Mutual influences between native and non-native vowels in production: evidence from short-term visual articulatory feedback training

Natalia Kartushina[1,2*], Alexis Hervais-Adelman[3], Ulrich Hans Frauenfelder[1], Narly Golestani[3]

[1]Laboratory of Experimental Psycholinguistics, FPSE, University of Geneva, 42 bd du Pont d'Arve, 1205 Geneva, Switzerland

[2]Basque Center on Cognition, Brain and Language, Paseo Mikeletegi, 69, 20009 San Sebastian, Spain

[3]Brain and Language Lab, Campus Biotech, Faculty of Medicine, University of Geneva, 9 Chemin des Mines, 1202 Geneva, Switzerland

ABSTRACT

We studied mutual influences between native and non-native vowel production during learning, i.e., before and after short-term visual articulatory feedback training with non-native sounds. Monolingual French speakers were trained to produce two non-native vowels: the Danish /ɔ/, which is similar to the French /o/, and the Russian /ɨ/, which is dissimilar from French vowels. We examined relationships between the production of French and non-native vowels before training, and the effects of training with non-native vowels on the production of French ones. We assessed the acoustic position and compactness of the trained vowels, and of the French /o/, /ø/, /y/ and /i/ vowels, which are acoustically closest to the trained vowels. Before training, the compactness of the French vowels was positively related to the accuracy and compactness in the production of non-native vowels. After training, French speakers' accuracy and stability in the production of the two trained vowels improved on average by

*Corresponding author: n.kartushina@bcbl.eu

19% and 37.5%, respectively. Interestingly, the production of native vowels was also affected by this learning process, with a drift towards non-native vowels. The amount of phonetic drift appears to depend on the degree of similarity between the native and non-native sounds.

Key words: L2 production, articulatory training, L1-L2 interactions, intra-speaker variability, production training, stability in production

1 INTRODUCTION

Second-language (L2) learners often experience considerable difficulty in producing non-native speech sounds, resulting in a foreign accent. Accents are largely attributed to a bias in the perception of L2 sounds, arising from the native (L1) phonology. While the effects of L1 on the production of foreign sounds are well established, less is known about the impact of L2 on the production of native sounds, particularly in novice L2 learners. This study aims to explore the effects of production training with the Danish /ɔ/ and Russian /ɨ/ vowels on the production of native French vowels in speakers with no previous experience with Danish or Russian.

1.1 The influence of L1 on the production of L2 sounds

The effect of the native language on second-language production has been widely documented (Flege, MacKay, & Meador, 1999; Goto, 1971; Long, 1990; Piske, MacKay, & Flege, 2001). For instance, Korean and Spanish learners of English have difficulty in producing the /i/-/ɪ/ contrast, whereas Italian learners experience more difficulty with the /ɚ/-/ʌ/ contrast (Flege, 2003). Here we will try to address how the native language affects L2 production.

It is assumed that at the beginning of L2 learning, L1 phonology influences that of the L2; L1 is used to process L2 sounds in terms of their similarity/dissimilarity[1] to native categories (Archibald, 1998; Flege, 1995). According to Flege's (1995) Speech Learning Model (SLM), for example, similar L2 sounds assimilate perceptually to L1 categories by a mechanism of equivalence classification. This mechanism can block L2 category formation, and native sounds may be used to produce similar L2 sounds. Dissimilar L2 sounds (i.e., those that are sufficiently phonetically different from the closest native category to be perceived as being different from it) do not perceptually assimilate to L1 categories, and novel categories are expected to be established for them.

Consider, for instance, Korean and Japanese learners of the Australian-English /e/-/æ/ contrast. While the English /e/ vowel is assimilated to the respective similar /e/ vowels in both Korean and Japanese, only in Japanese is the English /æ/ vowel phonetically distinct from the closest /e/ category, whereas in Korean it also assimilates to /e/. Consequently, many Korean learners of English do not produce the English /e/ and /æ/ vowels distinctly: the acoustic spaces for these vowels largely overlap, suggesting that they use one native-like category to produce both the English /e/ and /æ/ vowels. Japanese speakers, on the other hand, produce /e/ and /æ/ contrastively since a new category is created for the dissimilar English /æ/ vowel, and the existing L1 /e/ category is used to produce the English /e/ (Ingram & Park, 1997).

Even over extended learning periods, advantages for dissimilar over similar vowels seem to persist. For instance, Japanese speakers assimilate the English /ɹ/-/l/ contrast to one Japanese /ɽ/, with the English /l/ being perceptually more similar to the Japanese /ɽ/ than to the English /ɹ/. At the end of one year of formal learning, native Japanese children improved more

[1]Depending on the theoretical framework adopted, the similarity can be phonetic (Speech Learning Model [SLM] by Flege, 1995) or gestural (Perceptual Assimilation Model [PAM] by Best, 1995).

in their production of the English /ɹ/ than of /l/, as judged by native English speakers (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004).

The age of onset of L2 learning influences the production of similar (i.e., difficult) L2 sounds. Early bilinguals produce similar cross-language sounds distinctly, whereas late bilinguals don't (Guion, 2003; MacLeod, Stoel-Gammon, & Wassink, 2009). This suggests that they use the native category to produce similar L2 sounds (Baker & Trofimovich, 2005). These results show that early L2 learners are more likely to create novel categories for similar L2 sounds than are late L2 learners.

In addition to the acoustic similarity of L1 and L2 sounds, the distribution, and more specifically the compactness of L1 sound categories, has more recently been shown to influence the perception and production of L2 sounds (Kartushina & Frauenfelder, 2013, 2014). In these studies, the compactness of person-specific vowel productions was measured by quantifying the spatial distribution, that is, the spread of produced tokens in the F1-F2 vowel space (see methods for more on how this was calculated). It was shown that Spanish speakers whose L1 productions were more compact (i.e., less variable) produced French (L2) sounds more accurately than those whose productions were more variable. This was attributed to the fact that speakers with more compact L1 productions have a greater proportion of their acoustic space available for the formation of new L2 sounds. Conversely, speakers with more variable L1 productions have little acoustic space available for the formation of new L2 sounds, which are therefore more likely to fall within the space of existing L1 sounds and to be confused with them (Kartushina & Frauenfelder, 2014). These results are consistent with Flege's postulate regarding a shared L1-L2 space, which claims that "sounds are related to each other at a position-sensitive allophonic level" (1995, p. 239).

To summarize: (1) dissimilar L2 sounds are produced and acquired more easily than similar ones; (2) the detrimental effects of cross-language similarity on L2 production increase with the age of L2 acquisition, and (3) the compactness of L1 categories and their similarity to L2 sounds affect L2 production.

1.2 The influence of L2 on the production of L1 sounds

While the effects of L1 on L2 production have been extensively described, less is known about the effects of L2 learning on the production of native speech sounds. Grosjean (1989) argued that L1 and L2 coexist and interact constantly in bilinguals. For instance, the extent of use of either the native or second language has been shown to affect speakers' performance in the other language. Piske and colleagues (Piske et al., 2001) have shown that the strength of the L2 accent is affected by the amount of continuous use of the native language. In Anglophone areas of Canada, native Italian speakers who continuously and frequently used Italian in their everyday life (on average 53%) were perceived by native Canadian speakers as having a stronger Italian accent when speaking English than those who used Italian less frequently (on average 10%). Conversely, and more curiously, native productions may themselves become foreign-language accented after only a few months of immersion in an L2-speaking country. In a case study, Sancier and Fowler (1997) showed that, following a four-month stay in the US, productions of a native Brazilian Portuguese speaker were perceived by native Portuguese listeners as being American-English (AE) accented (for effects of L2 immersion on L1 production latencies, see Baus, Costa, & Carreiras, 2013; Ivanova & Costa, 2008; Linck, Kroll, & Sunderman, 2009).

So how, specifically, does L2 use affect L1 production? In the SLM model, Flege postulates that "phonetic categories established in childhood for L1 sounds evolve over the lifespan to reflect the properties of all L1 and L2 phones" (p. 239, 1995). Although to date there are no longitudinal studies available having tracked the evolution of L1 and L2 categories, some existing studies on L1 and L2 phonetic production in bilinguals support this claim (Barlow, 2014; Flege, 2003; Flege & Eefting, 1987a, 1987b; Flege, Schirru, & MacKay, 2003; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008; Guion, 2003; MacLeod et al., 2009; Mora, Keidel, & Flege, 2015; Mora & Nadeu, 2012; Sancier & Fowler, 1997; Sundara, Polka, & Baum, 2006). These studies reveal that experience with an L2 can have three possible effects on native categories (see Figure 1): (1) no change, (2) drift toward the L2 category and (3) deflection away from the L2 category, to maximize opposition with it.

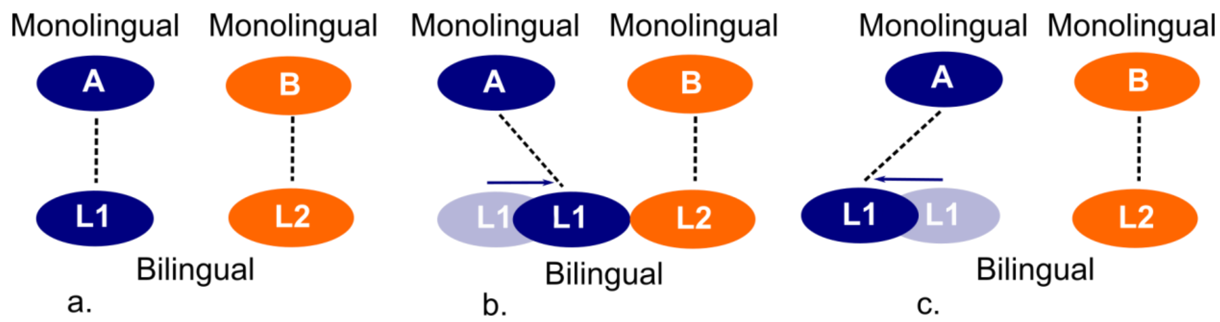

Figure 1. Modifications that L1 categories can undergo after learning L2 sounds in bilinguals as compared to the same phonetic categories of monolingual speakers of the respective languages (A and B) a) No change; b) L1 category drifts toward the similar, L2 category; c) L1 category deflects away from the L2 category to maximize the L1-L2 contrast. Note that the L2 sounds are also likely to undergo modifications, but for illustration purposes these are shown as being static.

The type of change that L1 sounds can undergo depends on several factors, mainly including: (1) the degree of (perceived) similarity to the closest L2 sound, (2) experience with the L2 and L1 (i.e., related to age of acquisition, amount of use, etc.), and (3) proficiency in the L2. These (and other) factors have recently been reviewed in a paper by Kartushina, Frauenfelder, & Golestani (accepted), and are briefly described in the next sections.

1.2.1 When does the native category change?

The few existing studies that have assessed phonetic production in simultaneous bilinguals suggest that in these individuals, production in both of the languages does not differ from those of monolingual speakers of the respective languages (Guion, 2003; MacLeod et al., 2009; Sundara et al., 2006; see, however, Fowler et al., 2008 for contradictory results). Moreover, simultaneous bilinguals are able to distinctly produce even sounds that are similar across their two languages (see Figure 2a). The situation is different in non-proficient bilinguals (see Figure 2b). Due to equivalence classification (or perceptual assimilation), they use native categories to produce similar L2 sounds (Flege, 1987; Flege & Eefting, 1987a), and the production of native categories seems to be unaffected, i.e., these categories remain unchanged and are used to produce both L1 and L2 sounds.

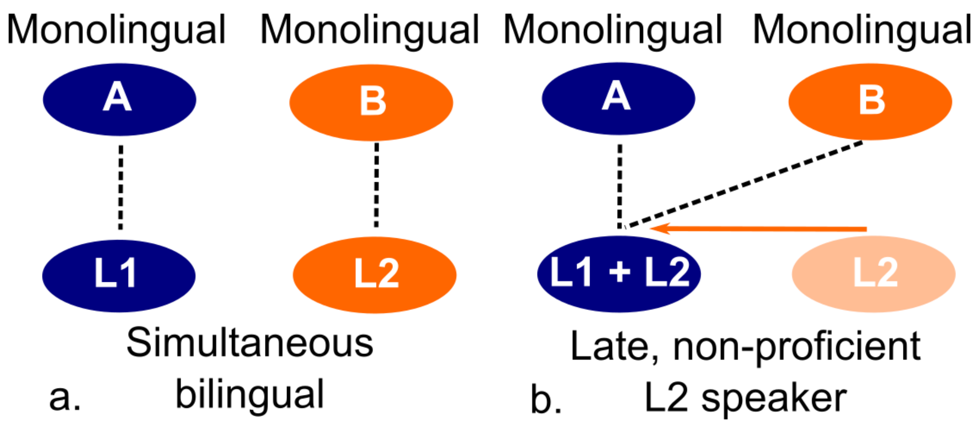

Figure 2. Production of L1 and L2 sounds in a) simultaneous bilinguals and in b) late, non-proficient L2 speakers as compared to monolingual speakers of the respective languages (A and B).

