NORTHWESTERN UNIVERSITY

Language Retention and Improvement after a Study Abroad Experience

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Linguistics

By

Caroline Engstler

EVANSTON, ILLINOIS

June 2012

ABSTRACT

Language Retention and Improvement after a Study Abroad Experience

Caroline Engstler

This dissertation examines the linguistic and extralinguistic factors that influence changes in performance in second language speech perception and speech production experienced by study-abroad returnees in the first nine months after the end of the study-abroad program.

English-French bilinguals who had recently returned from a four-month study abroad program in France participated in two speech perception and two speech production experiments at three points in time after their return home from the study-abroad stay (two, five, and nine months after their return). Two groups of native French speakers and native English speakers served as controls. A number of linguistic and extralinguistic factors were investigated with regards to their influence on performance changes across time.

The experiments revealed that bilinguals not only did not show signs of language attrition, but rather they retained their skills for all tasks across the first five to nine months after their return from France, and for some tasks (AX discrimination, lexical decision, picture naming) they even improved their performance across time. This was true for bilinguals who were still taking French classes, and even for those that were not taking classes and reported low overall exposure to French. In addition, it was shown that initial proficiency did not affect the amount of improvement across time, while amount of exposure made a difference in the phonological perceptual task (AX discrimination). Furthermore, highly motivated bilinguals with a positive attitude towards the French culture were shown to improve more across time than less

motivated bilinguals. Finally, high executive functioning was shown to lead to bigger improvements across time.

These results are important at a theoretical level, since they provide the possibility to extend second language speech acquisition models and can inform a multi-componential model of language attrition, which includes linguistic as well as extralinguistic factors for language changes across time. Furthermore, the results have practical implications, in that they can help language instructors and students identify which linguistic and behavioral factors to focus on in the language learning environment to aid in the long-term retention of second language skills.

ACKNOWLEDGEMENTS

I am grateful to the many people who have helped me throughout the various stages of preparing and writing this dissertation.

First and foremost, I would like to thank my advisor, Matt Goldrick, for the amount of time and energy he has put into guiding me through this project, for providing the fastest and most thorough feedback imaginable, and for showing me how to stay positive through a seemingly never-ending task.

Further thanks go to my two other committee members, Ann Bradlow and Ken Paller, for their helpful feedback on earlier versions of this dissertation as well as for challenging me to think more deeply about the questions this research has addressed. Special thanks also go to my 'honorary' committee member, Margaret Sinclair, who has helped me with the development of the French stimuli and made sure that all things French in this dissertation are indeed correct.

I also thank the faculty and staff of the Department of Linguistics as well as my fellow graduate students for their intellectual and social support. I have been privileged to receive financial support for this project from the Department of Linguistics, a Northwestern University Graduate Research Grant, as well as NSF grant BCS-0846147 to Matt Goldrick.

Thank you to the participants in this study for sitting through hours of not always exciting linguistics experiments and for sharing their stories from their time abroad with me along the way. Thanks also to Lucie Gauthier and Melissa Baese-Berk for recording the French and English stimuli respectively, and to Sarah Gregory, Kelly Kahle, and especially Sean Arn for

helping me with subject running and the tedious task of doing acoustic measurements. Many thanks go to Stewart Callner for proofreading all 260+ pages of this dissertation.

Thanks so much to my friends and family in Austria, the U.S., and the rest of the world for their support and encouragement. And finally, I am eternally grateful to Jeremy, who has put his own life on hold for the past few months for doing everything from teaching me how to code in Python and R (and troubleshooting everything I coded wrong) to running the household and being a full-time dad. You don't know how thankful I am to have you in my life, and now that it's your turn, please let me make up for it.

To Jeremy and Julius

– My favorite bilinguals

Table of Contents

List of tables

List of figures

Chapter 1: Introduction and literature review

1.1 Introduction

Most colleges and universities in the United States offer study-abroad programs and more and more students decide to go to another country for a few weeks, a semester or a year to study another language. For example, in 2008, around 240,000 American students received academic credit through study-abroad programs (Institute of International Education, 2009). While the factors that influence successful L2 acquisition during a study-abroad program have been investigated in a number of studies (e.g., Jiménez Jiménez, 2003; Sunderman & Kroll, 2009), systematic behavioral studies have paid less attention to what happens to the study-abroad returnees' language skills once they return to their home countries. In spite of the absence of such systematic studies, most learners of a foreign language (abroad or at home) know from experience that as soon as they stop using the language, their ability to use and understand it seems to diminish. While some studies on second language loss – or *L2 attrition* – have shown that learners' perception and production skills diminish once exposure to the language ends (e.g., Cohen, 1989; Dugas, 2000), other studies have suggested that learners' perceptions of language loss do not necessarily reflect their abilities when performing linguistic tasks (i.e., their abilities may have diminished less or not at all; e.g., Murtagh & van der Slik, 2004). Unfortunately, individual studies often cannot be compared directly since they looked at different groups of participants, different language, or used different tasks. Therefore, in order to get a clearer picture of the different aspects that influence L2 retention versus attrition in study-abroad returnees, a comprehensive study investigating the role of linguistic as well as extralinguistic

factors in L2 attrition and retention, ideally in a longitudinal design, is desirable. This is what the current study aims to provide.

The current study investigates the retention versus attrition of L2 French in a group of American students who have recently returned from a study-abroad program in France. The participants (henceforth *bilinguals*), are tested at three points in time across the first nine months after their return from France. To provide as clear a picture as possible on the different aspects influencing L2 retention versus attrition, a number of variables are taken into consideration.

Linguistic factors that may influence L2 retention include whether language comprehension or production is investigated, or which processing level one focuses on (e.g., lexical versus phonological processing). Furthermore, for phonological processing, different types of sounds may be influenced differently by changes across time (i.e., attrition, retention, or improvement). These sound types include vowels versus consonants, sounds that are similar in the two languages versus sounds that only exist in one of the languages, or sound contrasts that are similar in the two languages versus contrasts which do not map onto similar native-language (L1) contrasts. At the lexical level, changes across time may affect words differently based on their lexical status (cognates versus noncognates). The current study investigates the effect of attrition (or retention) on all of these linguistic factors. The focus of the study is on spoken language, i.e., speech perception and speech production.

