters can be quite brief, some clips in that setting were shorter than 25 words: mean length 22.2 words and 10.1 seconds.

Additionally, the clips contained complete phrases, so the exact length of clip in words varied (mean length 26.6 words and 13.0 seconds for all clips). Most of the clips from the family, friends, and university settings were uninterrupted; however, many of the services clips were sometimes interrupted by very short periods of silence as interlocutor turns had been edited out. To minimise the undue effect of individual differences between clips, three clips per setting per speaker were extracted, resulting in a total of 288 clips (24 speakers \* 4 settings \* 3 clips). The recordings were normalized to remove variation in volume.

To avoid an undue effect of grammatical inaccuracies and disfluencies on accentedness ratings, I made an attempt to choose clips without errors and hesitations. The clips did not contain names of persons, geographical locations, or any other extra-linguistic information that might draw attention to the speakers' foreignness.

### 2.4 *Experimental procedure*

This paper presents two perception experiments, both of which explored the effect of setting on accentedness perception but which also differed in a number of ways, making them partially independent of each other. Firstly, the two experiments employed different accentedness rating tasks. In the first experiment, four clips recorded in different settings from the same speaker were presented at the same time which allowed for direct comparison. This discrimination task was used first in order to assess the effect of recording setting. After a statistically significant effect was found in experiment 1, in the

second experiment, which is in essence an identification task, the clips were presented randomly one clip at a time.

Secondly, Experiment 1 employed the 216 clips from the NNEs only; Experiment 2 used all of the 288 clips from both NESs and NNEs. Because the goal was to assess the relative perceived accentedness of the clips recorded in different settings, the second experiment, which presented the clips individually, employed NES clips in order to create a range of English language proficiencies. In the first experiment the clips in the four different settings were presented at the same time allowing for a direct comparison of settings, so the creation of a proficiency continuum among the speakers was not seen as essential to experiment design.

Thirdly, there were also differences in the rating scales used: in Experiment 1 the speakers were rated on a scale which read 'I can hear a very strong foreign accent' and 'I cannot hear a foreign accent at all' at the two extremes and in Experiment 2 the scale read 'Definitely a first language speaker of English' and 'Definitely a second language speaker of English'. Thus, in the first experiment any deviation from the NZE target (even native to other varieties of English such as American English) would be considered a manifestation of an accent while in the second experiment only deviations due to an assumed L1 interference would result in a lower score. Consequently, Experiment 1 explores variation in NZE-accentedness (with implications for passing for a NS of the same dialect as listeners) and Experiment 2 explores variation in English-native-likeness (with implications for passing for a NS of any English variety). I elaborate on the methodological details for Experiment 1 and Experiment 2 below.

### 3. Experiment 1

#### 3.1 *Participants and procedure*

The listeners in Experiment 1 were twenty five native speakers of New Zealand English who were recruited through announcements posted around the University of Canterbury campus. There were sixteen females and nine males. The age, education, socio-economic class of the participants were comparable to those of the speakers: average age 27, age range 18–69, all had achieved or were studying towards a Bachelor's degree or above at the time of the study. Ten claimed no knowledge of a foreign language.

The listeners were seated individually in a quiet lab in front of the computer

with head-phones. Stimuli were presented electronically using the E-Prime 2.0 software (Psychology Software Tools 2012). Before starting the actual task the listeners read the instructions on the screen, completed a practice trial with comparable clips from a male NS of New Zealand English and if needed, adjusted the volume and clarified the procedure with the research assistant (the author). After that, the listeners were presented with eighteen sets of four clips, each set corresponding to a speaker with a random combination of clips from the four settings. In the task, the listeners were instructed to rate the presented clips and place them on a scale which read ‘Very strong foreign accent’ and ‘No foreign accent at all’ at the two extremes (Figure 1). At the top of the screen there were four symbols associated with the four clips recorded in the different settings. The listeners played the clips by clicking on each of the four symbols one at a time and indicated their accentedness rating by clicking on the scale below where an identical symbol then appeared (Figure 2). They could replay the clips for a given speaker and change the position of symbols on the scale until they moved on to the next speaker by clicking on ‘Done’.