1.2.2 When the native category drifts towards the non-native category

Drift of native towards non-native sounds (see Figure 1b) has been reported in late L2 speakers who have been immersed in an L2-speaking environment (Chang, 2012; Flege, 1987; Major, 1992; Sancier & Fowler, 1997) and in early bilinguals whose L2 has become their dominant language (Mora et al., 2015; Mora & Nadeu, 2012). This was found to be the case in the longitudinal case study described above, where after a 4-month stay in an English-speaking environment, a native Portuguese speakers VOTs increased to approach those of English during the production of Portuguese stop consonants (Sancier and Fowler, 1997) (see Figure 3a). Remarkably, these changes were not permanent and her Portuguese productions regained their previous VOTs after a two-month stay in Brazil, where Portuguese is spoken. Importantly, despite these category drifts, stop consonants in Portuguese and English continued to be produced distinctly.

Even a shorter stay in an L2-speaking country can affect the production of native sounds in naïve L2 speakers. It has been shown that in native speakers of American English learning Korean in Korea during a five-week stay, the acoustic properties of their English vowels and consonants approached those of Korean sounds. This drift of the L1 phonetic system toward L2 phonetic categories was attributed to assimilatory mechanisms explained by ‘global linkages’ (those occurring across groups of sounds and at different levels, i.e., segmental and suprasegmental) between English and Korean speech sounds (Chang, 2012). Note however, that for vowels, only the F1 drifted toward Korean F1 values, with no significant change in the F2. Interestingly, when the data were analyzed separately for the different vowels, significant or marginally significant drift in F1 was found for only six out of the eleven L1 vowels examined. This suggests that some vowels may be more affected by L2 experience than others. Phonetic drift toward non-native sounds has also been reported by Major (1992), who studied five native American-English speakers who had been living in

Brazil for 12-35 years since the ages of 22-36. These participants had shortened their VOTs (toward Portuguese realizations) during the production of native /p/, /t/ and/k/ stops, especially in casual speech. The amount of drift varied as a function of L2 proficiency; the better was their mastery of Portuguese, the more Portuguese-like their English productions were (but see Chang, 2013, for contradictory results). The drift of native categories toward the non-native ones has also been reported in early bilinguals; L1-Catalan-L2-Spanish bilinguals who used Spanish more frequently were found to show drift of their Catalan /ɛ/ vowel toward the similar Spanish /e/ vowel compared to those who spoke the L2 less frequently (Mora et al., 2015; Mora & Nadeu, 2012).

The studies described in this section suggest that well-established L1 categories can be malleable in L2 learners and in bilinguals, and that they can change even after relatively short amounts of immersion in a new linguistic environment.

When immersion in a non-native linguistic environment is prolonged and the L2 becomes dominant, it has been shown that the L1 and L2 categories move closer to one another, leading to merging of two sounds into an intermediate category. For example, Flege (1987) analyzed the VOTs of initial [t] and [t^h] stop consonants produced by French-English bilinguals (native French speakers living in Chicago), in French and English words. The results revealed that the VOTs for the French [t] (short-lag) and for the English [t^h] (long-lag) were very similar for both languages (51 and 49 ms for French and English, respectively), and not representative of prototypical /t/ sounds in either language (see Figure 3b). Their VOTs were longer in French (i.e., more English-like) and shorter in English (i.e., more French-like) compared to those of monolingual speakers of French and of American-English, these being 33 and 77 ms, respectively. These results suggest that in speakers who are immersed in an L2-speaking environment for a prolonged duration and who predominantly speak the L2, the drift of L1 categories towards L2 ones can be accompanied by a parallel drift of L2 categories towards L1 ones. Similar findings were reported in the study by Major (1992) mentioned above, in American-English speakers who had been living in a Portuguese-speaking environment for a prolonged period of time.

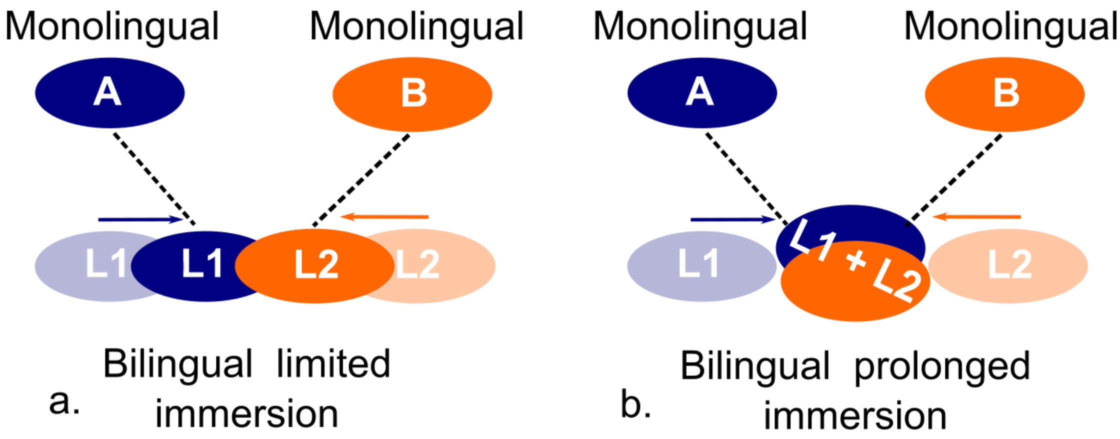

Figure 3. Production of L1 and L2 sounds in late L2 speakers who have been immersed in an L2 linguistic environment for limited versus prolonged periods of immersion as compared to monolingual speakers of the respective languages (A and B).

These two studies suggest that proficient bilinguals who use their L2 continuously in their everyday life create intermediate, possibly even merged phonetic categories that combine the relevant acoustic properties of both languages.

1.2.3 When the native category deflects away from an L2 category

Deflection of native categories away from L2 ones, or as Chang calls it, “dissimilatory drift” (2012, p.252), has been reported in late, proficient bilinguals who are able to perceive the phonetic differences between similar L1 and L2 sounds (Flege & Eefting, 1987a, 1987b; Major, 1992) (see Figure 4). In production, these speakers deflect L1 and L2 sounds away from one another relative to the productions of monolinguals of the respective languages, thereby creating contrast between similar sounds (i.e., exaggerating their dissimilarities) in a common L1-L2 interspace. Flege and Eefting assessed the production of the Dutch and English /t/ stops in Dutch speakers of English (Flege & Eefting, 1987a). The results revealed that only in proficient English speakers (i.e., those with the most native-like accent in English, as judged by native speakers), was the native Dutch VOT shorter (17 ms) than that of the prototypical Dutch /t/ (23 ms), which is itself shorter than the longer VOT of English (90 ms). Similarly, deflection of native from non-native vowels has been reported in Spanish speakers of English who started learning English at the age of 5-6 years: they produced the Spanish /p/, /t/ and /k/ with shorter VOTs (18 ms) than did monolingual Spanish speakers (23 ms) (Flege & Eefting, 1987b). In both studies, the authors attributed the shorter VOTs to the need for enhancing the phonetic contrast with the long VOTs of English.

Phonetic deflection has also been demonstrated for L2 sounds. Flege and colleagues (2003) assessed early Italian-English bilinguals' production of the English /e^{I}/ vowel, which is similar to the Italian /e/. It was found that they exaggerated their production of the /e^{I}/ vowel compared to monolingual English speakers such that acoustically, it moved even further from the Italian /e/. This effect was attributed to the dissimilation of the English /e^{I}/ from the Italian /e/. However, in this study, the production of native, Italian vowels was not assessed and therefore conclusions cannot be drawn about changes in L1 phonetic production.

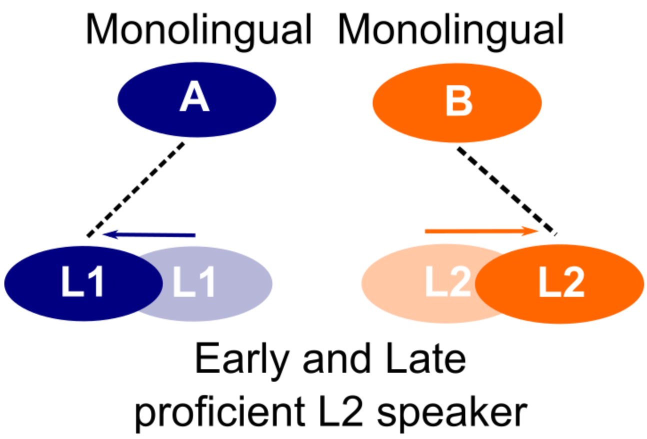

Figure 4. Production of L1 and L2 sounds in early and late proficient L2 speakers as compared to monolingual speakers of the respective languages (A and B).

1.3 Summary and limitations of previous L2 learning studies

An analysis of the literature reveals that changes to native categories (either shifts toward or away from non-native categories) have mainly been found in proficient L2 speakers, or in those who are immersed in an L2 environment. Conversely, native productions of non-proficient L2 speakers who are not immersed in an L2 environment remain unchanged relative to those of monolingual speakers. Although there are fewer studies on L1 vowels than on consonants, the results of at least one study that analyzed the production of L1 vowels and consonants suggest that the effects of an L2 on the production of L1 vowels are similar to the L2's effects on the production of consonants (Chang, 2012, 2013). The above conclusions, however, should be approached with caution for the following reasons: First, and crucially, since in previous studies the production of both L1 and L2 sounds was not assessed longitudinally (i.e., before and after L2 learning), one cannot make claims about pre-learning phonetic production and therefore about whether the changes were really due to learning. Secondly, very few studies compared bilinguals' productions in both languages (L1 and L2) to those of monolingual speakers of the respective languages. Therefore, the results provide only indirect evidence regarding L2 speech sound learning and its effects on the production of L1 sounds. In order to understand how the learning of non-native sounds affects the production of native categories, one would ideally (1) explore learning of L2 sounds in naïve listeners (so as to control for the amount of experience with the L2), and (2) compare the production of L1 and L2 sounds to those of native speakers of the respective native and non-native languages. Chang's (2012) study, described above, partially addresses these two issues. In this study, naïve English speakers who were learning Korean for five weeks were longitudinally assessed on their production of English consonants and vowels. Their production of English sounds before and after learning was compared to that of Korean sounds produced by native Korean speakers. Importantly, this study assessed L1 productions at multiple phonological levels (subsegmental, segmental, and global), and therefore provides a very rich analysis of L1 changes. This study, however, lacks important information regarding pre- and post-training production of the non-native Korean vowels to which the participants were exposed. Further, the speech production parameters that were analyzed did not take into account the natural dispersion of the native and non-native target categories in F1/F2 space, since mean F1 and F2 were examined separately.

1.4 Previous vowel training studies

The current study made use of vowels to address questions about the learning of speech sound production and the impact of such learning on native phonetic production. Vowels are of particular interest for exploring these questions since the establishment of new vowel categories can easily be visualized in the acoustic space of vowels, where native and non-native categories can be compared to each other. In this section we will provide a brief overview of some previous vowel production training studies.

In phonetic training studies, typically, acoustic representations of non-native productions are used to provide learners with feedback during training. For example, Dowd and colleagues (1998) used spectrograms of the resonance frequencies of the vocal tract (F1, F2) as feedback to train English speakers to produce isolated French oral vowels. This method was successful in improving the production of only those vowels having phonetic differences that are easily detected based on this feedback (e.g., F2 is noticeably different for the /y/-/u/ vowels). Carey (2004) trained native Korean speakers to produce three English vowels (/æ/, /ɜ:/ and /ɔ/) by providing them, on each trial, with graphic representations, in F1/F2 space, of their productions alongside those of native English speakers. Similar Korean vowels were also displayed. Improvement was reported for only one of the three trained vowels (/æ/), possibly due to the overloaded feedback display. Indeed, in order for feedback to be effective, it should be easily interpretable, and should provide concise information regarding the critical acoustic dimensions to be learned (Öster, 1997). In a study by Aliaga-García and Mora (2007), native Spanish speakers were trained not only to produce but also to perceive the English /i:/-/ɪ/ and /æ/-/ʌ/ vowel contrasts. Apart from the different production tasks that were administered (e.g., articulatory and visual descriptions, imitation, reading, native input), participants also underwent 15 minutes of training with a visual pronunciation software application called EyeSpeak, which graphically displays the position of vowels in the F1-F2 vowel space. It was found that training improved production along the F1 dimension for the /i:/-/ɪ/ vowels only. The limited training effects may be due to the very short amount of training (15 minutes) that was administered using this method. In another study on training vowel production by Kartushina and colleagues (Kartushina, Hervais-Adelman, Frauenfelder, & Golestani, 2015), a real-time analysis of the acoustic properties (i.e., F0, F1 and F2) of vowels produced by non-native speakers was implemented to provide them with immediate, trial-by-trial visual feedback about their articulation, alongside that of the same vowel produced by native speakers. Using this method, it was shown that one hour of training per vowel improved the production of four Danish /e/, /ɛ/, /y/ and /ø/ vowels by an average of 17% in naïve French-speaking participants (Kartushina et al., 2015). This method improved the production even of non-native vowels that are perceptually similar to native ones, most likely because the feedback provided was easily interpretable, and because it represented crucially contrastive information about the trained sounds.