While such a comprehensive look at the different aspects of a language and how they are influenced differently by changes across time will be informative, to provide a complete picture of the factors for language attrition and retention one also has to look at extralinguistic factors,

i.e., individual differences between speakers. The individual differences investigated in the current study include the fact whether the bilinguals were still taking French classes after their return from France, their L2 proficiency level at different points in time, the amount of exposure they received to spoken French since their return home (over and above the exposure in French class), the bilinguals' motivation to learn French, their attitudes towards the French culture, and their anxiety when using French, as well as cognitive factors, such as working memory span, phonological awareness, and executive functioning.

The remainder of this chapter first provides an overview of previous research on phonological and lexical L2 retention and attrition, and then reviews models of L2 acquisition to provide a framework to explain L2 retention versus attrition. Next, previous research on sound status as a factor in language acquisition will be reviewed. In this context, the specific sound status differences between the sound systems of French and English (and previous work on how such differences influence L2 French speech perception and production) will be reviewed. Finally, studies on the influence of lexical status on L2 acquisition will be reviewed to provide a framework within which the current study's results on language retention versus attrition can be evaluated.

1.2 Previous research on phonological and lexical L2 attrition, retention, and improvement

1.2.1 Phonological attrition, retention, and improvement

Few studies to date have investigated the retention versus attrition of phonological knowledge in the L2 of language learners. The results from this previous research are mixed,

with some studies finding attrition, others retention, and yet others improvement on phonological tasks. Studies have looked at both speech perception and production.

Dugas (2000) tested American college students who had participated in a study-abroad program in France on their pronunciation of French sounds. Participants performed a picture naming task in French with target words containing critical French sounds that were either similar to English phonemes (such as French /u/, which is similar to English /u/) and "new" French sounds for which no similar English phoneme exists (such as French /y/). The participants were tested once at the end of their study-abroad program and again 11 months later. From the participants' productions, word pairs were created that consisted of each target word produced at the first recording and the same word produced at the second recording 11 months later (the word from the first recording was the first word in the pair half of the time). These word pairs were played to native French speakers who judged whether the vowel in the first or the second word sounded more native-like. Dugas found that more speakers showed signs of attrition than improvement from the first to the second recording. No significant differences were found for similar versus new sounds.

Sancier and Fowler (1997) found signs of L2 attrition (as well as L1 attrition) and relearning in a case-study of a native-speaker of Brazilian Portuguese with L2 English. The speaker's voice onset times (VOTs, i.e. the duration of time between the release of the consonant to the onset of voicing of the following vowel) of the voiceless stops /p/, /t/, and /k/ were measured at three times: after a stay of 4.5 months in the United States, after a stay of 2.5 months in Brazil, and again after a stay of 4 months in the U.S. VOTs were shorter (i.e., more Portuguese-like) after the speaker's 2.5-month stay in Brazil, indicating that some attrition of

production abilities had taken place while in the L1 environment. However, the VOTs were longer (i.e., more English-like) when the speaker was recorded again after another 4 months in the U.S., suggesting that she reacquired her previously lost, more English-like production abilities. It should be noted, too, though, that monolingual English listeners did not deem the speaker's productions from the third recording (after the stay the U.S.) more English-like than the productions from the second recording (after the stay in Brazil), suggesting that while at an acoustic level, the productions had improved, the change may not have been substantial enough to be detected by native English listeners. No results were reported for native-English listener judgments of the speaker's productions at the first versus second recordings. Therefore, it is unclear whether the attrition found between those productions in the acoustic analyses would be detected by listeners.

Gonzalez Moncada (1995) tested English teachers in Colombia on their English productions a few years after the end of their teacher training. Specifically, VOTs of the voiceless stops /p/, /t/, and /k/ were measured. Colombian college students majoring in English (i.e., the same program that the teachers had followed) were tested to provide a baseline for the English level at the end of teacher training. The author found attrition in one group of teachers, but she found improvement in another group of teachers. Teachers in Colombian public schools had become less English-like in their VOTs compared to the college students whereas teachers in private language centers had improved compared to the college students, i.e., their VOTs were more native-like in English. Gonzalez-Moncada suggests that the major contributor to this difference in performance seemed to be the amount of English use since the end of teacher

training: teachers in private language centers reported using English more and Spanish less than the teachers in public schools.

Another study found improvement in L2 proficiency after the end of formal training. Weltens (1989) investigated retention of L2 French in Dutch high school and college students. The students had studied French either for four or six years, and the years of non-use at which participants were tested were either two or four years. In two sound perception tasks (a sound discrimination task and a sound-spelling matching task), the participants who had not been exposed to French for two to four years were more accurate than the control group that had just completed their French training (i.e., had zero years of non-use). Furthermore, participants who had studied French for six years were overall better than those who had studied French for four years, and participants improved more across time on consonant contrasts than vowel contrasts. Weltens' study suggests that even in the absence of exposure to an L2, further processing seems to take place that can lead to improvements in the language that are noticeable as late as six years after the end of training.

Murtagh and van der Slik (2004) found retention rather than attrition or improvement in a group of former secondary school students of L2-Irish when they were tested 18 months after the end of L2 instruction, compared to a test right at the end of instruction. Participants had studied Irish in school at three different levels (ordinary L2 classes in an English language school, higher level L2 classes in an English language school, Irish immersion school). When tested on 14 target sounds embedded in a 200-word text to be read aloud, all three groups of students had retained their pronunciation accuracy (sound productions were labeled as right or wrong by the author and a second rater) at the end of the 18-month period since the end of training. While

some of the participants in the study were still exposed to Irish, the study suggests that even without any exposure to an L2, speech production may not suffer any losses within the first year and a half after the end of training. The above studies suggest that L2 speakers can experience attrition, retention, or even improvement in their sound processing skills after the end of formal training in the L2.

While the two production studies that have found improvement across time (private center English teachers in Gonzalez-Moncada, and reacquisition of more English-like VOTs in Sancier and Fowler's speaker upon her return to the U.S.) involved exposure to the L2, Weltens found improvement without further exposure in two sound perception tasks. It may be possible that attrition is simply more likely to happen in production than perception, and that in order to improve in L2 production, exposure to the L2 is necessary (even if to a much more limited extent than during the training period). Studies on lexical L2 attrition, in fact, have shown that language production skills seem to be more susceptible to attrition than language perception skills (e.g., Tomiyama, 2000). However, since there were other differences between participants in the studies presented here, such as amount of training received, languages studied, time at which attrition testing took place, no definitive conclusions can be drawn about how attrition may affect production and perception differently, and what role further exposure may play in each. In the current study, therefore, phonological perception and production skills will be tested in the same group of participants, so that direct comparisons of attrition patterns and the effects of certain individual differences (such as proficiency or exposure) can be made.