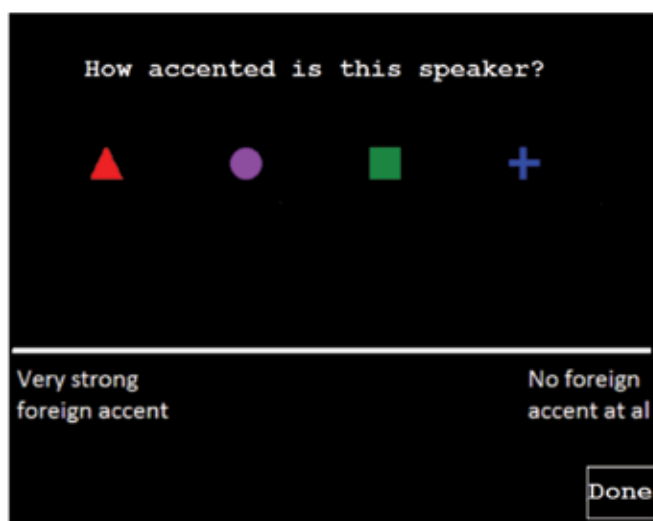


Figure 1: Slide presented to listeners in Experiment 1.



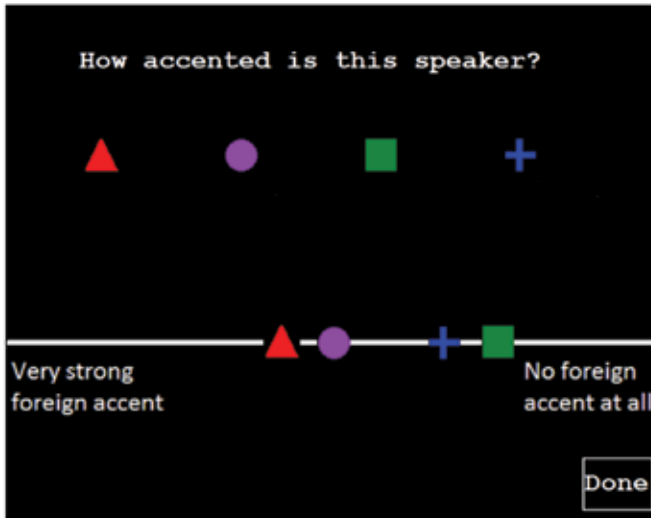


Figure 2: Slide with listener response in Experiment 1.

The order of the eighteen speakers in the experiment, clips (the choice of one of the three clips for each speaker in each setting), the order of the four symbols and the four settings on the screen were randomized. The task was self-paced and took about 30 minutes. At the end, the listeners completed a short biographical questionnaire. They were given a \$10 voucher for completing the task. The research was reviewed and approved by the University of Canterbury Human Ethics Committee.

### 3.2 Results and discussion

The position of the symbols on the scale was recorded as an accentedness rating from 1 (Very strong foreign accent) to 100 (No foreign accent at all) which was subsequently analysed in R (R Core Team, 2014). A linear mixed-effects model was fit to the data with the perceived accentedness rating as the dependent variable. The fixed effects in the full model included two-way interactions between *setting* and each of the other variables as well as their main effects (Table 1). *Speaker*, *clip*, and *listener* were included as random intercepts, and *setting* was introduced as a random slope for *listener* (Barr, Levy, Scheepers, & Tily 2013). If an interaction or a fixed effect was found

to be non-significant, I simplified the model by excluding the interaction or the variable from the model and then compared the previous and the current models with an ANOVA. The significantly better or simpler model was kept.

**Table 1: The effects included in the statistical model.**

SPEAKER-RELATED FIXED EFFECTS	LISTENER-RELATED FIXED EFFECTS	CLIP-RELATED FIXED EFFECTS	RANDOM INTERCEPTS
L1 Proficiency Sex	Age Sex L2 knowledge (binary) <sup>2</sup>	Setting Word frequency <sup>3</sup> Clip length in seconds Clip length in words Speech rate (words per second) Progression in the experiment (1 through 18)	Speaker Listener Clip

Table 2 represents the final model. The higher proficiency speakers in the university setting were chosen as the reference level (Intercept). The *estimate* and the *standard error* columns in the table give us the predicted accentedness rating and standard error for a level respectively. So for the Intercept (the higher proficiency speakers in the university setting), the predicted accentedness rating is 45.696. To calculate the predicted accentedness rating for a different level, the respective value in the estimate column is added or subtracted. For example, the higher proficiency speakers received a rating 5.081 higher in the services setting than in the university setting; this difference was significant, as indicated in the significance column. The difference in accentedness ratings between the friends and the university settings was not found to be significantly different. The difference between the family and the university settings did not reach the level of significance ( $p=0.05$ ); however, the trend was in the direction of the family setting being judged less accented.