1.5 Current study

In order to explore the impact of L2 learning on the production of native sounds, monolingual French speakers were trained to produce two foreign vowels, the Danish /ɔ/ and the Russian /ɨ/, using the articulatory feedback training technique described above (Kartushina et al., 2015). The Danish /ɔ/ is similar to the French /o/ vowel, and the Russian /ɨ/ is dissimilar from French vowels.

In the current study, we improved the method used in the study by Kartushina and colleagues (2015) (1) by increasing the phonetic variability of stimuli (more tokens and more speakers), and (2) by providing information, during the feedback, about target acoustic spaces derived from the productions of native Danish and Russian speakers, rather than based on individual target items (see Methods section). This made the stimuli and the feedback more representative of the non-native target vowel categories, so as to increase the effectiveness of training (Lively, Logan, & Pisoni, 1993; Wong, 2013).

The Danish /ɔ/, which is acoustically similar to the French /o/ vowel (and which we will henceforth call an 'L1-proximal' vowel), and the Russian /ɨ/, which is not similar to French vowels (i.e., and which we will call an L1-distant vowel) (see Figure 5) were selected for two reasons: (1) to test the SLM claim that dissimilar non-native sounds are easier to acquire than similar ones (Flege, 1995), and (2) to assess changes in L1 sound production as a function of its similarity to the newly-learned L2 sounds. More information regarding the selection of these vowels and their similarity to French vowels can be found in the 'non-native vowels' section (see Section 2.2.2.1 below).

We wanted to first assess the perceptual similarity of the non-native vowels to French vowels. In an initial, exploratory categorization task, the Danish /ɔ/ and Russian /ɨ/ vowels were therefore categorized in terms of their similarity to the acoustically closest French vowels by native monolingual French speakers who were different from those who participated in the training experiment.

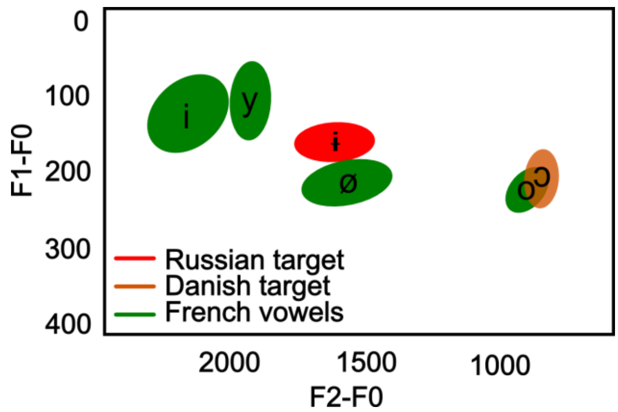

Figure 5. Formant frequencies of French and non-native target vowels (i.e., Russian /ɨ/ and Danish /ɔ/) produced in words by native speakers of French, Russian and Danish, recorded for the study. Ellipses are centered on the mean, and the circumference represents a Mahalanobis Distance of 1 (i.e., which represents a standard deviation of 1).

In the main experiment, we used vowel repetition tasks in order to explore the effects of training on the production of native and non-native vowels. The following native vowels were tested: the French /o/ vowel, which is acoustically very close to the Danish /ɔ/, and the French /ø/, /y/, /i/ vowels. The acoustic proximity of the three latter vowels to the Russian /ɨ/ ranges from close to distant, respectively (see Figure 5). We computed the acoustic positions of L1 vowels with respect to those of the trained vowels, as well as their compactness (i.e., acoustic stability between realizations) before and after training. Finally, in order to test for relationships between the production of native and trained vowels before training, we assessed the pre-training position and compactness of the French /o/, /ø/, /y/ and /i/ vowels, extracted from utterances during a word reading task.

We predicted that training would improve the production of foreign vowels, with greater improvement in the production of the dissimilar, L1-distant Russian /ɨ/ vowel compared to that for the similar, L1-proximal Danish /ɔ/ vowel. We further predicted that the production of foreign sounds before training would be related only to the acoustic properties of similar native vowels. More specifically, based on the results of Kartushina and Frauenfelder (2014), we predicted that speakers whose acoustically-similar French categories (i.e., /o/ and /ø/) are more compact and/or closer to the non-native target vowels (i.e., /ɔ/ and /ɨ/, respectively) would produce the latter more accurately and compactly (i.e., less variably). We therefore predicted such relationships between the Danish and the French /o/ vowels and possibly between the Russian /ɨ/ and the French /ø/ vowels, and no such relationships between the Russian /ɨ/ vowel and the French /y/ and /i/ vowels. Moreover, we expected that training of non-native vowels would affect the production of native vowels, either by deflecting them away from (i.e., dissimilatory shift) or by drifting them towards non-native vowels. Different French vowels (i.e., /ø/, /y/ and /i/) were included in order to test whether learning a non-native L1-distant vowel (i.e., the Russian /ɨ/) affects the production of different native French vowels equally, or whether effects on native vowels differ as a function of their degree of similarity/proximity to the trained vowels.

2 METHODS

2.1 Categorization task

In order to assess the perceptual similarity of the non-native to the native (French) vowels, 24 monolingual French speakers participated in vowel categorization and acceptability rating tasks. They were asked to transcribe, using French vowels, auditorily-presented, isolated Danish /ɔ/ and Russian /ɨ/ vowels, and to mark the degree of acceptability of these vowels in terms of how representative they are of French vowels. The instructions were as follows: “Vous entendrez des sons étrangers. Pour chaque son indiquez à quel son français il ressemble, et marquez votre degré de certitude sur une échelle de 1 à 10 où 1 est 'pas du tout d'accord que c’est comme un son français' et 10 est 'tout à fait d'accord que c’est comme un son français"’. (“You will hear foreign sounds. Please first indicate which French sound it is similar to. Second, rate your assessment of the acceptability of this vowel on a scale from 1 to 10, where 1 is ‘strongly disagree that it is French-like’ and 10 is ‘strongly agree that it is French-like’”). Two examples (i.e., one female and one male voice) of each non-native vowel were then presented. These sounds were taken randomly from the set of stimuli that were recorded for the training (see Stimuli section). Participants could listen to the sounds as many times as they wished until they had completed the task.

2.2 Main experiment

2.2.1 Participants

Twenty native monolingual French speakers (18 female and two male, mean age = 21.9 years) from the student body of the University of Geneva participated, but did not take part in the categorization task. None had any experience with Danish or Russian. Participants completed a language background questionnaire. Some participants had some knowledge of other languages (German, English, Spanish, and Italian), but none of these languages was reported as being spoken proficiently. On a scale from 1 to 5, where 5 is a proficient speaker, the maximum reported rating was 4. They reported no history of speech or hearing impairment. Participants received course credits for their participation. The experiment was conducted in accordance with the Declaration of Helsinki. Participants gave informed consent and were free to withdraw from the experiment at any time.

2.2.2 Stimuli

2.2.2.1. Non-native vowels

Two non-native vowels, the Danish /ɔ/ and the Russian /ɨ/, were chosen based on articulatory, acoustic and perceptual comparisons to French vowels. Vowels were chosen from these two non-native languages because these are not frequently spoken or taught in Switzerland or in France, and we were therefore able to ensure that our participants did not have any previous experience with these vowels. The Danish /ɔ/ vowel is described as an advanced raised open-mid back rounded vowel (Basbøll, 2005; Grønnum, 1997). Two French /ɔ/ and /o/ back rounded vowels are phonetically similar to the Danish one, and are described as open-mid and close-mid vowels, respectively (Georgeton, Paillereau, Landron, Jiayin, & Kamiyama, 2012). From an articulatory point of view, the Danish /ɔ/ vowel is intermediate between the French /ɔ/ and /o/ vowels. Specifically, although the IPA symbols that are used to transcribe the Danish /ɔ/ and the French /ɔ/ open-mid back rounded vowels are identical, these two vowels are different from an articulatory point of view, the Danish /ɔ/ being raised. Acoustically, the Danish /ɔ/ vowel is closer (but not identical) to the French /o/ than to the

French /ɔ/ vowel (Georgeton et al., 2012; Steinlen, 2005). Importantly, the results of the categorization task (see Results section), which are in line with the acoustic comparison, show that native French speakers perceive the French /o/ vowel as being similar to the Danish /ɔ/ vowel.

The Russian /ɨ/ vowel is a high central unrounded vowel (Jones & Ward, 2011). French lacks central unrounded vowels. From an articulatory point of view, the three closest French vowels are the French /ø/ vowel, which is a mid-close front rounded vowel, and the French /y/ and /i/, which are high front unrounded and rounded vowels respectively. The acoustic proximity of these vowels to the Russian /ɨ/ varies from closest to most distant, respectively (see Figure 5). Again, the results of the categorization task showed that the perceptual similarity of these French vowels to the Russian /ɨ/ vowel follows the same pattern, with the French /ø/ being perceived as the most similar and the French /i/ as the least similar of the three vowels.

Eight native speakers of Russian and Danish (two female and two male speakers for each language, mean age = 30 years) were recorded while reading sentences that started with "I say …", followed by three consecutive items: a Russian or Danish word, an hVd[2] pseudo-word and finally another word. The words and pseudo-words each contained the target vowel in the middle position. An example in Russian is: "Я говорю сыт, хыд*, был ("I say sated, pseudo-word*, was"). An example in Danish is: "Jeg siger måle, håde*, skåle ("I say measure, pseudo-word*, cheer"). The sentence was repeated seven times. Recordings were carried out in a quiet room, using a Marantz PMD670 portable recorder with a Shure Beta 58A microphone sampled at 22.05 kHz directly to 16-bit mono .wav files.

[2] The hVd [xVd] context is considered to allow for neutral production of the vowel, with minimal impact of co-articulation from the phonemic context (Peterson & Barney, 1952).

Participants were trained and tested (before and after training) using stimuli recorded by a speaker of the same sex as themselves. This was done to minimize differences between the vocal-tract characteristics of the French participants and the native Danish and Russian speakers.

The vowels that were produced in words and pseudo-words (84 per language) were analysed and extracted. Vowels were then trimmed to 350 ms relative to the midpoint using PRAAT, as follows: the beginning of a vowel was determined as starting at the offset of the /h/ noise, and the end was determined as corresponding to the point of F2 movement toward /d/. The midpoint of the vowel was located, and a 350 ms segment centered at the midpoint was extracted. This procedure allowed isolated vowels to be produced in a natural task (reading) that were as free as possible from co-articulation effects. A linear amplitude ramp of 20ms was then applied to the onsets and offsets of the trimmed vowels (as for the stimuli in (Kartushina et al., 2015)). Formant tracks were visualized using the Praat software package for acoustic analysis (Boersma & Weenink, 2010). Items having the most stable formant tracks, which were assessed visually, were retained. Five exemplars of each target vowel extracted from the pseudo-words, and ten exemplars extracted from words were retained for each speaker (i.e., 15 exemplars x two vowels x four speakers). Vowel exemplars extracted from pseudo-words were presented to listeners in the repetition task and during training. The first two formant frequencies (F1 and F2) of all the exemplars (words and pseudo-words) produced by the two speakers of the same sex were computed to construct representative sex-matched Russian and Danish target vowel spaces that were used (1) to assess non-native production accuracy before and after training, and (2) to give visual feedback during training. These will hereafter be referred to as the ‘target spaces”. There were two target spaces for each trained vowel, one for female and one for male speakers. The ten exemplars (five per speaker) of each target vowel extracted from the pseudo-words only were used as non-native stimuli for the vowel repetition tasks during pre- and post-training tests and during training.

2.2.2.2 Native vowels

In order to be able to compare the production (i.e., repetition) of native to non-native isolated vowels, we recorded French vowels under exactly the same conditions as for the non-native recordings. Four native monolingual speakers of French (two female and two male speakers who did not participate in the training study, mean age = 29.5 years) who lived in Geneva were recorded while reading a list of four sentences comprised of one sentence per French vowel: /o/, /ø/, /y/ and /i/. Each sentence was comprised of the opening context “I say” followed by a word, an hVde pseudo-word and another French word. An example is: “Je prononce lit, hide*, si” (“I say bed, pseudo-word*, yes”). The list of sentences was repeated seven times. The vowels that were produced in the pseudo-words (i.e., seven exemplars x four vowels x four speakers) were extracted, analysed and trimmed to 350 ms. These vowels underwent the same acoustic pre-processing as described above for the non-native stimuli. The ten best exemplars (five per speaker) of each French vowel extracted from the pseudo-words were used as native stimuli in the pre- and post-training tests.

