The Impact of L2 English Learners' Belief about an Interlocutor's English Proficiency on L2 Phonetic Accommodation

Fan Jiang[1] [ORCID] **· Shelia Kennison[2]**

Abstract

The current study investigates the influence of L2 English learners' belief about their interlocutor's English proficiency on phonetic accommodation and explores whether interaction-induced phonetic convergence could improve L2 English learners' vowel pronunciation. Results from two experiments show that when the subjects believed that their interlocutor was a native English speaker, they generally converged to her vowel pronunciation. When the subjects believed that their interlocutor was a non-native English speaker, they generally diverged from her vowel pronunciation. In addition, phonetic convergence enabled the subjects to improve their L2 English vowel pronunciation, leading to greater similarity to the native interlocutor. The findings are discussed in terms of the Communication Accommodation Theory, the Interactive Alignment Theory, and the Speech Learning Model.

Imitation is a ubiquitous human behavior. Previous research has demonstrated that people tend to imitate their interlocutors in interaction or in shadowing tasks, resulting in similar vowel quality (e.g. Babel, 2009; Pardo, 2010), speaking rate (e.g. Giles et al., 1991; Levitan & Hirschberg, 2011), pitch (e.g. Gregory & Webster, 1996; Goldinger et al., 1997; Babel & Bulatov, 2012; Gessinger et al., 2018), and voice onset time (VOT) (e.g. Shockley et al., 2004; Nielsen, 2007). Terminology regarding this phenomenon varies, with it being referred to as accommodation (e.g. Trudgill, 1986; Gregory et al., 1997; Babel, 2012), alignment (e.g. Pickering & Garrod, 2004; Trofimovich & Kennedy, 2014) or convergence

✉ Fan Jiang
  jiangfan@nbut.edu.cn

  Shelia Kennison
  shelia.kennison@okstate.edu

[1] School of Foreign Languages, Ningbo University of Technology, No. 201 Fenghua Road, Jiangbei District, Ningbo City, Zhejiang, China

[2] Oklahoma State University, Stillwater, USA

(e.g. Pardo, 2006; Kim et al, 2011; Pardo et al., 2012). In this study, we use term "accommodation" to refer to the general tendency for a speaker to adjust his/her speech on the phonetic dimension to an interlocutor.

A number of studies have investigated the influence of language factors (e.g. Kim, et al., 2011), social factors (e.g. Namy et al.2002; Pardo, 2010) and psychological factors (e.g. Gregory et al., 1997; Gregory et al., 2000) on phonetic accommodation. For example, Kim et al. (2011) found that phonetic convergence was mediated by the language distance between interlocutors in spontaneous speech: the closer the language distance between interlocutors was, the more convergence the interlocutors would exhibit towards each other. On the contrary, Walker and Campbell-Kibler (2015) demonstrated that greater language distance between the talkers triggered more convergence. It is yet unknown how language distance between interlocutors affects the magnitude of phonetic convergence.

The most studied sociopsychological factors include gender and attitudes towards an interlocutor. For instance, Namy et al. (2002) demonstrated with a shadowing task that the female participants converged more to the model speaker than the male participants, while Pardo (2006) found the opposite with a conversation task, i.e. the male participants converged more than the female participants. Abrego-Collier et al. (2011) manipulated the participants' attitude towards the speaker by asking them to listen to two versions of a narrative (one positive and one negative) recounted either by a straight man or a gay man. Results showed that the subjects with a positive attitude towards the talker converged towards the talker in terms of VOT, while those with a negative attitude diverged from the talker's VOT. Adopting a similar approach, Yu et al. (2013) also reported that in their study the participants with positive evaluation towards the speaker converged more than those with negative feelings. Similarly, Babel (2012) examined the role of gender and liking in phonetic convergence and found that the more attractive the female participants rated the white model talker, the more they converged to his vowels. For the male participants, the result was the reverse: the more attractive the white model talker was rated, the less the male participants converged to his vowels. In addition to studies examining the effect of attitude on phonetic convergence from short-term encounters in lab-based experiments, Pardo et al. (2012) investigated how attitude or closeness towards an interlocutor affected phonetic convergence in a long-term relationship. The results showed that the degree of phonetic convergence was related to the closeness of the male college roommates.

Phonetic Accommodation in L2 Interactions

Studies exploring phonetic accommodation between L2 learner interactions or native-to-nonnative interactions have yielded some interesting findings. For example, Trofimovich and Kennedy (2014) examined interaction between L2 English learners of different language backgrounds with a picture story completion task and a map task. The beginning and the end of the interactions were rated by native English speakers. Results revealed that the L2 learners converged towards each other in pronunciation. In addition, this study also found that similarity between the interlocutors, such as fluency, communicative effectiveness, anxiety and attractiveness are positively correlated with the degree of convergence. Zajac and Rojczyk (2014) investigated L2 phonetic imitation with Polish L2 English learners with a shadowing paradigm and found that the native v.s. non-native status of the model talker affected the degree of convergence. Specifically, the participants converged to the vowel duration of the native model talker and diverged from the vowel duration of the

nonnative model talker, indicating that social psychological factors affected L2 learners' phonetic accommodation. Zajac and Rojczyk hypothesized that "attitude towards L2 pronunciation may affect learners' convergence strategies" (Zajac & Rojczyk, 2014, p508). However, they didn't survey the participants' attitude towards L2 pronunciation. In addition, it has been proposed that L2 learners' familiarity with the interlocutor's language (Costa et al., 2008), proficiency (Young, 1988), phonetic talent (Lewandowski, 2011) and personality traits (Lewandowski & Jilka, 2019) may affect the magnitude of L2 phonetic convergence.

Phonetic convergence in native-to-nonnative conversational interactions has been proposed as an effective approach in promoting L2 phonetic acquisition (Trofimovich, 2013). Currently, most L2 English learners are struggling to eliminate foreign accent, especially for those who are studying in an L2 environment, as native-like pronunciation is often credited with trust and credibility (Lev-Ari & Keysar, 2010). For those immersed in an L2 environment, they have ample opportunities to interact native speakers. Therefore, is it possible for L2 learners to improve their L2 pronunciation during interaction with native speakers through phonetic convergence? In Experiment 1 of the study, we aim to probe into this question by examining two groups of Chinese L2 English learners studying on an American campus.

