Concordia Working Papers in Applied Linguistics

Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014 © 2014 COPAL

Sociophonetic Convergence in Native and Non-Native Speakers of French

Monique Murphy
The University of Texas at Austin

Abstract

Sociophonetic research on monolingual speakers has indicated that there are numerous unconscious shifts in an individual's phonetic space depending on social context. However, second language accent is largely assumed to be a static entity. This study examined production of Standard French's 10 oral vowels in native French speakers and non-native learners. Participants were paired in conversation with both a native and a non-native interviewer. Overall measurements indicated significant differences in vowel space depending on interviewer, mediated by proficiency level. It was also found that vowel shifts largely mirrored the differences in vowel space between interviewers in both direction and degree. This provides evidence that both natives and non-natives use similar strategies to collapse social distance through their speech. The findings of this study should compel evaluators of non-native accent to always consider to whom learners are speaking as well as phonological proficiency in the L2.

When it comes to speech production and accent, researchers to date have come to somewhat conflicting conclusions. On the one hand, it is generally accepted that a speaker will be the most understandable in a

language learned early, and will have an obviously "native" accent in that language, experiencing no difficulty in producing all necessary phonetic, phonological, and prosodic information necessary. Second language learners often struggle to achieve this level of easy production, and are easily identified as being "non-native" through some unorthodox pronunciation or prosodic contour. While factors that are within a speaker's control, such as motivation, integration with an L2 speech community, extent of formal study of the L2, practice, and phonetic training do contribute to greater success in nativelike speech production, they are generally found to be less predictive than factors of circumstance and luck such as age of acquisition, length of residency in the L2-speaking country, gender, and speech perception abilities (Flege et al., 1995; Piske et al., 2001; Birdsong, 2006; Best & Tyler, 2007). Nevertheless, some learners strive for an unnoticeable accent due to the social stigma and communication difficulties a non-native accent can produce, especially when the accent is a marker of an ill-favored ethnic or socioeconomic group (Kalin & Rayko, 1978; Munro & Derwing, 1995; Bresnahan et al., 2002; Derwing, 2003; Munro, 2003).

On the other hand, sociolinguistic research has established that one's accent, even if native, is by no means static. Speakers often phonetically accommodate to one another during the context of a conversation. They may converge to maximize comprehensibility and to decrease social distance when one speaker views the other favorably, and diverge when one speaker views the other with suspicion or dislike (Giles et al., 1991; Babel, 2009). These phonetic shifts are not conscious, and are easily affected by subtle priming on the part of researchers, even in the L2. Beebe (1981) found that Chinese-Thai learners produced their Thai with a more identifiable "Chinese" or "Thai" accent depending on the ethnicity of their interviewer, while Berkowitz (1989) found that Dominican L2 English speakers shifted their accents depending on how empathetic they perceived their interviewers to be. Zuengler (1982) found that L2 English speakers changed their output if their interviewer showed a negative attitude towards bilingual education.

If accent is at the same time strongly linked to biological and cognitive factors, but also affected by social context, what can be said about the evolution of second language accent over the course of one or many conversations? This study attempted to discover whether a single factor, the native language of an interviewer, would affect speech production in both native and non-native speakers of French. If it can be said that a speaker sounds "less" or "more" nativelike solely based on the knowledge

that the interviewer is a native or non-native speaker, it cannot be said that nativelikeness is a fossilized attribute, or that accent can be measured on a numeric scale.

METHOD

Participants

24 native and non-native French speakers of all levels of proficiency were recruited from an ad placed in a campus events calendar at the University of Texas. Of these 24 participants, 18 completed all sections of the experiment. Only the data from these 18 participants was used in the analysis. The participants ranged from 21-63 years of age (*M* = 31.4 years). There were 13 non-native speakers (9 female, 4 male) and 5 native speakers (4 female, 1 male). The non-native speakers were further divided into Intermediate (8) and Advanced (5) levels of proficiency (native speakers of French will be referred to throughout this paper as "natives"). Proficiency was based on a judgment by the native speaker interviewer, as well as their comprehension level and the variety of verb forms and vocabulary used by each participant during the experiment. Three native speakers spoke a hexagonal French dialect, two a northern African French dialect, and none had lived in the United States before age 18. All speakers described themselves as middle class and had completed at least some college. Non-natives began learning French from a range of 10-38 years of age (*M* = 16.9 years). All but one advanced speaker had lived for at least one month in a Francophone community (France, Belgium, Francophone Canada, or Burkina Faso). Participants were paid $10 at the end of each completed session.