2.2.2.3 Native words

A word reading task was used to assess pre-training native vowel production, in order to test for relationships with non-native vowel production. Native vowel production was assessed using a word reading task rather than during isolated vowel production. This was because we wanted to use a more ecologically valid measure of native vowel production, i.e., one which captures naturally occurring phonetic variability. Fifteen French [CV], [CVC] and [CVCV] words containing the target vowel in the first (for the CVs and CVCs) or in the first or second syllable (for the CVCVs) were selected for each vowel. These were selected such that the consonants that preceded and succeeded the target vowel would differ as much as possible across words, therefore allowing maximal phonetic variability.

2.2.3 Procedure

Participants were trained to produce non-native Danish and Russian vowels over three training sessions that were administered on alternating days (e.g., Monday, Wednesday and Friday). In order to assess training effects on the production of isolated native and non-native vowels, during the first (before training) and last (after training) session, participants performed native and non-native vowel repetition tasks. They also performed a French word reading task during the first session. The order of the pre- and post-training tasks (i.e., word reading and native and non-native vowel repetition before training, and only the last two tasks after training) was randomized across speakers.

All tasks were performed on a DELL computer, using Sennheiser PC-350 headphones fitted with a microphone. The pre- and post-training tests were administered using the DMDX software (Forster & Forster, 2003). The training procedure was administered using Matlab (MATLAB Release 2011a, The MathWorks, Inc., Natick, Massachusetts, United States) and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Throughout the entire experiment (testing and training), stimuli were presented at a comfortable listening level, which was adjusted on a participant-by-participant basis.

2.2.3.1 Non-native and native vowel production

In order to assess the production of native and non-native vowels before and after training, vowel repetition tasks were administered. In the non-native vowel repetition task, participants were asked to repeat the ten exemplars of the Danish /ɔ/ and of the Russian /ɨ/, three times each, for a total of 30 production trials per vowel. On each trial, an auditory stimulus was presented and participants were prompted by a visual cue to repeat the vowel they had heard as accurately as possible. The cue remained on the screen for 2000 ms, and responses were recorded during this period. The same procedure was used to assess the production of isolated native vowels, the only difference being that instead of two non-native (Danish and Russian) vowels, participants heard four French vowels (i.e., /o/, /ø/, /y/ and /i/). Each of them was repeated 20 times.

Accuracy of the non-native vowel productions was measured using the Mahalanobis Distance (for details see Kartushina & Frauenfelder, 2014; Kartushina et al., 2015) between each token produced by the French speakers and the target vowel space derived from the recordings of the native speakers (see non-native stimuli section). This metric was used in order to take into account the natural variability in speech production, as characterized by the target spaces. For each participant, thirty distance scores (DSs) were calculated per vowel before and after training.

The distances separating the native productions from the non-native target vowel spaces before and after training were also estimated using the Mahalanobis Distance. The following four DSs were calculated before and after training: distances from the French /o/ to the Danish /ɔ/, and from the French /ø/, /y/ and /i/ vowels to the Russian /ɨ/. For each participant and each French vowel, 30 distance scores were calculated before and after training.

2.2.3.2 Native word reading

In order to test whether there was a relationship between the acoustic properties of native vowels produced in words and non-native vowels before training, a French word reading task was administered. Fifteen words were selected for each of the four native vowels (see Stimuli section). On each trial, participants first saw a cross in the middle of the screen for 1000 ms, followed by a word. Participants were asked to read it aloud as naturally as possible at a moderate tempo. The word remained on the screen for 2000 ms, and responses were recorded during this period. Each word was repeated twice. In total, 30 words (15 words/vowel x two repetitions) were recorded per each of the four native vowels, in a randomized order.

Two acoustic properties of the native French vowels produced in the words were computed for each individual and for each vowel. First, we computed the mean Mahalanobis Distance (taken over the 30 Mahalanobis Distances) between each native vowel and its respective non-native vowel target space. Second, we computed the compactness (CS) of each native vowel across the 30 stimuli that were produced. For the former, four distances were calculated for each participant: the distance from the French /o/ vowel to the Danish /ɔ/ target space, and from the French /ø/, /y/ and /i/ vowels to the Russian /ɨ/ target space. The CS was calculated as the area of an ellipse (the distribution of the produced tokens in the F1-F0/F2-F0 space was assumed to be elliptical) which had major and minor axes having a length of one standard deviation along the given axis. This area was then scaled as a proportion of the compactness of the targets. Namely, the CS of the French vowel /o/ was scaled to that of the Danish /ɔ/, and the CSs of the French /ø/, /y/ and /i/ vowels were scaled to that of the Russian /ɨ/. Thus, a lower CS corresponded to more stable productions, with a CS of 1 being equal to the compactness of native speakers of that vowel (for a complete description see Frauenfelder & Kartushina, 2014).

2.2.3.3 Training

At the beginning of the first session, all participants received instructions explaining the nature of the feedback and its correspondence to the articulators (i.e., tongue position with respect to the vertical, or vowel height, and horizontal, or vowel backness, axes) during production. Briefly, the visual feedback consisted of a two-dimensional visual display showing F1 (which corresponds to tongue height) along the y-axis, and F2 (which corresponds to the front-back position of the tongue) along the x-axis. More details regarding the feedback are provided in the next section. Participants were told that they would be trained to produce two vowels. Since no feedback was given on F3, participants were explicitly instructed that of the two vowels to be learned, the one that would appear in the middle of the screen (i.e., the Russian /ɨ/) would be unrounded. This instruction was necessary because we wanted participants to carry out the task based on the feedback relating to F1/F2, without attempting to produce a rounded vowel. Participants were familiarized with this feedback while reading the French /a/, /e/, /i/, /o/ and /u/ vowels, three times each. The familiarization phase lasted five minutes and was immediately followed by the training.

There were three training sessions. The first and the last sessions consisted of five training blocks and the second session consisted of ten training blocks. Furthermore, each training block consisted of two mini-blocks, i.e., one for the Danish and one for the Russian vowel, the order of which was randomised. Within each mini-block, 10 same-sex tokens of the given vowel were presented three times each, resulting in 30 presentations of each vowel per mini-block. The same amount of training was administered per vowel; in total, participants produced each vowel 600 times. There were pauses between blocks, and the duration of these was controlled by the participants.

Trials began with the appearance of a white screen for 250 ms. At the offset of this screen, a vowel was presented over the headphones. We prompted our participants with a visually-presented countdown signal (‘3, 2, 1’) that lasted 1000 ms in total. This was followed by a “!” sign, at which point a recording lasting 500 ms began. Participants were instructed to repeat the vowel while the “!” sign was on the screen. Then, 250 ms later, feedback was presented on-screen for 2000 ms. The countdown signal was used in order to limit the number of utterances that were initiated too late and thus not recorded.

Visual feedback during training. The articulatory feedback was based on an immediate, trial-by-trial acoustic analysis of the vowels produced by participants. On each trial, participants heard a non-native vowel and repeated it. They then saw the position of their vowel production in F1 (y-axis) / F2 (x-axis) space, which correspond to the degree of openness and to the back-to-front position of articulation, respectively. The distribution of the entire target vowel acoustic space was displayed in order to include information about the variability of non-native targets in the feedback display. Moreover, in the second and throughout all subsequent trials, the visual feedback also included information about the position of the non-native vowel produced on the previous trial (see Figure 6).

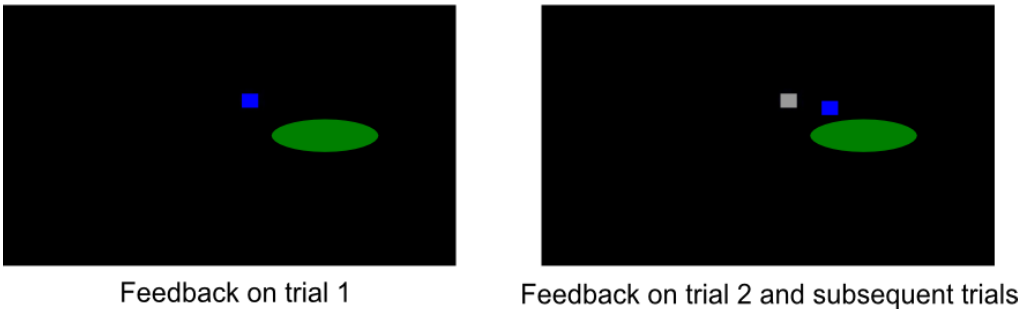

Figure 6. Example of visual feedback provided to participants regarding their vowel production on the first and second (and all subsequent) training trials (blue square = participant’s production, green ellipse = target vowel, grey square = vowel produced on preceding trial).

The produced vowels were recorded on a hard-disk as 16-bit .wav files, sampled at 22.05 kHz. The onsets and the offsets of the vowels were automatically estimated and the formant values (F0, F1 and F2), averaged over the duration of the tokens, were computed in Matlab using the same procedure as in Kartushina et al. (2015). The F1 and F2 were computed by solving for the roots of the Linear Predictive Coding (LPC) polynomial. For this purpose, we adapted the scripts from the COLEA for speech analysis software (COLEA is a suite of tools that are a subset of the COchLEA Implants Toolbox; Loizou, 1998). An LPC order of 24 was used, based on the rule of thumb that the LPC order should be equal to 2+ (the sampling frequency/1000). The F0 was analysed using the cepstrum method. Importantly, since the visual feedback regarding F1 and F2 was adjusted for F0, the target acoustic space was also adjusted for F0 (F0 was subtracted from F1 and from F2). The feedback image was presented on a 43 cm screen, in a window of 1280 pixels x 1024 pixels. The range of values was 2500 Hz—400 Hz on the x-axis, and 500 Hz—50 Hz on the y-axis. A unit shift in the x direction represented a change in Hz equivalent to a 0.25 unit shift in the y direction.

3 RESULTS

3.1 Categorization task

The results revealed that the Danish /ɔ/ vowel was consistently (90%) categorized as a French /o/. In the remainder of the cases it was categorized as a French /ɔ/. The average acceptability rating was 7.03, indicating that this vowel was perceived as being relatively French-like. The Russian /ɨ/ vowel, in comparison, was inconsistently categorized: in 52.5% of the responses it was categorized as a French /ø/, in 22.5% as a French /y/ and in 17.5% as a French /i/, with acceptability ratings of 4.65, 3.5, and 2.67, respectively.

These results suggest that native French speakers perceived the Danish /ɔ/ vowel as a French /o/ with a relatively high degree of certainty (i.e., they heard it as being a relatively prototypical French /o/). The Russian /ɨ/ vowel, on the other hand, was not perceived as being a prototypical example of any native French category. This makes it what Flege terms a 'new' (i.e., L1-distant) non-native sound. Nevertheless, French speakers more often perceived it as being similar to a French /ø/, less often to a French /y/, and least often to a French /i/. These results are consistent with the relative acoustic similarity between the Danish /ɔ/ and Russian /ɨ/ vowels and the respective French vowels (see Figure 5).

3.2 Non-native vowel production

3.2.1 Pre/post training performance

The following analyses aimed at examining the effects of training on the production of the non-native (i.e., trained) vowels. Recordings from the pre- and post-training tests were verified for intensity and absence of noise (e.g., coughs, sneezes, sighs, etc.). A small percentage (0.8%) of recordings was discarded on this basis. For each recorded token, the silent portions preceding and following the vowel were removed using Praat. The F0, F1 and F2 values were extracted using the same procedure as during the training, and used to calculate the DSs (Mahalanobis Distance based on acoustic values). Outliers (making up 3% of the data) were detected using Quantile-Quantile plots and were excluded from the analyses. The remaining DSs ranged from 0.01 to 7.81. Statistical analyses were run using the R software package (R Core Team, 2012). The significance level was set to a t value of +/-2 based on the T-distribution for df>20.

The DSs were fitted to a general linear mixed-effects model (R Core Team, 2012). The effects of Time (before vs. after training) and Vowel (Danish vs. Russian) were included as fixed factors (with “before” and “Danish” as the reference levels for Time and Vowel, respectively). The ‘maximal’ random structure with correlation parameters between the fixed factors and random slopes was used, it included by-subject and by-vowel random slopes adjusted for Time and Vowel and for Time, respectively (Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013; Winter, 2013). This structure most effectively generalizes the results (Barr et al., 2013).