Theoretical Accounts of Accommodation

Three major theoretical accounts have been proposed to explain the mechanism of speech accommodation. First, according to the Communication Accommodation theory (CAT), people adjust their language on different levels according to their communicative needs, beliefs, attitudes and socialcultural conditions (Giles et al., 1991). Interlocutors accommodate their speech in interaction, triggering two main directions of accommodation, i.e. convergence and divergence (Coupland & Giles, 1988). Convergence refers to the phenomenon that people adjust their speech towards their interlocutor so that their use of sounds, words and structures becomes similar, while divergence refers to the fact that a speaker can adjust his/her speech features away from the interlocutor. In addition to convergence and divergence, interlocutors may also maintain their speech features or conversational styles, which is termed maintenance. In interaction, different types of accommodation results from the goal of communication perceived by the interlocutors. Convergence is usually triggered by a need for social approval or identification.

Another theoretical framework, the Interactive Alignment Theory (IAT) posits that for successful communication to occur, interlocutors must align their language at the lexical, syntactic and phonological level to reach a common situational model or "common ground" during interaction (Pickering & Garrod, 2004). Once engaged in conversational interactions, speakers will align to each other automatically as a result of priming. After hearing a word or an utterance produced by an interlocutor, speakers' mental representations associated with that word or utterance will be activated and they tend to use these representations in subsequent language production. According to Pickering and Garrod (2004), linguistic alignment normally leads to the alignment of situation modes. According to ITA, phonetic convergence works as an implicit and automatic mechanism.

Studies to date have found conflicting results, supporting either CAT or IAT. Consequently, a three account, a hybrid one has been proposed. According to Babel (2012), speech accommodation primarily results from an automatic alignment mechanism, and it

is also likely to be affected by social factors. In the current study, we want to examine the nature of L2 phonetic accommodation and test whether it is automatic, subject to social psychological factors or to a combination of both.

So far, most studies exploring the influence of sociopsychological factors on phonetic accommodation were conducted with native English speakers (e.g. Nielsen, 2007; Babel, 2009, 2012; Pardo et al., 2010). Only a handful of studies targeted native to nonnative interaction (Kim et al., 2011). Additionally, previous studies on L2 phonetic convergence mainly involved subjects whose native language is Polish (Zając & Rojczyk, 2014), Korean (Kim et al., 2011), Spanish (Berry & Ernestus, 2018) or German (Lewandowski, 2011). As it is still unknown how language distance affects phonetic accommodation, it is necessary to extend the range of subjects to involve those whose L1 is significantly distant from English, like Mandarin Chinese speakers. The overall goal of the current study is twofold. Theoretically, we aim to explore the mechanism of L2 phonetic accommodation. Pedagogically, we attempt to examine the effectiveness of interactive alignment as a means to promote L2 speech learning.

The Present Study

Motivated by the above consideration, we intend to conduct a study to test how Chinese L2 English learners' belief about the interlocutor's English proficiency affects phonetic accommodation and to what extent phonetic convergence contributes to improvement in L2 vowel pronunciation. To answer these questions, the current study compares Chinese L2 English learners' vowel pronunciation with that of their interlocutor. The subjects' interlocutor is a female native American English speaker. We manipulated the participants' belief about their interlocutor's English proficiency by introducing her as a native English speaker from the U.S. to the experimental group and as a non-native English speaker from Switzerland to the control group. The native English speaker was not informed of this manipulation prior to or during the experiment.

Two English vowels /ɛ/ and /æ/ were chosen as the target, as they do not exist in the Mandarin phonetic system and Chinese L2 English learners tend to merge them into one phonological category (Thomson, 2008). Phonetic convergence is operationalized as the tendency for the subjects to pronounce the vowels in the way that is produced by the native English speaker, and it is measured by the Euclidean distance between the same token of a target vowel produced by a subject and the native speaker (Babel, 2012).

Previous studies of speech accommodation mainly adopted interaction tasks and shadowing tasks. Interaction tasks usually involve participants in information exchange tasks, just as map descriptions (Anderson et al., 1991) or picture descriptions (Van Engen et al., 2010). On the other hand, shadowing tasks ask participants to listen and repeat the target words after a model speaker (Babel & Bulatov, 2012; Goldinger, 1998; Goldinger & Azuma, 2004; Shockley et al., 2004). Though shadowing may capture the immediate effect of repetition and unravel the mechanism of the perception and production loop, this kind of decontextualized repetition is unnatural and uncommon in everyday communication, and it may not capture the socio-psychological motivations such as attractiveness or attitudes towards the talker.

In the present study, as we were interested in how L2 English learners' belief about their interlocutor's English proficiency may affect their conversational behavior, we opted for an interaction task, i.e. a picture description task which is similar to the map description

task used by Pardo (2006), because it helps to gauge the social psychological motivations of L2 phonetic accommodation in a more communicative context. As participants interact face to face, an interaction task may have high ecological validity (Felker et al., 2018). Additionally, repetition is a common phenomenon in communication, where interlocutors tend to repeat each other's words in conversational turns. A picture description task asks the participants to take turns asking each other questions, when they naturally repeat words and sentence patterns.

Two hypotheses are engendered from the research questions. First, according to IAT (Pickering & Garrod, 2004), it can be hypothesized that the participants in both the experimental group and the control group will converge phonetically to the native speaker irrespective of the manipulation of the subjects' belief about the interlocutor's English proficiency. However, if the participants in the experimental group demonstrate more phonetic convergence than those in the control group, the results will support CAT, which argues for a social and interpersonal motivation for convergence (Giles & Coupland, 1991). Second, according to the Speech Learning Model (Flege, 1995), which holds that the ability to learn sounds absent from one's native language will persist into adulthood, it is hypothesized that the subjects could improve their pronunciation of the target vowels due to phonetic convergence. The study may contribute to testing the theoretical models of IAT and CAT. In addition, an examination of the efficacy of phonetic convergence in the promotion of L2 English pronunciation may provide us with deeper insights into the nature and development of L2 speech learning.

Method

Experiment1

Participants Twenty visiting scholars from China (male $n = 8$, female $n = 12$) at Oklahoma State University (OSU) and a native American English speaker were recruited to participate in the study. The mean age of the subjects is 36.5 years ($SD = 2.40$). At the time of the study, the Chinese subjects have been studying at OSU for an average of 5.6 months ($SD = 2.35$) and they have all passed College English Test (Band 4) in China. Ten were randomly assigned to the experimental group (3 males and 7 females) and the other ten were assigned to the control group (5 males and 5 females). The native English speaker was a 21-year-old female student majoring in English at OSU. She had little experience talking with Chinese students on campus. In addition, no subject had known the native speaker personally prior to the study and none of the participants had speech, language or hearing disorders. The study had been approved by the Institutional Review Board at OSU, and the subjects were given a little gift (3 dollars' worth) as a reward for their participation.

