Article

Language *and* Speech

Language *and* Speech
1–31

DOI: 10.1177/0023830917737111
journals.sagepub.com/home/las
SAGE

Effects of Exposure and Vowel Space Distribution on Phonetic Drift: Evidence from American English Learners of French

Benjamin Lang
New York University, USA

Lisa Davidson
New York University, USA

Abstract

Recent work by Chang has shown that even at the very earliest stages of second language (L2) acquisition, the phonetic implementation of speakers' native English phoneme categories is slightly modified by contact with L2 Korean, which is referred to as "phonetic drift." This study investigates whether rapid phonetic drift generalizes to another pairing of languages. We examined naïve American English learners of French, who were recorded producing both American English and French vowels after one and six weeks of a study abroad program in Paris. In addition, the Study Abroad group is compared with proficient American English L1 speakers of French who have been residents of Paris for at least five years, to investigate the impact of long-term use of an L2 on the vowel categories of L1. Whereas the Study Abroad group showed no evidence of phonetic drift after six weeks, the Paris Residents' American English vowel space shifted along F1 and several English vowels demonstrated clear movement toward French monolingual norms. A closer look at the high vowels provides insight into how phonetic categories are influenced both by drift and by a pressure to keep vowel categories distinct between the languages. The results are also discussed with respect to potential effects of the size of the vowel inventory and the amount of input required to cause phonetic drift.

Keywords

Second language acquisition, phonetic drift, vowels, speech production

Corresponding author:
Lisa Davidson, Linguistics, New York University, Department of Linguistics, 10 Washington Place, New York, NY 10012, USA.
Email: lisa.davidson@nyu.edu

Introduction

1.1 Phonetic drift over shorter and longer time periods

Deriving from the critical period hypothesis (e.g., Lenneberg, 1967; Patkowski, 1990; Scovel, 1988), language acquisition research often focuses on age effects in acquiring the phonetic categories of the second language (L2). As a speaker ages past a certain point, certain phonemes and phonotactic combinations become more difficult to produce and perceive accurately if they are not part of the first language (L1). This limitation contributes to the assumption that the properties of the L1 retain a substantial influence over how a second language is learned and produced, and may even lead to the assumption that a speaker's L1 is essentially unchanging after a certain age in a speaker's life.

However, more recent research has shown that the phonetic categories of both L1 and L2 are to some extent malleable, and that the implementation of L1 categories can be affected by experience with an L2 (Bergmann, Nota, Sprenger, & Schmid, 2016; Chang, 2012, 2013; de Leeuw, Mennen, & Scobbie, 2012; de Leeuw, Schmid, & Mennen, 2010; Flege, 1987, 2003; Flege & Hillenbrand, 1984; Mayr, Price, & Mennen, 2012; Peng, 1993; Ulbrich & Ordin, 2014). One term that has been used to describe the effect of L2 phonetic properties on the realization of L1 categories is "phonetic drift," as introduced by Chang (2012). Phonetic drift occurs when the establishment of L2 phonetic categories evokes a shift in the production of L1 phonetic categories along some acoustic measure. Chang (2012, 2013) argues that this influence is pervasive and can shape speech production in L1 broadly, even from minimal contact.

Chang (2012) investigated phonetic drift in American English speakers learning elementary L2 Korean by assessing their consonant and vowel productions of American English and Korean weekly for the first five weeks of immersive second language study in Korea. Results demonstrated that the entire vowel space of American English speakers experiences a generalized acoustic shift in F1 in the direction of the vowel space of Korean after five weeks of immersive contact with Korean. Small increases in voice onset time (VOT) were also observed in American English voiceless stop aspiration over time. These VOTs appeared to move in the direction of Korean aspirated stops, which also lengthened in VOT significantly, arguably "over-aspirating" beyond native Korean norms (presumably to differentiate themselves from the lengthening VOT of American English voiceless stops).

While Chang's research is novel for showing that there can be very rapid effects of phonetic drift after short periods of immersion, several other studies have established that in learners who have had longer and more extensive contact with their L2, some L1 sounds indicate the effect of phonetic drift while others seem to remain unchanged. In a paper that is particularly relevant to the current study, Flege (1987) examines English formant values for high vowels and VOT in native American English learners of French. Learners with an average of 11.7 years of living in Paris produced American English /t/ with considerably shorter, French-like, VOT values than American English monolinguals. The investigation of production differences for American English /u/, French /u/, and French /y/ in the same groups of native speakers of American English with varying levels of experience in French showed a different pattern, however. Results showed that regardless of their experience with French, from just nine months of study abroad experience to teaching French in the US to an average of 11 years in France speaking French as their primary language, the F2 values for American English /u/ did not significantly differ from monolingual American English F2 standards. As for French /u/, the students and teachers produced a vowel that was not significantly different from American English /u/, but the American residents of Paris did begin to move toward French /u/ (though they did not reach the monolingual standard). Flege concludes

that the groups with less experience and/or input seem to be equating the French /u/ vowel with their American English category. The most experienced group shows a difference between American English and French /u/, and they achieve separation by lowering the French F2 toward French norms while their American English category does not change relative to American English monolinguals.

Other studies have also shown evidence of phonetic drift as a result of long-term language contact. Russian L1 speakers in Dmitrieva, Jongman, and Sereno (2010) with long exposure to American English in an L2 environment demonstrated that the lack of word-final devoicing in American English led to a significant difference in vowel length and voicing into closure duration between devoiced variants of voiced stops and voiceless stops in their L1. Mayr et al. (2012) studied Dutch-speaking twins, one of whom moved to an English-speaking environment at age 32. Compared with the monolingual twin, the bilingual twin continued to produce a voicing lead for voiced stops that is more like Dutch, but VOT values for the voiceless stops that were intermediate between English and Dutch values. For vowels, the L1 Dutch vowel space had shifted in the direction of her L2 English. A production study with Spanish-Quechua bilinguals (Guion, 2003) showed that simultaneous, early, and some mid-bilinguals (i.e., L2 learners who began learning around 11 years of age) produced distinct Spanish and Quechua vowels, though mid-L2 bilinguals produced monolingual-like Spanish vowels along with Quechua vowels that demonstrated an effect of the L2 Spanish on the production of L1 Quechua. Lastly, a study of a Brazilian Portuguese speaker of L2 American English demonstrated that phonetic drift can be affected by the most recent type of input a speaker has received (Sancier & Fowler, 1997). Before and after 2–4 months spent in each respective language community, the speaker's VOTs for coronal stop productions had shifted toward the values of the dominant language in her environment.

The goal of the present study is both to examine whether the early, rapid phonetic drift observed in Chang (2012, 2013) for English speakers learning Korean extends to other pairs of languages with different phoneme inventories, and to further evaluate the effect that long-term L2 exposure has on the vowels of the native language. The focus of this study is an analysis of phonetic drift in complete novice learners before and after six weeks of exposure to French as L2. We compare these speakers with those who have lived in Paris for five years or longer. In the next section, we review parts of Flege's Speech Learning Model (SLM), which serves as a framework for understanding why L1 and L2 categories may have a bidirectional influence on the phonetic implementation in each language.

1.2 Speech Learning Model and phonetic category formation

Following Chang (2012), the current study will take as a starting point Flege's SLM (Flege, 1995, 2003, 2007) to provide a basis for this cross-linguistic interaction. Because of its focus on both perception and production, the SLM is a particularly applicable framework for the current study on vowel production, though there are competing models, for example, the Perceptual Assimilation Model (PAM) (Best, 1995; Best, McRoberts, & Goodell, 2001) and PAM-L2 (Best & Tyler, 2007), which focus mainly on the perception of non-native sound categories.

The SLM provides a model of cross-linguistic interactions that addresses age-related changes in the interaction between the L1 and L2 phonetic subsystems, and maintains that the capacity for speech learning remains intact across the life span (Flege, 2003, 2007). Two important mechanisms from the SLM pertaining to category formation in L1 and L2 phonetic systems are category assimilation and category dissimilation. Category assimilation may occur when a new category fails to be established for an L2 speech sound despite present yet unperceived audible differences between the L2 sound and a similar L1 category (e.g., American English speakers may be unable to form a

category for French /u/ that is distinct from American English /u/). This mechanism potentially results in a modified category somewhere between the original L1 category and the similar L2 category, but with greater weight given to the original L1 category (Flege, 2003).

Conversely, category dissimilation can occur when a new category is established for an L2 sound. Some research has shown that when a new L2 category is formed but dissimilation occurs, L2 learners can produce a phonetic implementation that in some way overshoots the L1 native category values (Flege & Eefting, 1988; Flege, Schirru, & MacKay, 2003). For example, Flege et al. (2003) show that when one group of Italian speakers who were also early learners of English produced diphthongal English /eɪ/, they produced an off-glide that was higher and farther front than the native monolingual English speakers did. Flege et al. argue that the overshoot allows them to better distinguish English /eɪ/ from Italian /e/.

Category assimilation and dissimilation are important components of the framework of equivalence classification, which pertains to the relationship between a speaker's age and their ultimate attainment of a second language. According to Flege (1987), "equivalence classification is a basic cognitive mechanism which permits humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category" (49). L1 categories eventually become attractors of L2 consonants and vowels (Flege, 2003, 2007). However, it is precisely this same mechanism that then prevents older speakers from separating similar L1 and L2 categories, as the adults have less effective use of sensory input and are influenced by these strong attractor categories in their L1 (Flege, 1987, 1988), which inhibit the formation of new categories and subsequently produce phonemes with non-native phonetic properties (Flege, 2003, 2007).

Though L1 phonetic drift could potentially refer to any type of shift in the phonetic implementation of a sound, previous findings suggest that drift is a type of assimilation in that an L1 category modified by phonetic drift demonstrates the influence of an attracting L2 category. However, Chang's (2012, 2013) conception of phonetic drift is a refinement of Flege's (2003, 2007) description of assimilation, because it does not require that L1 and L2 categories merge into an single phonetic implementation for both languages that is influenced by both languages, only that the L1 category moves toward the phonetic implementation for a corresponding L2 sound (a situation which is in fact reported in other studies by Flege, such as Flege & Hillenbrand, 1984). In phonetic drift—as opposed to full assimilation—corresponding sounds should still be statistically distinguishable in both languages.

Due to these interactions in the phonetic category space for different languages, the SLM establishes a mechanism for not only the influence of an L1 on an L2 as new and similar phonemes are assimilated and dissimilated, but also for the effect of an L2 on an L1 for both production and perception. Specifically, the pressures of equivalence classification and the hypothesized drive of bilinguals to maintain separate phonetic structures for each language predict that L2 may well influence the phonetic categories of L1, giving rise to phonetic drift.

1.3 Research question and rationale

Previous accounts of the interaction between L1 and L2 phoneme categories provide the inspiration for the present study: do L1 American English vowels begin to drift as a result of influence from French vowels in an immersive L2 environment? We seek to investigate this question with two extensions from Flege (1987), based on the experimental design of Chang (2012). Special emphasis is placed on the high vowels, as some of these contrasts have been shown to be particularly difficult for American English speakers learning French in past research (e.g., Darcy et al., 2012; Flege, 1987; Levy, 2009; Levy & Law, 2010; Levy & Strange, 2008).

1.3.1 Amount of exposure. The first expansion concerns the fact that while Flege (1987) compared two groups with varying amounts of French experience, the earliest stages of L2 acquisition were not represented in his study. In the current study, our low-experience participants are students tested twice over the course of their study in France (Study Abroad participants), before and after the first six weeks of immersion, with no prior French language experience. Moreover, our participants were recorded while still in Paris, unlike Flege's student participants who had not used French for the 3–6 months prior to his study.