This finding suggests that listeners' accentedness perception may vary by setting of recording. As predicted by Piller (2002), who found that L2 speakers believed they passed for a native speaker more commonly in short service encounters, the clips in the services setting in this experiment were

**Table 2: Model summary for accentedness ratings of non-native English speakers in different settings.**

	Estimate	Standard error	df	t value	Pr(> t )	Significance
(Intercept)	45.696	4.951	34	9.230	0.000	-
setting_family	2.456	1.307	169	1.879	0.062	
setting_friends	0.600	1.325	1567	0.451	0.652	
setting_services	5.081	1.309	169	3.881	0.000	***
proficiency_lower	-19.789	7.147	19	-2.769	0.012	*
progression	-0.132	0.067	1717	-1.972	0.049	*
setting_family : proficiency_lower	-0.368	2.263	169	-0.162	0.872	
setting_friends : proficiency_lower	0.697	2.267	166	0.307	0.759	
setting_services : proficiency_lower	-4.561	2.264	169	-2.015	0.046	*

Note: \*  $p < 0.05$ ; \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

rated as less accented. However, contrary to her prediction, the friends setting was not among the less accented ones. Admittedly, the friends setting is very broad as the only variable defining it is audience, and there is much potential for variation, for example, in topic, which could have an effect on perception. Additionally, some speakers spoke to NSs of NZE, and others to NNEs of a shared or not shared L1. Speaking to non-NZE-speaking friends could have resulted in convergence to more accented speech, which is compatible with the current findings. This study focuses on the speakers as a group; however, future analysis of the specific effects of different audiences would allow to disentangle the effects of audiences that have different relationships with the speaker.

As expected, proficiency was found to have a significant effect such that the speakers who I had assigned to the lower proficiency group received a stronger accentedness score compared to the higher proficiency group. This finding is not meant to be an important focal point but simply is reported to lend support to my division of the speakers into the two groups. Furthermore, I found a significant interaction between setting and proficiency, such that lower proficiency speakers in the services setting did not receive the advantage that higher proficiency speakers did and were judged more accented

in that particular setting. This interaction is plotted in Figure 3. It suggests that it is only the higher proficiency speakers who receive a setting advantage and are judged less accented in the services setting. When the model was re-run with levels of proficiency re-leveled and lower proficiency as the Intercept, no significant main effect of setting was found. This means that lower proficiency speakers were judged similarly foreign-accented irrespective of setting.

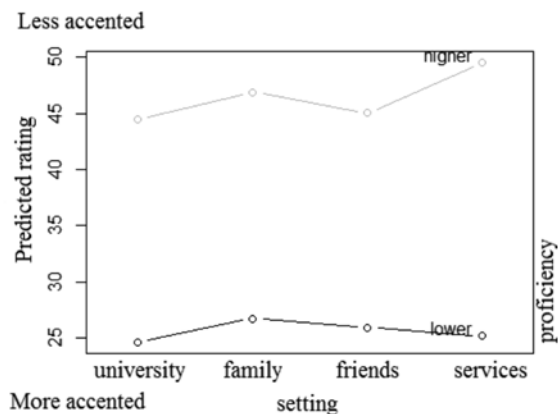


Figure 3: Model prediction for accentedness rating in the four settings (from model in Table 2).

Finally, there was a significant effect of progression in the experiment such that the clips presented to the listeners later in the experiment were judged to be more foreign-accented and received a stronger accentedness score. I argue that this effect is due to the listeners' expectation to hear accented speech (Lindemann & Subtirelu 2013). This experiment did not employ NESs as controls, so the listeners may have noticed that the range of accentedness they heard was from light to strong and realized that the experiment included NNEs only. This may have brought them to expect to hear foreign accented speech and rate it in accordance with their expectations.

## 4. Experiment 2

### 4.1 *Participants and procedure*

The listeners were thirty native speakers of New Zealand English, similar to the speakers and the listeners from Experiment 1, with the age range 18–50, age mean 24, fourteen females and sixteen males. The majority had achieved or were studying towards a Bachelor's degree. Fourteen listeners claimed no knowledge of a foreign language.