Materials Eight target words for common objects were chosen, with 4 words containing the same vowel (/ɛ/ or /æ/), as shown in Table 1. Another 4 words were chosen as filler words. All the words are monosyllabic. Finally, the 12 objects were drawn at different locations in two pictures. The English words for the objects were not printed on the pictures.

In addition, a 5-point Likert scale questionnaire with 4 questions was designed, with "1" to "5" representing from "strongly disagree" to "strongly agree". The questions inquire about the subjects' willingness to sound like native English speakers (Question 1) or nonnative English speakers (Question 2) and preference to talk with native (Question 3) or nonnative English speakers (Question 4).

Table 1 Target vowels and words used in the picture description task

/ɛ/	/æ/	Filler words
Pet	Bat	Bird
Set	Cat	Cup
Bed	Map	Book
Web	Mat	Phone

Experimental Design A 2*2 mixed design is adopted, with the subjects' belief about their interlocutor's English proficiency (native proficiency vs. non-native proficiency) as the between-subject variable and vowel (/ɛ/ and /æ/) as the within-subject variable. The dependent variable is the formants (F1 and F2) of the vowels produced by the native speaker and the subjects in the first and second round of picture description.

Procedures The experiment was conducted in a cognitive science lab, where each participant talked with the native English speaker for approximately 5 min. Before the experiment, the native English speaker was introduced to the subjects in the experimental group as a native English speaker from America and to the subjects in the control group as a non-native English speaker from Switzerland. The subjects were also told a cover story for the experiment, i.e. to measure how speakers cooperate with each other in communication. First, the native speaker and a subject were given two pictures and they took turns asking questions about the location of each object only using the question prompt *"Where is …?"* and the answer prompt *"The … is on/in/at/behind/near…"* so that the target words could be embedded in each utterance they produced. The native speaker always asked the subject a question first. Then the subject answered that question and asked the native speaker the same question in return (Round 1). In this way, the native speaker acted as a model talker for the subject, who naturally repeated the target words after the native speaker in interaction. When they finished asking about all the objects in their pictures, they exchanged pictures and followed the same procedure again (Round 2). During the experiment, a microphone (CAD Audio U37) was placed approximately 30 cm from the subjects' mouth and their productions were digitally recorded through Audacity on a computer at a 44.1 K sampling rate and a 16-bit quantization rate. Immediately after the experimental session, each subject filled out the questionnaire. Finally, the subjects were debriefed about the research purpose and the manipulation, and they all expressed their understanding and support.

Speech analysis. During interaction, each subject and the native speaker produced 32 tokens of the target words (16 tokens containing /ɛ/ and 16 containing /æ/). First, acoustic analyses were performed on the speech samples using Praat (Boersma & Weenink, 2016). A span of 30 ms in the middle of the vowel on the spectrogram was first identified and the mean F1 and F2 values for the time window were obtained for all the target vowels produced by the subjects and the native speaker. The spectrogram for each target vowel was visually inspected for accuracy. In the case of a formant hallucination, we adjusted formant settings and changed the number of formants to be searched from the default 5 to 4. If the F1 and F2 were too close together (within 100 Hz), we adjusted the number of formants to be searched

Vowel formants	Token 1(SD)	Token 3(SD)	t (78)	p
/ɛ/ F1	701.50(53.98)	696.01(80.93)	0.36	0.72
F2	1806.67(133.22)	1782.01(96.90)	0.95	0.35
/æ/ F1	936.89(124.78)	881.32(115.14)	2.07	0.04*
F2	1709.23(94.39)	1681.05(85.36)	1.4	0.17

Table 2 Native speaker's vowel formants (in Hertz) in the experimental group

*An asterisk means it is statistically significant

from 5 to 6. These were done to guarantee the correctness of the formant measurement. We only considered modal voicing for the vowels and did not analyze cases of other phonation types such as creaky or breathy voice, as voice quality is usually correlated with multiple physiologic and psychological factors such as age, height, gender, mood, stress and health (Kreiman et al., 2005), which were difficult to control in this study. In addition, what we were interested in was how L2 learners adapted their vowel pronunciation in response to the vowel pronunciation of the native English speaker. Then, the F1 and F2 of the target vowels produced by all the participants were normalized following Labov's method (Labov et al., 2006) in order to normalize all speakers to a standard vowel space while preserving individual differences. We adopted Labov's method, because it is able to scale the original Hertz values as a part of its normalization process.

Next, in order to demonstrate whether the subjects' vowel pronunciation converged towards that of the native speaker as a result of auditory exposure and oral repetition, we followed the approach used by Babel (2012), who calculated the Euclidean distance between the same word produced by the subject and the model talker. In the present study, the Euclidean distance in the F1*F2 space was calculated for the same word produced by the subject and the native speaker in Round 1 (token 1) and Round 2 (token 3), because the subjects produced token 1 and token 3 immediately after the native speaker. The formula for measuring the Euclidean distance in F1 and F2 is[1]:

$$\sqrt{\left(W_N F1 - W_S F1\right)^2 + \left(W_N F2 - W_S F2\right)^2}$$

Specifically, the Euclidean distance (ED1) between the first token produced by the native speaker and the first token produced by the subject was calculated. Then the Euclidean distance (ED2) between the third token produced by the native speaker and the third token produced by the subject was calculated. Finally, ED1 was subtracted from ED2, which resulted in the difference in Euclidean distance between the baseline and after repetition. The difference in the Euclidean distance is the dependent measure in the statistical analysis. Three possibilities may be obtained, a negative value, a positive value or a zero value. A negative value suggests a decrease in phonetic distance, which indicates convergence; a positive value suggests an increase in phonetic distance, which implies divergence; a zero value means no change in phonetic distance, which demonstrates speech maintenance.

---

[1] W = word, N = native speaker, S = subjects.