The high-experience participants (Paris Residents) have lived in Paris nearly continuously for at least five years and speak French regularly. Since several studies have shown that speakers who are long-term residents in an L2 environment show evidence of phonetic drift, it is likely that our Paris Residents will also show drift in their vowels (Bergmann et al., 2016; de Leeuw et al., 2010, 2012; Mayr et al., 2012). On the other hand, Chang (2013) showed that American English speakers participating in the same immersion program as the participants in Chang (2012) who did have previous experience with Korean, either through heritage exposure or classroom instruction, did not show the same drift effects that the novice learners did after five weeks of immersion. He attributed this finding to a novelty effect; that is, at the earliest stages of acquisition, the phonetic differences between a learner's L1 and L2 may be especially salient and prone to influencing L1, whereas greater familiarity with both languages at later stages attenuates the effect of phonetic drift. Moreover, Flege (1987) did not find evidence of drift for American English /u/ among his American speakers living in Paris and using French regularly.

Therefore, we test for three possible outcomes. First, the results of Chang (2012, 2013) suggest that if there is phonetic drift for the Study Abroad participants in the current study, it may be more pronounced than it is for the Paris Residents since a novelty bias should no longer be present for the latter group, though this outcome may be particular to people who have not been immersed in speaking their L2 prior to being tested. Alternatively, our speakers could follow Flege's (1987) participants and show little effect of French on their American English at any time point. Lastly, our experienced learners of French could be more like the Dutch speaker in Mayr et al. (2012) or the Quechua speakers in Guion (2003), for example, who showed substantial shifts in their L1 vowels after acquiring L2. These conflicting results indicate that the reasons for any particular outcome may be due to the details of the specific language contact situation, which will be addressed with respect to our English learners of French.

1.3.2 Examination of broader vowel space. The second extension of the present study concerns testing the American English vowel space beyond the limited set of vowels in Flege (1987). First, this study includes a holistic analysis of the vowel space, represented by nine American English vowels and 10 French vowels, in order to investigate whether something like the global shift in F1 in the native vowel space of American English learners of Korean in Chang (2012) is also evident for American English learners of French. F1 was not directly reported on in Flege's discussion, though it is noted in the appendix that the most experienced group of American English learners had a significantly higher F1 value for /u/ than all of the other American English groups. Since Chang (2012) reports that generalized phonetic drift occurs over time as a global decrease in F1 for American English learners' vowel space after contact with Korean while individual vowels are also reported to move along F2, it is possible that F1 changes are also important to the French–American English comparison as well if there are global differences between the F1 of the English and French vowel space.

One notable difference between the Korean and English pairing in Chang (2012, 2013) and the current study is that French has a more crowded and less peripheral vowel space, with 11 oral vowels in Parisian French (Fougeron & Smith, 1999), as compared with the eight vowels of Seoul

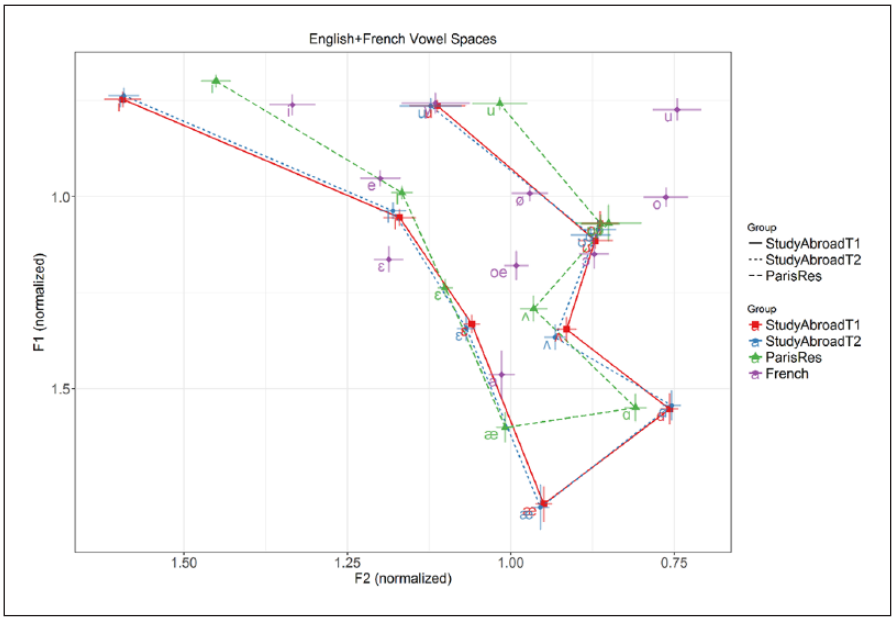

Figure 1. American English vowel space (using Nearey normalization) for Study Abroad T1, Study Abroad T2, and Paris Resident speakers, with native French vowels (unconnected) as references for visualizing how Paris Resident vowels compare to both monolingual American English and French norms. In all plots, error bars represent standard error of the mean.

Korean (Lee, 1999). Moreover, while the global Korean vowel space differs from the English vowel space in being expanded along F2 for both front and back vowels (see Chang 2012, Figures 1 and 12), the comparison plot of the Study Abroad speakers/monolingual American English speakers and Parisian French in Figure 1 shows that French non-low back vowels are further back than the English ones, but the low vowel /a/ (the only low vowel in French) is much higher than English low vowels. This means that the French vowel space is more compact than the English one especially on the height dimension. This raises the question of whether phonetic drift of a speaker's first language might be affected by the size and distribution of the inventory of the L2 they are acquiring.

Though we are not aware of current studies that examine this, there does exist some related evidence from the perception and discrimination of L2 vowels by L1 speakers of languages with both large and small vowel inventories. Iverson and Evans (2007, 2009) show that L2 learners of English from languages with crowded vowel inventories (German, Norwegian) are better at identifying English vowels than listeners from medium and smaller vowel inventories (French and Spanish). While German and Norwegian speakers map more English vowels to unique native vowels than the Spanish and French listeners do, there is also evidence that none of the L1 groups are using purely native phonetic cues to identify the English vowels (Iverson & Evans, 2007). The high variability training study in Iverson and Evans (2009) confirms that German speakers are better than Spanish speakers at identifying English vowels after training, but at the same time, even after training the German speakers have not changed their identification of what they consider the best English exemplars for each vowel to be. Iverson and Evans hypothesize that the improvement for German speakers from training arises because "auditory training makes the categorization process more efficient and automatic in a way that is long lasting, but does not generally change the representation of the categories" (875).

The relationship between the effect of training on the perception of L2 vowels (Iverson and Evans' studies) and the effect of immersion on the production of L1 vowels (Chang 2012, current study) is very indirect, but the Iverson and Evans studies at least provide some evidence that it could actually be easier for speakers to learn to differentiate L2 from L1 vowels when both languages have crowded spaces. This is perhaps counterintuitive, since Flege's SLM seems to suggest that learners from sparse vowel spaces would perform better since there is less likelihood of assimilating L1 and L2 vowels. However, we do not know what the opposite effect is; that is, is learning a more crowded L2 vowel space more or less likely to cause phonetic drift of a speaker's L1 categories? It is also unclear what role inventory size plays in the rapid immersion/early learning situation that Chang (2012) and the current study investigate, but the comparison of his results for Korean and ours for French may shed light on this question. Nominally, based on Iverson and Evans, we might expect that American English learners of French would be better able to maintain a distinction between the two languages as compared with Korean. However, it may not be the case that production behaves the same as perception, and there are still other factors to take into consideration, such as amount of L2 exposure. These issues will be further addressed in the discussion.

1.3.3 Organization of high vowels. Finally, we use the full complement of high vowels (/i,u/ for American English and /i, y, u/ for French) as a case study for investigating the contributions of equivalence classification and vowel space dispersion for our participants. The analysis in Flege (1987) ignored two important elements: American English /i/ and F1. In Flege's study, it is possible that the F2 movements in American English /u/ were complemented by a shift in at least one other vowel category like American English /i/, but such data were not reported. To examine whether an overall reorganization of the high vowel space on both the F1 and F2 dimensions occurs in order to keep the French and American English phoneme categories distinct, or whether even the experienced Paris Resident speakers assimilate one or more of their categories, the current study investigates the relationship among all of the high vowels in both languages.

One prediction is that L1 American English learners of French show signs of moving the categories of their American English vowels toward the categories of the ambient L2 French. Particularly in the case of high vowels, American English /i/ and American English /u/ may decrease relatively in F2 toward the phonetic norms of French /i/ and French /u/ (as shown for the French vowel means in Figure 1), and to equally distribute the distance between all of the vowels in the high vowel space with the introduction of a new vowel category, French /y/.

2 Methods

2.1 Participants

The participants in this study included 29 speakers. Twenty-two participants were classified as native American English speakers by means of a linguistic background survey administered at the time of testing. Of these 22 speakers, 11 participants (eight female, three male; mean age of 19 years, range of 18–21) were students with no family, classroom, or immersion exposure to French prior to the present study, beginning a semester abroad at New York University in Paris (NYUP) for the fall of 2014. From here on, this group is referred to as the Study Abroad group. The remaining 11 speakers of American English (four female, seven male; mean age of 38 years, range of 28–56) had both classroom and immersion exposure to the French language prior to the testing period and were instructors at NYUP or American friends of instructors at NYUP; this group is referred to as the Paris Resident group. The final seven participants (six female, one male; mean age of 19, range of 19–20) were native speakers of French and formed the representative control

sample of French for this study. This group is referred to as Native French. These participants were students at the Institut universitaire de technologie Paris Descartes. All individuals were paid for their participation.

The 11 Study Abroad individuals were tested at Time 1 (T1) and Time 2 (T2). Time 1 pertains to all 11 Study Abroad participants who were tested once at the end of the first week of arrival for the Study Abroad program. The American English productions of the Study Abroad T1 participants are also used as monolingual norms for American English where such comparisons are required.[1] Time 2 pertains to the same 11 Study Abroad participants who were then tested a second time at the end of 6 weeks. Seven out of 11 Study Abroad participants had exposure to a language other than American English or French (most often intermediate to advanced level Spanish study). American English dialect varied for participants, which will be discussed further in the results. Information about hometown and other details pertinent to their linguistic background for all speakers are in Appendix A.

At the time of data collection, Study Abroad participants were beginning a semester abroad in Paris and starting an immersive course in Elementary French with native speakers while taking at least two other courses in English. Each week, Study Abroad participants had six total hours of class in French with their native L1 French instructor (1.5 h per day, Monday–Thursday). This time in class constituted a majority of their contact with the language according to surveys administered to Study Abroad participants at the second session. Study Abroad participants reported speaking French with their friends and other Parisians for at most 3 h a week outside of the classroom after six weeks in Paris.

The 11 remaining native American English speakers comprise the Paris Resident group. These individuals have had at least five years of nearly continuous, immersive contact with French by living in the country and using it for a job, at home, or for school; two participants were also married to a native French speaker. Four out of 11 participants had exposure to a language other than American English or French (most often elementary to intermediate Spanish classroom study). American English dialect varied for participants. Paris Resident speakers had a wide range of exposure to French (9–48 years, $M = 24.5$ years). All participants reported using French for their job and all but one reported using French at home. All participants had begun learning French in a classroom environment, but ultimately reported learning the language from immersive exposure in a job setting or with routine visits to France prior to moving to Paris. The average daily L2 French usage among these participants was 65%.