The participants completed the task on a computer with E-Prime 2.0 (Psychology Software Tools 2012) in a quiet room. They heard the audio stimuli via head-phones, saw the instructions on the screen, and entered their responses on the keyboard. Each listener was presented with 1 clip from each speaker in every setting (24 speakers x 4 settings x 1 random clip = 96 clips). The task consisted of four quarters with three breaks in between. Each quarter used one clip from each speaker in one of the settings, with the order of speakers and settings randomized. After a clip presentation, the listener rated the speaker on an accentedness scale from 1 (Definitely a first language speaker of English) to 7 (Definitely a second language speaker of English), then guessed where the speaker was from, and commented on what influenced their decisions in the first and second questions. After the task, which took about one hour, the listeners completed a biographical questionnaire and received a \$10 coffee voucher for their time. The research was reviewed and approved by the University of Canterbury Human Ethics Committee. This paper focuses on the listeners' responses to the first question (perceived accentedness); for a discussion of the responses to the second and third questions see (Gnevsheva 2016, under review).

### 4.2 *Results and discussion*

A mixed-effects regression model was fit to the NNES data in R (R Core Team 2014) with perceived accentedness rating as the dependent variable. As setting was the primary variable of interest, the fixed effects were its two-way interactions with all the other variables and their main effects (Table 3). *Setting* was introduced as a random slope for *listener*, but the model did not converge and the random slope was excluded (Barr, Levy, Scheepers, & Tily 2013).

If an interaction or a main effect did not reach significance, the model was re-run without it, and the older and the newer models were compared with an ANOVA. The better or the simpler model was kept, and the process was

**Table 3: The effects included in the statistical model.**

SPEAKER-RELATED FIXED EFFECTS	LISTENER-RELATED FIXED EFFECTS	CLIP-RELATED FIXED EFFECTS	RANDOM INTERCEPTS
L1 Proficiency Sex	Age Sex L2 knowledge (binary)	Setting Word frequency Clip length in seconds Clip length in words Speech rate (words per second) Progression in the experiment (1 through 96)	Speaker Listener Clip

repeated with the remaining interactions and main effects. The final model in Table 4 contains fixed effects which were significant or improved model fit.

The reference level (Intercept) is the clips recorded in the university setting by females of higher proficiency judged by listeners with no L2 knowledge; this allowed for a better comparison with Experiment 1 where the base level was higher proficiency speakers in the university setting. The estimate rating for this level is 3.781. Their counterparts in the family or services setting were judged significantly less accented. This finding partially supports Piller's (2002) claims about different settings and is aligned with the first experiment's results which investigated foreign accentedness with a slightly different method. Another statistically significant main effect was speaker proficiency. Unsurprisingly, speakers of lower proficiency were judged as more accented which shows that NSs of New Zealand English generally agreed with my assignment of the speakers to the two proficiency groups.

The main effect of listener L2 knowledge did not reach significance, but it participated in an interaction with setting such that listeners with L2 knowledge judged clips in the friends and services setting as less accented compared to the university setting (Figure 4). The interaction with the family setting did not reach significance at  $p=0.05$  level, but the trend was in the same direction ( $p<0.1$ ). This experiment was not designed to explain this finding, but it is possible that listeners with L2 knowledge use cues that are different from those used by listeners with no L2 knowledge. Previous research has

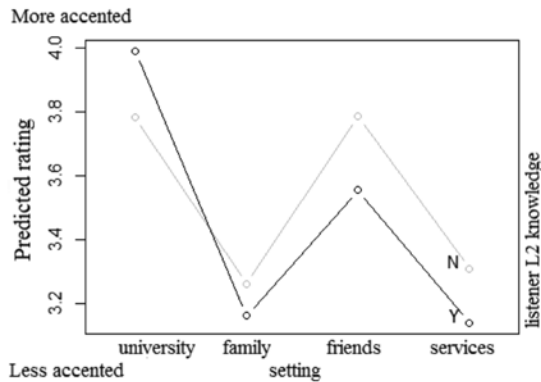
**Table 4: Model summary for accentedness ratings of non-native English speakers in different settings.**

	Estimate	Standard error	df	t value	Pr(> t )	Significance
(Intercept)	3.781	0.389	43.0	9.709	0.000	-
setting_family	-0.522	0.221	264.1	-2.359	0.019	*
setting_friends	0.005	0.221	263.5	0.023	0.982	
setting_services	-0.473	0.222	263.0	-2.133	0.034	*
speaker.sex_male	-0.843	0.433	25.6	-1.947	0.063	
L2.knowledge_yes	0.208	0.297	38.7	0.698	0.489	
proficiency_lower	2.143	0.398	17.9	5.384	0.000	***
setting_family: speaker.sex_male	0.622	0.284	188.4	2.194	0.030	*
setting_friends: speaker.sex_male	0.170	0.283	187.3	0.600	0.550	
setting_services: speaker.sex_male	0.430	0.284	187.9	1.514	0.132	
setting_family: L2.knowledge_yes	-0.305	0.171	1977.0	-1.788	0.074	
setting_friends: L2.knowledge_yes	-0.440	0.171	2003.0	-2.569	0.010	*
setting_services: L2.knowledge_yes	-0.377	0.171	1996.0	-2.202	0.028	*