The above studies also suggest that different types of sounds, such as consonants versus vowels and "new" versus "similar" sounds may be affected differently by attrition. These factors will also be included in the current study.

1.2.2 Lexical attrition, retention, and improvement

Previous research on lexical attrition, similar to the phonological studies, has found mixed results, with some studies finding attrition, retention, or improvement, and other studies reporting mixed results for different groups of participants or for different measures. Studies have looked at language comprehension and production.

Cohen (1989) reported a case study of two children (9 and 13 years old at the onset of the study; L1 Hebrew, L2 English) whose production vocabulary was tested in their L3 Portuguese at 1, 3, and 9 months after the end of exposure to Portuguese. Vocabulary production was measured through the number of discrete words used in a storytelling task. Cohen reported that no significant changes in vocabulary use were found at the first two test sessions, but significantly decreased numbers of words were produced at the third test, 9 months after the end of exposure to Portuguese. Furthermore, attrition was more pronounced for the younger of the two children. These results suggest that lexical attrition in speech production is likely to happen within a year after the end of exposure, but may not be noticeable within the first few months.

In two studies investigating L2 attrition in North American secondary school students acquiring English, Cohen (1986) and Gardner, Lalonde, Moorcroft and Evers (1987) found improvements in lexical production skills (in a lexical fluency task in Gardner et al. and a picture description task in Cohen) after the end of a summer break relative to tests performed at the

beginning of the break. This was the case despite the absence of input during the break. Cohen suggests that the improvements may be the outcome of a 'settling-in process', i.e., during the summer break, when no new L2 input was received by the students, further L2 processing took place during which L2 knowledge was consolidated, and during which errors that had been acquired earlier may even have been 'unlearned', leading to overall better performance after the hiatus in L2 input. One should note that in Cohen's 1989 study described above, no attrition was found, either, until 9 months after the end of exposure to the L3. These two studies, therefore, suggest that in the first three months after the end of exposure lexical production skills may be fairly robust, and may even improve, and only later (i.e., sometime before 9 months post-exposure) attrition may set it.

Bahrick (1984) reported results that seem to indicate the opposite of Cohen's studies, i.e., he reported initial attrition followed by a long period of retention. However, Bahrick's study used a very different timeline from Cohen's. Bahrick tested former L2 students of Spanish in lexical perception and production (among other linguistic measures). The students were native English speakers who had studied Spanish at the secondary school level, and who had not received any considerable amount of input in Spanish for 1 to 50 years. Bahrick found patterns of attrition and retention that suggest that an initial period of attrition happens until about three to six years after the end of exposure to the L2. After this period, L2 knowledge seems to remain fairly stable until much later (i.e., about 30 years after the end of exposure), when another drop in L2 lexical performance may happen. Since Bahrick did not look at L2 performance in speakers any earlier than 12 months after the end of exposure, it is unclear whether the participants in his study also did not experience attrition until about 9 months after the end of

exposure, parallel to Cohen's participants. Furthermore, the patterns of attrition versus retention may be different for child- versus adult L2 speakers (all participants in Bahrick's study had studied Spanish in high school or college). Bahrick, furthermore, found that perceptual lexical skills seemed to be more resistant to attrition than productive lexical skills.

Tomiyama (2000) also found different patterns of attrition versus retention for lexical comprehension and lexical production. In a case study of an 8-year old Japanese child returning to Japan after seven years in the United States, Tomiyama assessed English lexical perception and production skills between two months and 33 months after the child's return to Japan. Receptive vocabulary was assessed through the Peabody Picture Vocabulary Test. In this test, a word is named by the experimenter and the participant has to select, out of four choices, the picture that correctly illustrates the word. Productive vocabulary was assessed by naming objects in a picture book. Tomiyama found that the child started showing lexical retrieval difficulties in speech production at eight months after the return to Japan, but that the child's productive vocabulary stabilized from about 19 months after the return. In fact, while no new attrition of productive vocabulary took place until the last testing session at 33 months, some words that appeared to have been lost at the 19-month test seemed to have been reactivated, and were in the child's productive vocabulary again by the 33-month test. In lexical perception, no significant changes were found across time. Tomiyama's results suggest that lexical perceptual and productive skills are affected by patterns of attrition differently. Perceptual knowledge remains stable across approximately the first three years after the end of immersion. In contrast, there may be an initial decline in productive vocabulary; this eventually levels out, after which there may even be improvements. This is in line with Cohen's settling-in period. Some lexical items

may have been unavailable to the child during the settling-in period, after which they became available again. Of course, the timeline in Cohen's and Tomiyama's studies are different, with the settling-in period in Cohen's study taking place within two months after the end of exposure, but in Tomiyama this period seems to not have happened until about a year and a half after the end of immersion. Initial proficiency levels as well as contrasting modes of acquisition (limited L2 exposure in an L1 environment in Cohen's study versus an L2-immersion situation in Tomiyama's study) may be responsible for the different time course of attrition versus retention and improvement in the two studies.

Mehotcheva (2010) investigated productive lexical skills through a picture naming task in five German and Dutch study-abroad returnees who had spent five to twelve months in Spain. Her results for L2 Spanish lexical production were also mixed. Reaction time data showed that all participants were slower at naming pictures in Spanish at the time of the second test (twelve months after the end of their study-abroad program) than at the first test (about 1 month after the end of the program). In addition, participants' reaction times increased across time especially for low- and medium-frequency words, while they increased less for high-frequency words. In terms of accuracy rates, no clear pattern of attrition versus retention was found for the five participants. Two of the participants were more accurate at Time 1 than at Time 2, two participants were more accurate at Time 2 than at Time 1, and one participant performed similarly at both times. There were also no clear patterns of how word-frequency affected the changes across time. Overall, therefore, Mehotcheva found some evidence for lexical attrition in that the participants were slower at Time 2 than at Time 1, but the accuracy data showed that it was also possible for some

study-abroad returnees to improve between the two tests. These results were not influenced by the amount of exposure reported by participants.