There was a significant effect of Time (β = -0.42, SE = 0.16, t = -2.63) at the reference level of Vowel, indicating that productions for the Danish vowel were more accurate after (i.e., decreased DSs) than before training. There was no effect of Vowel (t<1) at the reference level of time. A separate general linear mixed-effects analysis revealed a significant effect of Time for the Russian vowel too (β=-0.43, SE= 0.07, t=-6.02). For each trained vowel and for each subject, the average DSs for the pre- and post-training tests were used to calculate the improvement in production performance following training. There was a 19% overall improvement, with improvements of 18% and 20% for the Danish /ɔ/ and Russian /ɨ/ vowels, respectively (see Figure 7).

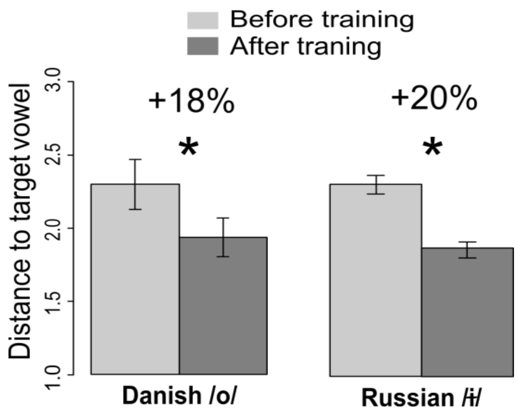

Figure 7. Effect of training on production accuracy for trained vowels. Mean distance scores are shown, and error bars represent ±1 standard error of the mean. The percentage of improvement in production for each vowel is indicated.

We also examined training-related changes in the compactness of vowel productions. A compactness score, based on a joint analysis of the F1-F0 and F2-F0 of the non-native vowels (CS_{NN}), was estimated for each participant and for each vowel. Bartlett's test showed a violation of homogeneity of variances [$\chi^2(1) = 14.02$, $p < .001$] for the CS_{NN}. The logarithm transformation was therefore applied to the CSs. After this transformation, Bartlett's showed no violation of homogeneity of variances [$\chi^2(1) = 0.95$, $p = .32$]. An ANOVA was performed on the $logCS_{NN}$, with within-group factors of Time and Vowel. There was a significant effect of Time on the $logCS_{NN}$ ($F_{(1,19)} = 10.95$, $p = .003$), which showed that productions after training were more compact than before. There was no effect of Vowel ($p > .1$), nor was there a Time*Vowel interaction (p>.1). Separate one-tailed paired tests revealed a significant effect of Time for both the Russian ($t_{(1,19)} = 2.99$, p=.0035) and Danish ($t_{(1,19)} = 1.86$, p=.038) vowels, respectively. Before training, the Russian and Danish vowels were produced with a mean compactness (standard deviation) of 3.91 (4.64) and 2.89 (1.72) units, respectively, whereas after training, they reached 1.91 (2.16) and 2.19 (1.59) units, respectively. For each vowel, the mean CS_{NN} for the pre- and post-training tests were used to calculate the improvement in vowel compactness relative to pre-training levels. This revealed improvements of 25% for the Danish /ɔ/ vowel and of 50% for the Russian /ɨ/ vowel (see Figure 8).

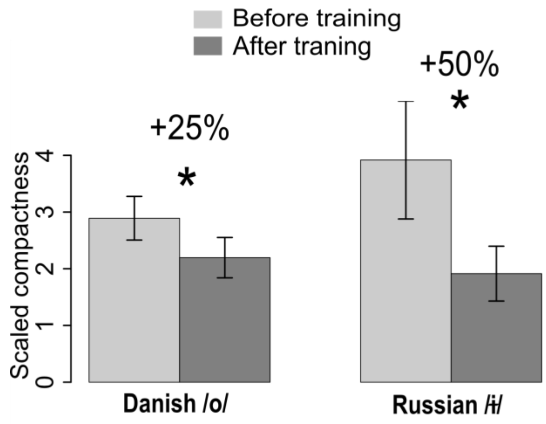

Figure 8. The compactness of non-native vowels produced by French speakers before and after training. Mean scaled compactness scores are shown, and error bars represent ±1 standard error of the mean. The percentage of improvement in compactness for each vowel is indicated.

3.2.2 Performance during training

In order to explore the production of the Danish and Russian vowels during training, recordings from 20 training blocks (24,000 tokens in total) underwent the same pre-processing procedures as did the pre/post-training test recordings, starting with acoustic quality checks and finishing with removal of outliers. As a result, 12% of the data was discarded. In the following analyses, we tested for changes in the production accuracy and compactness of the Danish /ɔ/ (i.e., L1-proximal) and Russian /ɨ/ (i.e., L1-distant) vowels during the training.

The DSs were fitted to a general linear mixed-effects regression model. The effects of Block (from one to twenty) and Vowel (Danish vs. Russian) with "Danish" as the reference level were included as fixed factors. The 'maximal' random structure, with correlation parameters between the fixed factors and random slopes was used: it included by-subject random slopes adjusted for Block and Vowel and by-vowel random slopes adjusted for Block. There was a significant effect of Block (β=-0.03, SE=0.006, t=-4.75), indicating that the more training blocks participants performed, the more accurate their production of the Danish vowel became (i.e., decreased DS). There was also a significant effect of Vowel (β=-4.45, SE=0.15, t=-2.95), with the Russian vowel being produced on average more accurately than the Danish one (see Figure 9). A separate general linear mixed-effects analysis with Block as fixed factor and maximal random structure revealed a significant effect of Block for the Russian vowel too (β=-0.02, SE= 0.009, t=-2.52).

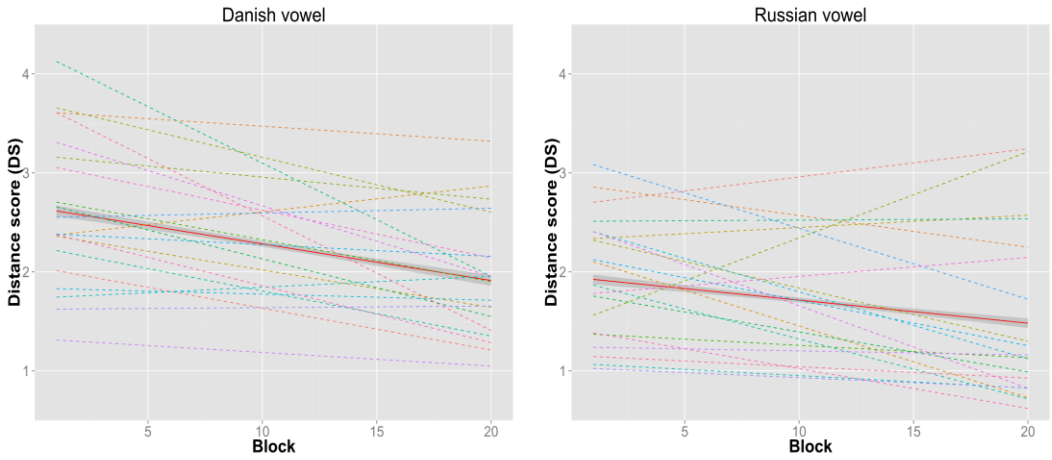

Figure 9. Mean DSs (solid red line) across 20 training blocks for the Danish and Russian vowels. Individual learning slopes are indicated by coloured dashed lines.

3.3 Native vowel production

In the following analyses we tested whether training with the non-native, L1-proximal Danish /ɔ/ vowel affected the production of the closest native /o/ vowel, and whether training with the non-native, L1-distant Russian /ɨ/ vowel affected the production of the closest native /ø/, /y/ and /i/ vowels. Recordings from the pre- and post-training vowel repetition tests underwent the same pre-processing procedures as did the non-native vowel recordings, including acoustic quality checks, trimming of silences, formant extraction, DS calculations and checking for outliers. The post-training recordings of one of the participants were not considered for the analyses since the participant accidentally switched off the microphone during the task. Some data (6 %) were discarded as a result of the above-mentioned analyses. The remaining DSs ranged from 0.24 to 6.77.

We examined at the following measures: the position and the compactness of the native vowels with regard to those of the non-native target vowels produced by native speakers. For the analyses of DS, the effects of Time (before vs. after training, with 'before' as the reference level), Condition (Russian vs. Danish vowel, with 'Danish' as the reference level) and the Time*Condition interaction were included as fixed factors. The random structure included by-subject and by-item random slopes adjusted for Time and Condition and for Time, respectively. There was a significant effect of Condition (β=0.43, SE=0.21, t=2) at the reference level of Time, indicating that before training the French /o/ vowel was overall closer to the Danish target vowel /ɔ/ than were the French /ø/, /y/ and /i/ vowels to the Russian target vowel /ɨ/. There was no effect of Time ($t<1$) at the reference level of Condition. There was a significant Time*Condition interaction (β=-0.3, SE=0.089, t=-3.47).

Separate general linear mixed-effects model analyses were run for each Condition to determine in which Condition the effect of Time was most marked. For the Russian /ɨ/ vowel Condition (i.e., effects of training on the French /ø/, /y/ and /i/ vowels), the fixed factors were Time (with 'before' as the reference level), Vowel (French /ø/, /y/ and /i/ vowels, with the French /ø/ as the reference level) and Time*Vowel interaction. The random structure included by-subject and by-item random slopes adjusted for Time and Vowel and for Time, respectively. There was a significant effect of Time (β=-0.39, SE=0.13, t=-3.04) at the reference level of Vowel, indicating that the French /ø/ vowel was closer to the Russian /ɨ/

vowel after compared to before training. There was a significant effect of the French /i/ vowel (β=1.59, SE=0.22, t=7.3), indicating that it was produced overall farther from the Russian /ɨ/ vowel than the French /ø/ vowel. There was no effect of the French /y/ vowel (t<1), indicating that it was not produced overall closer to the Russian /ɨ/ vowel than the French /ø/ vowel. There was a significant Time*Vowel interaction for the /ø/ and /i/ vowels (β=0.26, SE=0.087, t=2.9) and for the /ø/ and /y/ vowels (β=0.19, SE=0.089, t=2.16).

In order to better understand the interaction between Time and Vowel for the French /ø/, /y/ and /i/ vowels in the Russian /ɨ/ vowel Condition, separate general linear mixed-effects model analyses with Time as a fixed factor (with 'before' as the reference level) and by-subject and by-item random slopes adjusted for Time as a random structure were run for each vowel. There was a significant effect of Time for the /ø/ vowel (β=-0.45, SE=0.18, t=-2.41), and a marginally significant effect of Time for the /y/ vowel (β=-0.17, SE=0.10, t=-1.68, p=.09[3]). In both cases, the DSs were smaller after compared to before training, demonstrating that the French /ø/ and /y/ vowels became closer to the Russian /ɨ/ vowel after training. There was no effect of Time for the French /i/ vowel (t<1).

For the Danish /ɔ/ vowel condition (i.e., impact of training on the French /o/ vowel), the fixed factor was Time (with 'before' as the reference level). The random structure included by-subject and by-item random slopes adjusted for Time. There was no effect of Time for the French /o/ vowel (t<1). See Figures 11 and 12 for illustrations of the results in terms of DSs and formant values, respectively.

[3]Exact p-values of the effects were computed using Markov Chain Monte Carlo (MCMC) sampling (10000 simulations), implemented in the 'LanguageR' package of the R software (Baayen, Davidson, & Bates, 2008).

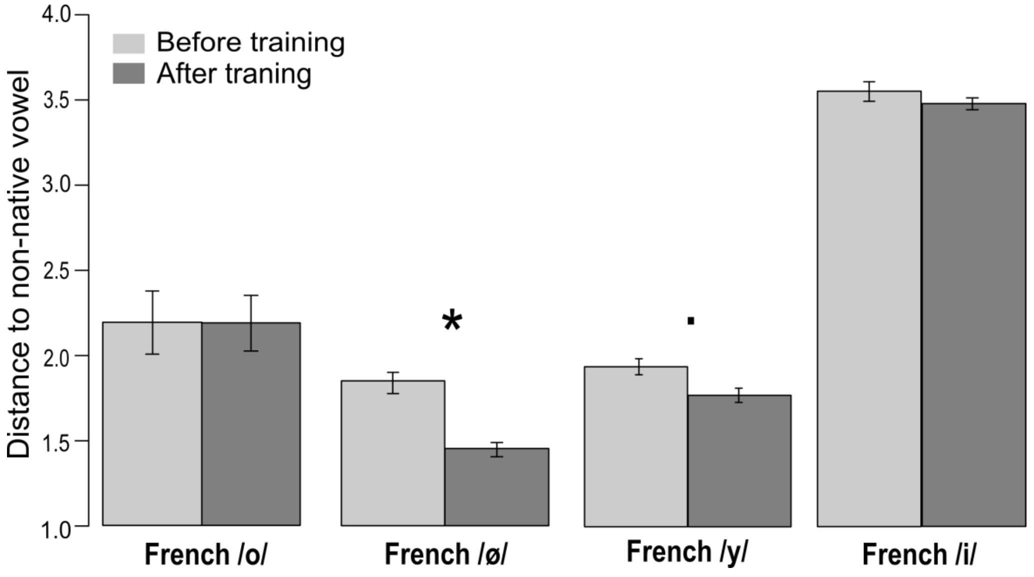

Figure 11. Effect of training on the production of native French vowels. Mean distance scores are shown, and error bars represent ±1 standard error of the mean. * denotes significant results (p<.05), • denotes marginally-significant results (p=.09).