Note: \*  $p < 0.05$ ; \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

found that listeners who rarely interacted with NNEs perceived a stronger accent in an accentedness rating task (Kraut & Wulff 2013). My finding may be reflective of a similar influence as I expect that listeners with L2 knowledge are more likely to interact with NNEs on a daily basis; however, I did not collect the listeners' social network information and cannot be certain of that. Three three-way interactions of (1) L2 knowledge, setting, and speaker sex, (2) L2 knowledge, setting, and word frequency, and (3) setting, speaker sex, and listener sex were tested post-hoc but were not significant.

Speaker sex did not reach significance at the level of  $p = 0.05$ ; however, the trend was in the direction of male speakers being rated less accented. These results are reminiscent of the finding by Kraut & Wulff (2013) that some

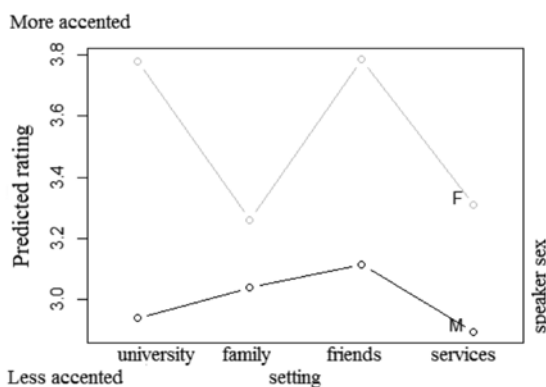


**Figure 4: Model prediction for accentedness rating of listeners with and without L2 knowledge in the four settings (from model in Table 4).**

groups of female NNEs received a higher accentedness score compared to male speakers of the same proficiency. Speaker sex also participated in a significant interaction with setting such that male speakers in the family setting were judged significantly more accented compared to the university setting in comparison to how less accented women were judged in the family setting compared to the university setting (Figure 5). When I re-ran the model with male speakers as the reference level, no significant difference was found between the settings.

On the one hand, this may be reflective of variation in speaker speech by setting (as in Sharma, 2011); on the other, listeners may react differently to male and female speakers in different settings. The Gnevsheva (2015b) production study did not find variation by speaker sex in the production of monophthongal vowels, but only a thorough investigation of male and female production of other features would be able to ascertain that. Although future research will be needed to explore this further, I can tentatively suggest that listeners perceive a different degree of accent when men and women speak on different topics. Psychology literature has shown that different sexes elicit a different expectation of expertise and women are often perceived to be less knowledgeable in a male-associated task (Thomas-Hunt & Phillips 2004 and references therein). As will be discussed in the next section, the family and university settings differed in formality and technicality of the vocabulary used. My data were not gathered to test this hypothesis specifically, but I





**Figure 5: Model prediction for accentedness rating of male and female speakers in the four settings (from model in Table 4).**

tentatively suggest that scientific or professional vocabulary produced in a female voice was perceived to be more ‘foreign’ than that produced in a male voice because of listener expectation of male expertise in a professional environment. Listener expectations and past experiences with males and females in different situations may affect perceived accentedness as other assumed social information (e.g., speaker ethnicity) has been found to do so as well (e.g., Rubin 1992; Gnevsheva in press).

## 5. General discussion

A number of variables were tested in the two accentedness perception experiments described above. The experiments differed in their methodology (see Table 5). By way of reminder, in Experiment 1 the listeners were presented with four clips at a time (one clip from each setting for NNEs only) and were asked to place them on a foreign accentedness scale. In Experiment 2 the participants listened to one clip at a time with four clips from an individual speaker (the four settings from NESs and NNEs) and rated them on a second language accentedness scale.