Table 3 Subjects' vowel formants (in Hertz) in the experimental group

Vowel formants	Token 1(*SD*)	Token 3(*SD*)	*t* (78)	*p*
/ɛ/ F1	790.26(131.33)	720.31(108.82)	2.6	0.01*
F2	1814.46(245.75)	1776.8(196.62)	0.76	0.45
/æ/ F1	811.22(139.92)	823.1(89.8)	-0.45	0.65
F2	1822.71(195.87)	1733.59(142.44)	2.33	0.02*
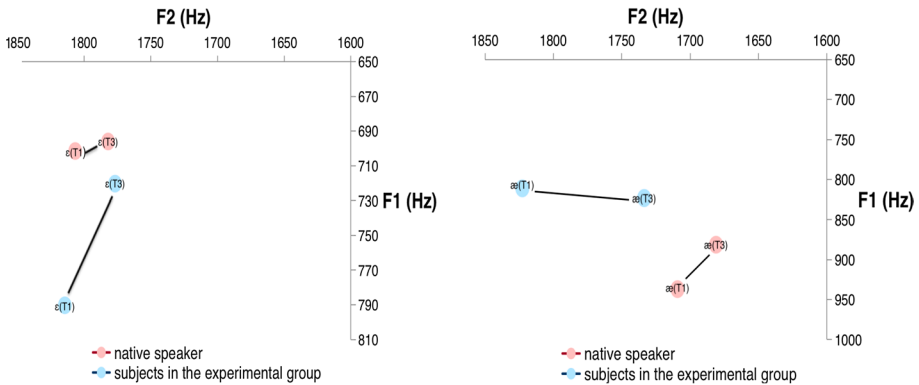
Fig. 1 Vowel distance change for the subjects in the experimental group for /ɛ/(left) and /æ/(right) (T1=token 1, T3=token 3)

Results First, following Babel (2012) the F1 and F2 values which were more than 3 standard deviations away from the mean were identified as outliers and removed from the data set. In total, 12 outliers were identified and removed from the final data set. The F1 and F2 of /ɛ/ and /æ/ for the native speaker and the subjects in the experimental group are shown in Tables 2 and 3:

As shown in Table 2, except for a marginally significant change in F1 for /æ/ from token 1 to token 3, the native speaker's vowel formants did not differ significantly from the token 1 to token 3, indicating that the native speaker' pronunciation of /ɛ/ and/æ/ remained stable throughout the experiment.

As shown in Table 3, for /ɛ/ produced by the subjects in the experimental group, F1 decreased significantly from token 1 to token 3. F2 also decreased from token 1 to token 3, though it was not a significant decrease. For /æ/, F2 decreased significantly from token 1 to token 3. F1 increased from token 1 to token 3. However, it was not a significant increase.

To better demonstrate the vowel distance shift, data from Tables 2 and 3 are exhibited in Fig. 1 below:

Figure 1 illustrates the vowel distance between the native speaker and the subjects for token 1 and token 3. As shown in Fig. 1, the native speaker's production of the two vowels did not differ significantly from token 1 to token 3. For the subjects, their pronunciation of the two vowels became closer to that of the native speaker after repetition. This result suggests that after hearing the target words produced by the native speaker for 3 times, the subjects converged towards the native speaker's pronunciation and improved their vowel pronunciation.

Table 4 Native speaker's vowel formants (in Hertz) in the control group

Vowel formants	Token 1(*SD*)	Token 3(*SD*)	*t*(78)	*p*
/ɛ/ F1	712.53(66.13)	708.67(56.24)	0.28	0.78
F2	1780.29(88.01)	1767.67(110.17)	0.57	0.57
/æ/ F1	844.73(63.86)	866.05(96.8)	1.02	0.31
F2	1662.87(65.44)	1650.74(71.64)	0.79	0.43

Table 5 Subjects' vowel formants (in Hertz) in the control group

Vowel formants	Token 1(*SD*)	Token 3(*SD*)	*t*(78)	*p*
/ɛ/ F1	722.33(134.87)	713.28(105.98)	-0.33	0.74
F2	1823.75(213.42)	1822.09(301.21)	0.3	0.98
/æ/ F1	800.19(116.12)	782.37(140.79)	0.62	0.54
F2	1844.37(200.31)	1819.10(297.51)	0.45	0.66
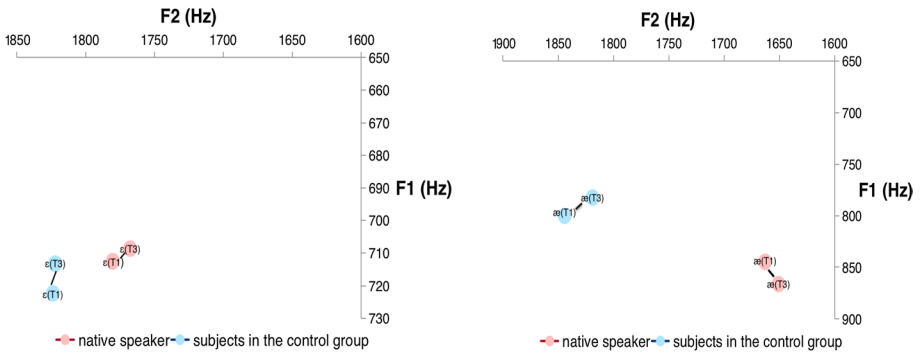
Fig. 2 Vowel distance change for the subjects in the control group for /ɛ/(left) and /æ/(right) (T1 = token 1, T3 = token 3)

Next, the F1 and F2 of /ɛ/ and /æ/ for the native speaker and the subjects in the control group are displayed in Tables 4 and 5:

As shown in Table 4, the F1 and F2 for both /ɛ/ and /æ/ did not change significantly from token 1 to token 3, indicating that the native speaker's target vowel pronunciation remained stable across the two tokens.

As shown in Table 5, the F1 and F2 for both /ɛ/ and /æ/ did not change significantly from token 1 to token 3, indicating that the subjects' target vowel pronunciation remained relatively stable across the two tokens.

To better illustrate the vowel distance shift, data from Tables 4 and 5 are plotted in Fig. 2 below:

As shown in Fig. 2(left), the subjects' F1 and F2 for /ɛ/ were similar to those of the native speaker. However, after hearing the target words containing /ɛ/ for 3 times, the subjects' F1 and F2 for /ɛ/ even diverged a little from the native speaker. For /æ/, the subjects' F1 and F2 remained almost unchanged from token 1 to token 3.

Then, Euclidean distance was calculated for /ɛ/ and /æ/ in the experimental group and control group. In the experimental group, the participants' average difference in distance

for /ɛ/ and /æ/ between the first repetition and the second repetition (after hearing the same tokens spoken by the native speaker for 3 times) are -110.22 ($SD = 124.00$) and -113.50 ($SD = 149.81$) respectively. A one-sample t-test shows that the differences in distance for /ɛ/ and /æ/ significantly differ from 0 in the experimental group ($t(9) = -5.84$, $p = 0.00$, $t(9) = -5.89$, $p = 0.00$). As a negative value implies a decrease in phonetic distance, the result suggests that the subjects in the experimental group converged to the vowel pronunciation of the native English speaker, so that the subjects' vowel pronunciation became similar to that of the native speaker.