Finally, similar to their results in phonological retention, Murtagh (2003) and Murtagh and van der Slik (2004) found no changes across time in an object description task. Scoring of the task included both vocabulary adequacy as well as grammatical correctness. All three groups of participants (i.e., the ordinary and high level L2 learners from English language schools, and the L2 learners from Irish immersion schools) showed no significant losses or improvements in their Irish production skills across the first eighteen months after the end of formal instruction.

Similar to the results reported for phonological attrition, retention, and improvement, the results reported by lexical L2 attrition studies are mixed. Attrition, retention, and improvement have been found in both lexical comprehension and production. Since results cannot be compared directly across studies using different participant groups, languages, and time windows investigated, in the current study, lexical attrition versus retention will be investigated in the same group of participants in both perception and production.

1.3 Theoretical frameworks for language attrition, retention and improvement: Models for second language speech acquisition

One question that comes up again and again in research on language attrition and retention is whether attrition can be seen as a reversal of language acquisition. This question is the main focus of regression theory (Jakobson 1941/1968). But even if one does not believe that attrition is simply a reversal of acquisition, understanding the processes underlying language acquisition and cross-language interaction is necessary in order to make sense of attrition

phenomena. The next section will therefore turn to models of second language acquisition and interpret their claims with regard to their relevance to L2 retention versus attrition. The three most prominent L2 speech models are discussed below.

1.3.1 The Speech Learning Model (SLM)

Flege's Speech Learning Model (SLM; Flege & Hillenbrand, 1984; Flege, Schirru, & MacKay, 2003) deals with phoneme categorization in the L1 and the L2; it characterizes the influences of the L1 on the L2 and influences of the L2 on L1 sound categorization. According to the SLM, L2 acquisition can be easy or difficult depending on the relationship between the sounds across the two languages. It will be difficult to acquire L2 sounds that are similar to L1 sounds, because they will be *assimilated* to the L1 sound category. This means that a sound encountered in the L2 is interpreted as a member of the L1 category. This makes it harder for a speaker to perceive or produce the L2 sound in a native-like fashion. In the case of L1 English speakers acquiring L2 French, a French sound such as the /u/ in *mouton* (/mutõ/, English 'sheep') would be difficult to acquire to a native-like degree because of its similarity to the English sound /u/ as in *moose*. If an L2 sound, on the other hand, is different enough from any L1 sounds, it will not be classified as a member of an L1 phoneme category, and it will be easier for an L2 learner to acquire this sound in a native-like fashion. The reason for this is that a separate category has to be formed for the sound in the L2 (rather than incorporating it into an existing L1 category). In this case we see *dissimilation* of the L1 and the L2 sounds. An example for this would be the French sound /y/ as in *rue* (English 'street') which has no close equivalent in English. An L2 learner would thus create a new phoneme category for /y/.

Assimilation and dissimilation have different consequences for L1 and L2 representations. Assimilation changes both the L1 and the L2 phoneme category so that they both fall somewhere in between native L1 and L2 categories. With dissimilation, on the other hand, the L1 and the L2 categories move apart from each other, making each category more extreme than a native L1 or L2 category.

In the current study, patterns of phoneme-level language attrition in an L2 word repetition task will be analyzed in the light of the SLM. The SLM predicts that *new* sounds (such as French /y/) will be subject to dissimilation, whereas *similar* sounds (such as French /u/) will be subject to assimilation. Because the former are predicted to have been acquired more fully than the latter, the SLM predicts more native-like production of and less attrition for new vs. similar sounds.

1.3.2 The Perceptual Assimilation Model (PAM)

The Perceptual Assimilation Model (PAM; Best, 1995) was developed to explain non-native sound *contrast* perception. The focus in this model is on naïve listeners of a non-native language. Specifically, this model claims that the ability to contrast non-native sound pairs depends on the similarity of these sound contrasts to L1 phoneme contrasts. For example, the distinction of a non-native pair will be easy if there exists a similar phoneme contrast in the L1, but it will be difficult if both non-native sounds fall within a single phoneme category in the L1. Similarity in this model is defined as the perceived similarity of the articulatory gestures involved in producing the sounds in question, rather than as acoustic similarity.

Recently, the model has been expanded to include second language speech perception (PAM-L2; Best & Tyler, 2007). Like PAM, PAM-L2 also focuses on the similarity of non-native sound contrasts to L1 sound categories. The model addresses several different types of non-native sound contrasts. The most relevant types for L2 retention versus attrition are two category assimilation contrasts, category goodness assimilation contrasts, and single category assimilation contrasts.

An example for two category assimilation of a French contrast for a native English learner would be /d-t/. In both French and English these two sounds differ in their voice onset times, i.e., the time from the release of the consonant to the onset of the voicing of the following vowel. The French consonant pair is phonetically distinct from the English equivalent, with overall shorter voice onset times in French than English, but the contrast does exist in both languages and the two sounds in the French contrast can therefore be mapped directly onto their English equivalents to form a comparable contrast. On the other hand, the French /u-y/ contrast, as in the French minimal pair *pour* [puʁ] – English 'for' – versus *pur* [pyʁ] – English 'pure', is an example of a category goodness assimilation contrast. Since there is no /y/ in English, naïve listeners tend to interpret the French /y/ as an exemplar of the perceptually closest English sound, which is /u/. In fact, both the French /u/ and /y/ sounds will be mapped onto the single /u/ phoneme in English, but French /u/ will be interpreted as a better exemplar (show a better goodness of fit) of the phoneme category /u/ than French /y/. Finally, the French contrast /y-œ/ is an example for single-category assimilation. Both sounds are likely to be interpreted as equally poor exemplars of /u/ in English and will be the most difficult contrast to discriminate for native English speakers (cf. Levy, 2009). In addition to this explanation of perceptual difficulties with

L2 contrasts, the PAM also posits that if an L2 sound is articulatorily very similar to an L1 sound (e.g., /d/ and /t/), no new phonetic and phonological L2 category is formed. In contrast, if an L2 sound is phonetically distinct from an L1 sound but perceived as an exemplar of an L1 category (e.g., /y/), a new phonetic category is added to the existing L1 phonological category.

PAM-L2 predicts that two category assimilation contrasts are the easiest to acquire, category goodness assimilation contrasts are more difficult to acquire, and single category assimilation contrasts are the most difficult to acquire.