Despite these methodological differences between the two experiments, a number of similar trends emerged in the results. First, the effect of proficiency was found to be significant in both experiments and the speakers who I had

**Table 5: Details about the two experiments.**

	EXPERIMENT 1	EXPERIMENT 2	
Dependent variable	Foreign accentedness rating	Second language accentedness rating	
Stimulus presentation	4 at a time	1 at a time	
Speakers	NNESs	NESs and NNESs	
Variables retained in the best-fit model	Setting	Setting	
	Proficiency	Proficiency	
	Progression	Speaker sex	
	Proficiency:setting		L2 knowledge
			Setting:speaker sex
		Setting:L2 knowledge	

assigned to the lower proficiency group were rated significantly more accented in both experiments. This supports my division of the NNESs into two groups by proficiency and is not meant to be the focus of the study.

Setting, which was the main independent variable of interest, was found to be significant as a main effect and in interactions. The results of this quantitative study support some of the Piller's (2002) claims about passing based on qualitative data. According to her, NNESs are more likely to pass for a native speaker in short service encounters or in communication with friends. If extrapolated to the four settings in my experiments, this may predict a less accented rating in the friends and services settings. Both experiments found that the clips in the services setting were judged significantly less accented compared to the university setting (for both males and females in the first experiment and for females only in the second experiment). Additionally, the family setting was found to be rated significantly less accented than the university setting for female speakers in Experiment 2; it failed to reach significance in Experiment 1, but the trend was in the same direction with  $p < 0.1$ . My findings support Piller's claims about the services setting; however, the friends setting was not among the less accented ones. I acknowledge that the clips in the friends setting varied dramatically in terms of topic, interlocutors, and conversation flow, so it is problematic to make such a generalization about all communication with friends. What the participants in Piller (2002) might be noticing is their increased confidence in communication

with people and topics that they are familiar and comfortable with. Under this interpretation my speakers' communication with the interviewer in the family setting may fit this criterion despite it being an interview: the speakers were speaking on an informal topic to someone they knew as part of their social circle in a comfortable environment (their home).

One reason for variation in perception can be listeners' reaction to variation in production. If the speaker style-shifts in a certain manner, it is important to know whether it is salient for the listener and if variation in production reflects the variation in perception. The variation by setting in perception partially matched the variation by setting in production found in Gnevsheva (2015b). The services setting was most native-like in the speakers' production of the vowels and the listeners' ratings of accentedness. This relative lighter accentedness in production and perception would, presumably, make it easier to pass for a native speaker in short service encounters as predicted by Piller (2002). The inter-relationship between the family and university settings in production and perception was not exact. Korean L1 participants were more native-like in their production in the university setting compared to the family setting while German L1 speakers' production was not different between the two settings. In perception, on the other hand, the university setting was rated as more accented for females of both language groups (Experiment 2) or no significant difference at  $p < 0.05$  was found (Experiment 1). Purnell (2010) and Munro & Derwing (2015) argue that the mapping of acoustic and perceptual cues is not exact, so while the speakers may be signalling nativeness in some elements, if the listeners are focusing on a different set of elements, the signal may not be noticed. Moreover, a more nativelike variable may be interpreted to signal other, non-nativeness-related social information as the same cue may activate a number of potential meanings in the indexical field (Eckert 2008). Finally, of course, this analysis compares variation in accentedness ratings of clips to variation in accentedness in production of vowels and there are many other features that may vary among settings in production and affect accentedness perception: consonants, prosody, etc. For example, because the university setting is more likely to contain longer, polysyllabic words, it could have provided more opportunities for non-nativeness for the speakers.

On the other hand, we can suspect that listeners react differently to different settings even when production is similar. The obvious difference between the family and services settings and the university setting is the topic, which was more professional, dealing with research and innovation and requiring the speakers to use more technical language, terms, and jargon in the university

setting while in the family and services settings the topic was more mundane and the language was less industry-specific as in the following examples.

1. *all cellulose composites are monocomposites and a monocomposite just um means that the fibre and the matrix are made of the same material which means they're chemically very very similar if not identical* (Jack. University interview. Perception clip 2)
2. *I would say that I'm closest to my mom because she's the one I I can have the most personal um discussions or conversations with* (Jack. Family interview. Perception clip 2)
3. *er no er the light is is fine I only need one but I need the bulb and the bulb that fits in there wasn't there do you have any other shelf or* (Jack. Services self-recording. Perception clip 3)

The topic of the university setting is less familiar and more 'foreign' to a non-specialized audience which could lead to a stronger accentedness perception. If we assume a usage-based account of accentedness rating, we presuppose that listeners compare the accent in the clip to an ideal representation based on their multiple experiences with other NSs. A more technical topic would be more novel than a family-related account resulting in stronger accentedness ratings. In a similar fashion, previous research has found an effect of word frequency on accentedness ratings such that the lower the word frequency the more accented the speaker was rated (Levi et al. 2007). Applying the word frequency hypothesis to the two interviews in this experiment, one could hypothesize that the university setting clips with their technicalities could produce a similar effect and attract higher accentedness ratings. For that reason, the mean word frequency in the clips was calculated and entered into the model. The effect did not reach significance or improve model fit and was dropped from the final model, but it is possible that this measure did not capture the word frequency effect well and a different one may prove a better prediction.