In the control group, the average difference in distance for /ɛ/ and /æ/ between the first repetition and the second repetition (after hearing the same token spoken by the native speaker for 3 times) are 62.48 ($SD = 141.44$) and 78.70 ($SD = 116.05$) respectively. A one-sample t-test shows that the differences in distance for /ɛ/ and /æ/ significantly differ from 0 in the control group ($t(9) = 3.05$, $p = 0.01$; $t(9) = 3.77$, $p = 0.00$). As a positive value suggests increase in phonetic distance, the result implies that the subjects in the control group did not converge, but diverged, though not much,[2] from their interlocutor so that the subjects' vowel pronunciation did not become similar to that of the native speaker.

Next, a mixed model ANOVA was performed on the data, with the subjects' belief about the interlocutor's English proficiency (native vs. non-native) and vowel (/ɛ/ vs. /æ/) as independent variables and differences in Euclidean distance as the dependent variable. The results show a significant main effect for the subjects' belief about the interlocutor's English proficiency, $F(1,18) = 60.66$, $p < 0.001$, partial $\eta^2 = 0.791$, which indicates that the subjects in the experimental group converged towards the native speaker interlocutor ($M = -111.86$, $SD = 58.72$), while the subjects in the control group diverged from the "non-native speaker" ($M = 67.86$, $SD = 56.17$). The main effect for vowel is not significant, $F(1,18) = 0.01$, $p > 0.05$, partial $\eta^2 = 0.001$. Finally, no interaction is found between the subjects' belief about the interlocutor's English proficiency and vowel, $F(1,18) = 0.30$, $p > 0.05$, partial $\eta^2 = 0.019$.

Finally, results from the questionnaire analysis show that the subjects in both the experimental group ($M = 4.7$, $SD = 0.48$) and the control group ($M = 4.8$, $SD = 0.42$) were more willing to sound like native English speakers and there was no significant difference between the two groups ($t(18) = -0.49$, $p > 0.05$), indicating that both groups had a higher level of preference for learning pronunciation from native English speakers rather than non-native English speakers. In addition, all of the 20 participants voiced preference to talk to native English speakers instead of non-native English speakers.

Discussion

The first aim of the current study is to explore whether the subjects' belief about the interlocutor's English proficiency mediates phonetic accommodation. Results from the data analysis show a robust effect of interlocutor identity manipulation on phonetic accommodation, when the same research confederate talks with two randomly assigned groups of subjects. Overall, the subjects to whom the confederate was introduced as a native English speaker converged to the confederate's vowel pronunciation, while the subjects to whom the confederate was introduced as a non-native English speaker showed no convergence except for only one subject (No. 5), who converged to the /ɛ/ pronunciation of

---

[2] F1 and F2 usually vary even for the same words spoken by the same speakers at different times of a day.

the "non-native speaker". This result is in line with the Communication Accommodation Theory (CAT) (Giles et al., 1991), which argues for a social and interpersonal motivation for convergence. According to CAT, in order to achieve their communicative goal more easily and gain social approval, people tend to accommodate to their interlocutors at all levels of speech, phonetically, lexically and syntactically (Giles et al., 1991). As shown by the questionnaire results, the subjects unanimously preferred to learn English pronunciation from native English speakers rather than from non-native English speakers. When the subjects in the experimental group knew that they would talk to a native English speaker, they might be sufficiently motivated to imitate the way words were produced by the native speaker. While to the subjects in the control group, their lack of interest in their interlocutor may have demotivated them in imitating her English pronunciation, resulting in speech divergence.

The second aim is to test whether speech convergence could induce L2 learners to produce more native-like vowels. As illustrated in Fig. 1, in the experimental group, the subjects' baseline productions of /ɛ/ and /æ/ were almost overlapping, suggesting that Chinese L2 English learners tended to conflate the two vowel categories, as found in previous studies (e.g. Thomson, 2008). In the current study, after listening to the same token of words produced by the native English speaker 3 times and repeating them 2 times, the subjects in the experimental group converged to the speaker's pronunciation, and separated the two vowels, resulting in a more native like vowel space.

One limitation of Experiment 1 is the small sample size (10 participants in each group) due to the availability of Chinese visiting scholars on campus, thus making the effect size not robust and reliable enough. In addition, though we found the participants converged towards the native speaker while diverged from the "nonnative speaker", we were not sure whether this effect would still have held if the research confederate had been introduced as "Australian" versus "Swiss", since the participants were studying in America and were more familiar with Americans. Before we can draw any reliable conclusion, we need to prove that the results obtained in Experiment 1 were not due to another construct, i.e. familiarity with the speaker's accent. Therefore, we conducted a second experiment.

Experiment 2

Experiment 2 aims to delineate the effect of speaker accent familiarity from speaker identity with a similar design with Experiment 1. In Experiment 2, a research confederate, a native American English speaker was introduced either as a native American English speaker or as a Swiss English speaker to two groups of college English learners at a university in Northeastern China. The research confederate was a female teacher (aged 27) newly employed at the university and she was not familiar to any of the participants. Sixty college freshmen (25 males and 35 females) at the same university were pooled to participate in the study. The average age of the Chinese students is 18.23 ($SD = 0.43$). They major in Chinese, physics, math, geography and education. Their average English score in the national college entrance exam (Gaokao) is 128.82 ($SD = 4.34$) (full score: 150) and they reported no or very little prior experience interacting with English speaking people. The subjects have been mainly exposed to American English and British English through textbook recordings and other listening materials at school. The subjects were randomly assigned to an experimental group ($n = 30$) and a control group ($n = 30$). The materials,

Table 6 Native speaker's mean vowel formants (in Hertz) in the experimental group

Vowel formants	Token 1(*SD*)	Token 3(*SD*)	*t*(29)	*p*
/ɛ/ F1	707.36(37.97)	719.30(39.75)	-1.82	0.08
F2	1719.19(37.56)	1728.61(28.11)	-2.00	0.06
/æ/ F1	820.00(52.64)	835.94(36.23)	-1.88	0.07
F2	1712.17(90.60)	1699.43(79.12)	1.22	0.23

Table 7 Subjects' mean vowel formants (in Hertz) in the experimental group

Vowel formants	Token1(*SD*)	Token 3(*SD*)	*t*(29)	*p*
/ɛ/ F1	734.61(29.57)	719.17(32.42)	2.20	0.04
F2	1739.78(42.36)	1713.82(18.81)	3.22	0.00
/æ/ F1	779.80(52.42)	821.38(35.97)	-4.75	0.00
F2	1656.24(90.13)	1692.54(75.57)	-2.95	0.00

Table 8 Native speaker's mean vowel formants (in Hertz) in the control group

Vowel formants	Token1(*SD*)	Token 3(*SD*)	*t* (29)	*p*
/ɛ/ F1	718.76(21.52)	726.38(18.91)	-1.81	0.08
F2	1713.78(41.33)	1720.29(22.71)	-0.86	0.40
/æ/ F1	798.50(25.00)	796.60(32.47)	0.30	0.77
F2	1780.96(33.39)	1788.66(56.72)	-1.01	0.32

procedures, experimental design and method for data analysis used in Experiment 2 are identical with those in Experiment 1.