The final group contains seven native French speakers. None qualified as being fluent in American English by means of a survey administered at the time of testing. Six participants reported classroom exposure to a language other than American English or French, and one of these participants also had Portuguese input at home. All spoke standard Parisian French. The vowel space of native French speakers was collected in order to assess whether the English vowels of both the Study Abroad and Paris Resident speakers were drifting toward the French vowels that could potentially be considered the phonemic counterparts of the English vowels. That is, if there are significant differences between the two time points of the Study Abroad group, or between Study Abroad Time 1 and the Paris Resident group, reference to the French vowels allows us to assess whether phonetic drift is consistent with influence of the French vowel space on the English vowel space.

2.2 Stimuli

The stimuli for the French condition consisted of six tokens of the 10 French vowels /i, y, e, ø, ɛ, œ, a, ɔ, o, u/. The American English stimuli consisted of six tokens of the nine American English

vowels /i, ɪ, ɛ, æ, ɑ, ʌ, o, ʊ, u/. Each target vowel was contained within a word with a consonant–vowel–consonant (CVC) sequence and this word was then embedded in the middle of a phrase. This contextual environment allowed for ease of segmentation. For example, the French word "poche" [pɔʃ] ("pocket") contains the target vowel /ɔ/. This word was then embedded in the middle of the phrase "son chapeau a une *poche* secrète" ("His hat has a secret pocket") in order to elicit more natural speech. In some instances, the CVC pattern was maintained across a word boundary instead of within a single word due to the constraints of where certain vowels and words can occur in each language. However, as a rule, vowels were never in initial or final positions in phrases and all vowels were flanked on both sides by either a stop or a fricative, regardless of whether the CVC sequence was within a word or contained a word boundary (e.g., "I'll have *two* cups" or "she *sits* on the sofa"). In the American English stimuli, affricates were also used. In the full data set, the chosen vowels often had a phonemic counterpart between the languages (e.g., French /i/ and American English /i/) while other vowels unique to each language like French /ø/ and American English /ʊ/ were chosen to ensure coverage of the whole vowel space. All stimuli were checked with multiple native speakers prior to their usage in the experiments. One American English stimulus was removed from data analysis, "one doe at the lake," as it was later found to be in violation of the CVC pattern. The full set of stimuli is included in Appendix B.

Lastly, except for the high vowels for the Paris Resident speakers, only the American English speakers' productions of English were analyzed. Most non-native productions in French are not reported here, as the main focus for analysis of the whole vowel space is to investigate phonetic drift in the L1 of American English speakers.

2.3 Procedure

Each American English speaker participated in the production of both French and American English stimuli. For French speakers, just the French condition was run. Instructions were given verbally in the speaker's native language and the same experimenter (the first author) ran all experiments. Participants first filled out a survey on their linguistic background. They then performed a reading task in which they read aloud each of the phrases presented to them at a self-selected pace. To accommodate the Study Abroad T1 participants who did not yet know how to read French, an additional sound sample of only the French phrases, recorded in Paris by a 28-year-old female native speaker of French, was provided. All participants in Study Abroad T1, Study Abroad T2, and Paris Resident groups were required to listen to the sample at least once but no more than three times to keep the procedure consistent across groups. No such audio sample was used for the American English condition for any of the groups. As a reminder, since the French data for the Study Abroad group were not the focus of the current study, they are not discussed further. For Study Abroad T1 and Study Abroad T2, the French condition and the American English condition were run once per session. The two sessions were run six weeks apart starting on the second week of the semester. For the Paris Resident group, the French condition and the American English condition were recorded in one session.

For each participant, the French condition was presented first, followed by the American English condition (where applicable). Within each condition, stimuli were presented in the form of PowerPoint presentations on an Apple Macbook Pro laptop computer. The English condition was equally divided into four PowerPoint blocks, and the French condition into three PowerPoint blocks. The order of the within-language blocks was randomized for each participant, but within a block, the stimuli were in the same arbitrary order for each participant. Subjects were asked to advance the slides themselves after pronouncing each phrase. After one PowerPoint ended, the

experimenter would switch to the next PowerPoint. Experiments were run in a quiet room at the NYUP building. Recordings were made with a TASCAM DR-40 4-track handheld digital audio recorder with a built-in condenser microphone that was positioned on a table 6–8 in away from the participant's chin. Sampling rate was set to either 48k or 96k in mono. Files created were in WAV 24-bit format. Most participants finished all tasks within 30 min.

2.4 Acoustic analysis

For the acoustic analysis, F1 and F2 measurements were taken from manually segmented recordings in the Praat waveform editor. Tokens were the target vowel in each of the 114 phrases of the stimuli. Formant frequencies were measured at the midpoints of each vowel. Vowels were segmented using the criteria of the onset and offset of F2 adjacent to the silence or burst of a stop, or the frication of a fricative. Trials where the vowel could not be accurately measured because of glottalization, ambient noises, or unclear onsets were excluded from the statistical analysis (N = 139, or 7% of the data).

Once the F1 and F2 midpoints were extracted using a Praat script, they were normalized using the Nearey1 formula implemented in the R package vowels (Kendall & Thomas, 2010; Nearey, 1978). This normalization technique was chosen because it was shown to be one of the best performing ones for comparing across speakers with different vocal tract properties by Adank, Smits and van Hout (2004). While we recognize that there is some doubt about whether normalization is appropriate when comparing across different languages (Clopper, 2009), we felt that it was necessary since the number of male and female speakers in each group differed due to the inherent limitations of finding appropriate participants. Moreover, because the Paris Resident participants are the same speakers producing both languages, the concerns about normalization may be diminished for these particular comparisons.

2.5 Statistical analysis

Formant midpoint measurements in both experiments were analyzed with mixed-effects linear regression using the *lme4* and *lmerTest* packages in R (Bates, Maechler, Bolker, & Walker, 2014; Kuznetsova, Brockhoff, & Christensen, 2013). All regression models had fixed effects of Vowel and Group, and included random intercepts for word and participant and random slopes for individual vowels. Group analyses include comparisons of the Study Abroad participants in the first week (Study Abroad T1) and the sixth week (Study Abroad T2), and Paris Residents. Again, when comparisons with monolingual American English speakers are required, we use the Study Abroad T1 group, since they had no exposure to French before arriving in Paris. When relevant for an analysis, we also include native French speakers as a group to provide a baseline for the norm for French vowel categories.

3 Results

3.1 Phonetic drift in the American English vowel space

This section examines whether there is an effect of the length of time living in Paris on the potential phonetic drift of the participants' American English vowel spaces and individual vowels. Before we turn to the analyses of the Study Abroad and Paris Resident groups, we first establish whether or not the global vowel space differs on either the F1 or F2 dimensions in the comparison of American English and Parisian French. As in the subsequent analyses, the Study Abroad T1 English

vowels are treated as the monolingual American English condition, and the native French speakers' vowels are used for the French vowel space. F1 and F2 are the dependent variables in two separate analyses. The language factor has two levels, English and French, with English used as the reference. There are random intercepts for word and participant. Results for F1 (difference = 0.176 for the Nearey normalized values, or 114 Hz) show that there is a significant effect of language ($\beta = -0.182$, $t = -3.36$, $p < 0.001$), but results for F2 (difference = 0.011, or 93 Hz) are not significant ($\beta = -0.018$, $t = -0.43$). These results establish a difference between the global vowel space along F1 for English and French (which can also be seen visually in Levy and Strange 2008 and Levy 2009), and lead to the prediction that global phonetic drift is more likely to happen for F1 than for F2. However, as is clear from Figure 1, even though there is no global difference in F2, there is a substantial difference in French /u/ as compared with American English /u/; this will be examined in the individual vowel comparisons.

For the comparison of the three groups of language learners (Study Abroad T1 and T2, and Paris Residents), the dependent variable in the first analysis is F1. The Vowel factor is sum coded, since there is no a priori reason to choose any individual vowel as the baseline. The Group factor is treatment coded with Study Abroad T1 as the baseline, since the relevant comparisons are between Study Abroad T1 and the other two levels of the group. Figure 1 demonstrates the vowel spaces for both Study Abroad and Paris Resident groups, including the native French vowels as a reference for the direction of phonetic drift.

For F1 overall, the results in Table 1 indicate that there is a significant difference between Study Abroad T1 and Paris Residents, but no difference between Study Abroad T1 and T2. There are no

Table 1. Statistical results for group and interactions between group and vowel for F1. Individual statistical results for vowels (here and below) are not of interest here and are not included to conserve space. Asterisk indicates a significant effect. PR = Paris Resident, SA-T2 = Study Abroad Time 2.

	Estimate	Std. Error	*t* value	*p* value
(Intercept)	1.202	0.012	100.64	
Group-PR	–0.058	0.015	–3.96	$p < 0.001$
Group-SA-T2	0.000	0.005	0.06	n.s.
Vowel-ɛ:Group-PR	–0.036	0.022	–1.59	n.s.
Vowel-æ:Group-PR	–0.145	0.052	–2.81	$p = 0.01$
Vowel-ɑ:Group-PR	0.052	0.034	1.52	n.s.
Vowel-ɪ:Group-PR	–0.006	0.020	–0.29	n.s.
Vowel-i:Group-PR	0.010	0.031	0.33	n.s.
Vowel-ʊ:Group-PR	0.016	0.024	0.64	n.s.
Vowel-ʌ:Group-PR	0.009	0.024	0.40	n.s.
Vowel-o:Group-PR	0.059	0.039	1.53	n.s.
Vowel-u:Group-PR	–0.006	0.019	–0.30	n.s.
Vowel-ɛ:Group-SA-T2	0.012	0.015	0.78	n.s.
Vowel-æ:Group-SA-T2	0.005	0.015	0.36	n.s.
Vowel-ɑ:Group-SA-T2	–0.009	0.015	–0.62	n.s.
Vowel-ɪ:Group-SA-T2	–0.018	0.015	–1.24	n.s.
Vowel-i:Group-SA-T2	–0.010	0.015	–0.70	n.s.
Vowel-ʊ:Group-SA-T2	–0.015	0.015	–1.00	n.s.
Vowel-ʌ:Group-SA-T2	0.021	0.015	1.44	n.s.
Vowel-o:Group-SA-T2	0.017	0.016	1.04	n.s.
Vowel-u:Group-SA-T2	0.000	0.015	0.03	n.s.

Table 2. Statistical results for group and interactions between group and vowel for F2. PR = Paris Resident, SA-T2 = Study Abroad Time 2.

	Estimate	Std. Error	*t* value	*p* value
(Intercept)	1.036	0.012	89.31	
Group-PR	–0.007	0.011	–0.58	n.s.
Group-SA-T2	0.005	0.004	1.36	n.s.
Vowel-ɛ:Group-PR	0.046	0.014	3.20	$p < 0.001$
Vowel-æ:Group-PR	0.066	0.015	4.41	$p < 0.001$
Vowel-ɑ:Group-PR	0.056	0.020	2.73	$p < 0.01$
Vowel-ɪ:Group-PR	0.001	0.020	0.06	n.s.
Vowel-i:Group-PR	–0.137	0.030	–4.54	$p < 0.001$
Vowel-ʊ:Group-PR	–0.001	0.016	–0.04	n.s.
Vowel-ʌ:Group-PR	0.056	0.013	4.28	$p < 0.001$
Vowel-o:Group-PR	–0.003	0.002	–0.05	n.s.
Vowel-u:Group-PR	–0.088	0.035	–2.53	$p < 0.01$
Vowel-ɛ:Group-SA-T2	0.003	0.011	0.25	n.s.
Vowel-æ:Group-SA-T2	0.000	0.011	0.01	n.s.
Vowel-ɑ:Group-SA-T2	–0.010	0.011	–0.87	n.s.
Vowel-ɪ:Group-SA-T2	0.002	0.011	0.16	n.s.
Vowel-i:Group-SA-T2	–0.008	0.011	–0.73	n.s.
Vowel-ʊ:Group-SA-T2	–0.002	0.011	–0.21	n.s.
Vowel-ʌ:Group-SA-T2	0.012	0.012	1.05	n.s.
Vowel-o:Group-SA-T2	–0.002	0.011	–0.14	n.s.
Vowel-u:Group-SA-T2	0.005	0.011	0.47	n.s.

significant interactions for vowel and Study Abroad T2, indicating that there are no F1 shifts for individual vowels for the Study Abroad participants after six weeks. The interactions between vowel and Paris Residents show that there is a significant difference for /æ/, which has a lower F1 than it does for monolingual American English speakers.