Experiment

Each participant completed two interviews with a single interviewer in a private, quiet office on campus at the University of Texas. The first interview was conducted by a female L1 English, L2 French speaker of advanced proficiency, and the second interview was conducted by a female native speaker of hexagonal French with a high proficiency of L2 English. During pre- and post-recording, interviewers were instructed to use only their L1 with the participants. The interview consisted of biographical and personal questions (where the participant had lived, when they began learning French, daily habits, hobbies, etc.), and was

terminated by the interviewer at a natural stopping point sometime between 15 and 20 minutes after starting the first question.

Audio was recorded using a TASCAM DR-05 digital voice recorder at 44.1 kHz in .wav format and was imported onto an Asus M50Vm laptop computer. All data was analyzed using PRAAT. For each interview file, the first third was eliminated, and F1 and F2 values were extracted from the midpoint of the steady state from no more than 35 tokens of each of the 10 French oral vowels /i/, /y/, /e/, /ø/, /ɛ/, /œ/, /u/, /o/, /ɔ/, and /a/. 1-2 tokens of these vowels were also taken from each interviewer over the course of all interviews until 35 of each was reached. Due to the low frequency of /œ/ in the data, as well as its frequent coarticulation with /ʁ/, this vowel was not analyzed in participant/researcher comparisons, but is included in all other sections of the analysis. Data was coded for F1, F2, group membership (whether the conversation occurred with an English or French native speaker), gender, age, proficiency level, and age at first learning.

All data were analyzed using R (R Development Core Team, 2013) and the R packages *lme4* (Bates & Maechler, 2009) *vowels* (Kendall & Thomas, 2012), and *languageR* (Baayen, 2011). The data were analyzed using linear mixed effects models, with speakers (and for global analysis, vowels) as random effects. Condition (whether the data came from a conversation with a native English or French speaker), proficiency, age, age of first learning, gender, and L1 were included as fixed effects. Measurements were separated by vowel after a global model was created, as global comparisons of F1 and F2 masked whether certain vowels were more or less affected by group membership. Model fit was assessed using likelihood ratio tests. This paper presents t-values from the mixed effects models with degrees of freedom calculated as the number of observations minus the number of fixed effects, as well as MCMC-estimated p-values that are considered significant at the 0.05 level.[1]

[1] According to Winter (2011), reporting test statistics for mixed models is a controversial issue, since degrees of freedom cannot be accurately estimated, and considers this method a "shorthand."