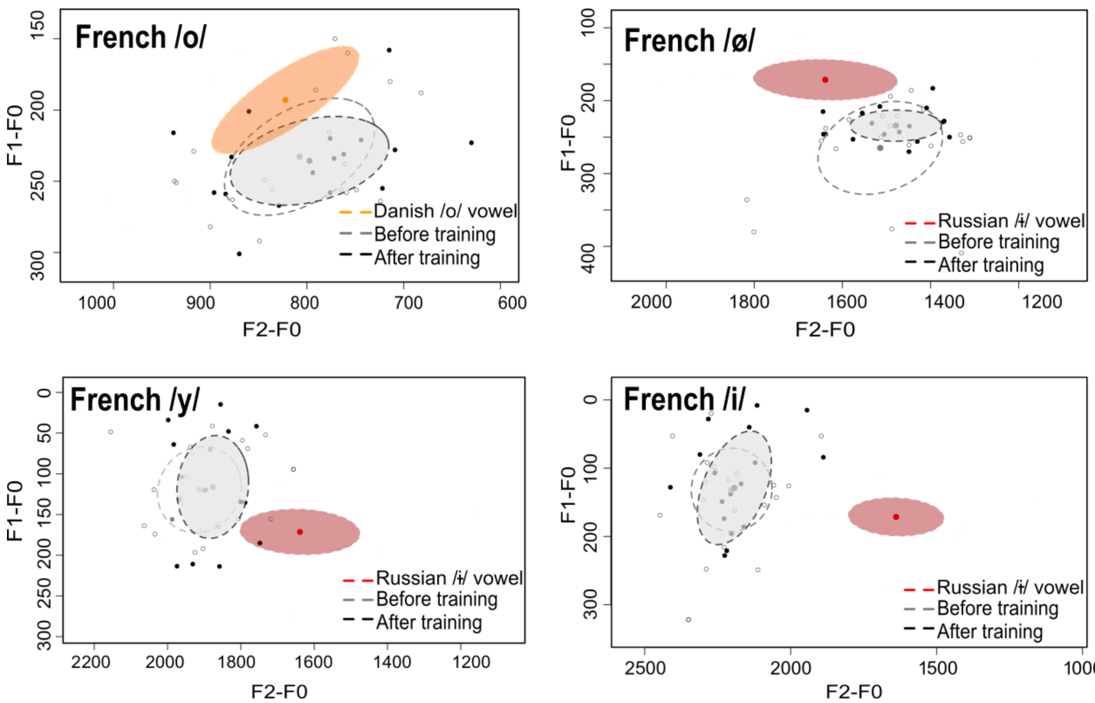

Figure 12. Formant values for native French vowels before (light grey) and after (dark grey) training with the Russian and Danish non-native vowels. These are shown with respect

to the formant values of the non-native target vowel spaces (filled ellipses). Ellipses represent one standard deviation. Mean formant values for each speaker are also shown, with the dark and light grey dots representing individuals' productions before and after training, respectively.

We also tested for training-related changes in the compactness of the native French vowels. For each French vowel, participant and time point, a compactness score (CS_{FR}) was estimated based on a joint analysis of the F1-F0 and F2-F0. As before, the CSs were scaled as a proportion of the CS of the non-native target space. Bartlett's tests showed no violation of homogeneity of variances [$\chi^2_{(1)}$=0.39, p =.52]. An ANOVA was performed on the CS_{FR}, with within-group factors of Time and Vowel. There was no significant effect of Time on CS_{FR} (p>.1), meaning that overall, training did not affect the compactness of native vowel production, and there was no Time*Vowel interaction (p>.1). There was a significant effect of Vowel ($F_{(1,3)}$=26.7, p=.04), and Tukey post-hoc tests revealed that the only significant difference was between the French /o/ and /y/ vowels, the former being less compact than the latter (p=.02).

3.4 Training-related changes in F1 and in F2 for the native and non-native vowels

We wanted to better understand training-related changes in the acoustic location of native and non-native vowel production. To this end, we examined F1 and F2 separately so as to have a pure measure of vowel position in the acoustic space, independent of (i.e., not calculated with respect to) the target vowel position. We therefore computed the difference between pre- and post-training formant values (i.e., $F1_{POST}$-$F1_{PRE}$, $F2_{POST}$-$F2_{PRE}$) for each speaker and for each vowel. This measure of training-related change in F1 and F2 was also computed for the non-native, trained vowels. The relationship between the training-related changes in the formant values (i.e. F1 and F2, separately) between native and non-native vowels was examined using two-tailed Pearson correlation tests. There were significant positive correlations in the changes in F1 (r=.68, *p*<.01) and in F2 (r=.61, *p*<.01) between the Danish /ɔ/ and the French /o/ vowels, demonstrating that participants who obtained a larger, positive pre-/post-test difference score for the Danish /ɔ/ also obtained a larger, positive pre-/post-test difference score for the French /o/ (and vice versa for negative difference scores). Thus, although we cannot comment on the global direction of movement of F1 (and of F2), we can say that, within subjects, the movement was in the same direction in the native and non-native vowels. In other words, we have shown that within participants, after training, the F1 and the F2 values of the native vowels moved in the same direction as did the F1 and F2 values of the non-native vowels (see Figure 13). For the Russian /ɨ/ and French /ø/, /y/, and /i/ vowels, a Holm-Bonferroni correction was applied to the p-value. There were no correlations in training-related changes either in F1 or in F2 between the Russian /ɨ/ and the French /ø/ and /y / vowels. There was a trend towards a positive correlation in the change in F2 only between the Russian /ɨ/ and the French /i/ vowels (r=.45, *p*=.05).

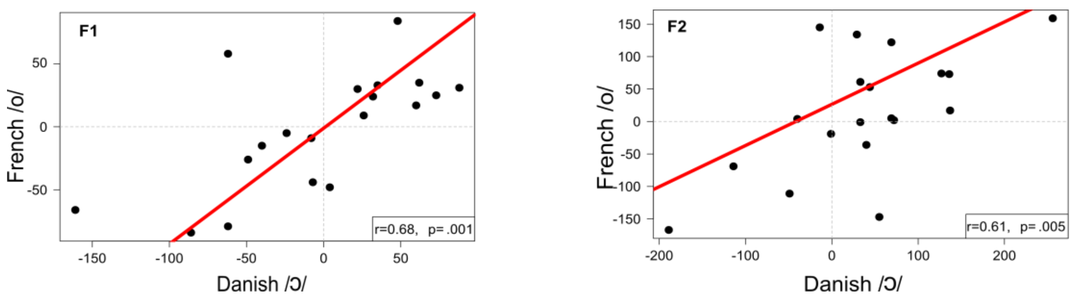

Figure 13. Training-related changes in F1 (on the left) and in F2 (on the right) for the Danish /ɔ/ and the French /o/ vowels.

3.5 Native word reading: pre-training relationships between the production of native and non-native vowels

We wanted to test for pre-training relationships between native and non-native vowel production. Native recordings from the word reading task underwent the same pre-processing procedures as did the native productions from the pre/post-training tests. 7% of the data was discarded based on the exclusion criteria described above. Recordings of two participants were lost because they accidentally switched off the microphone. For each participant, the mean compactness and the distance (DS) to the respective non-native, target vowels were computed for each French vowel.

Based on the results of Kartushina & Frauenfelder (2014), we performed one-tailed Pearson correlation analyses between native and non-native vowels on (1) their position in the acoustic space (DS), and (2) their compactness. A Bonferroni correction was applied to the number of tests performed per data set (i.e., four); consequently, the significance threshold was set to p=.0125.

There was a marginally-significant correlation between the position of the French /o/ and the Danish /ɔ/ vowels (r=0.51, p=.0153), suggesting that across individuals, the closer the French sound was to the target Danish vowel, the greater was the production accuracy for this Danish vowel. Regarding compactness, there was a significant correlation between, on the one hand, the compactness of the French /o/ vowel, and on the other hand, the compactness and the production accuracy of the Danish /ɔ/ vowel: speakers whose native French vowel was more compact before training produced the Danish vowel more compactly (r=0.64, p=.004) and more accurately (r=0.49, p=.019).

No relationship was found between the position of the French /ø/ vowel and the production accuracy of the Russian /ɨ/ vowel (r=0.16, p>.1). Higher compactness, however, of this acoustically closest French vowel was related to better (r=0.70, *p*=.0012) and to less variable (r=0.78, *p* <.001) pre-training production of the Russian vowel.

None of the relationships between the Russian /ɨ/ vowel and the French /y/ and /i/ vowels were significant (*p*>.1).

4 DISCUSSION

Twenty monolingual French speakers were trained, using a visual articulatory feedback technique, to produce two unfamiliar, non-native vowels: the Danish /ɔ/, which is similar to the French /o/, and the Russian /ɨ/, which is dissimilar from French vowels. This training approach implements a real-time analysis of the acoustic properties (F0, F1 and F2) of the vowels produced by non-native speakers, and provides them with immediate visual feedback about their articulation, together with information about target vowel acoustic spaces derived from the productions of native Danish and Russian speakers. The effectiveness of training was assessed by computing the Mahalanobis Distances (Distance score, DS) in F1/F2 space between each token produced by the French participants and the non-native target vowel space derived from the recordings of native speakers.

First, an assessment of perceptual similarity of the non-native and native French vowels by native French speakers revealed that the Russian /ɨ/ vowel was indeed perceived as being a ‘different’, L1-distant vowel, and that the Danish /ɔ/ vowel was perceived as being a ‘similar’, L1-proximal vowel. Thus, there was overall convergence between the perceptual and the acoustic similarity of the non-native and native vowels (see Figure 5).

4.1 Production of non-native vowels

4.1.1 Effectiveness of training

One hour of articulatory feedback training per vowel improved French speakers' production of the non-native Russian and Danish vowels by 20% and 18%, respectively, as reflected by decreased DSs. These results are in line with other L2-training studies in adults showing a benefit of articulatory feedback training in the production of non-native sounds (Akahane-Yamada, McDermott, Adachi, Kawahara, & Pruitt, 1998; Carey, 2004; Kartushina et al., 2015; Massaro et al., 2008).

The absence of a difference in the amount of improvement between the L1-distant (Russian) and L1-proximal (Danish) vowels contrasts with the results of previous studies showing that L1-distant foreign speech sounds benefit more from exposure/learning than do L1-proximal sounds (Aoyama et al., 2004; Flege, Bohn, & Jang, 1997; Major & Kim 1996 in Major, 2008). Although these were not laboratory training studies, the authors attribute the learning disadvantage for L1-proximal sounds to equivalence classification, which blocks the formation of new L2 categories (Flege, 1995). In our study, it is possible that the articulatory feedback that French learners received helped them to overcome limitations that equivalence classification imposes on the establishment of new phonetic categories for L1-proximal non-native sounds. Thus, trial-by-trial visual feedback regarding the phonetic features of the L1-proximal non-native Danish /ɔ/ vowel likely enabled French participants to modify their articulatory patterns to match those required for production of the target vowel.

Training decreased the variability in the production of non-native vowels: after training, the compactness score for Danish and Russian vowels produced by French participants decreased by 25% and 50%, respectively (with lower scores corresponding to more stable productions). Thus, this articulatory feedback training technique both improved and stabilized French participants' production of non-native sounds. High pretraining variability in production for the L1-distant Russian /ɨ/ vowel suggests that, before training, more heterogeneous articulatory patterns were used to produce it. These patterns are most likely similar to those required for producing the French /ø/, /y/, and /i/ vowels, as suggested by the results of the categorization task. Thus, the training-related reduction in variability may reflect the establishment of a new speech sound representation.

Although we did not assess the establishment of new non-native categories in perception, some studies suggest that the phonetic properties of speech sounds used in production correspond to those used in perception. Perkell and colleagues (2004) have shown that accurate discrimination of the native English /i/-/ɪ/ contrast is related to more distinct production of these sounds. The authors interpreted these results within the framework of the DIVA model which states that articulatory movements for speech sounds are primarily planned in auditory space, suggesting that speakers whose auditory target sound spaces are ‘tighter’, or more precise, will produce the corresponding speech sounds more compactly and distinctly. Another study, however, found only marginal correlations between participants’ distinctness in the production of and in the discrimination of a Dutch /ɪ/-/ɛ/ phonetic contrast (Franken, McQueen, Hagoort, & Acheson, 2015). More research is needed to understand the relationship between the phonetic properties of speech sounds used in production and those underlying perceptual representations, in both native speakers and L2 learners.