For F2, shown in Table 2, there are no significant main effects of group differences between Study Abroad T2 or Paris Residents as compared with Study Abroad T1. There are no significant interactions between vowel and Study Abroad T2 for F2 either, confirming that there are no shifts for any vowels after six weeks for either formant measurement. However, for the interactions between Vowel and Group for Paris Residents, there are significant differences for /i/, /ɛ/, /æ/, /ɑ/, /ʌ/, and /u/. These results indicate that /i/ and /u/ have a lower F2, or are more back than the Study Abroad T1 (monolingual American English) speakers and /æ/, /ɛ/, /ɑ/, and /ʌ/ have a higher F2, or are more fronted.

The results of these analyses indicate that while there is no general drift of the F1 or F2 space between Study Abroad T1 and T2, there is an overall difference in the vowel space on the F1 dimension between the Study Abroad groups and the Paris Resident group. Specifically, the F1 space has shifted up, and there are individual differences for two vowels. For F2, there is no significant global difference between the Paris Residents and the Study Abroad groups, though on this dimension there are many more differences between groups for individual vowels.

While the results so far indicate that there are significant differences for global F1 values and for individual vowels for the Paris Resident group as opposed to the monolingual American English speakers (Study Abroad T1), there are several factors that could be contributing to this result. Because we do not have pre-exposure data for Paris Resident groups, a pre- and post-immersion

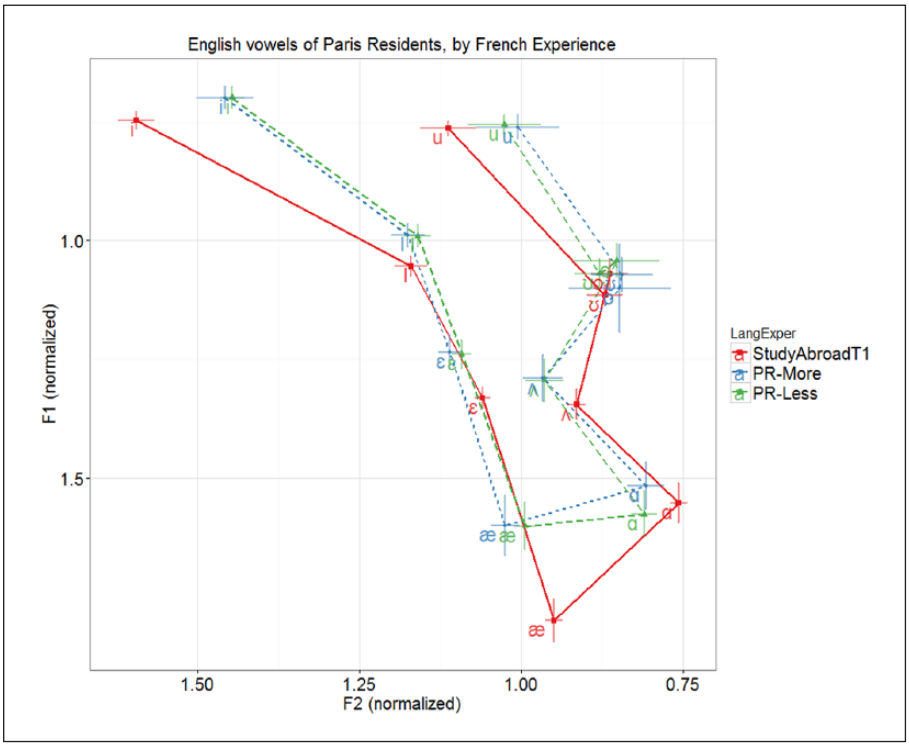

Figure 2. American English vowel space for Study Abroad T1, More Experienced and Less Experienced Paris Resident speakers.

analysis is not available for the Paris Resident speakers. Instead, we delve more deeply into the Paris Resident pattern by examining two potential factors that could interact with the difference between monolingual English and Paris Resident English: years of exposure to L2 and the American English dialect background of the Paris Resident speakers.

3.1.1 Effect of amount of exposure to L2. We first consider whether years of exposure to French, despite generally high rates of exposure and amount of use, impact the results found for Paris Resident speakers. For this analysis, we divide the Paris Resident group into those who have less than 20 years of cumulative experience with French, including classroom instruction (range: 9–18 years, $M = 14.8$ years, $N = 6$) and those with 20 years of experience or more (range: 20–48 years, $M = 28.8$, $N = 5$). We compare the divided More Experienced and Less Experienced Paris Resident groups with the monolingual American English speakers, as illustrated in Figure 2. The Vowel factor was sum coded, and the Group factor was treatment coded, with monolingual American English as the reference group since we are interested in deviation from the vowel values for this group.

Results for overall F1 are in Table 3. Both the More and Less Experienced Paris Residents groups are significantly different from the monolingual American English speakers. For the interaction between group and vowel, in the comparison of the More Experienced Paris Residents and the monolingual American English speakers, the same individual vowels /æ/ and /u/, have significant or nearly significant interactions as when the Paris Residents are taken together as a group. For the comparison of Less Experienced Paris Residents and monolingual American English speakers for F1, there was a significant interaction for the vowel /æ/, but no effect for /u/. When the

Table 3. Statistical results for group and interactions between group and vowel for F1. The reference level for group is monolingual American English speakers. MoreExpPR = More Experienced Paris Residents, LessExpPR = Less Experienced Paris Residents.

	Estimate	Std. Error	*t* value	*p* value
(Intercept)	1.202	0.012	111.24	
Group-LessExpPR	–0.057	0.016	–3.65	$p = 0.002$
Group-MoreExpPR	–0.060	0.016	–3.65	$p = 0.002$
Vowel-ɛ:Group-LessExpPR	–0.033	0.027	–1.23	n.s.
Vowel-æ:Group-LessExpPR	–0.145	0.061	–2.4	$p = 0.001$
Vowel-ɑ:Group-LessExpPR	0.078	0.047	1.63	n.s.
Vowel-ɪ:Group-LessExpPR	–0.008	0.022	–0.35	n.s.
Vowel-i:Group-LessExpPR	0.009	0.035	0.24	n.s.
Vowel-ʊ:Group-LessExpPR	0.014	0.029	0.48	n.s.
Vowel-ʌ:Group-LessExpPR	0.013	0.029	0.44	n.s.
Vowel-o:Group- LessExpPR	0.033	0.051	0.64	n.s.
Vowel-u:Group-LessExpPR	0.047	0.028	1.67	n.s.
Vowel-ɛ:Group-MoreExpPR	–0.038	0.028	–1.35	n.s.
Vowel-æ:Group-MoreExpPR	–0.143	0.064	–2.22	$p < 0.001$
Vowel-ɑ:Group-MoreExpPR	0.015	0.050	0.3	n.s.
Vowel-ɪ:Group-MoreExpPR	–0.003	0.024	–0.13	n.s.
Vowel-i:Group-MoreExpPR	0.012	0.037	0.33	n.s.
Vowel-ʊ:Group-MoreExpPR	0.018	0.031	0.57	n.s.
Vowel-ʌ:Group-MoreExpPR	0.005	0.031	0.18	n.s.
Vowel-o:Group-MoreExpPR	0.092	0.054	1.70	n.s.
Vowel-u:Group-MoreExpPR	0.058	0.030	1.92	n.s.

reference level is changed to More Experienced Paris Residents, there is no significant effect of group or interactions with the Less Experienced Paris Residents, indicating that the two types of Paris Residents do not show differences for F1 among themselves.

The overall differences in F2 for the comparison of both Less and More Experienced Paris Residents as compared with the monolingual American English speakers also mirror those for the combined group. The results are shown in Table 4. There are no significant differences for either More or Less Experienced Paris Residents for F2. However, the significant interactions for the Less Experienced Paris Residents and vowel are /i/, /ɛ/, /æ/, /ɑ/, and /ʌ/, and for the More Experienced Paris Residents they are /i/, /ɛ/, /æ/, /ɑ/, /ʌ/, and /u/, which are the same six vowels that show significant interactions for F2 when the Paris Resident group is taken as a whole. Again, when the reference level is changed to More Experienced Paris Residents, there is no significant effect of group or interactions with the Less Experienced Paris Residents, indicating that the two subgroups of Paris Residents do not differ from each other for F2.

Taken together, these results indicate very few differences between the More Experienced and Less Experienced Paris Resident groups, except that the More Experienced Paris Resident group has a more backed /u/ compared with the monolingual American English speakers than the Less Experienced group does. It is possible that the slightly greater effect for the More Experienced Paris Resident group for /u/ indicates that an increase in input and longer exposure to an L2 continues to cause drift in the L1 vowel system even after 20 years of exposure. However, considering the great similarities between the vowel spaces shown in Figure 2 comparing the two levels of experience with the monolingual speakers, it is clear that the effect of long-term experience in

Table 4. Statistical results for group and interactions between group and vowel for F2. MoreExpPR = More Experienced Paris Residents, LessExpPR = Less Experienced Paris Residents.

	Estimate	Std. Error	*t* value	*p* value
(Intercept)	1.029	0.015	70.38	n.s.
Group-LessExpPR	–0.007	0.014	–0.49	n.s.
Group-MoreExpPR	–0.007	0.014	–0.46	n.s.
Vowel-ɛ:Group-LessExpPR	0.038	0.018	2.13	$p = 0.02$
Vowel-æ:Group-LessExpPR	0.051	0.018	2.77	$p = 0.002$
Vowel-ɑ:Group-LessExpPR	0.057	0.023	2.46	$p = 0.001$
Vowel-ɪ:Group-LessExpPR	–0.005	0.023	–0.24	n.s.
Vowel-i:Group-LessExpPR	–0.142	0.037	–3.83	$p < 0.001$
Vowel-ʊ:Group-LessExpPR	0.015	0.017	0.86	n.s.
Vowel-ʌ:Group-LessExpPR	0.055	0.017	3.21	$p < 0.001$
Vowel-o:Group-LessExpPR	0.004	0.036	0.12	n.s.
Vowel-u:Group-LessExpPR	–0.076	0.044	–1.71	n.s.
Vowel-ɛ:Group-MoreExpPR	0.055	0.019	2.89	$p = 0.001$
Vowel-æ:Group-MoreExpPR	0.082	0.019	4.21	$p < 0.001$
Vowel-ɑ:Group-MoreExpPR	0.053	0.025	2.17	$p = 0.01$
Vowel-ɪ:Group-MoreExpPR	0.010	0.024	0.41	n.s.
Vowel-i:Group-MoreExpPR	–0.131	0.039	–3.33	$p < 0.001$
Vowel-ʊ:Group-MoreExpPR	–0.020	0.018	–1.06	n.s.
Vowel-ʌ:Group-MoreExpPR	0.056	0.018	3.1	$p < 0.001$
Vowel-o:Group-MoreExpPR	–0.008	0.018	–0.22	n.s.
Vowel-u:Group-MoreExpPR	–0.103	0.047	–2.18	$p = 0.001$

French on the speakers' (native) English vowels are similar in the two groups of Paris Residents, resulting in the same pattern of movement.