Similar to the SLM, PAM-L2 predicts that L2 sound contrasts that include sounds that have developed their own phonetic and/or phonological category can be expected to be more resistant to attrition than L2 sounds that have been assimilated to L1 categories (because they have been acquired better in the first placed). Specifically, PAM-L2 predicts that in two category assimilation contrasts both sounds will be assimilated to existing L1 phonological categories. For example, for native English learners of French, French /t/ vs. /d/ will be mapped onto English /t/ and /d/. This also implies that little or no perceptual learning of either sound takes place (since each sound is simply assimilated to an existing L1 category). Since these contrasts are fairly easy to discriminate even for naïve listeners of a language, it is expected that L2 learners (such as the bilinguals in the current study) would perform at a high level – almost native-like or somewhat lower. As for attrition, since little perceptual learning took place in the first place, it is possible that no more changes occur across time, and that bilinguals simply retain their high-performance levels at this discrimination task regardless of whether they are still exposed to French or not. Alternatively, if listeners did in fact experience some perceptual learning of the sounds making up the two category assimilation contrasts, they may show attrition of their discrimination

abilities due to the fact that the sounds in each pair were not acquired in depth and may therefore be vulnerable to attrition. In either case, no improvement is expected, even if listeners are still exposed to the L2. Since little if any perceptual learning is expected for these sound pairs is expected in the first place, it is unlikely that this learning would take place after the immersion period rather than during it.

PAM-L2's predictions for category goodness assimilation contrasts are somewhat different. More perceptual learning is expected for the one sound in the contrast that is less similar to an existing L2 category. So, in the case of the French /u-y/ contrast, PAM-L2 predicts perceptual learning of /y/, resulting in an added phonetic and phonological category for this sound, whereas little or no perceptual learning is predicted for /u/, which is simply assimilated to the English /u/. Listeners should initially have more difficulty with this contrast than with two category assimilation contrasts, but after they have reached a sufficiently high L2 proficiency, they may perform at a high level for this contrast as a result of the perceptual learning that has taken place for /y/. There may even be improvement across time if participants are still exposed to the language. The bilinguals in the current study fall within the window where PAM-L2 predicts continued improvements. PAM-L2 predicts that most perceptual learning takes place within the first six to twelve months after second language immersion; bilinguals in the current study on average started their 4-month immersion experience six months prior to the first testing session. So, some improvements can still be expected with continued exposure. A critical caveat is that the bilinguals' exposure to the L2 is bound to be much more limited than during the participants' stay in France. Therefore, substantial improvements are less likely to be observed.

Finally, it is possible that attrition could be observed, but this is unlikely given that robust perceptual learning for /y/ is predicted to have taken place during the bilinguals' stay in France.

In conclusion, if any differences are found across participant groups and across time, the bilinguals' performance for two category assimilation contrasts is expected to be as good as or more native-like than category goodness assimilation contrasts at Time 0. Attrition is more likely to be found in two category assimilation contrasts than category goodness assimilation contrasts. Improvement, though unlikely to be found for either contrast, is more likely to occur for category goodness assimilation contrasts than for two category assimilation contrasts.

1.3.3 The Native Language Magnet Theory (NLM)

A third model of speech sound acquisition, the Native Language Magnet Theory (NLM; Kuhl, 1991) will not be used to interpret the data in the current study, but it will be briefly covered here for completeness' sake. The NLM was developed to explain first language acquisition, but it can also be applied to second language acquisition. The theory revolves around the claim that our native language influences how we categorize speech sounds. The best exemplar of a certain sound (*prototype*) "attracts" similar sounds like a magnet, so that they are perceived as members of that particular sound category. This L1-dependent categorization mechanism is developed during the first year of life and can make it difficult for adults to acquire non-native phonemes (Iverson & Kuhl, 1995).

While the NLM theory states that L2 sounds will be influenced by similar sounds in the L1, the theory does not speculate about the influence the L2 in a proficient speaker might have on L1 sound categories. Neither does the NLM theory discuss the possibility for changes across

time in the strength of the magnet effect (such as when an L2 is used much more frequently than an L1).

With regard to language attrition, the question arises whether the acquisition of an L2 can alter L1 phoneme categories or the strength of the magnet effect on L2 sounds (as signs of L1 attrition). Furthermore, once L2 prototypes have been acquired, can a decrease in use of the L2 lead to L2 sounds being attracted more strongly again by L1 prototypes? However, since the NLM theory does not include the possibility for dynamic changes in phoneme categorization across time, it is difficult to explain language attrition phenomena in light of this theory. The NLM theory will, therefore, not be used to try to explain language attrition processes in the current study.

1.4 Sound status as a factor in language retention

Differences in the processing of vowels and consonants (i.e., sound status) are a well-known phenomenon. Previous research has shown that listeners have shorter reaction times for consonants than vowels in sound identification tasks (Studdert-Kennedy, Liberman, & Stevens, 1963) and phoneme monitoring tasks (Cutler & Otake, 1994). Differences in vowel versus consonant processing have also been found in ERP (event related potential) studies, where a larger negativity was associated with consonant than vowel processing in a speech perception task (Carreiras, Gillon-Dowens, Vergara, & Perea, 2007) and in studies with language production-impaired populations (Caramazza, Chialant, Capasso, & Miceli, 2000) where contrasting selective difficulties in producing vowels and consonants have been found in aphasics.

More directly relevant to the current study, previous research on L2 acquisition and attrition has suggested that bilinguals treat vowels and consonants differently. For example, Birdsong (2007) found that highly proficient native-English L2-speakers of French were more native-like in their productions of French consonants than vowels. In perception, a contrasting result has been found with native Dutch speakers. Weltens (1989) found that Dutch L2-speakers of French achieved higher proficiency in discriminating French vowels than consonants, and that this vowel advantage was still present four years after the end of exposure to French. These perceptual findings might be attributable to acoustic differences between vowels and consonants. Bohn (1995) suggests that some acoustic cues may be more salient than others to L2-learners, regardless of their native language. For example, duration seems to be more salient to listeners than spectral cues. Since spectral cues play a more important role in the perception of vowels than consonants (at least stop consonants, for which VOT is the primary cue), a difference in the discrimination of vowels compared to consonants is not surprising. The current study will therefore also compare the bilinguals' performance in discriminating and identifying vowels versus consonants.