In our study, the variability of the Danish /ɔ/ vowel, on the other hand, was reduced by only 25%, possibly suggesting that a less variable (more target-like) pattern of articulation was already in use before training, and that it was only somewhat stabilized by training. The qualitatively greater stabilization in the production of the Russian compared to the Danish vowel is compatible with predictions of the SLM model, which claim that L1-distant sounds are easier to acquire than L1-proximal sounds.

The results revealed no differences in pre-training production accuracy between the Danish and Russian vowels. The average pre-training performance for both vowels was about 2.3 DS units from the corresponding target mean, indicating that French speakers produced both vowels with comparable accuracy. This suggests that although the Russian /ɨ/ vowel was produced more variably, mean F1/F2 values were as close to the target vowel as were those of the Danish /ɔ/ vowel. Based on the SLM we might have expected the L1-distant Russian /ɨ/ vowel to be more accurately produced than the L1-proximal Danish /ɔ/, but other studies show that at early stages of learning, L1-distant sounds may be produced less accurately than L1-proximal sounds, presumably due to the fact that the former require new patterns of articulation to be learned, whereas the latter can be accomplished using patterns similar to those used for L1 (Aoyama et al., 2004; Major & Kim, 1996 in Major, 2008).

4.1.2 Production during training

In contrast to our pre- versus post-training results, analyses of performance during the training periods showed that during training, French speakers produced the Russian vowel more accurately (i.e., closer to the target vowel on average) and learned it better (closer to the target vowel on the last training block) than the Danish vowel. This suggests that the training was more effective/rapid during the training period for the distant than for the proximal vowel, which might in turn suggest that the former vowel is more easily learnable. This result is compatible with the above-described SLM claim concerning the relatively greater difficulty in establishing a new phonetic category for a similar than for a dissimilar L2 sound (Flege, 1987; 1995). Our finding of greater training benefits during training itself for the L1-distant compared to the L1-proximal vowel suggest that the predictions of the SLM, which were initially developed to explore L2 acquisition in the context of longer-term language immersion, can be extended to short-term learning contexts.

Other studies, albeit not laboratory training studies, have also shown different learning rates for L1-proximal versus L1-distant L2 sounds (Flege, 1987; Major & Kim, 1996 in Major, 2008). For example, in their study of Korean learners of English, Major and Kim showed that the rate of learning for the L1-distant sound /z/ was faster than for the L1-proximal sound /ǰ/. Both sets of results are in line with the Similarity Differential Rate Hypothesis according to which dissimilar phenomena are acquired at faster rates than similar ones (Major & Kim, 1999).

In sum, our results suggest that after training with articulatory feedback, French speakers improved to a similar degree in their production of L1-proximal and L1-distant non-native vowels (18% and 20% improvement, respectively). They differed, however, in two other aspects of learning: (1) during training, in the presence of articulatory feedback, production of the L1-distant Russian vowel was more accurate than that of the L1-proximal Danish vowel and it reached higher levels of accuracy by the last training block; (2) training induced a qualitatively greater decrease in the variability of the production of the Russian than the Danish vowel.

4.2 Production of native vowels

In this section we will consider the effects of training with the Danish /ɔ/ vowel on the production of the similar French /o/ vowel, and of training with the Russian /ɨ/ vowel on the production of the three closest French vowels: /ø/, /y/ and /i/. In the case of the latter, remember that the perceptual and acoustic similarity of the three French vowels to the Russian /ɨ/ vowel ranged from more to less similar in the following order: /ø/, /y/ and /i/.

4.2.1 Effect of training with the Danish /ɔ/ vowel on production of the similar French /o/ vowel

The results revealed no global change in the French /o/ category: it did not become closer to the Danish /ɔ/ after training. This result may be due to the position of the French /o/ vowel before training; acoustically, it partially overlapped with the Danish /ɔ/, leaving little room to move closer. Mean productions of the French /o/ for some French speakers before training were within the Danish target space. Moreover, the Danish /ɔ/ was categorized perceptually as being a good prototype of the French /o/ vowel. It is therefore not surprising that the overall position of the similar French vowel did not move further toward the Danish /ɔ/ vowel.

However, examination of the position of the Danish /ɔ/ and the French /o/ vowels revealed that within subjects, there were strong and moderately strong positive correlations in the post- versus pre-training changes in F1 and F2, respectively.. This indicates that *within individuals,* the native and non-native categories underwent similar training-related changes, with partial drift in the same direction after training.

4.2.2 Effect of training with the Russian /ɨ/ vowel on the production of the French /ø/, /y/ and /i/ vowels

French /ø/ vowel

The overall position in F1/F2 of the French acoustically closest /ø/ vowel moved closer to the Russian /ɨ/ after training. This interesting result suggests that short-term articulatory training with L1-distant sounds modifies the production of the closest native sound. Note that French speakers did not perceive the Russian /ɨ/ vowel as being prototypical of the French /ø/, and that they clearly perceived the difference between the two vowels.

Importantly, these results suggest that, taken together with the above results showing no drift in the French /o/ vowel, two conditions need to be met for a drift in the position of a native category toward a non-native one to take place. First, the non-native sound should be perceptually close to, but not prototypical of, the native one. Second, the native and non-native categories should not overlap acoustically, so that there is room for one category to move relative to the other. These two conditions are met in the case of the French /ø/ vowel, which after training with the Russian /ɨ/ vowel, drifts towards it (see Figure 1 for examples of the modifications that native categories can undergo). This finding confirms Flege’s hypothesis about the ‘reorganization of native space’ (see below) (Flege, 1995), and is in line with the results of Chang’s (2012) recent study on the effects of L2-immersion on the production of native categories.

The F3 of the French /ø/ vowel, a critical parameter which is related to roundedness and which thereby perceptually distinguishes it from the Russian /ɨ/ (the French vowel being rounded and the Russian one being unrounded), was not modified by training (p>.1). This suggests that training induced change only in the trained acoustic dimensions (‘partial feature change’) in the production of native vowels.

French /y/ vowel

There was a marginally significant effect of training on the position of the French /y/ vowel, with a drift toward the Russian /ɨ/. This weaker effect of training with the Russian /ɨ/ vowel on the overall position of the French /y/ vowel as compared to that of the French /ø/ vowel can be explained within the framework proposed above. As can be seen from the acoustic plots (Figure 5), the French /y/ vowel falls between the French /ø/ and the French /i/ vowels relative to the Russian /ɨ/. It has room to move towards the trained, non-native Russian vowel, and is sufficiently similar to it to allow its position to be influenced by the training. However, given that it perceived as being more different from the non-native target vowel compared to the French /ø/, its overall position is only marginally influenced by training with the Russian /ɨ/. Again, the F3 of the French /y/ vowel was not modified by training (p>.1).

French /i/ vowel

The overall position of the French perceptually and acoustically furthest /i/ vowel was not modified by training with the Russian /ɨ/. However, when we looked for relationships within individuals as to how F1 and F2 changed between the native and non-native vowels, we found a moderate trend for a positive correlation between changes in F2 between the French /i/ vowel and the trained Russian /ɨ/ vowel. This result shows that although training with the Russian /ɨ/ does not result in an overall drift in the global position of this acoustically and perceptually more distant native French /i/ vowel, within individuals it results in similar changes in one of the trained dimensions (i.e., the front-backness) of the native and the non-native vowels. This is once again in line with the above-proposed interpretation for why only *certain* native vowels are influenced by learning to produce new non-native vowels. In this case, the native /i/ vowel is perceptually and acoustically too distant from the trained one to be influenced in its overall position. However, an influence on native vowel production is still detectable at the individual subject level (with movement of the two vowels in the same direction within subjects), and when examining specific trained acoustic dimensions (i.e., F1 or F2, as assessed independently of the target acoustic space). The F3 of the French /i/ vowel was again not modified by training (p>.1).

4.2.3 Phonetic drift

In sum, our results show that under certain circumstances, one hour of articulatory training with non-native vowels changes the production of native French vowels, with a drift toward non-native ones. There was no change in the compactness of the native vowels, suggesting that the distribution of native vowels is not affected by learning new non-native sounds, at least after short periods of training as in our study. The observed training-related drift in the production of certain native sounds towards non-native ones is in line with the SLM postulate according to which L1 categories are not fossilized but continue to develop over the lifespan "to reflect the properties of all L1 or L2 phones identified as a realization of each category" (Flege, 1995, p. 239). However, our results suggest that the changes that native categories undergo depend on the degree of acoustic and perceptual similarity to the non-native sounds: i.e., the more acoustically and perceptually similar they are, the more their production is affected, unless the native vowel is so similar that it overlaps acoustically with the non-native one, which it, is categorized as being prototypic of the native category. Phonetic drift of native toward non-native sounds has been previously reported in speakers who had experienced immersion in an L2-speaking country (Chang, 2012, 2013; Flege, 1987; Major, 1992; Sancier & Fowler, 1997).

Importantly, the phonetic drift observed in our study cannot be attributed to a lack of L1 use (L1 attrition, e.g., Schmid, 2010), as has been proposed in some previous studies with bilinguals (see Chang, 2012, for a discussion of this). Participants in our study had only specific, time-limited exposure to the foreign vowels, amounting to one hour per day for three days. Outside of the training, they continued to use French as their principal language. Our results show that even brief articulatory practice with foreign speech sounds influences the production of native sounds. Also, the observed effects of L2 training on native vowels are unlikely to be due to general phonetic imitation (Babel, 2012; Sato, Grabski, Garnier, Granjon, Schwartz, & Nguyen, 2013) since only those acoustic dimensions (i.e., F1 and F2, but not F3) that were trained were modified.

4.3 Pre-training relationships between the production of native and non-native vowels

First, we will consider the relationship between the production of the similar Danish /ɔ/ and French /o/ vowels, and then that between the dissimilar Russian /ɨ/ and French /ø/, /y/ and /i/ vowels.

The results revealed that, before training and across individuals, the closer the Danish /ɔ/ is to the native French /o/ vowel, the better is its production accuracy, suggesting that for perceptually assimilated vowels, the closest native category is used to produce the non-native sound. This result is consistent with the SLM claim according to which late speakers use close L1 sounds to produce similar L2 ones (Flege, 1995). Moreover, we found that speakers whose native French vowel was more compact produced the Danish vowel more compactly and accurately before training. These results are consistent with those of Kartushina and Frauenfelder (2014), which showed that speakers whose native vowels are acoustically more compact produce similar and dissimilar L2 vowels more accurately, since they are more likely to differentiate between L1 and L2 sounds in production and in perception.

Before training, there was no relationship between the acoustic distance from the acoustically closest French /ø/ to the Russian /ɨ/ vowel and the production accuracy for this Russian vowel by French speakers. Greater compactness, however, of this acoustically closest French vowel was strongly related to better and less variable production of the Russian /ɨ/ vowel before training, again consistent with Kartushina & Frauenfelder (2014). These results suggest that the acoustically closest French vowel was not used to produce the Russian vowel, but that the variability or distribution of this native vowel was related to the production accuracy and variability of the non-native vowel. This is consistent with the SLM claim that new categories will be established to produce dissimilar L2 sounds, and that the likelihood of this depends on the availability of 'blank slots' in the speaker’s phonetic space (Escudero & Boersma, 2004; Flege, 1995).

None of the pretraining relationships between the Russian /ɨ/ vowel and the French /y/ and /i/ vowels were significant. This suggests that non-native speech sound production is not related to the production of acoustically and perceptually more distant (or dissimilar) native speech sounds.