Experiment 1 found a significant interaction between setting and proficiency such that lower proficiency speakers in the services setting were not rated significantly less accented compared to the university setting. As opposed to Experiment 1, Experiment 2 revealed significant interactions of setting with listener L2 knowledge and speaker sex. The different interactions in the two experiments are possibly due to differences in methods used. In Experiment 2 the listeners were presented with one clip at a time, so four

clips from the same speaker were rated independently exhibiting an effect of setting in perception; in Experiment 1, however, the four clips were presented at the same time, and a listener could make direct comparisons between the settings. These differences may be a task effect. On the one hand, phonetic perception may vary in different environment as listeners have been shown to behave differently in discrimination and identification perception tasks. For example, Waylard (2007) found discriminability of non-native contrasts (Korean and Thai stop consonants) was better predicted by identification than discrimination data. On the other hand, it is possible that the accentedness ratings in Experiment 1 are more representative of the variation inherent in the clips than Experiment 2 where listener-dependent factors may play a larger role. Therefore, the significant interactions with speaker sex and L2 knowledge emerge when the listeners do not realize that they listen to the same speaker more than once and their stereotypes and expectations play a larger role in the assignment of accentedness ratings.

Progression was found to have a significant effect on perceived accentedness in Experiment 1. The listeners were more likely to perceive a stronger foreign accent closer to the end of the experiment. I argued that this effect is due to the listeners' expectation to hear accented speech. The listeners may have noticed that the range of accentedness they heard was from light to strong and realized that the experiment included NNESs only. This may have brought them to expect to hear foreign accented speech and rate it in accordance with their expectations. To sum up, the role of listener expectation in perception surfaces in these experiments several times and confirms earlier observations of its profound effect on foreign-accented speech perception (Lindemann & Subtirelu 2013).

## 6. Conclusion

To sum up, the two experiments reported on in this paper investigated the effect of setting on accentedness perception and found that the same speakers in different settings (e.g., university and services) may receive a different accentedness rating. The found variation was partially in line with Piller's (2002) predictions and the speakers were judged less accented in the services and, sometimes, family setting. Such variation in perception may be listener-dependent, speaker-dependent, or both. The significant interactions of setting with speaker sex and listener L2 knowledge have highlighted the role of

listener expectation and experience, and it may prove an interesting path for future research.

## Notes

- 1 This research was partially supported by a University of Canterbury research grant. I am thankful to Kevin Watson, Jen Hay, and two anonymous reviewers for helpful and encouraging comments on an earlier draft. All the remaining shortcomings are naturally mine.
- 2 The listeners were asked whether they spoke any language besides English and how well. Because some participants only listed the languages and did not comment on their proficiency, it was impossible to make more minute distinctions. Listener L2 knowledge in this paper divides participants into two groups: those who reported any L2 knowledge and those that did not.
- 3 Mean log CELEX frequency of CELEX content words in the clip (Baayen et al. 1995).

## References

- Baayen, R. H., R. Piepenbrock and L. Gulikers. 1995. 'The CELEX lexical database (CD-ROM).' Philadelphia: Pennsylvania: Linguistic Data Consortium, University of Pennsylvania.
- Barr, D. J., R. Levy, C. Scheepers and H. J. Tily. 2013. 'Random effects structure for confirmatory hypothesis testing: Keep it maximal.' *Journal of Memory and Language*, 68(3), 255–278.
- Chung, T. M. and P. Nation. 2003. 'Technical vocabulary in specialised texts.' *Reading in a Foreign Language*, 15(2), 103–116.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453–476.
- Eckert, P. 2000. *Linguistic variation as social practice: the linguistic construction of identity in Belten High* (Vol. 27.). Malden, Mass: Blackwell Publishers.
- Flege, J. E. and K. L. Fletcher. 1992. 'Talker and listener effects on degree of perceived foreign accent.' *Journal of the Acoustical Society of America*, 91(1), 370–389.
- Gnevsheva, K. 2015a. 'Acoustic analysis in Accent of Non-Native English (ANNE) corpus.' *International Journal of Learner Corpus Research* 1(2), 256–267.
- Gnevsheva, K. 2015b. 'Style-shifting and intra-speaker variation in the vowel production of non-native speakers of New Zealand English.' *Journal of Second Language Pronunciation* 1(2), 135–156.
- Gnevsheva, K. 2016. 'Beyond the language: The social cues in accentedness judgments.'
- Gnevsheva, K. 2017. 'Within-speaker variation in passing for a native speaker.'