Results

First, consistent with the approach in Experiment 1, the F1 and F2 values which were more than 3 standard deviations away from the mean were identified as outliers and removed from the data set. In total, 17 outliers were identified and removed. The F1 and F2 of /ɛ/ and /æ/ for the native speaker and the subjects in the experimental group are shown in Tables 6 and 7:

As exhibited in Table 6, for both /ɛ/ and /æ/, the native speaker's F1 and F2 did not change significantly from token 1 to token 3, suggesting that the native speaker's pronunciation remain stable throughout the experiment.

As shown in Table 7, the subjects in the experimental group significantly decreased their F1 and F2 for /ɛ/, while significantly increased their F1 and F2 for /æ/. The mean Euclidean distance for /ɛ/ and /æ/ between the first repetition and the second repetition are -40.16 (*SD* = 33.92) and -66.29 (*SD* = 40.20) respectively. A one-sample *t*-test shows that the differences in distance for /ɛ/ and /æ/ significantly differ from 0 in experimental group ($t(29) = -6.49$, $p = 0.00$; $t(29) = -9.03$, $p = 0.00$). The decrease in the Euclidean distance attests to the fact that the subjects in the experimental group converged to the vowel pronunciation of the native English speaker.

Table 9 Subjects' mean vowel formants (in Hertz) in the control group

Vowel formants	Token1(*SD*)	Token 3(*SD*)	*t* (29)	*p*
/ɛ/ F1	709.67(35.33)	757.65(26.33)	-7.69	0.00
F2	1726.20(34.67)	1756.69(25.47)	-3.89	0.00
/æ/ F1	821.51(30.43)	839.11(35.23)	-2.40	0.02
F2	1807.52(29.60)	1825.36(49.17)	-2.24	0.03

Secondly, the F1 and F2 of /ɛ/ and /æ/ for the native speaker and the subjects in the control group are shown in Tables 8 and 9:

As displayed in Table 8, the native's F1 and F2 for both /ɛ/ and /æ/ did not differ significantly between token 1 and token 3, suggesting that the native speaker's vowel pronunciation remained relatively stable throughout the experiment.

As illustrated in Table 9, the subjects in the control group significantly increased their F1 and F2 for both /ɛ/ and /æ/. The mean Euclidean distance for /ɛ/ and /æ/ between the first repetition and the second repetition are 11.12 ($SD=24.32$) and 18.79 ($SD=21.45$) respectively, indicating that the subjects in the control group widened their difference in vowel pronunciation from the native speaker. A one-sample t-test shows that the differences in distance for /ɛ/ and /æ/ significantly differ from 0 in control group ($t(29)=2.50$, $p=0.02$; $t(29)=4.80$, $p=0.00$).

A mixed model ANOVA was performed on the data, with the subjects' belief about the interlocutor's English proficiency and vowel as independent variables and differences in Euclidean distance as the dependent variable. The results show a significant main effect for speaker identity, $F(1,58)=102.46$, $p<0.001$, partial $\eta^2=0.639$. There is a significant difference in the direction of speech accommodation. Specifically, the subjects in the experimental group converged to the research confederate ($M=-53.22$, $SD=39.15$), while the subjects in the control group diverged from the research confederate ($M=14.96$, $SD=23.07$). There is a significant main effect for vowel, $F(1,58)=4.66$, $p=0.035$, partial $\eta^2=0.074$. /æ/ ($M=-23.75$, $SD=53.48$) is subject to more convergence than /ɛ/ ($M=-14.52$, $SD=39.05$). There is a significant interaction between the subjects' belief about the interlocutor's English proficiency and vowel, $F(1,58)=15.64$, $p<0.001$, partial $\eta^2=0.212$. Simple effect analysis shows that in the experimental group, /æ/ ($M=-66.29$, $SD=40.19$) is subject to more convergence than /ɛ/ ($M=-40.16$, $SD=33.92$), $t(29)=4.21$, $p<0.001$; In the control group, both /ɛ/ ($M=11.12$, $SD=24.32$) and /æ/ ($M=18.79$, $SD=21.46$) are subject to divergence and there is no significant difference between them, $t(29)=-1.31$, $p=0.201>0.05$.

Finally, the same questionnaire from Experiment 1 was administered to the subjects in Experiment 2 and yielded similar findings to Experiment 1. Results show that the subjects in both the experimental group ($M=4.8$, $SD=0.41$) and the control group ($M=4.83$, $SD=0.38$) voiced preferences to sound like native rather than nonnative speakers of English and there is no significant difference between the two groups, $t(58)=-0.33$, $p>0.05$. Taken together, the results from the questionnaire reflect a psychological preference for L2 learners of English to learn pronunciation from native English speakers.

Discussion

In order to test if the results obtained in Experiment 1 were confounded by another construct, i.e. familiarity with the speaker's accent, we conducted Experiment 2. The subjects in Experiment 2 have been exposed to English input mainly from CD or tape recordings by native English speakers from Britain and America and they were not familiar with the research confederate, either. Results from Experiment 2 suggest that the subjects converged towards the "native speaker" rather than the "nonnative speaker". This result replicates the findings in Experiment 1, indicating a tendency for Chinese L2 English learners to converge to native English speakers and this effect is not confounded by the fact that the subjects in Experiment 1 were all studying at an American university and, therefore, were more familiar with the American English accent or identify more with Americans.