3.1.2 Effect of American English dialect region for Paris Residents. The next question is whether the dialectal differences among the speakers can account for any of the differences between the American English monolinguals and the (combined) Paris Resident speakers. This is tested in two ways. Because we could not choose the dialectal background of the participants and still obtain enough speakers, speakers from a variety of locations are represented. However, both in the Study Abroad T1 (monolingual) and in the Paris Resident groups, there are enough speakers from areas in the Northeast of the United States (those from New York, Connecticut, Massachusetts, Pennsylvania) to be able to use those speakers in two different comparisons. First, within the Paris Resident group, we compare the speakers from the Northeastern US ($N = 6$) with a heterogeneous group from other parts of the United States ($N = 5$) to determine whether they differ significantly for any vowels. Second, we can isolate the Northeast speakers in the Paris Resident group and the monolinguals ($N = 4$) to see if the pattern that was reported for the combined set of all speakers is mirrored in a comparison between monolinguals and Paris Residents just within speakers from the Northeast. For all analyses, the factor of vowel is sum coded and group is treatment coded (Northeast vs. Other Regions in the first analysis, and Northeast monolingual English vs. Northeast Paris Resident in the second analysis).

Within the Paris Resident group, the results for F1 for the comparison between Northeast and Other Regions show that dialect region does not quite reach significance ($\beta = -0.029$, $t = -1.91$, $p = 0.08$). The only interaction for F1 between region and vowel that is also nearly significant is for

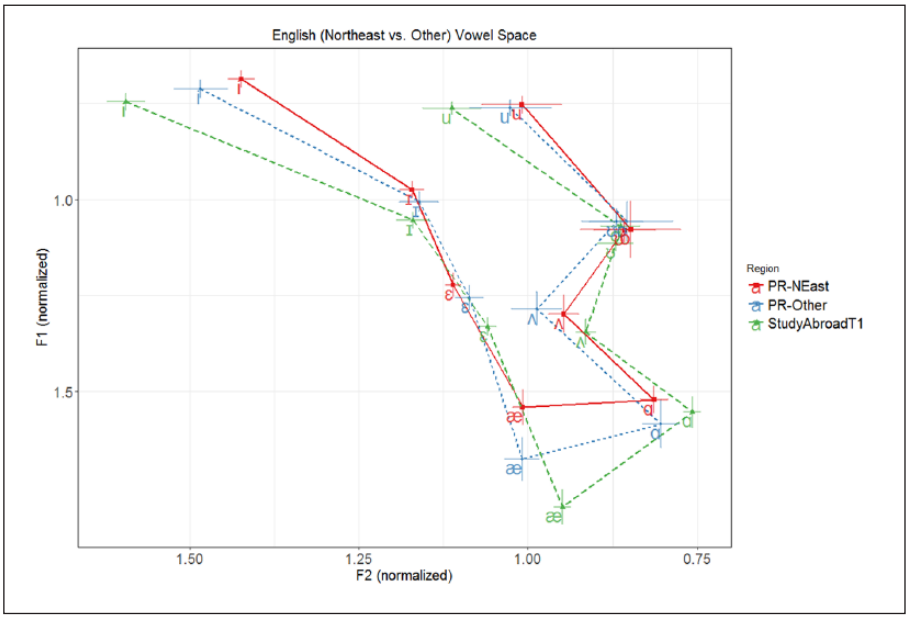

Figure 3. American English vowel space for Study Abroad T1, Paris Resident-Northeast and Paris Resident-Other Regions speakers.

/æ/ ($\beta = -0.105$, $t = -1.96$, $p = 0.08$). There are no significant main effects for F2 for region, nor are there any significant interactions between region and vowel. Figure 3 shows the plot of the vowel space differences for both Northeastern and Other Region Paris Residents, including the monolingual American English speakers for comparison.

The comparison of monolinguals and Paris Residents for the subset of speakers from the Northeast mimics the same pattern as the larger set. For F1, there is a significant group difference between Paris Resident and monolingual speakers, and a significant interaction between group and vowel for /æ/, as shown in Table 5. For F2, there is a significant difference for group, and there are significant or near significant interactions between group and vowel for the same six vowels (/i, ɛ, æ, ɑ, ʌ, u/) that were significant for the whole group of Paris Residents, as shown in Table 6.

The analyses taking into account dialectal differences suggest that there may be some contribution of the original American English dialect of the Paris Resident speakers. As shown in Figure 3 and the statistical comparison between Paris Residents from the Northeast versus Other Regions, for some vowels there is a greater distance between the Paris Resident speakers from the Northeast and the monolingual American English speakers. The nearly significant global difference in F1 between the two Paris Resident groups appears to be due mainly to the low vowels /æ/ and /ɑ/, since the lower boundary of the vowel quadrilateral is shifted up (lower in F1) for the Paris Residents from the Northeast. However, as indicated in Figure 3, the difference between the two Paris Resident groups and the monolingual American English speakers is a matter of degree, since both Paris Resident groups differ from the monolinguals along the same direction for both F1 and F2 for all vowels that show a difference. The comparison between the Paris Residents and monolingual American English speakers from the Northeast confirms that the differences shown, when all speakers from all groups are included, are still present even when dialect is controlled for.[2]

Table 5. Statistical results for group and interactions between group and vowel for F1 for monolingual American English and Paris Resident speakers from the Northeast.

	Estimate	Std. Error	*t* value	*p* value
(Intercept)	1.201	0.015	78.43	
Group-PR	–0.069	0.018	–3.83	$p = 0.005$
Vowel-ɛ:Group-PR	–0.047	0.045	–1.05	n.s.
Vowel-æ:Group-PR	–0.246	0.078	–3.17	$p < 0.001$
Vowel-ɑ:Group-PR	0.029	0.051	0.58	n.s.
Vowel-ɪ:Group-PR	–0.025	0.023	–1.13	n.s.
Vowel-i:Group-PR	0.007	0.030	0.23	n.s.
Vowel-ʊ:Group-PR	0.039	0.032	1.22	n.s.
Vowel-ʌ:Group-PR	0.051	0.039	1.33	n.s.
Vowel-o:Group-PR	0.147	0.075	1.95	n.s.
Vowel-u:Group-PR	0.002	0.017	0.10	n.s.

Table 6. Statistical results for main effect of group, and interactions between group and vowel for F2 for both monolingual American English and Paris Resident speakers from the Northeast.

	Estimate	Std. Error	*t* value	*p* value
(Intercept)	1.028	0.013	79.98	
Group-PR	–0.003	0.013	–0.22	n.s.
Vowel-ɛ:Group-PR	0.063	0.022	2.84	$p = 0.005$
Vowel-æ:Group-PR	0.058	0.022	2.64	$p < 0.01$
Vowel-ɑ:Group-PR	0.064	0.022	2.86	$p = 0.005$
Vowel-ɪ:Group-PR	0.017	0.022	0.75	n.s.
Vowel-i:Group-PR	–0.177	0.022	–7.99	$p < 0.001$
Vowel-ʊ:Group-PR	–0.001	0.022	–0.03	n.s.
Vowel-ʌ:Group-PR	0.052	0.022	2.36	$p = 0.02$
Vowel-o:Group-PR	0.013	0.025	0.55	n.s.
Vowel-u:Group-PR	–0.090	0.022	–4.00	$p < 0.001$

3.2 Changes in high vowels among Paris Residents

The relationship between French and American English high vowels in particular has been an area of interest in previous research since some studies have shown that the high rounded vowel /y/ is difficult for American English speakers to produce and perceive accurately (e.g., Baker & Smith, 2010; Darcy et al., 2012; Levy, 2009; Levy & Law, 2010; Levy & Strange, 2008), and because the back vowel /u/, though typically treated as the same phonemic category across both languages, has very different phonetic properties. We also turn to the high vowels to examine whether the Paris Residents have been able to develop separate categories for their two languages, or whether they have assimilated some categories between the two languages and are implementing them with the same phonetic characteristics. As shown in Figure 4, categories of interest include the high front vowel /i/, back vowel /u/, and the relationship between French /y/ and American English /u/, which appear to cluster together in a central space in the vowel plot in the figure.

The first analysis focuses on the high front vowel /i/ and its implementation in both American English and French. This analysis is a linear model in which the factor of group is defined by language: monolingual American English (Study Abroad T1), native French, Paris Resident American

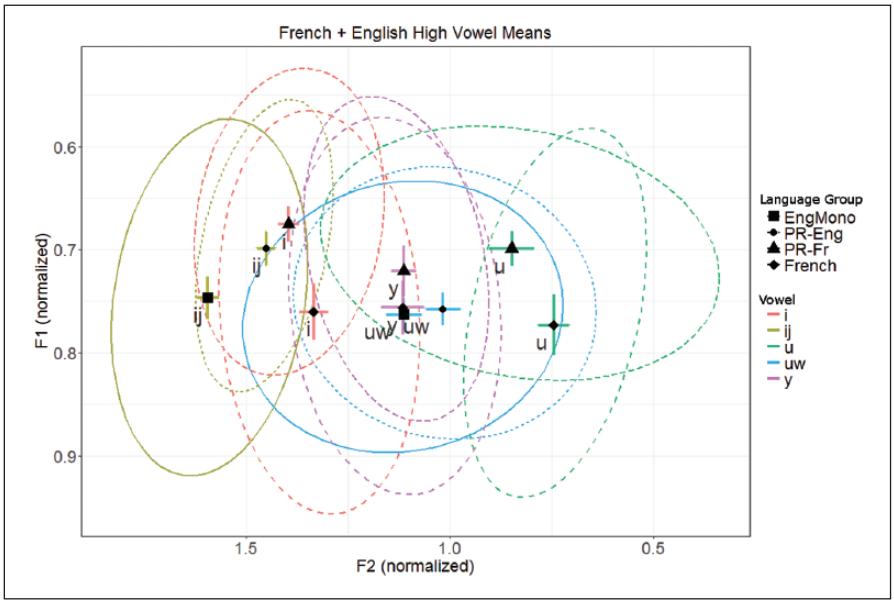

Figure 4. High vowel means for both American English and French for monolinguals and Paris Residents. The phonemes of each language are differentiated by different symbols. For French, IPA symbols /i, y, u/ are used. For English, "ij" = /i/ and "uw" = /u/. Ellipses represent 95% confidence intervals and error bars represent standard error of the mean.

English, and Paris Resident French speakers. The reference level for this factor was Paris Resident American English, since the comparisons with this language group are the ones that are most relevant to the research question. Results for F2 show that there are significant differences for all language groups: monolingual American English ($\beta = 0.144$, $t = 3.64$, $p = 0.001$), native French ($\beta = -0.119$, $t = -2.59$, $p = 0.01$), and Paris Resident French ($\beta = -0.058$, $t = -3.82$, $p = 0.001$). Tukey HSD tests confirm that in addition to these significant comparisons, there are also significant differences for native French and monolingual American English ($\beta = -0.262$, $z = -5.71$, $p = 0.001$), and Paris Resident French and monolingual American English ($\beta = -0.201$, $z = -4.92$, $p = 0.001$). Native French and Paris Resident French are not significantly different ($\beta = 0.062$, $z = 1.37$, $p = 0.49$). The results for this statistical analysis and all of the following analyses are summarized in Table 7.

For F1 for /i/, there is a significant difference between Paris Resident American English and native French ($\beta = 0.062$, $t = 2.10$, $p = 0.04$). Tukey HSD tests indicate that there are further differences in the comparisons between Paris Resident French and monolingual American English ($\beta = -0.071$, $z = -2.69$, $p = 0.03$) and Paris Resident French and native French ($\beta = -0.085$, $z = -3.02$, $p = 0.01$).