- International Journal of Bilingualism*. Advance online publication DOI: 10.1177/1367006915616197
- Gnevsheva, K. in press. 'The expectation mismatch effect in accentedness perception of Asian and Caucasian non-native speakers of English.' *Linguistics*
- Isaacs, T. and P. Trofimovich. 2011. 'Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of second language speech.' *Applied Psycholinguistics*, 32(1), 113–140.
- Johnson, K. 1997. 'Speech perception without speaker normalization: An exemplar model.' In K. Johnson and J. W. Mullennix (eds), *Talker variability in speech processing*. San Diego: Academic Press. 145–165.
- Kraut, R. and S. Wulff. 2013. 'Foreign-accented speech perception ratings: a multifactorial case study.' *Journal of Multilingual and Multicultural Development*, 34(3), 249–263. doi: 10.1080/01434632.2013.767340
- Levi, S. V., S. J. Winters and D. B. Pisoni. 2007. 'Speaker-independent factors affecting the perception of foreign accent in a second language.' *Journal of the Acoustical Society of America*, 121(4), 2327–2338.
- Lindemann, S. and N. Subtirelu. 2013. 'Reliably biased: The role of listener expectation in the perception of second language speech.' *Language Learning*, 63(3), 567–594.
- Munro, M. J. 1993. 'Productions of English vowels by native speakers of Arabic: acoustic measurements and accentedness ratings.' *Language and Speech*, 36(1), 39–66.
- Munro, M. J. and T. M. Derwing. 1995. 'Foreign accent, comprehensibility, and intelligibility in the speech of second language learners.' *Language Learning*, 45(1), 73–97.
- Munro, M. J. and T. M. Derwing. 2015. 'A prospectus for pronunciation research in the 21st century: A point of view.' *Journal of Second Language Pronunciation*, 1(1), 11–42.
- Munro, M. J., T. M. Derwing and A. K. Holtby. 2011. 'Evaluating individual variability in foreign accent comprehension.' Paper presented at the the 3rd annual pronunciation in second language learning and teaching conference, Iowa State University.
- Piller, I. 2002. 'Passing for a native speaker: identity and success in second language learning.' *Journal of Sociolinguistics*, 6(2), 179–206.
- Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from <http://www.pstnet.com>.
- Purnell, T. 2010. 'Phonetic detail in the perception of ethnic varieties of US English.' In D. R. Preston and N. Niedzielski (eds), *A Reader in Sociophonetics*. New York: Walter de Gruyter. 289–326.
- R Core Team. 2014. 'R: A language and environment for statistical computing.' R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Rubin, D. L. 1992. 'Nonlanguage Factors Affecting Undergraduates' Judgments of

- Nonnative English-Speaking Teaching Assistants.' *Research in Higher Education*, 33(4), 511–531.
- Rubin, D. L. and K. A. Smith. 1990. 'Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants.' *International Journal of Intercultural Relations*, 14, 337–353.
- Sharma, D. 2011. 'Style repertoire and social change in British Asian English.' *Journal of Sociolinguistics*, 15(4), 464–492.
- Thomas-Hunt, M. C. and K. W. Phillips. 2004. 'When what you know is not enough: Expertise and gender dynamics in task groups.' *Personality and Social Psychology Bulletin*, 30(12), 1585–1598. doi: 10.1177/0146167204271186
- Vishnevskaya, G. 2008. 'Foreign accent: Phonetic and communicative hazards.' In E. Waniek-Klimczak (ed), *Issues in accents of English*. Newcastle upon Tyne: Cambridge Scholars Publishers. 235–251.
- Waylard, R. P. 2007. 'The relationship between identification and discrimination in cross-language perception: The case of Korean and Thai.' In O.-S. Bohn and M. J. Munro (eds), *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins Publishing Company. 201–220.