RESULTS

Overall Differences in Formants Between Conditions

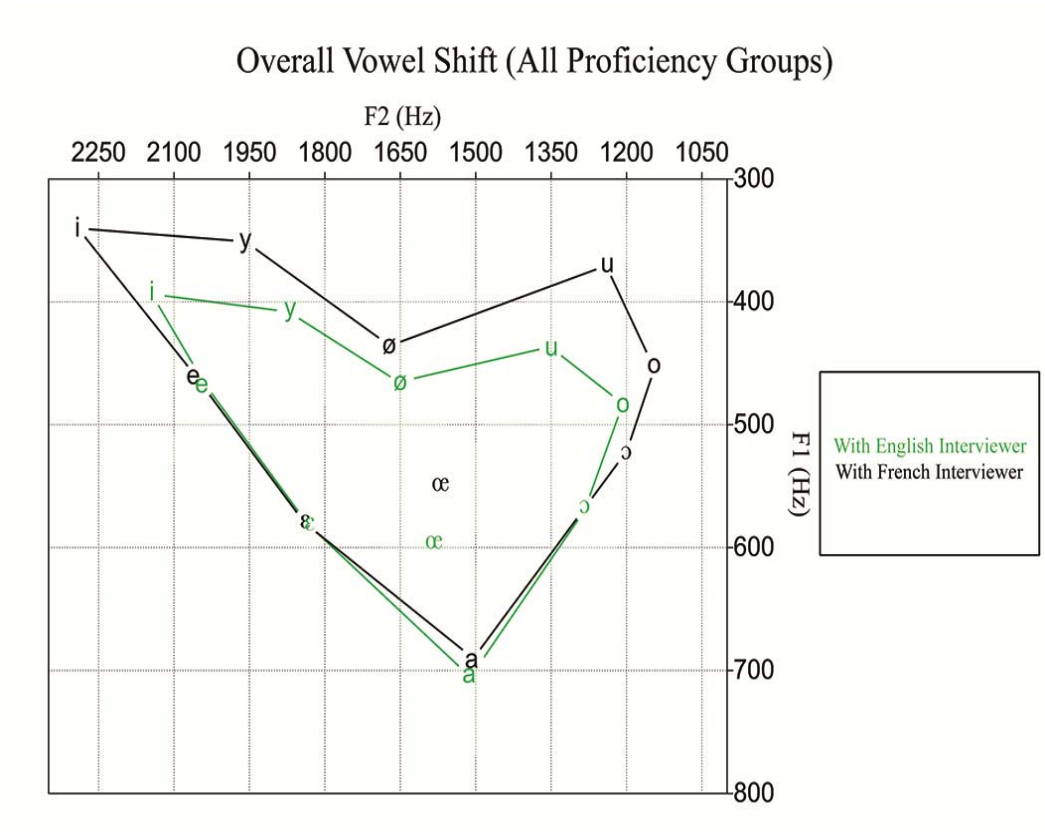

Figure 1. Vowel Shift Depending on Interviewer, Collapsed Proficiency Levels

The interaction between condition and proficiency provided the model of best fit for both F1 and F2 data when speaker and vowel were treated as random effects. Overall differences between conditions are shown in Figure 1. Participants produced an average F1 of 32 Hz lower in the French condition than in the English condition. This difference was significant ($t_{(10015)} = -18.6$, $p<.001$). F2, however, was an average of 6.6 Hz higher in the French condition than in the English condition, a non-significant difference ($t_{(10015)} = 1.36$, $p>.05$). The fact that overall F2 level was not significantly different between conditions is likely due to the more peripheral vowel space in the French condition, with movement forward by front vowels and movement backward by back vowels obscuring any differences.