The second analysis examines /u/. Again, F2 results show significant differences for all language groups compared with the Paris Resident American English baseline: monolingual American English ($\beta = 0.096$, $t = 2.66$, $p = 0.01$), native French ($\beta = -0.282$, $t = -3.54$, $p = 0.002$), and Paris Resident French ($\beta = -0.154$, $t = -2.17$, $p = 0.05$). Tukey HSD tests show significant differences for all of the other pairwise comparisons as well.

The results for F1 indicate a significant difference between Paris Resident American English and Paris Resident French ($\beta = -0.05$, $t = -4.17$, $p < 0.001$). Further significant effects from the Tukey HSD tests include the comparisons between Paris Resident French and monolingual

Table 7. Summary of statistical results for the comparison of high vowels for English monolingual, native French, Paris Resident English and Paris Resident French speakers. '*' indicates a significant result and '–' is not significant.

Language pairs	/i/ comparison		/u/ comparison		French /y/-English /u/ comparison	
	F1	F2	F1	F2	F1	F2
EngMono / FrNative	–	*	–	*	–	–
PR-Eng / EngMono	–	*	–	*	–	*
PR-Eng / FrNative	*	*	–	*	–	–
PR-Eng / PR-French	–	*	*	*	*	–
PR-French / EngMono	*	*	*	*	–	–
PR-French / FrNative	*	–	*	*	–	–

American English ($\beta = -0.064$, $z = -3.01$, $p = 0.01$) and Paris Resident French and native French ($\beta = -0.070$, $z = -3.26$, $p = 0.006$).

The final analysis examines the sounds that appear in the central part of the F2 space: American English /u/ and French /y/. These vowels are coded as a four-way factor with the following levels: monolingual American English /u/, native French /y/, Paris Resident American English /u/, and Paris Resident French /y/. The baseline is Paris Resident American English /u/. Results show that only monolingual American English /u/ is significantly different from the baseline Paris Resident American English /u/ ($\beta = 0.096$, $t = 3.02$, $p = 0.005$), and this is confirmed by Tukey HSD contrasts which find no other significant results.

The results for F1 indicate a significant difference between Paris Resident American English /u/ and Paris Resident French /y/ ($\beta = -0.046$, $t = -2.83$, $p = 0.01$). No other significant differences are revealed by the Tukey HSD contrasts.

Taken together, the F2 results from this section indicate that Paris Residents' ability to separate their American English and French high vowels differs by backness, but it is not always the case that Paris Resident American English and French match their monolingual norms even when French and American English are significantly different within the Paris Resident speakers. For the front vowels, both of Paris Residents' /i/ vowels differ from other groups in all comparisons, except that their French /i/ has achieved the native French F2 value, which is more backed than American English. For the back vowels, the significant difference between Paris Resident American English and monolingual American English /u/ suggests that there has been a backing effect of French on Paris Resident American English /u/.

As for the relationship between French /y/ and American English /u/, the only significant F2 difference is the one between monolingual American English /u/ and Paris Resident American English /u/, which has already been noted. While Paris Resident speakers have attained native-like values for French /y/, this result seems attributable to a lack of difference between monolingual American English /u/ and native French /y/. Thus, Paris Residents may be using their "original" American English /u/ values for French /y/, while moving their own current American English /u/ back a bit and their French /u/ back even more. The idea that Paris Resident speakers are moving American English /u/ to differentiate it from the crowded central vowel space additionally comes from the F1 results. In general, the Paris Resident speakers' vowels *both* in American English and in French are shifted up. The fact that, with the exception of /u/, both their American English and French vowels have similar F1 values suggests that differences from the monolinguals may be due to characteristics of these particular speakers that were sampled, but it is worth noting that Paris

Resident American English /u/ is the only vowel that differs in F1 from a close Paris Resident French counterpart—in this case /y/. In other words, the Paris Residents may be shifting their production along F1 just in this particular situation where two phonemes from the two languages are competing, in order to dissimilate their American English /u/ from both French /u/ and French /y/.

4 General discussion

4.1 Phonetic drift and the whole vowel space

This study examined phonetic drift in two populations of L1 American English learners of French: Study Abroad speakers, classified as inexperienced learners of French tested at the end of week one and week six in immersion, and Paris Resident speakers, who had extensive French immersion experience. Drawing primarily from Chang (2012, 2013) and Flege (1987), this study investigated both rapid and long-term phonetic drift in the global vowel space and specifically the high vowel space in American learners of French.

Unlike Chang's (2012) results demonstrating a rapid shift in vowel categories for American learners of Korean, the Study Abroad speakers in this study had no significant difference in the production of their American English vowel space. From week one to week six, neither individual vowel movement nor a global shift in formant frequencies was observed in this group. These results suggest that the short-term but rapid effects of phonetic drift in Chang (2012) may not be reproduced for all language pairings. A few ideas are presented as possible explanations for the failure to replicate the same phonetic drift effect in these speakers.

The first possibility is that the vowel space of French is considerably more crowded and in some ways more compact than that of Korean, with differences in the distribution of vowels. Based on the comparisons of each language with American English, French tends to be more condensed, especially in F1 (see Figure 2), while the vowel space of Korean is generally more expansive in F2 (back vowels and some front vowels are both more extreme) and there is a slight shift toward a lower boundary for the vowel space (higher in F1) (Chang, 2012, Figure 12; see also Yang, 1996). Though Iverson and Evans' (2007, 2009) results for perception of L2 vowels may provide some evidence that a more crowded L1 vowel space may actually facilitate identification of L2 in the long term, the results for the Study Abroad group for French suggest that there is no congruent effect for rapid L1 phonetic drift in vowel production. Instead, the results seem to show the opposite, since Korean learners showed drift for their L1 vowels while French learners did not. The drift for the American learners of Korean is compatible with the assimilation predictions of Flege's SLM model, though whether or not these speakers converge on single merged vowel categories for both English and Korean would require data from native English speakers who have had longer contact with Korean.

As for the learners of French, if assimilation between vowels across the two languages is to occur, the more crowded vowel space of French may make it difficult for English learners to determine which vowels are the most appropriate counterparts for the purpose of assimilation. In some cases, phonological categories and phonetic implementation may conflict (as they do perceptually for the English and French vowel space in Levy & Strange (2008)), and learners may need more evidence both from phonological distributions and phonetic variability before the vowel system shows phonetic drift. For example, a vowel like /ʌ/ in English does not seem to have an obvious counterpart in French, but the Paris Residents in this study seem to be moving it toward French /œ/ (see Figure 1). This may be because these are the most central vowels for each language, causing them to drift closer in F1 and F2, though they do not seem to be a phonological "match" with respect to rounding (and this study does not provide acoustic detail about the rounding in /œ/.)

Moreover, to the extent that the French vowel space is actually contracted in some ways relative to English (notably in F1 for the low vowels), it may be that drift driven by vowel space contraction requires more input than a rapid six weeks of immersion before it can be observed. It is obviously the case that vowel space narrowing can occur, since the lowest vowels in the English vowel space for Paris Residents do contract toward the French vowels (i.e., /æ/ raises toward French /a/, and /ɑ/ also fronts). Further exploration of the influence of both the size of the L2 vowel inventory and the relative compactness of the space on L1 phonetic drift requires information from more language pairings to better understand these effects.

Further addressing the issue of input, the level of L2 immersion outside of the 6 h of L2 classroom time per week in Study Abroad speakers during the first six weeks only equates to less than 5% of their time. With such a minimal amount of their time spent using the languages in an immersive environment, learning the language in class, or speaking with native speakers, it follows that their immersion experience may not have been extensive enough to motivate phonetic drift. Interestingly, this is different from the similar group in Chang (2012) of American English learners of Korean. This similar group had a total of 20 weekly hours of class time in an L2 environment, but also reported very little usage outside of the classroom. It may be that the lower amount of classroom time for the French learners as compared with the Korean learners is at least partly responsible for the failure to manifest some sort of phonetic drift in the present study. If this amount of exposure in the Study Abroad speakers is compounded with the complications of a more crowded and compacted vowel space, and as compared with the Paris Resident speakers who do display phonetic drift, it suggests that a rapid shift in these first six weeks of immersion in the vowel space of American English learners of French is not possible under these circumstances and would require more in-depth immersion to potentially be observed within the same timeframe as Chang (2012).

In comparison with the Study Abroad speakers, the level of L2 French immersion for Paris Resident speakers is markedly different in terms of variety and quantity of exposure. For the Paris Residents, both global F1 and individual American English vowels differed from the monolingual American English values. The high vowels /i/ and /u/ were shown to be primarily backer in comparison with monolingual American English speakers while vowels /ɛ, ɑ, ʌ/ were shown to be significantly higher. In addition, the low vowel /æ/ was also shown to be higher and fronter. For /i/ and /u/, this movement is in the direction of the French category norm for these vowels, and while the Paris Residents' English F2 values seem pulled by the French ones for the similar phonemes, they do not go as far back as the French categories (see Figure 1). For /æ/ and /ɛ/, a shift upward and forward demonstrates movement toward the closest French categories of /a/ and /ɛ/. As evidenced by Figure 2, French /a/ is quite fronted compared with American English /ɑ/, and is closer to American English /æ/ (e.g., Strange et al., 2007). With significant F2 differences for both /æ/ and /ɑ/, French /a/ may be influencing both of these American English vowels for the Paris Residents. We note here that though there were some differences between the Paris Resident speakers from the Northeast and those from other regions in the US, the differences between the groups were gradient versions of change in the same direction. This suggests that the influence of L2 French on the English vowel categories is stronger than whatever phonetic differences might have been present before the participants began learning French.

Vowels such as American English /ɪ/, /ʊ/ and /o/ are not affected by drift. In the case of the first two, this may be because there is no French phoneme category that would lead to an equivalence classification for American English speakers (Flege, 1987; Flege & Hillenbrand, 1984). As for American English /o/, it is notable that French /ɔ/ overlaps the American English production, which may prevent the Paris Resident speakers from equating American English /o/ with French /o/ and showing backing and/or raising of /o/.

A further explanation for the difference between the Study Abroad Time 2 and Paris Resident groups in measurable phonetic drift may be found in Chang's (2013) discussion of the role of ongoing L2 engagement. Chang points out that in the studies which have shown either the perception of foreign accentedness or substantial amounts of phonetic drift for L1 phonetic categories, the speakers are typically living in the country of their L2, are regular users of that language and/or have had prolonged and continuous exposure to it (de Leeuw et al., 2010; Dmitrieva et al., 2010; Herd, Walden, Knight, & Alexander, 2015; Mayr et al., 2012; Mennen, 2004; Sancier & Fowler, 1997). Our Paris Residents have had prolonged contact with L2 and are currently living in an L2 ambient environment, whereas the Study Abroad speakers—though in France at the time of recording—are likely not using enough French for their L1 vowel system to experience drift at this point.

An obvious question, then, is why Flege's (1987) speakers do not show phonetic drift for F2 for /u/, whereas our comparable group did. Though there is already quite a difference between the average values for native French and American English F2 for /u/ in Flege (approximately 1200 Hz for French and 1650 Hz for American English), that difference grows in the current study (M = 1288 Hz for French, 1809 Hz for American English), possibly due to an expansive fronting trend for /u/ in many American English dialects (e.g., Fridland, Kendall, & Farrington, 2014; Grieve, Speelman, & Geeraerts, 2013; Oder, Clopper, & Ferguson, 2013). In addition, the error bar around Flege's American English back vowel in his Figure 2 is quite large, suggesting substantial variability in the American English /u/ category. Though it is purely speculative, the increased difference between the phonological back vowels seen 30 years after the original Flege study may provide more distance between the vowels and, therefore, more room for drift to occur.