In the current study phonological retention versus attrition in the L2 of native English speakers who are proficient in French (after studying in France for a semester) are investigated. To study attrition of French phonology by English-native bilinguals, it is necessary to know what problems naïve English speakers as well as English-native learners of French have with French phonology. A number of previous studies have addressed these problems. In the section below, the main differences between French and English vowels and consonants that previous research

has focused on are presented; results from previous studies on speech perception and production are then reviewed.

1.4.1 Differences between French and English consonants and vowels

Previous research on French-English *consonant* contrasts has focused on the voice onset time (VOT) differences between the two languages. Voice onset time is the duration from the release of the constriction when producing a stop consonant, such as /p/, or /d/, to the onset of voicing (i.e., as resulting from periodic vibration of the vocal folds). French voiced stops are normally produced with lead VOT, i.e., voicing begins before the stop burst release, whereas English voiced stops generally are produced with short-lag VOT, i.e., voicing begins at the burst release or shortly thereafter (ranging from 0 msec to approximately 30 msec). French voiceless stops are normally produced with short-lag VOT, whereas English voiceless stops are produced with long-lag VOT (i.e., VOT greater than approximately 30 msec; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973). The French voiceless stops and the English voiced stops are, therefore, acoustically quite similar, in spite of corresponding to contrasting phonological categories across languages.

In addition to this duration contrast, French and English /d/ and /t/ differ in place of articulation, with French coronals being articulated as dentals, and English coronals being produced as alveolars. Furthermore, English has consonants in both dental (e.g., /ð/) and alveolar position (e.g., /d/), whereas French has no alveolar consonants (Polka, Colantonio, & Sundara, 2001).

The most studied problems for English speakers with French *vowels* concern the front rounded vowels /y/, /ø/, and /œ/, which are not part of the English sound inventory (e.g., Gottfried, 1984; Levy & Strange, 2008). Furthermore, French, unlike English, has a phonemic contrast between oral and nasal vowels, e.g., /ɛ-ɛ̃/.

1.4.2 Perception of French sounds by native English speakers

Studies on the perception and production of French and English vowels and consonants usually have found clear differences between the two monolingual groups, confirming that the consonant and vowel features investigated were indeed different for the two languages. Furthermore, bilinguals and, to a somewhat more limited extent, L2-learners seem to develop separate phoneme systems for the sounds in the two languages that tend to be intermediary to the two monolingual categories. In the following paragraphs, studies of French phoneme perception and production by English speakers are described in more detail.

Consonant perception

In an early cross-linguistic study, Caramazza, Yeni-Komshian, Zurif, and Carbone (1973) compared the perception and production of French stop consonants by monolingual Canadian French speakers, monolingual Canadian English speakers, and French-English bilinguals. In the perception task, participants were asked to label synthetic voiced-voiceless VOT continua of the stop consonants /b-p/, /d-t/, and /g-k/ presented in consonant + vowel (CV) syllables. In these continua the VOT of the consonants was manipulated so that they ranged from -150 msec to 150 msec in steps of 5 to 10 msec. As expected, the two monolingual groups differed significantly. Compared to English speakers, the French speakers exhibited an earlier crossover point, i.e., the point along the voiced-voiceless continuum where listeners tend to switch from being more

likely to judge a sound as voiced to being more likely to judge it as voiceless. The bilinguals' crossover points fell in between the native speakers' both in English and in French. These results confirm that the consonants in the two languages are indeed different acoustically, and furthermore, that becoming fluent in both languages influences how one perceives the consonants in the two languages.

Polka, Colantonio, and Sundara (2001) and Sundara, Polka, and Genesee (2006) investigated the discrimination of place of articulation in the English /d-ð/ contrast. The authors tested English- and French- monolingual and bilingual infants, children, and adults and found that the discrimination of English-native participants improved with age, while French-natives showed poorer discrimination than English-natives, and the (French dominant) bilinguals performed similarly to French-native speakers. These results show that the L1 influenced nonnative sound perception and suggest that in the case of the bilinguals perceptual patterns were determined by their dominant language.

French versus English /d/ was the focus of another study by Sundara and Polka (2008). Canadian French and Canadian English monolingual adults, simultaneous bilinguals, and advanced early L2 learners of English and French performed an AXB discrimination task containing tokens of French versus English /d/. English monolinguals performed better than French monolinguals and early L2 learners, reflecting the English listeners' experience with a phonemic dental-alveolar contrast for fricatives (/θ/ and /ð/ versus /s/ and /z/). However, while L2 learners performed better than French monolinguals, their performance was worse than that of simultaneous bilinguals (who performed similarly to the English monolinguals). This pattern

suggests that L2 learners merged the /d/ categories for their two languages whereas simultaneous bilinguals were able to keep them separate.

These studies suggest that bilinguals' and L2-learners' perception of consonants in their two languages is different from the respective monolingual speakers'. With regard to language attrition, this raises the possibility that bilinguals' or L2 learners' perception of the consonants in their two languages could change once they stop being exposed to one of the languages.

Vowel perception

While the above studies looked at consonant perception, Gottfried (1984) was the first to perform a comprehensive cross-linguistic perception study of French vowels. Native French listeners, English-native learners of French, and English monolinguals were tested on the discrimination of eight French vowel contrasts in consonant-vowel-consonant (CVC) syllables and in isolation. For CVC syllables, English monolinguals made more errors than French monolinguals and learners of French. Learners of French, on the other hand made more errors than native French speakers only for contrasts including front rounded vowels, but performed similarly to the native French speakers for contrasts not including front rounded vowels. When vowels were presented in isolation, all English speakers made more errors than the French monolinguals for contrasts with or without front rounded vowels. These results suggest that some vowels (front rounded vowels) are more difficult to acquire than others for L2 learners of French.

More recently, Levy and Strange (2008) had naïve American English listeners as well as native-English advanced learners of French perform an AXB categorization task involving six French vowel contrasts. Generally, the learners of French performed better than the naïve

listeners, but both groups had a similarly high error rate for the /u-y/ contrast. Furthermore, the naïve listeners performed better for vowels in a bilabial consonant context than in an alveolar context, whereas there was no consonantal context effect for the learners of French. From this the authors conclude that context can affect the outcome of L2 sound acquisition. The results, furthermore, confirm the difficulty of front rounded vowels for speakers of English.