Overall Differences in Formants and in Shift Trend by Proficiency

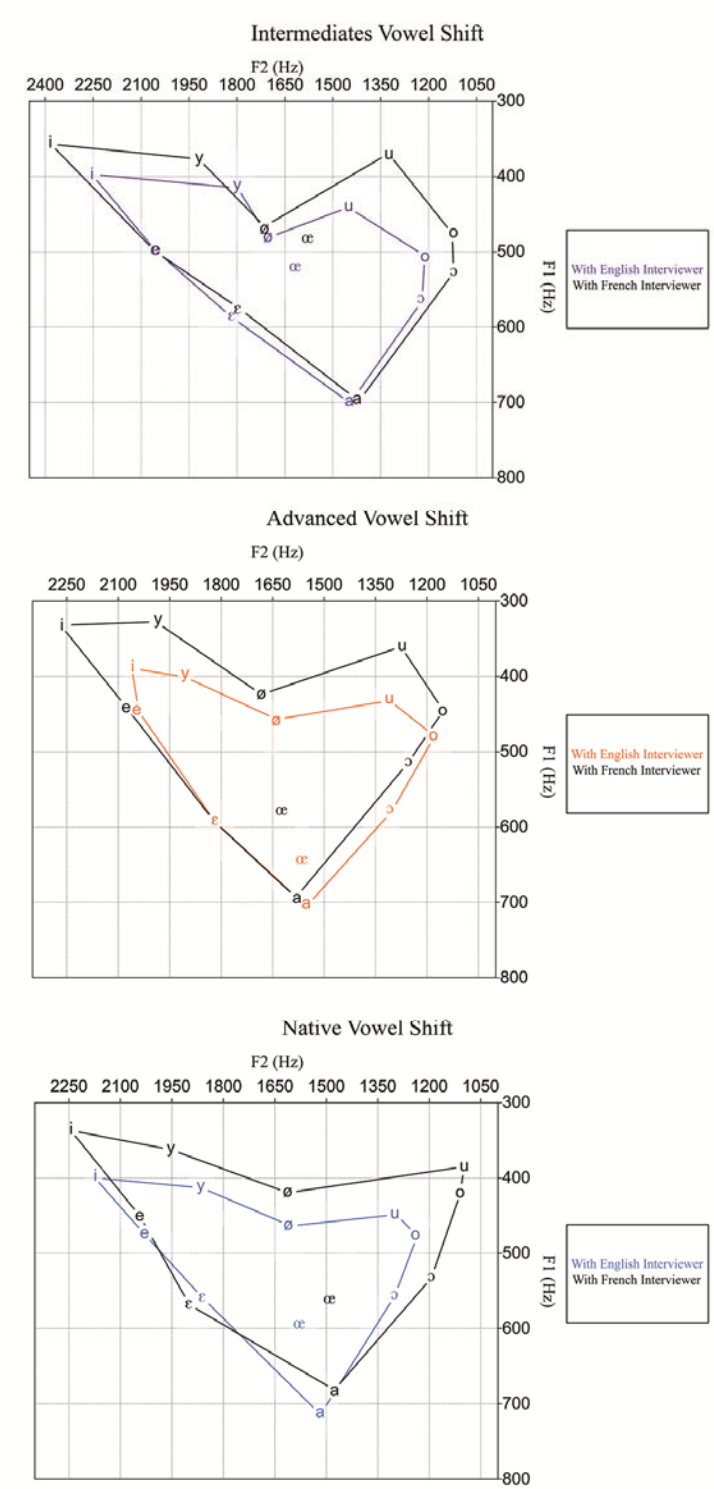

Figure 2. Vowel Shift Depending on Interviewer, By Proficiency Level

Figure 3. Vowel Shift Depending On Proficiency, By Interviewer

All proficiency levels differed significantly on F1 depending on condition, indicating that the vowel space was shifted upward in the French condition for every group by varying degrees. See Figures 2 and 3 for more detailed comparisons of the shifts in formant values by proficiency group. The mean F1 for the intermediate group lowered by 23 Hz with a native French speaker, $t_{(10015)} = -11.65$, $p<.001$. The mean F1 for the advanced group lowered by 35 Hz, $t_{(10015)} = -13.21$, $p<.001$.The mean F1 for the native group lowered by 37 Hz, $t_{(10015)} = -11.64$, $p<.001$.

There was no significant difference between proficiency levels on mean F1 in the English condition. (intermediate and advanced: $t_{(10015)} = -0.672$, $p>.05$. Intermediate and native: $t_{(10015)} = -0.428$, $p>,05$. Advanced and native: $t_{(10015)} = 0.197$, $p>.05$). In the French condition, however, intermediate-level learners showed significantly higher mean F1 than both advanced and native speakers (intermediate and advanced: $t_{(10015)} = -2.71$, $p<.01$. Intermediate and native: $t_{(10015)} = -3.06$, $p<.01$) , but advanced and native speakers did not differ significantly: $t_{(10015)} = -0.644$, $p>.05$.

For F2, only advanced speakers differed significantly between conditions. The mean F2 for advanced speakers was greater by 56 Hz when with a French native speaker, indicating a more fronted vowel space, $t_{(10015)} = 4.45$, $p<.001$. Intermediate speakers raised their F2 by 16 Hz and native speakers lowered their F2 by 5 Hz, but these differences were not significant (for intermediates: $t_{(10015)} = 1.01$, $p>.05$. For natives: $t_{(10015)} = -0.35$, $p>.05$).

In the English condition, there were no significant differences between any of the proficiency levels on F2 (for intermediate and advanced: $t_{(10015)} = 0.55$, $p>.05$, for intermediate and native: $t_{(10015)} = 0.60$, $p>.05$, and for advanced and native: $t_{(10015)} = 0.12$, $p>.05$). In the French group, the difference in mean F2 was not significant between intermediate and native ($t_{(10015)} = ,-0.97$ $p>.05$), but was significant between advanced and native ($t_{(10015)} = -3.08$, $p<.01$), and near significant for advanced and intermediate ($t_{(10015)} = -1.94$, $p=.05$).

Differences in Formants and Shift Trend By Vowel, Proficiency

In the following sections, it is often the case that the best fit model for F1 was not the same as for F2. For example, if Condition alone provided the model of best fit for the data, it can be safely said that Proficiency did not make a significant impact on the data, and that it can be assumed that in both conditions all groups produced similar F1/F2 values.