4.2 High vowel case study

The high vowels of French and American English are an interesting set to examine whether learners show evidence of attempting to separate their two languages, and what phonetic implementation is used, because this is a case where a specific set of F1 and F2 values map onto two different phoneme categories in the two languages, namely /u/ in American English and /y/ in French. In the comparison between American English and native French in this study shown in Figure 4, the means for American English /u/ and French /y/ are entirely overlapping, and as already noted, previous research has confirmed that American English learners of French have shown substantial difficulty in separating these vowels in a variety of perceptual, lexical, and production tasks (e.g., Baker & Smith, 2010; Darcy et al., 2012; Gottfried, 1984; Levy, 2009; Levy & Law, 2010; Levy & Strange, 2008).

To take one of these studies into further consideration, Levy and Strange (2008) examined the perception of the French vowels /i/, /y/, /œ/, and /u/ in both inexperienced and experienced American learners of French in the discrimination of minimal pairs. Results indicated that American English learners were not equally accurate on all of the pairs. The high vowel pair /y-u/ tended to be the most difficult, consistently confusing even the experienced French language professionals in discrimination tasks. Levy and Strange (2008) argue that the perception of French high vowels is directly affected by level of experience, but still retains some ambiguity even in experienced learners. With regard to production, Flege (1987) demonstrated that American English /u/ was still comparable with the monolingual standard even for the most experienced learner, but the experienced group did produce French /u/ closer to the French norm than the inexperienced learners did (though they did not achieve native-like values). As for /y/, the only American English group that did not match the native French speakers in F2 were the American students who had lived in France for only nine months, and who were not residing in France at the time of the study. The groups with more experience and exposure did not show significant differences from the native speakers.

In the present study, Paris Resident speakers were shown to have a complementary shift in their high vowel space. As shown in Figure 4, both American English /i/ and /u/ for Paris Resident speakers have significantly decreased F2, indicating that their native L1 category demonstrates phonetic drift toward the French norm. Though Paris Residents have retracted their French /i/ to achieve native norms, and their American English /i/ has drifted toward the French values, the speakers still keep their two languages separate. For /u/, results show that Paris Residents have retracted their French and American English /u/ compared with monolingual American English. A further observation about /u/ in both languages for the Paris Resident speakers is they are produced with very large variability (see the very large ellipses for Paris Resident French and American English /u/ in Figure 4). Since American English /u/ is fronted in many dialects, and an altogether different English phoneme is not moving into the high back space, it may be that wide variability in general is characteristic of the production of American English /u/. It is possible, then, that Paris Resident speakers are essentially transferring their American English variability to their French productions, which have a lower F2 mean, but which still span a larger F2 range than the other high vowels do. Another reason that the Paris Residents may be retracting French /u/ is to make space in the French inventory for /y/. Paris Residents are making no attempt to distinguish American English /u/ from French /y/ in F2, though it may be relevant that they do distinguish these phonemes in F1.

The bulk of this evidence suggests that the Paris Residents' vowels in this study do show substantial effects of phonetic drift, but they also show evidence of not fully assimilating vowels that do have counterparts in both languages, like /i/ and /u/. Unlike Flege (1987), the French vowel /y/ produced by Paris Residents is not significantly different from American English /u/ in F2, but it is distinguished in F1, and in fact it is the only Paris Resident vowel that has a higher F1 than the other high vowels. Again, the difference between Flege's findings and those reported in this study may have to do with different starting values for American English /u/ in his speakers. The current results confirm the SLM's predictions based on equivalence classification (Flege, 1988, 2007), including the expectation that there is a bidirectional influence between the proficient learners' L1 and L2 phonetic categorizations, though the addition of a new phoneme (here, French /y/) can lead to the movement of other phoneme categories in order to keep at least some space between them (Flege, 2003). However, category assimilation does not necessarily occur between the most likely categories that could be equated, and speakers are not necessarily forming a single merged category that stands in for both American English and French /i/ or /u/ (cf. Flege, 2007).

These results suggest that a fuller picture of when to expect phonetic drift, full assimilation, or dissimilation can benefit from consideration of a broader phonetic space beyond individual vowel pairs (as in Chang, 2012; Guion, 2003), or of the range of variability present in L1. As both Chang (2012) and Guion (2003) point out, many shifts of the type seen in the current study are consistent with a drive toward dispersion of the vowel space to maintain as many distinctions as possible (cf. Adaptive Dispersion Theory, Flemming, 2004; Liljencrants & Lindblom, 1972; Lindblom, 1986). It seems that the dual forces of a propensity toward assimilation (as Flege 2003, 2007 points out, native categories are powerful attractors) and dispersion in order to maintain at least some category separation—albeit not perfectly—are at play in our Paris Resident speakers.

5 Conclusion

While the English learners of Korean in Chang (2012) showed a shift in the global vowel space after just five weeks of immersion, the American English learners of French in this study did not show the same shift after six weeks. This is likely a combination of the differences in vowel space between French and Korean (more crowded versus sparser), and the relatively low amount of L2 input that the American English learners in the current study experienced. On the other hand, the

experienced speakers in our study do show substantial amounts of phonetic drift in their vowels. There are still open questions in the research program investigating phonetic drift, especially pertaining to the amount of input, ambient language and length of time that is necessary for drift to begin occurring. In the case of Korean it was rapid, but other language pairings may require a much greater proportion of L2 input at the initial stages, or a longer time with L2 input. More generally, the presence of phonetic drift under any circumstances—and the fact that it can change as input changes (e.g., Sancier & Fowler, 1997)—confirms that the interaction between phonetic subspaces is bidirectional. Pressures such as overall dispersion, or assimilation or dissimilation may be active in continually updating phonetic categories as the amount of input for both languages changes over time.

Acknowledgements

We would like to thank Alec Marantz, John Moran, Lena Warnke, Clio McConnell, Mathilde Ugarte-Alizon and Faeren Adams for the input on this project during various stages. We would also like to thank the PEP lab at NYU and the audience at the 2016 Linguistic Society of America meeting in Washington, DC for their feedback.

Funding

This research was funded by the Dean's Undergraduate Research Fund at New York University.

Notes

1. A potential point of concern is that the Study Abroad Time 1 speakers are used to examine whether both short-term and long-term exposure to French leads to phonetic drift in American English vowels, even though they are the same speakers at both time points in the short-term case. In order to confirm that the native American English vowels of these speakers are an adequate comparison set for both the Study Abroad and Paris Resident speakers, it is useful to compare the current results to comparable data reported by Strange et al. (2007). Strange et al. provide a comparison of Parisian French and American English vowels as produced by speakers living in New York, which is very similar to the Study Abroad group in this study. Many of the differences shown in the comparison of Study Abroad and the native French speakers in this study in Table 1 are also evident in the Strange et al. (2007) results. In particular, their Figure 2 shows that English /i, e/ are fronter than their French counterparts by about 1 Bark, and French /u/ is backer by about 1 Bark. Likewise, English low vowels /æ, ɑ/ are lower than French /a/, though the difference is more prominent for female speakers than for male speakers. English /æ/ is considerably fronter than French /a/ by 1–1.5 Bark for both males and females. Given these similarities, our Study Abroad T1 group represents a reasonable benchmark for native American English.
2. In addition to dialect, another potential concern in the data is whether the American English back vowel /u/ is in a coronal or non-coronal context. Results in Levy and Strange (2008) show that in their comparison of American English and French vowels, when American English /u/ is surrounded by coronal consonants, F2 is significantly higher (the vowel is more fronted) than when /u/ is surrounded by non-coronals. If a preceding coronal has a similar effect in our data, then /u/ should be treated separately depending on consonantal context since any potential influence of the French /u/ on phonetic drift may differ depending on whether or not the American English /u/ is in a coronal context. To investigate this, we classified the American English /u/ tokens in our data as to whether or not they were preceded by a coronal consonant. This vowel, as well as /i/ as a control, were then analyzed for F2 values with a linear mixed effects model with Phoneme (/i/ vs. /u/), Coronality ([+coronal], [-coronal]), and Group (Study Abroad T1 and Paris Resident) as factors. Results for F2 showed no significant effect of coronality ($\beta = 0.008$, $t = 0.112$), and no interactions between coronality and phoneme ($\beta = 0.083$, $t = 0.803$) or coronality and group ($\beta = -0.002$, $t = -0.052$). Therefore, we are confident that a preceding coronal consonant does not lead to a fronter /u/ than in other contexts and we continue to collapse across all tokens of /u/.

References

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*, *116*, 3099–3107.

Baker, W., & Smith, L. C. (2010). The impact of L2 dialect on learning French vowels: Native English speakers learning Québécois and European French. *Canadian Modern Language Review*, *66*(5), 711–738.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-6. Retrieved from http://CRAN.R-project.org/package=lme4

Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics*, *58*, 71–86.

Best, C. (1995). A direct-realist view of cross-language perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore: York Press.

Best, C., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, *109*(2), 775–794.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro & O.-S. Bohn (Eds.), *Second Language Speech Learning: The Role of Language Experience in Speech Perception and Production* (pp. 13–34). Amsterdam: John Benjamins.

Chang, C. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268.

Chang, C. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, *41*, 520–533.

Clopper, C. G. (2009). Computational methods for normalizing acoustic vowel data for talker differences. *Language and Linguistics Compass*, *3*(6), 1430–1442.

Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., & Scott, J. H. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English– L2 French acquisition. *Second Language Research*, *28*(1), 5–40.

de Leeuw, E., Mennen, I., & Scobbie, J. M. (2012). Singing a different tune in your native language: first language attrition of prosody. *International Journal of Bilingualism*, *16*(1), 101–116.

de Leeuw, E., Schmid, M., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, *13*(1), 33–40.

Dmitrieva, O., Jongman, A., & Sereno, J. A. (2010). Phonological neutralization by native and non-native speakers: the case of Russian final devoicing. *Journal of Phonetics*, *38*(3), 483–492.

Flege, J. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47–65.

Flege, J. (1988). Effects of equivalence classification on the production of foreign language speech sounds. In A. James & J. Leather (Eds.), *Sound Patterns in Second Language Acquisition* (pp. 9–39). Dordrecht: Foris.

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. 229–273). Timonium, MD: York Press.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities* (pp. 319–355). Berlin: Mouton de Gruyter.

Flege, J. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. I. Hualde (Eds.), *Laboratory Phonology 9* (pp. 353–380). Berlin: Mouton de Gruyter.

Flege, J., & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America*, *83*(2), 729–740.

Flege, J., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, *78*(3), 708–721.

Flege, J., Schirru, C., & MacKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*, 467–491.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner & D. Steriade (Eds.), *The Phonetic Bases of Markedness*. Cambridge: Cambridge University Press.

Fougeron, C., & Smith, C. (1999). *French Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.

Fridland, V., Kendall, T., & Farrington, C. (2014). Durational and spectral differences in American English vowels: Dialect variation within and across regions. *The Journal of the Acoustical Society of America*, *136*(1), 341–349.

Gottfried, T. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, *12*, 91–114.

Grieve, J., Speelman, D., & Geeraerts, D. (2013). A multivariate spatial analysis of vowel formants in American English. *Journal of Linguistic Geography*, *1*(1), 31–51.

Guion, S. (2003). The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*, *60*, 98–128.

Herd, W. J., Walden, R. L., Knight, W. L., & Alexander, S. N. (2015). Phonetic drift in a first language dominant environment. *Proceedings of Meetings on Acoustics*, *23*(1), 060005.

Iverson, P., & Evans, B. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America*, *122*, 2842–2854.