Using the same stimuli as in the above experiment, Levy (2009) also tested American English speakers with varying degrees of exposure to French (no exposure, formal instruction, formal instruction plus immersion) on their classification of the French front rounded vowels /y/ and /œ/. Participants of all levels of French tended to perceive these vowels as English back vowels. While the patterns of classification for /y/ were fairly consistent across language experience, participants with different language experience perceived /œ/ as closest to different English vowels. The author concludes that perceptual classification of nonnative sounds does not depend on acoustic similarities between vowels alone, but that native-language allophonic variation influences perceptual patterns developed during L2-learning. Since these studies found that some vowels tend to be more difficult to categorize and label than other vowels, it would be interesting to see if vowel type influences patterns of retention versus attrition in learners of French who are no longer exposed to their L2.

1.4.3 Production of French sounds by native English speakers

Consonant production

In addition to the perception study described above, Caramazza et al. (1973) measured the VOTs of voiced and voiceless stops produced by monolingual speakers of Canadian English,

monolingual speakers of Canadian French, and French-English bilinguals. As expected, the monolingual English speakers' productions showed a clear voicing contrast (short-lag versus long-lag VOT), whereas the contrast between voiced and voiceless was less pronounced for the French speakers (lead VOT versus short-lag VOT). The bilingual participants differed in their productions of French and English VOTs. Their patterns in English mirrored the contrast shown by the English monolinguals, while their patterns in French were similar to the French monolinguals, with more overlap between the voiced and voiceless categories. However, their productions were more similar to the monolinguals' (i.e., more native-like) in French than in English–likely due to the fact that they were native speakers of French who had acquired English as their L2. This study shows that in speech production, too, the native, or dominant, language exerts a strong influence on bilinguals' productions.

In a complementary study to the perception study by Sundara and Polka (2008) described above, Sundara, Polka, and Baum (2006) analyzed productions of /d/ and /t/ by native-speakers of Canadian English and Canadian French as well as simultaneous bilinguals for VOT values and measures of place of articulation differences (dental for French and alveolar for English). The productions by the two monolingual groups differed in VOT, relative intensity, and in several distributional measures of the burst frequency (mean, standard deviation and kurtosis) which provide cues to place of articulation differences. Across languages, the productions of /d/ versus /t/ differed with respect to VOT (long- vs. short-lag) and distributional measures of burst frequency. The productions of the bilinguals fell in between the monolinguals' productions along both of these measures. In English, the bilinguals' productions of /d/ and /t/ differed from monolinguals in VOT and measures of place of articulation, whereas in French they only

differed in VOT. As in perception, performance by simultaneous bilinguals showed clear differences between their two languages, whereas in both languages bilinguals' productions fell in between the monolingual productions.

Similarly, Fowler, Sramko, Ostry, Rowland, and Hallé (2008) tested whether early and late bilinguals create a single phonological system that lies in between the respective monolingual systems for the two languages. The authors compared the VOTs of /p, t, k/ produced by monolingual speakers of French and English to the VOTs produced by simultaneous French-English bilinguals and by late English-French and French-English bilinguals. They found that each group of bilinguals' productions in both languages were intermediary to the respective monolingual VOTs. Furthermore, late bilinguals producing VOTs in their L1 were the most similar to the monolinguals, followed by the simultaneous bilinguals, and the late bilinguals in their L2 were the most different from the monolinguals (but still different from the monolinguals of the other language). Importantly, all three groups of bilinguals produced distinct voiceless stops in both of their languages that differed from monolingual productions.

With all of these studies showing that bilinguals' productions fall somewhere along a continuum between the productions by monolingual speakers of the two languages and that proficiency in the two languages influences where on the continuum bilingual productions fall, it is, again, not clear whether bilinguals' productions would move along the continuum if they were no longer exposed to one of their languages.

Consonant and vowel production

Flege and Hillenbrand (1984) and Flege (1987) were interested in whether perceptual similarity between L1 and L2 sounds influences the production of these sounds. Native-English and native-French participants with various levels of proficiency in the L2 (ranging from no experience to high proficiency and having lived in the L2 country for several years) produced phrases in English and French containing the syllables /tu/ and /ti/ in English and /tu/ and /ty/ in French. VOT values of the consonants and F2 values of the vowels were measured. In the two studies VOT values of /t/ and F2 values of /u/ of the native-English speakers of French and of the native-French speakers of English fell in between the values of the monolingual English and French speakers. The higher the L2 proficiency, the more native-like the L2 productions were. For /y/, however, the authors found relatively native-like productions by both the native-English speakers of French and the French-native speakers of English. This study shows that, in addition to differences between speakers due to varying proficiency, different sounds themselves may be produced more or less native-like by bilinguals.

Birdsong (2007) compared consonants and vowels produced by advanced English-native learners of French. VOTs of the voiceless consonants /p, t, k/ were analyzed as well as vowel durations of /i, e, o, u/. The French learners' productions were compared to monolingual French productions. Furthermore, Birdsong analyzed the learners' global accent as determined by native speaker ratings of a read passage. Analyses showed that while overall, vowels produced by the French learners were longer than those produced by native speakers, a few of the learners produced French vowels with native-like duration. The same was true for VOT, with longer VOTs by French learners overall, but some learners performing similarly to native French

speakers. The analyses also revealed that more learners were able to produce native-like consonants than vowels. In addition, the learners who performed native-like at the segment level were the most likely to exhibit native-like pronunciation at the global level. The author concludes that while the L1 influenced L2 productions of the majority of the participants, it was possible for some late learners of French to perform native-like at the segmental as well as the global level of pronunciation.

With these studies showing that L2 learners may be native-like in their productions of some L2 sounds, but not necessarily all, it is also possible that different sounds would be affected to varying degrees if a speaker no longer receives input from one of the languages. Furthermore, individual differences between speakers can affect their performance in the L2. For example, Flege (1987) found that L2 speakers' nativelikeness of French vowels and consonants differed depending on the amount of exposure the participants had to French. Birdsong (2007) also found that some late learners were able to produce native-like L2 French vowels and consonants, while other learners were not able to do this. Birdsong, however, does not provide information about individual differences between speakers.

1.5 Lexical status as a factor in language retention

In the current study, the focus is not only on *phonological* retention and attrition. Another question to be investigated is whether the *lexical* status of a word influences its resistance to attrition. In other words, are words that are similar in form and meaning across the two languages affected differently by attrition than L2 words that have no close equivalent in the L1 in terms of shared form and meaning?