/i/. Condition alone provided the model of best fit for the data on F1 for /i/, and the interaction between condition and proficiency provided the model of best fit for the data on F2 for /i/.

The mean F1 in the English condition was 53 Hz lower when speakers tested with a French native. This difference was significant, $t_{(1188)} = -14.74$, $p<.001$.

For F2, all levels of proficiency showed significantly higher F2 in the French condition than in the English condition. Intermediates showed a difference of 130 Hz ($t_{(1187)} = 5.09$, $p<.001$), Advanced speakers showed a difference of 206 Hz ($t_{(1187)} = 10.26$, $p<.001$), and Natives showed a difference of 71 Hz ($t_{(1187)} = 2.80$, $p<.01$).

/y/. The interaction between condition and proficiency provided the model of best fit for the data on F1 for /y/, and condition differences alone provided the model of best fit for the data on F2 for /y/.

All proficiencies showed significantly lower mean F1 for /y/ in the French condition than in the English condition: Intermediates differed by 35 Hz between conditions ($t_{(882)} = -3.17$, $p<.001$), Advanced speakers by 71 Hz ($t_{(882)} = -9.45$, $p<.001$) and Natives by 50 Hz ($t_{(882)} = -5.69$, $p<.001$).

The French condition significantly differed from the English condition on F2, showing an average F2 for /y/ that was 89 Hz higher in the French condition ($t_{(882)} = 5.07$, $p<.001$).

/e/. The factors of Condition and Proficiency (but not the interaction between them) provided the model of best fit for the data on F1 for /e/, and Condition differences alone provided the model of best fit for the data on F2 for /e/.

There was no significant difference in F1 between the English and French conditions (difference of 9 Hz, $t_{(1219)} = -1.79$, $p>.05$).

There was no significant difference in F2 between the English and French conditions ($t_{(1219)} = 1.75$, $p>.05$).

/ø/. The interaction between Condition and Proficiency provided the model of best fit for the data on F1 for /ø/, and Condition differences alone provided the model of best fit for the data on F2 for /ø/.

Advanced and Natives showed a significant decrease in F1 for /ø/ between conditions. Advanced speakers differed F1 by 35 Hz between conditions ($t(1197) = -5.25$, $p<.001$) and Natives differed by 44 Hz ($t(1197) = -5.28$, $p<.001$). Intermediates differed by 11 Hz, which was not a significant difference ($t(1197) = -1.38$, $p>.05$).

Participants showed a difference of 22 Hz in F2 between the English and French conditions for /ø/, which was significant: t(1197) = -2.04, p<.05.

/ɛ/. Condition differences alone provided the model of best fit for the data on both F1 and F2 for /ɛ/. There were no significant differences between English and French conditions on F1 for /ɛ/ ($t_{(1158)}$ = -0.35, *p*>.05).

There were no significant differences between English and French conditions on F2 for /ɛ/ ($t_{(1158)}$ = 0.88, *p*>.05).

/œ/. The factors of Condition and Proficiency (but not the interaction between them) provided the model of best fit for the data on F1 for /œ/, and condition differences alone provided the model of best fit for the data on F2 for /œ/.

The French condition had an F1 of 47 Hz lower than the English condition, which was a significant difference ($t_{(134)}$ = -2.79, *p*<.01).

F2 did not significantly differ between conditions for /œ/ ($t_{(134)}$ = -0.37, *p*>.05). In addition, Intermediates and Natives had lower F2 in the French condition, but Advanced speakers had higher F2 in the French condition.

/u/. Condition differences alone provided the model of best fit for the data on F1 for /u/, and the interaction between Condition and Proficiency provided the model of best fit for F2 for /u/.

The French condition had a mean F1 68 Hz lower than the English condition for /u/. This difference was significant: $t_{(1072)}$ = -13.21, *p*<.001.

Intermediates differed by 126 Hz ($t_{(1072)}$ = -3.18, *p*<.01), between conditions for F2, and Natives differed by 202 Hz ($t_{(1072)}$ = -5.21, *p*<.001). Advanced speakers did not significantly differ on F2 between conditions ($t_{(1072)}$ = -1.17, *p*>.05).

/o/. The factors of Condition and Proficiency (but not the interaction between them) provided the model of best fit for the data on F1 for /o/, and the interaction between Condition and Proficiency provided the model of best fit for the data on F2 for /o/.