Iverson, P., & Evans, B. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America*, *126*, 866–877.

Kendall, T., & Thomas, E. (2010). Vowels: Vowel Manipulation, Normalization, and Plotting in R (Version 1.1). Retrieved from http://lingtools.uoregon.edu/norm/

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (Version 2.0-33) [R package].

Lee, H. B. (1999). *Korean Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.

Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.

Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *Journal of the Acoustical Society of America*, *126*(5), 2670–2682.

Levy, E. S., & Law, F. F. (2010). Production of French vowels by American-English learners of French: Language experience, consonantal context, and the perception-production relationship. *Journal of the Acoustical Society of America*, *128*(3), 1290–1305.

Levy, E. S., & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, *36*(1), 141–157.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel systems: the role of perceptual contrast. *Language*, *48*, 839–862.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. Ohala & J. Jaeger (Eds.), *Experimental Phonology* (pp. 13–44). Orlando: Academic Press.

Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch–English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition*, *15*(4), 687–700.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, *32*(4), 543–563.

Nearey, T. (1978). *Phonetic Feature Systems for Vowels*. Bloomington, IN: Indiana University Linguistics Club.

Oder, A. L., Clopper, C. G., & Ferguson, S. H. (2013). Effects of dialect on vowel acoustics and intelligibility. *Journal of the International Phonetic Association*, *43*(1), 23–35.

Patkowski, M. (1990). Age and accent in a second language. Reply to James Emil Flege. *Applied Linguistics*, *11*, 73–89.

Peng, S.-h. (1993). Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. *Phonetica*, *50*, 245–260.

Sancier, M., & Fowler, C. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, *25*, 421–436.

Scovel, T. (1988). *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. New York: Newbury House/Harper and Row.

Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America*, *122*(2), 1111–1129.

Ulbrich, C., & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics*, *45*, 26–42.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, *24*(2), 245–261.

Appendix A: Speaker Data

Study Abroad

Gender	Age	Hometown	Language other than FR or EN	Other information
Female	19	Tampa, FL	n/a	
Female	20	Bluebell, PA	Korean L2 fluency (classroom)	Parents spoke Korean
Female	20	Los Angeles, CA	Spanish L2 intermediate (classroom)	
Female	18	Nashville, TN	Mandarin Chinese L2 fluent (classroom)	
Male	18	Pittsburgh, PA	German L2 proficiency (classroom)	
Male	21	Boston, MA	Hebrew L2 proficiency (classroom)	Parents spoke Hebrew
Female	18	Tampa, FL	n/a	
Male	20	Los Altos, CA	n/a	Parents spoke Vietnamese
Female	20	Alameda, CA	Spanish L2 intermediate (classroom)	n/a
Female	18	New York, New York / Farmington, CT / Greenwich, CT	Spanish fluency (classroom) + 1 year of immersion in Spain	Parents spoke Arabic
Female	20	Long Island, NY / Chicago, IL	n/a	n/a

Paris Resident

Gender	Age	Starting age of French acquisition	% of day speaking French	Hometown	Language other than FR or EN	Other information
Female	33	16	75%	Milledgeville, GA	n/a	French Partner
Female	52	38	65%	Boston, MA	Small amount of Spanish L2 and German L2 (classroom)	French Partner
Male	28	14	10%	Bucks County, PA	n/a	
Male	35	13	90%	Long Island and Manhattan, NY / Bucks County, PA	Basic-Intermediate Modern Greek L2, Spanish L2, German L2, Italian L2 (classroom)	Mother spoke Greek
Female	56	34	65%	Los Angeles, CA	Spanish L2 proficiency (classroom)	Parents spoke Yiddish
Female	48	20	50%	Southwest USA	Some Spanish and Arabic	
Male	30	13	75%	New York, NY	n/a	
Male	36	27	70%	Ithaca, NY	Basic/Intermediate Italian and German L2	
Male	40	14	80%	Leominster, MA	n/a	French Partner
Female	34	14	65%	NJ, FL, GA, Germany, UK	n/a	
Male	33	15	70%	Birmingham, AL	n/a	

Parisian French

Gender	Age	Hometown (now all in Paris)	Language other than FR or EN
Female	19	Versailles, France	n/a
Female	19	Seine-et-Marne, France	Spanish L2 intermediate
Female	19	Orleans, France	Spanish L2 proficiency, minimal Chinese L2 (classroom)
Male	19	Paris area, France	Portuguese bilingual / Spanish L2 proficiency; minimal Chinese L2 (classroom)
Female	19	Paris, France	Minimal German L2, minimal Spanish L2 (classroom)
Female	19	Houilles, France	Spanish L2 basic-intermediate (classroom)
Female	20	Picardy, France	German L2 proficiency

Appendix B: Stimuli

French stimuli used in production task. The French sentence (in orthography) is on the left, and the CVC target sequence in IPA is on the right.

/i/		/y/	
il fit deux bâtiments	/fid/ - fit deux	je n'ai pas bu de jus	/byd/ - bu de
les tortues vivent longtemps	/viv/ - vivent	j'ai mangé du fromage	/dyf/ - du fromage
le régime de Vichy pendant la guerre	/viʃ/ - Vichy	il n'y a pas de puce sur ma carte	/pys/ - puce
il fume la pipe à tabac	/pip/ - pipe	il faut que tu comprennes	/tyk/ - tu comprennes
vous dîtes qu'elle voulait nous voir	/dit/ - dîtes	elle fut paresseuse	/fyp/ - fut paresseuse
c'est la vie des enfants	/vid/ - vie des	on peut le faire vu qu'il nous comprend	/vyk/ - vu qu'il
/e/		**/ɛ/**	
il a des enfants	/dez/ - des enfants	je veux qu'il sèche ses vêtements	/sɛʃ/ - sèche
tu as tes affaires	/tez/ - tes affaires	les enfants jettent des pierres	/ʒɛt/ - jettent
je connais la fée verte	/fev/ - fée verte	je voudrais cette pomme	/sɛt/ - cette
il veut ses vêtements	/sev/ - ses vêtements	le chef de cuisine	/ʃɛf/ -chef
tout ce que j'ai préparé	/ʒep/ - j'ai préparé	elle ne cesse de fumer	/sɛs/ - cesse
laisse-le chez toi	/ʃet/ - chez toi	j'ai la tête ailleurs	/tɛt/ - tête
/ø/		**/a/**	
il a deux voitures	/døv/ - deux voitures	tu connais sa sœur	/sas/ - sa sœur
c'est peu de fromage	/pød/ - peu de	je veux ta pomme	/tap/ - ta pomme
voilà ceux qui m'aiment	/søk/ - ceux qui	mon chat joue avec mes amis	/ʃaʒ/ - chat joue
j'ai eu trois feu verts sur la route	/føv/ - feu verts	mot de passe oublié	/pas/ - passe
je fais le vœu de ne plus fumer	/vød/ - vœu de	je veux qu'il sache la vérité	/saʃ/ - sache
les jeux d'enfants	/ʒød/ - jeux d'enfants	je n'ai pas d'espèces	/pad/ - pas d'espèces
/œ/		**/ɔ/**	
son chef d'œuvre est formidable	/dœv/ - d'œuvre	un petit os du corps humain	/tɔs/ - petit os
je voudrais un petit œuf bio	/tœf/ - petit œuf	le chien mange un petit os de dinde	/tɔs/ - petit os
la coquille d'œuf fait partie de la recette	/dœf/ - d'œuf	je vais aller voir mon pote ce weekend	/pɔt/ - pote
une petite œuvre artistique	/tœv/ - petite œuvre	allez à la poste pour envoyer votre lettre	/pɔst/ - poste
elles peuvent danser ce soir	/pœv/ - peuvent	son chapeau a une poche secrète	/pɔʃ/ - poche
ils peuvent acheter quelque chose	/pœv/ - peuvent	c'était mon rêve de gosse de le voir	/gɔs/ - gosse

(continued)

/o/		/u/	
passons par la côte d'Azur	/kot/ - côte	il faut que tous ces hommes mangent	/tus/ - tous ces
un sot qui fait rire	/sok/ - sot qui	au bout du couloir	/bud/ - bout du
j'aime Peau d'Âne	/pod/ - Peau d'Âne	j'aime les petites pousses de laitue	/pus/ - pousses
mon beau-frère	/bof/ - beau-frère	c'est vous qui l'aimez	/vuk/ - vous qui
le chat saute sur moi	/sot/ - saute	la soie est très douce et fine	/dus/ - douce
le taux de chômage	/tod/ - taux de	c'est le fou qui s'est trompé	/fuk/ - fou qui

American English stimuli used in production task. The English sentence (in orthography) is on the left, and the CVC(C) target sequence in IPA is on the right. The sentence "one doe at the lake" was removed from analysis.

/i/		/ɪ/	
I'll have tea cups	/tik/ - tea cups	he fits the bill	/fɪts/ - fits
they see four criminals	/sif/ - see four	remove the pits from the peaches	/pɪts/ - pits
four keys on a ring	/kiz/ - keys	she hits two homeruns at every game	/hɪts/ - hits
the feed is outside	/fid/ - feed	there are first aid kits in all rooms of the building	/kɪts/ - kits
let the tea bag steep for five minutes	/tib/ - tea bag	she sits on the sofa	/sɪts/ - sits
two bees on the flower	/biz/ - bees	he had a few zits on his chin	/zɪts/ - zits
/ɛ/		/æ/	
he bets two hundred	/bɛts/ - bets	he bats two hundred	/bæts/ - bats
she said we could eat	/sɛd/ - said	the tack in the wall fell out	/tæk/ - tack
a small peck on the cheek	/pɛk/ - peck	put the cap by the recycle bin	/kæp/ - cap
they built a shed in the backyard	/ʃɛd/ - shed	a small pack on the ground	/pæk/ - pack
the dog jets down the hallway	/dʒɛts/ - jets	the back of the store	/bæk/ - back
two debts to resolve	/dɛts/ - debts	she is sad about the breakup	/sæd/ - sad
/ɑ/		/ʌ/	
the pots are in the kitchen	/pɑts/ - pots	place the cup on the counter	/kʌp/ - cup
there are dots on the floor	/dɑts/ - dots	she putts down the green	/pʌts/ - putts
there are cots for the children	/kɑts/ - cots	he butts heads with his brother	/bʌts/ - butts
a small dock in the harbor	/dɑk/ - dock	she shuts both doors at night	/ʃʌts/ - shuts
she spots four criminals	/spɑts/ - spots	the rock juts out from the mountain	/dʒʌts/ - juts
the court reporter jots down the case	/dʒɑts/ - jots	he guts two fish per day	/gʌts/ - guts

(continued)

/o/		/ʊ/	
one doe at the lake	/do.æ/ - doe at	the book from the store	/bʊks/ - books
say Joe's name	/dʒoz/ - Joe's	she foots the bill every time	/fʊts/ - foots
she will sew patches onto the vest	/sop/ - sew patches	two cooks in the kitchen	/kʊks/ - cooks
the archer's bow tightens	/bot/ - bow tightens	he took four apples from the store	/tʊk/ - took
hand the hoe to the gardener	/hot/ - hoe to	watch out for the fish hooks on the dock	/hʊks/ - hooks
they are calling the tow truck	/tot/ - tow truck	she shook the tree	/ʃʊk/ - shook
/u/			
I'll have two cups	/tuk/ - two cups		
the dew point comes in at dusk	/dup/ - dew point		
we are building a chicken coop tomorrow	/kup/ - coop		
the food is outside	/fud/ - food		
they sue four criminals	/suf/ - sue four		
I bought five shoes yesterday	/ʃuz/ - shoes		