However, in Experiment 2 not every subject in the experimental group converged to the vowel pronunciation of the native speaker. For /ɛ/, the subjects' Euclidean distance ranges from -105.11 to 20.00. Even though the majority (26 subjects out of 30) of the subjects decreased their Euclidean distance from the native speaker, still 3 subjects (No. 4, No. 6, No. 25, No. 26) increased their Euclidean distance. While for /æ/, the subjects' Euclidean distance ranges from -158.64 to 15.86. Even though the vast majority (29 subjects out of 30) of the subjects decreased their Euclidean distance from the native speaker, still 1 subject (No.6) increased her Euclidean distance. We hypothesize this result is obtained due to the subjects' different phonetic talent or psychological traits, which were found to impact the degree of phonetic convergence (Lewandowski, 2011; Lewandowski & Jilka, 2019). In the present study, the fact that we did not measure the subjects' phonetic talent or psychological traits makes it only possible for us to speculate that the reason why some subjects in the experimental group did not converge towards the native speaker may be accounted for by their differential phonetic talent or psychological traits. Those who did not converge probably have less phonetic talent or are more conservative.

Similarly, not every subject in the control group diverged from the vowel pronunciation of the native speaker. For /ɛ/, the subjects' Euclidean distance ranges from -37.69 to 71. 95. Even though the majority (20 subjects out of 30) of the subjects increased their Euclidean distance from the native speaker, still 10 subjects decreased their Euclidean distance. While for /æ/, the subjects' Euclidean distance ranges from -14.22 to 72.27. Even though a majority (23 subjects out of 30) of the subjects increased their Euclidean distance from the native speaker, still 7 subjects decreased their Euclidean distance. Especially for subjects No. 14 and No. 16, they decreased their Euclidean distance both for /ɛ/ and /æ/.

Taken together, the above findings lead us to believe these results point to a very important realization that the automatic and social-psychological mechanisms of phonetic convergence are not mutually exclusive, i.e. phonetic convergence may occur automatically in interaction (Pickering & Garrod, 2004) and the extent of convergence is moderated by social-psychological factors (Giles & Coupland, 1991). According to IAT, interlocutors need to align their language at the lexical, syntactic and phonological levels to reach a common situation model or "common ground" in the course of interaction in order to have successful communication, i.e. alignment or convergence functions as an implicit and automatic mechanism. However, the fact that a majority of subjects in the control group diverged phonetically from their interlocutor runs counter to the implicit and automatic account. Thus, we believe phonetic convergence might occur both as a bottom-up process, which is a natural response to the auditory information in the speech stream and as a top-down process, which receives influences from other factors, be they social, psychological,

or cognitive. For the handful of subjects in the control group who converged towards their interlocutor despite their willingness to learn pronunciation from native English speakers rather than non-native speakers, we are not yet sure whether they converged because they consciously chose to do so or because the automaticity of convergence was simply more powerful to be subdued.

General Discussion

Findings from Experiment 1 and Experiment 2 consistently demonstrated a tendency for Chinese L2 English learners to converge phonetically to native English speakers rather than to non-native speakers. As demonstrated by the questionnaire findings, the subjects expressed preference for native English pronunciation over nonnative English pronunciation. Indeed, L2 learners with native-like pronunciation are usually associated with higher intelligence and greater credibility (Lev-Ari & Keysar, 2010) and tend to be more readily accepted as in-group members rather than out-group members. This recognition may motivate L2 learners of English to converge to native English speakers while diverge from non-native English speakers.

The current results are in line with Zając and Rojczyk (2014), who also found in his study that English learners tended to converge phonetically to native English speakers rather than to non-native English speakers, though different parameters (vowel duration in their study and vowel formant in the current study) are used for comparison. Overall, the results support findings from previous research indicating that psychological factors mediate the direction and magnitude of speech accommodation (e.g. Abrego-Collier et al., 2011; Yu et al., 2013).

One thing to note is that the subjects in Experiment 1 and 2 are entirely different in terms of their age, L2 experience and proficiency. In Experiment 1, the subjects were studying at an American university with higher L2 proficiency and more exposure to native L2 input compared to the subjects in Experiment 2, who were studying at a Chinese university with lower proficiency and less exposure to L2 input. Despite the subjects' different background, they all expressed preference for native English speakers over non-native English speakers. The effect of the subjects' belief about the interlocutor's English proficiency might have resulted from the Halo Effect or preferential processing (Namy et al., 2002), which is closely related to stereotyping. The Halo Effect (Nisbett & Wilson, 1977) refers to the phenomenon that people tend to associate more merits with things or people they already like. When the research confederate was introduced differently to the subjects in the two groups, the stereotype that non-native English speakers cannot speak English as well as native speakers, especially in terms of pronunciation, might be activated, which results in different levels of preference and motivation.

In addition, baseline findings in Experiment 1 suggest that Chinese L2 English learners did not distinguish /ɛ/ and /æ/, as their F1 and F2 for the two vowels were almost overlapping with each other. However, the findings from both experiments suggest that the subjects' L2 vowel pronunciation became more native-like through phonetic convergence, though we don't know if this improvement will persist after some delay. The fact that the adult L2 speakers could modify the vowel space after a short interaction with a native speaker attests to the Speech Learning Model (SLM) (Flege, 1995). According to SLM, L2 learners possess the ability to learn sounds not existent in their native phonetic repertoire by constructing new L2 speech categories. The more similarities an L2 sound bears with an L1 sound, the more difficult it is for L2 learners to construct a new sound category.

Consequently, L2 learners will assimilate the L2 sound into a similar phonetic category in the L1. In addition, this speech learning competence persists throughout L2 learners' life. In the present study, the English vowels /ɛ/ and /æ/ do not exist in the Mandarin Chinese phonetic repertoire. This relative ease for new speech category formation enabled the subjects to modify their vowel space during interaction with a native English speaker, even when the subjects have all entered adulthood.

token together, the finding that L2 learners of English demonstrate similar vowel spectra after interacting with native English speakers testifies to the proposal that phonetic convergence can be an effective approach in promoting L2 phonetic acquisition (Trofimovich, 2013). As convergence could take place without social motivations (Namy et al., 2002) or within mere contact and after a short period of non-native to native interaction, it may be advisable for L2 learners to be exposed to more native rather than nonnative input both in and out of class. In addition, it may be speculated that long-term immersion programs in a native language environment might better assist L2 learners in improving their L2 pronunciation. In native to nonnative interactions, phonetic convergence may help L2 learners modify their pronunciation and reduce foreign accent. However, the current study does not examine whether the improvement in L2 vowel pronunciation could carry over to other speech styles, or how long it could sustain. It remains a promising field for future studies to investigate the long-term effect of phonetic convergence on the improvement of L2 pronunciation.