All proficiency groups had lower F1 in the French condition than in the English condition, and the difference was significant for all groups (Intermediates differed by 32 Hz, $t_{(953)}$ = -3.69, *p*<.001, Advanced by 33 Hz, $t_{(953)}$ = -4.75, *p*<.001, and Natives by 56 Hz, $t_{(953)}$ = -7.19, *p*<.001).

All proficiency groups had lower F2 in the French condition than in the English condition as well, but the difference was only significant for Intermediates (difference of 89 Hz, $t_{(953)}$ = -3.18, *p*<.01) and Natives

(difference of 130 Hz, $t_{(953)} = -5.06$, $p<.001$). Advanced speakers did not differ significantly between conditions ($t_{(953)} = -1.19$, $p>.05$).

/ɔ/. The interaction between Condition and Proficiency provided the model of best fit for the data on F1 for /ɔ/, and Condition differences alone provided the model of best fit for the data on F2 for /ɔ/.

All proficiency groups had lower F1 in the French condition than in the English condition, and the difference was significant for all groups (Intermediates differed by 36 Hz, $t_{(982)} = -2.51$, $p<.05$, Advanced by 63 Hz, $t_{(982)} = -6.21$, $p<.01$, Natives by 24 Hz, $t_{(982)} = -2.07$, $p<.05$).

F2 differed by 83 Hz between the English and French conditions. This was a significant difference ($t_{(982)} = -6.40$, $p<.001$).

/a/. Condition differences alone provided the model of best fit for the data on F1 for /a/, and the interaction between Condition and Proficiency provided the model of best fit for the data on F2 for /a/.

All proficiency groups had lower F2 in the French condition than in the English condition, and the difference was significant for all groups (Intermediates differed by 103 Hz, $t_{(982)} = -3.79$, $p<.001$, Advanced by 54 Hz, $t_{(982)} = -2.78$, $p<.01$, Natives by 106 Hz, $t_{(982)} = -4.76$, $p<.001$).

DISCUSSION & CONCLUSION

It is clear from this data that sociophonetic influences, such as conversational partner, can have an impact on the output of a speaker, even in the speaker's non-native language. Native French speakers showed a more centralized, slightly lower oral vowel space when speaking with an L1 English, L2 French speaker, than they did when speaking to another L1 French speaker. Surprisingly, this pattern was quite similar to Intermediate and Advanced L2 French speakers'. Even students of French with somewhat limited experience conversing with native speakers were capable of shifting their production in the presence of the native French interviewer in this experiment in the same direction as native speakers. Though L2 French speakers did not have nativelike production on some vowels, it was these "troublesome" vowels that appeared to shift the most. High front /y/ and high back /u/ are often difficult to produce in a nativelike way for L1 American English speakers. It is possible that speakers were more conscious in their speech or more tense in their articulation when they believed that they were being judged

on their proficiency and accent by the native French interviewer, and less careful with a non-native French interviewer.

However, an alternative explanation is that speakers were not accommodating to an internally-represented norm of French accent at all, but to the individual with whom they spoke. In fact, this explanation is attractive in view of the native French speakers' performance. It is not likely that L1 French speakers felt self-conscious about their French accent, but they still moved their vowels in a similar way to L2 French speakers. When the participants' vowel spaces in each condition were compared with the vowel spaces of the interviewers, a stronger case is made for this explanation. In Figure 4, participants' and interviewers' vowel spaced are compared. In Table 1, the difference between participants' F1/F2 values between interviews and the difference between interviewers' F1/F2 values is shown (excluding /œ/, which did not have a sufficient number of tokens in interviewers' speech to compare). If participants were accommodating to the individual to whom they were speaking, rather than an internal representation of the vowel space, one would expect the difference between conditions to be similar to the difference between interviewers. This appears to be the case for the majority of vowels. For example, /u/ was 64 Hz higher and 268 Hz further back in the French interviewer's vowel space than in the English interviewer's vowel space, and participants' /u/ was 68 Hz higher and 111 Hz further back in the French condition than in the English condition. Participants seemed to hew more closely to the L1 French interviewer's vowels than to the L1 English interviewer's vowels. This could be due to the English interviewer's somewhat unusual vowel space, or it could be that participants, consciously or not, *wished* to emulate the native speaker's French pronunciation more, since she served as a kind of linguistic role model.