Although our study does not allow us to establish the direction of causality between native compactness and non-native production accuracy, the results of a previous study (Kartushina & Frauenfelder, 2014) suggest that the more compact a native category is, the more likely it is that dissimilar and similar L2 vowels will be produced correctly. A brain imaging study which examined the relationship between L1 and L2 phonetic processing also found evidence for relationships between individual differences in L1 and L2 phonetic discrimination (Díaz, Baus, Escera, Costa, & Sebastián-Gallés, 2008). Other recent work has explored individual differences in native speech sound production variability through development and between males and females (DiCanio, Nam, Amith, García, & Whalen, 2015; Romeo, Hazan, & Pettinato, 2013). While the factors that determine compactness (or distribution) have previously been studied for specific phonetic categories (e.g., phonetic context, the size of the vowel inventory, for details see (Recasens & Espinosa, 2006, 2009)), it is unclear what the factors are that underlie individual differences in the variability of phonetic production. Here we propose three possible explanations for this phenomenon. A first explanation is that the acoustic compactness in native phonetic production reflects individual differences in the precision of articulatory gestures. If so, native phonetic compactness can be characterized as a task-independent, articulatory-motor skill. A recent study on vowel variability in Yoloxóchitl Mixtec by DiCanio and colleagues supports this idea (DiCanio et al., 2015). Within speakers of Yoloxóchitl Mixtec, there are relationships between the degree of vowel variability across spontaneous and elicited speech styles, suggesting that individuals' variability in compactness is consistent across different speech tasks. A second explanation is that compact native phonetic production arises from accurate perceptual representations of sounds. For example, Perkell and colleagues (2004) have shown that accurate discrimination of the native English /i/-/ɪ/ contrast is related to more distinct production of these sounds. According to the DIVA model, speakers primarily plan articulatory movements in auditory space. This suggests that individuals with 'tighter', or more precise auditory phonetic representations will produce speech sounds more compactly, or distinctly. A third explanation for individual differences in compactness is related to general motor or cognitive skills. For example, a study by Yu (2010) has demonstrated a relationship between individual differences in cognitive processing style (e.g., in particular, traits related to the Broad Autism Phenotypes, e.g., social skills, communication, etc.) and speech processing (e.g., listeners' ability to perceptually compensate for vocalic context during consonant identification). Also, a recent study by Lev-Ari and Peperkamp (2013) has shown that individuals with low inhibitory control show greater drift in their L1 phonetic production. These explanations need to be explored in future studies. We performed additional analyses which showed that the amount of drift of the French /ø/ vowel was related to the individual compactness of the French /ø/ vowel before training: i.e., those speakers whose French /ø/ was more compact showed little or no drift toward the Russian /ɨ/ vowel after training ($r=.60$, $p<.05$). This suggests that speakers whose productions of L1 sounds are more compact are less likely to modify their production of those very sounds when learning novel non-native sounds. Further research is needed to assess the relationship between L1 and L2 phonetic processing and the directionality of such effects.

4.4 Limitations and future directions

A recent study has shown that individual differences in auditory discrimination abilities influence the degree to which speakers adapt their production to altered auditory feedback (Villacorta, Perkell, & Guenther, 2007). It is possible, therefore, that individual differences in perceptual abilities could partly predict articulatory learning in the current study. Also, given that we used a vowel repetition task which contained a perceptual component, it is possible that some of the difficulties in production were due to inaccurate perception rather than to production difficulties per se. Although we did not test for perception of the non-native vowels in our participants prior to the experiment, in our previous training study there was no evidence for a relationship between pre-training perception and pre-training production, or for a relationship with the amount of training-related improvement in L2 production (Kartushina et al., 2015).

In our study, we interpret the improved production of and the reduced variability in the non-native vowels as reflecting the emergence of new phonetic categories. To our knowledge, direct measures of category formation per se for L2 production do not exist. However, in future studies, category-goodness ratings by native speakers and acoustic analysis of L2 vowel productions recorded in variable phonetic contexts could be used to more directly assess the formation of new categories.

Our results show that under certain circumstances, one hour of training with a non-native vowel induces a drift in native categories toward the non-native, trained one. This finding is especially interesting when taking into account that our participants continued using their native language, i.e., French, outside the laboratory, and were exposed to non-native sounds only for one hour a day. In his recent study, Chang (2013) showed that there was less drift of native toward non-native categories in experienced compared to inexperienced L2 speakers. The author argues that greater experience with an L2 progressively reduces its impact on L1 production. This claim, however, is not supported by the findings of Flege (1987), Major (1992) and Mora and colleagues (2015), which showed that more experience with an L2 led to more L2-like native speech in early and late bilinguals who were immersed in an L2-speaking country. Further studies, therefore, are needed to address the effects of longer-term training on L1 production, and also to understand what happens when exposure to foreign sounds ceases, or when L2-proficiency reaches native-like levels.

To conclude, previous work may lead one to expect that influences of non-native phonetic learning on native phonetic production (i.e., drift) are observed under certain circumstances, including a non-negligible amount of exposure to and learning of a non-native language, diminished exposure to and use of L1, and later age of acquisition of a second language. Here, we report that native phonetic production can be modified even under the most unlikely circumstances and after only a brief period of training with non-native speech sounds. The longevity of these modifications remains to be established. These results suggest that even native phonetic production is a dynamic, plastic and 'live' system which is susceptible to very brief influences and possibly to continual modification as a function of new learning. More research is needed to shed light on the particular circumstances under which L2-acquisition can cause phonetic drift in the L1.

Acknowledgements

We would like to thank three anonymous reviewers for their helpful comments and suggestions. Narly Golestani and Alexis Hervais-Adelman are supported by the Swiss National Science Foundation (PP00P3_133701).

REFERENCES

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250. http://doi.org/10.1016/S0095-4470(03)00036-6

Archibald, J. (1998). Second language phonology, phonetics, and typology. *Studies in Second Language Acquisition*, *20*(02), 189–211.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A practical introduction to statistics using R*. New York: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. http://doi.org/10.1016/j.jml.2007.12.005

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*(1), 177–189. http://doi.org/10.1016/j.wocn.2011.09.001

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. *Language and Speech*, *48*(1), 1–27.

Barlow, J. A. (2014). Age of acquisition and allophony in Spanish-English bilinguals. *Frontiers in Psychology*, *5*. http://doi.org/10.3389/fpsyg.2014.00288

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. http://doi.org/10.1016/j.jml.2012.11.001

Basbøll, H. (2005). *The Phonology of Danish*. The Phonology of the World's Languages. Oxford: Oxford University Press.

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402–409. http://doi.org/10.1016/j.actpsy.2013.01.010

Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). Baltimore: York Press.

Boersma, P., & Weenink, D. (2010). *Praat: doing phonetics by computer [Computer program]. Version 5.2*. Retrieved from http://www.praat.org

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Carey, M. (2004). CALL visual feedback for pronunciation of vowels: Kay Sona-Match. *CALICO Journal*, *21*(3), 571–601.

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*(2), 249–268. http://doi.org/10.1016/j.wocn.2011.10.007

Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, *41*(6), 520–533. http://doi.org/10.1016/j.wocn.2013.09.006

Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(42), 16083–16088. http://doi.org/10.1073/pnas.0805022105

DiCanio, C., Nam, H., Amith, J. D., García, R. C., & Whalen, D. H. (2015). Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec. *Journal of Phonetics*, *48*, 45–59. http://doi.org/10.1016/j.wocn.2014.10.003

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*(04), 551–585.

Flege, J. E. (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. E. (1995). Second Language Speech Learning Theory, Findings, and Problems. In Strange, Winifred (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (Vol. 6, pp. 319–355). Berlin: Walter de Gruyter.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437–470.

Flege, J. E., & Eefting, W. (1987a). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*, 185–202.

Flege, J. E., & Eefting, W. (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Flege, J. E., MacKay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, *106*(5), 2973–2987.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491. http://doi.org/10.1016/S0167-6393(02)00128-0

Forster, K., & Forster, J. (2003). DMDX: A Windows Display Program with Millisecond Accuracy. *Behavior Research Methods Instruments and Computers*, *35*(1), 116–124.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*, *36*(4), 649–663. http://doi.org/10.1016/j.wocn.2008.04.001

Franken, M. K., McQueen, J. M., Hagoort, P. & Acheson, D. J. (2015). Assessing the link between speech perception and production in speech motor control through individual differences. In *Proceedings of the 18th International Congress of Phonetic Sciences*. (pp. 1-5).

Georgeton, L., Paillereau, N., Landron, S., Jiayin, G., & Kamiyama, T. (2012). Analyse formantique des voyelles orales du français en contexte isolé: à la recherche d'une référence pour les apprenants de FLE. In *JEP-TALN-RECITAL* (Vol. 1, pp. 145–152). Grenoble. Retrieved from http://www.aclweb.org/anthology/U/U12/F12-1019.pdf

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds 'L' and 'R'. *Neuropsychologia*, *9*(3), 317–323. http://doi.org/10.1016/0028-3932(71)90027-3

Grønnum, N. (1997). Danish vowels: the psychological reality of a morphophonemic representation. In *Proceedings of Journées d'Études Linguistiques* (pp. 91–97).

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. http://doi.org/10.1016/0093-934X(89)90048-5

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals. Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*, *60*(2), 98–128. http://doi.org/71449

Ingram, J. C., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, *25*(3), 343–370. http://doi.org/10.1006/jpho.1997.0048

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. http://doi.org/10.1016/j.actpsy.2007.06.003

Jones, D., & Ward, D. (2011). *The Phonetics of Russian*. Cambridge: Cambridge University Press.

Kartushina, N., & Frauenfelder, U. H. (2013). On the role of L1 speech production in L2 perception: Evidence from Spanish learners of French. In *Proceedings of Interspeech* (pp. 2118–2122).

Kartushina, N., & Frauenfelder, U. H. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology*, *5*. http://doi.org/10.3389/fpsyg.2014.01246

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, *138*(2), 817–832. http://doi.org/10.1121/1.4926561

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*(5), 320–331. http://doi.org/10.1016/j.wocn.2013.06.002

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing Access to the Native Language While Immersed in a Second Language Evidence for the Role of Inhibition in Second-

Language Learning. *Psychological Science*, *20*(12), 1507–1515. http://doi.org/10.1111/j.1467-9280.2009.02480.x

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3 Pt 1), 1242–1255.

Long, M. H. (1990). Maturational Constraints on Language Development. *Studies in Second Language Acquisition*, *12*(03), 251–285. http://doi.org/10.1017/S0272263100009165

MacLeod, A. A. N., Stoel-Gammon, C., & Wassink, A. B. (2009). Production of high vowels in Canadian English and Canadian French: A comparison of early bilingual and monolingual speakers. *Journal of Phonetics*, *37*(4), 374–387. http://doi.org/10.1016/j.wocn.2009.07.001

Major, R. C. (1992). Losing English as a First Language. *The Modern Language Journal*, *76*(2), 190. http://doi.org/10.2307/329772

Major, R. C. (2008). Transfer in second language phonology: A review. In J. G. H. Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (Vol. vi, pp. 63–94). Amsterdam ; Philadelphia: John Benjamins Publishing.

Major, R. C., & Kim, E. (1999). The similarity differential rate hypothesis. *Language Learning*, *49*(s1), 151–183.

Massaro, D. W., Bigler, S., Chen, T. H., Perlman, M., Ouni, S., & others. (2008). Pronunciation training: the role of eye and ear. In *Proceedings of Interspeech* (pp. 2623–2626). Brisbane.

Mora, J. C., Keidel, J. L., & Flege, J. E. (2015). Effects of Spanish use on the production of Catalan vowels by early Spanish- Catalan bilinguals. In J. Romero & M. Riera (Eds.), *The Phonetics-Phonology Interface: Representations and methodologies*. (pp. 33-53). Amsterdam: John Benjamins.

Mora, J. C., & Nadeu, M. (2012). L2 effects on the perception and production of a native vowel contrast in early bilinguals. *International Journal of Bilingualism*, *16*(4), 484–500. http://doi.org/10.1177/1367006911429518

Öster, A.-M. (1997). Auditory and visual feedback in spoken L2 teaching. *Reports from the Department of Phonetics, Umea University, PHONUM*, *4*, 145–148.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America*, *116*(4), 2338. http://doi.org/10.1121/1.1787524

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

R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria. Retrieved from http://www.R-project.org.html

Recasens, D., & Espinosa, A. (2006). Dispersion and variability of Catalan vowels. *Speech Communication*, *48*(6), 645–666.

Recasens, D., & Espinosa, A. (2009). Dispersion and variability in Catalan five and six peripheral vowel systems. *Speech Communication*, *51*(3), 240–258.

Romeo, R., Hazan, V., & Pettinato, M. (2013). Developmental and gender-related trends of intra-talker variability in consonant production. *The Journal of the Acoustical Society of America*, *134*(5), 3781–3792. http://doi.org/10.1121/1.4824160

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*(4), 421–436.

Sato, M., Grabski, K., Garnier, M., Granjon, L., Schwartz, J.-L., & Nguyen, N. (2013). Converging toward a common speech code: imitative and perceptuo-motor recalibration processes in speech production. *Frontiers in Psychology*, *4*. http://doi.org/10.3389/fpsyg.2013.00422

Schmid, M. S. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, *13*(01), 1–7.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism, 9*(01), 97. http://doi.org/10.1017/S1366728905002403

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306–2319. http://doi.org/10.1121/1.2773966

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. Retrieved from http://arxiv.org/pdf/1308.5499

Wong, J. W. S. (2013). The effects of perceptual and/or productive training on the perception and production of English vowels/ɪ/and/iː/by Cantonese ESL learners. In *Proceedings of Interspeech* (pp. 2113–2117).

Yu, A.C.L. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change, *PLoS ONE* 5(8): e11950. doi:10.1371/journal.pone.0011950.