Another limitation of the present study is not taking the subjects' personality type (extroversion or introversion) or psychological traits into account, as they may mediate the result of phonetic accommodation (Lewandowski & Jilka, 2019). Moreover, future work is also needed to investigate the mechanism of phonetic convergence and its role in L2 pronunciation learning. Especially for the large number of L2 English learners who do not have opportunities to interact with native English speakers, it is crucial to investigate whether they could improve their English pronunciation through massive exposure to native English input from multimedia or from interacting with native English speakers online.

Conclusion

Overall, the present study reveals that L2 learners' belief about an interlocutor's English proficiency mediates phonetic accommodation and interacting with a native English speaker enables the subjects to improve their English vowel pronunciation through phonetic convergence. The results of the study support a more social-psychological based account of speech accommodation and provide evidence for the SLM. The findings might contribute to our understanding of the role social psychological factors play in phonetic convergence and the potentials of utilizing phonetic convergence in L2 pronunciation learning. However, future work is needed to investigate the long-term effect of phonetic convergence on the improvement of L2 pronunciation and more complex research designs are needed to tap into the potential influence of personality and psychological traits on phonetic convergence.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Informed consent Informed consent was obtained from all the participants included in the study.

References

Abrego-Collier, C., Grove, J., Sonderegger, M., & Yu, A. (2011). Effects of speaker evaluation on phonetic convergence. In *Proceedings of the international congress of phonetic sciences XVII* (pp. 192–195). Hong Kong, International Congress of Phonetic Sciences.

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., & Sotillo, C. (1991). The Hcrc map task corpus. *Language and Speech, 34*(4), 351–366.

Babel, M. (2009). *Phonetic and social selectivity in speech accommodation*. PhD Dissertation of University of California, Berkeley.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*, 177–189.

Babel, M. E., & Bulatov, D. (2012). The role of fundamental frequency in phonetic accommodation. *Language and Speech, 55*, 231–248.

Berry, G. M., & Ernestus, M. (2018). Phonetic alignment in English as a lingua franca: coming together while splitting apart. *Second Language Research, 34*, 343–370.

Boersma, P., & Weenink, D. (2016). Praat, v. 6.0.2.1. http://www.praat.org.

Coupland, N., & Giles, H. (1988). Introduction: the communicative contexts of accommodation. *Language & Communication, 8*(3–4), 175–182.

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes, 23*, 528–556.

Felker, E., Troncoso-Ruiz, A., Ernestus, M., & Broersma, M. (2018). The ventriloquist paradigm: Studying speech processing in conversation with experimental control over phonetic input. *The Journal of the Acoustical Society of America, 144*(4), EL304–EL309.

Flege, J. (1995). Second-language speech learning, theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience, issues in cross-language research* (pp. 233–277). York Press.

Gessinger, I., Schweitzer, A., Andreeva, B., Raveh, E., Mobius, B., & Steiner, I. (2018). Convergence of pitch accents in a shadowing task. In *International Conference on Speech Prosody* (pp. 225–229).

Goldinger, S. D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review, 11*(4), 716–722.

Giles, H., & Coupland, N. (1991). *Language, contexts and consequences*. Open University Press.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory, communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation, developments in applied sociolinguistics* (pp. 1–68). Cambridge University Press.

Goldinger, S. D. (1997). Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 33–66). Academic Press.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279.

Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior, 21*, 23–43.

Gregory, S. W., Green, B. E., Carrothers, R. M., Dagan, K. A., & Webster, S. (2000). Verifying the primacy of voice fundamental frequency in social status accommodation. *Language and Communication, 21*, 37–60.

Gregory, S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology, 70*, 1231–1240.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Journal of Laboratory Phonology, 2*, 125–156.

Kreiman, J., Vanlancker-Sidtis, D., & Gerratt, B. (2005). Perception of voice quality. In D. Pisoni & R. Remez (Eds.), *Handbook of speech perception* (pp. 338–362). Blackwell.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English, phonetics, phonology, and sound change*. Mouton de Gruyter.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology, 6*, 1093–1096.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. In *INTERSPEECH*-2011 (pp. 3081–3084).

Lewandowski, N. (2011). *Talent in nonnative phonetic convergence*. PhD Dissertation, Universität Stuttgart.

Lewandowski, N., & Jilka, M. (2019). Phonetic convergence, language talent, personality and attention. *Frontiers in Communication, 4*, 1–18.

Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal and Language and Social Psychology, 21*, 422–432.

Nielsen, K. Y. (2007). Implicit phonetic imitation is constrained by phonemic contrast. In *Proc. 16th ICPhS* (pp. 1961–1964).

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect, Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250–256.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of Acoustic Society of America, 119*, 2382–2393.

Pardo, J. S. (2010). Expressing oneself in conversational interaction. In E. Morsella (Ed.), *Expressing oneself/expressing one's self, communication, cognition, language, and identify* (pp. 183–196). Lawrence Erlbaum Associates.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics, 40*, 190–197.

Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception & Psychophysics, 72*(8), 2254–2264.

Pickering, M. J., & Garrod, S. (2004). towards a mechanistic psychology of dialogue. *Behavioural & Brain Sciences, 27*, 169–226.

Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception & Psychophysics, 66*(3), 422–429.

Thomson, R. I. (2008). L2 English vowel learning by Mandarin speakers: Does perception precede production. *Canadian Acoustics, 36*(3), 134–135.

Trofimovich, P. (2013). Interactive alignment, Implications for the teaching and learning of second language pronunciation. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–9). Iowa State University Press.

Trofimovich, P., & Kennedy, S. (2014). Interactive alignment between bilingual interlocutors, evidence from two information-exchange tasks. *Bilingualism, Language and Cognition, 17*(4), 822–836.

Trudgill, P. (1986). *Dialects in contact (Language in Society)*. Oxford: Blackwell.

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R. (2010). The wildcat corpus of native- and foreign-accented English: communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech, 53*(Pt 4), 510–540.

Walker, A., & Campbell-Kibler, K. (2015). Repeat what after whom? Exploring variable selectivity in a cross-dialectal shadowing task. *Frontiers in Psychology, 6*, 1–18.

Young, R. (1988). Variation and the interlanguage Hypothesis. *Studies in Second Language Acquisition, 10*, 281–302.

Yu, A. C., Abrego-Collier, C., & Sonderegger, M. (2013). Phonetic imitation from an individual-difference perspective, Subjective attitude, personality and 'autistic' traits. *PLoS ONE, 8*(9), 1–13.

Zając, M., & Rojczyk, A. (2014). Imitation of English vowel duration upon exposure to native and non-native speech. *Poznan Studies in Contemporary Linguistics, 50*(4), 495–514.