As for proficiency effects, while all groups moved their vowel spaces in similar directions, native French speakers tended to shift more vowels than non-native speakers, and shifted greater distances. Advanced speakers shifted more vowels than Intermediate speakers, as well. Thus it is likely that accommodation in a second language is a skill that develops over time with greater comfort in expression. However, it is noteworthy that for several formant groups, proficiency and even age of first learning did not contribute to the best-fitting model at all, leaving the best predictor to be the native language of the interviewer.

Though this study was limited to oral vowels and two interviewers, lingering questions could easily be researched further by focusing on different cues (VOT, speech rate, pitch) or comparing productions with a

wide range of interviewers. It is of interest whether the objective differences found in these data would be discriminable to the human ear (in other words, would a listener be able to tell which interview a certain token was extracted from? Are participants judged as having a "more nativelike accent" with a native interviewer?). In addition, this study lacked data from true beginners of French learning. Due to the nature of the task (a 15-20 minute interview exclusively in French), beginning students may have felt discouraged from participating.

Despite these limitations, it is sincerely hoped that this study will provide not only motivation for further studies on sociophonetic influences on L2 speech, but hope for second language learners who have aspirations of nativelikeness in accent. In short, interlocutor identity may affect accent in a particular conversation more than any other factor. This is somewhat at odds with accepted views of "foreign language accent," which is often assumed to fossilize early during learning and be relatively resistant to change. If sounding more like a particular target is the goal of a learner, she would be advised to speak to a wide range of speakers whom she wishes to emulate, rather than practice particular sounds alone or lament a loss of neural plasticity. Furthermore, judgments of "nativelikeness" by researchers should always be approached with caution and appropriate nuance if the mere choice of data collector can skew this perception.

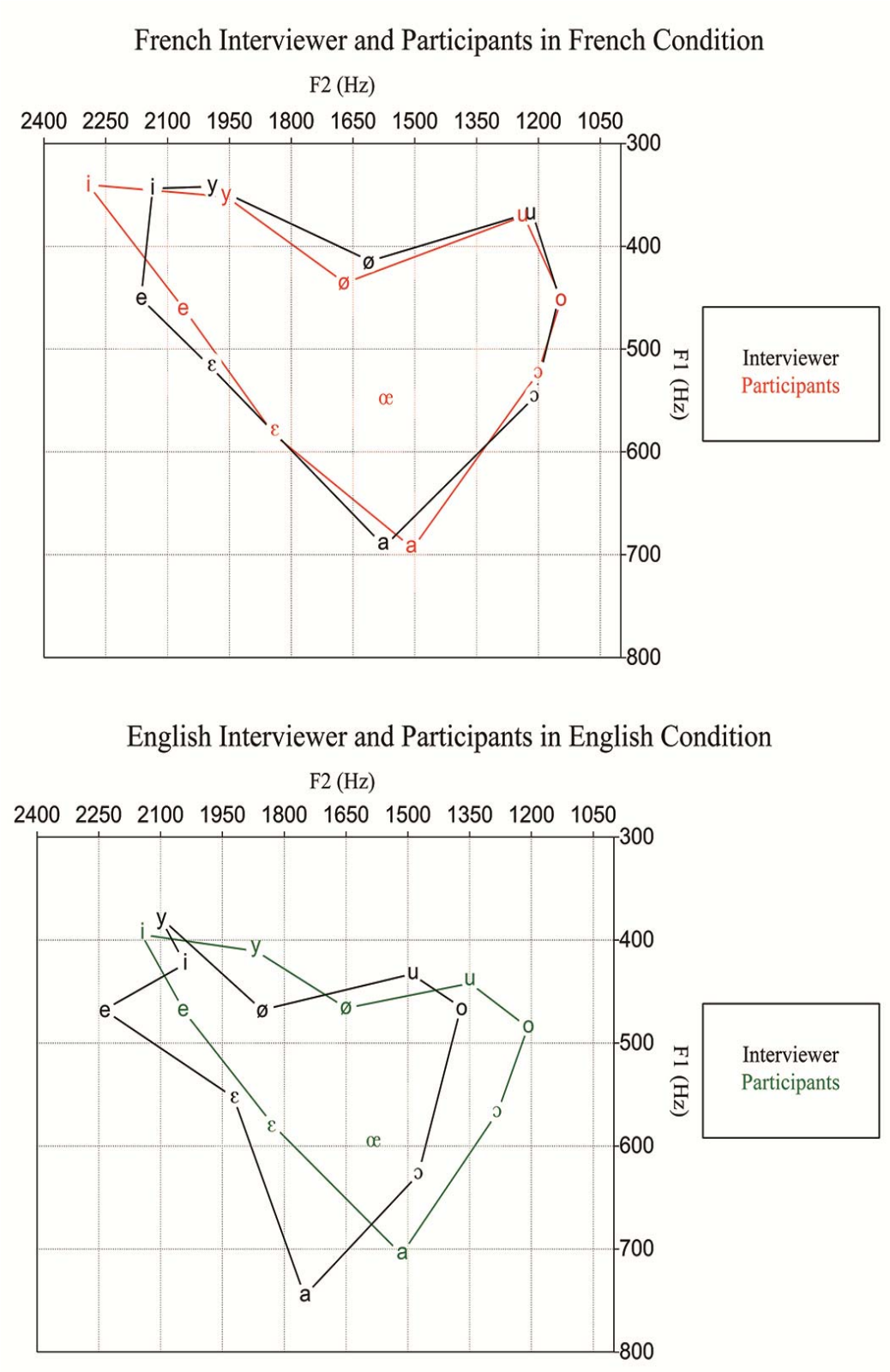

Figure 4. Comparison of Interviewers' and Participants' Vowels

Table 1. Differences in Formant Value by Participant Groups and Interviewer

Vowel	Participant F1 (English-French) (Hz)	Interviewer F1 (English-French) (Hz)	Participant F2 (English-French) (Hz)	Interviewer F2 (English-French) (Hz)
/i/	-52	-79	+148	+96
/y/	-56	-38	+89	-107
/e/	-7	-18	+17	-72
/ø/	-30	-53	+22	-241
/ɛ/	-2	-37	-9	+72
/u/	-68	-64	-111	-268
/o/	-40	-15	-76	-222
/ɔ/	-44	-81	-83	-265
/a/	-12	-81	-5	-173

REFERENCES

Baayen, R. H. (2011). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". R package version 1.4.

Babel, M.E. (2009). *Phonetic and Social Selectivity in Speech Accommodation*. (Doctoral dissertation). Retrieved from University of California, Berkeley archives: http://www.linguistics.berkeley.edu/phonlab/annual_report/documents/2009/Babel_dissertation_2009.pdf

Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes [Computer software manual]. Available from http://CRAN.R-project.org/package=lme4

Beebe, L.M. (1981) Social and situational factors affecting communicative strategy of dialect code-switching. *International Journal of the Sociology of Language, 32*, 139-149.

Berkowitz, Diana. (1989). The effects of cultural empathy on L2 phonological production. In M. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in L2 variation* (pp. 101-114). New York: Plenum Press.

Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception. In O. Bohn & M. J. Munro (Eds.) *Language Experience in Second Language Speech Learning*, (13-34). Philadelphia: John Benjamins.

Birdsong D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*(1), 9-49.

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W.Y., & Shearman, S. M. (2002). Attitudinal and affective response toward accented English. *Language & Communication*, *22*(2), 171-186.

Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59*(4), 547-567.

Flege, J.E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*(5), 3125-3134.

Giles, H., Coupland, J., & Coupland, N. (eds.). (1991). *Contexts of accommodation: Developments in applied sociolinguistics.* Cambridge: Cambridge University Press.

Kalin, R. & Rayko, D. S. (1978). Discrimination in evaluative judgments against foreign-accented job candidates. *Psychological Reports*, *43*, 1203-1209.

Kendall, T., & Thomas, E. R. (2009). Vowels: Vowel manipulation, normalization, and plotting in R. http://cran.r-project.org/web/packages/vowels/index.html.

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal, 20*(2), 38-51.

Munro, M. J. & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*(2), 191-215.

R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

Winter, B. (2011). A very basic tutorial for performing linear mixed effects analyses. Retrieved from http://menzerath.phonetik.uni-frankfurt.de/teaching/R/bw_LME_tutorial_HR.pdf.

Zuengler, J. (1982). Applying accommodation theory to variable performance data in L2. *Studies in Second Language Acquisition, 4,* 181-192.