This question more generally deals with the idea of co-activation of a language not currently in use by a bilingual speaker: A bilingual's processing of the language in use (the target language) can be influenced by the co-activation of the bilingual's other language (the non-target language). This cross-language influence can result in interference or facilitation in processing of words in the target language. Interference would lead to slower target processing, and possibly more errors, while facilitation would result in faster target processing, and possibly fewer errors.

One way to search for co-activation of a non-target language is to compare the processing of cognates or false friends to noncognates. Cognates are words that share both meaning and form (orthography and/or phonology) across two languages (e.g., English *bus*, French *bus*); whereas noncognates share the same meaning, but there is no resemblance in form across the two languages (e.g., English *car*, French *voiture*). False friends, finally, are also called interlingual homographs or interlingual homophones and share form but not meaning across languages (e.g., the French *pain* is *bread* in English). False friends are homo*graphs* if they overlap orthographically in the two languages, and they are homo*phones* if they overlap phonologically. Many studies have investigated the influence of lexical status across bilinguals' two languages (and even across trilinguals' three languages).

1.5.1 Language production

In language production cognates have been found to facilitate target processing (Costa, Caramazza, & Sebastián-Gallés, 2000) while interlingual homographs seem to lead to interference from the nontarget language (Jared & Kroll, 2001; Jared & Szucs, 2002). Costa et al. found that Catalan-Spanish bilinguals were faster at naming pictures whose names were cognates in their two languages than pictures that showed noncognate words in the target language. This

effect was greater when participants performed the task in their L2 than when they performed the task in their L1. While in both cases both languages are activated in parallel, a bilingual gets a bigger boost from their co-activated native language when they are performing a task in their L2 than vice versa. This is likely the case because processing in the L1 is already faster and co-activation of the nonnative language may not be fast enough to boost L1 activation even more for cognates. In the weaker L2, however, a cognate co-activated in the L1 has more potential to boost the weaker L2 activation.

1.5.2 Language comprehension

In language comprehension both cognate facilitation and cognate interference have been found, depending on task demands and the type of linguistic overlap. A number of studies have looked at masked priming (where participants are not aware they were briefly shown a word before an experimental trial) to investigate cross-language activation and cognate effects. Generally, these studies have found facilitation in a lexical decision task if a cognate prime was shown just prior to the target trial. This was the case whether the lexical decision task was performed in participants' L1 or L2 (e.g., De Groot & Nas, 1991; Grainger & Frenck-Mestre, 1998). Facilitation with noncognate primes, however, was only found if participants performed the task in their L2 and were presented with an L1 prime (Gollan, Forster, & Frost, 1997).

Cognate facilitation was also found in a study by Lemhöfer, Dijkstra, and Michel (2004), where bilinguals and trilinguals were faster at recognizing cognates than noncognates. Importantly, in this study the bilinguals were asked to identify words in either of their two languages (i.e., language non-specific lexical decision). Cognate facilitation in this non-specific lexical decision task occurred for cognates that overlapped orthographically or phonologically in

the two languages of the bilinguals or in two of the three languages of the trilinguals. Proficiency in the languages influenced the size of the effect. If the target item was a cognate between the trilinguals' native language and their L2, cognate facilitation was stronger than if the cognate was between the L1 and the L3. Furthermore, the effects were stronger when participants performed the task in their L2 rather than their L1.

In a language-specific lexical decision task, Dijkstra, Grainger and Van Heuven (1999) showed that cross-language influence can vary greatly depending on whether cognates overlap orthographically or phonologically. In fact, the authors found that Dutch-English bilinguals were faster at identifying cognates relative to English noncognates in an L2 English lexical decision task when they cognates overlapped orthographically, i.e., they showed cognate facilitation. If the cognates, however, overlapped phonologically only, the bilinguals were no faster at recognizing them as English words than they were for noncognates. Cognates that overlapped orthographically and phonologically also showed a cognate facilitation effect. Dijkstra et al. explain this as the combined facilitatory effect of orthography and semantics being stronger than the single inhibitory effect of phonology.

Woutersen, De Bot and Weltens (1995) furthermore found that cognate effects are comparable in visual and auditory lexical decision, which is relevant for the current study, since it includes an auditory lexical decision task in which cognates overlap orthographically, and partially also phonologically, across languages, and which will test whether bilinguals experience any kind of facilitation or interference during cognate versus noncognate processing.

1.5.3 Language comprehension and production

Jared and Kroll (2001) and Jared and Szucs (2002) used interlingual homographs and word-body neighbors (that share orthography of the word body, but not necessarily the onset) to investigate whether phonological representations of only one or both languages of a bilingual are activated when reading words aloud in one language (referred to as 'word naming'). Note that unlike picture naming (which relies on non-linguistic visual perceptual process), this task requires the engagement of orthographic perceptual processes. Jared and Kroll had English-French bilinguals name English target words that had word-body neighbors in French that differed in pronunciation (e.g., English *trance,* has (among others) the French word-body neighbors *France* and *chance* whose pronunciation in French is different from the pronunciation in English (English /-æns/ versus French /-ãs/). Fluent bilinguals were slower at naming English words with French neighbors than words without French neighbors, and they made more errors. This was especially the case if they had been exposed to French right before naming the English words. Jared and Szucs used a similar experimental setup, where the target words were interlingual homographs across English and French, i.e., they overlapped completely orthographically but were pronounced differently in the two languages (e.g., *pain* = English /pen/, French /pɛ̃/). English-French bilinguals (who performed the task in their L1) were slower at homograph naming and made more errors after having been exposed to French as compared to before exposure, whereas French-English bilinguals (who performed the task in their L2) suffered interference from the nontarget phonology both before and after exposure to French.

It should also be noted that the two above studies not only differed in the lexical status of the target items (i.e., cognates versus interlingual homographs) but also in the fact that the

cognates in Costa et al.'s study overlapped phonologically in the two languages, whereas the homographs in Jared and Kroll and Jared and Szucs, of course, overlapped orthographically. Jared, Friesen, and Haigh (2008) provided a third scenario by asking English-French bilinguals to perform a word naming task in both their languages involving cognates that overlap orthographically rather than phonologically in the two languages (e.g., 'train', which in French is pronounced /tʁɛ̃/). Their results were in fact, that there was a cognate facilitation effect for those participants who performed the task in their L2 French, whereas the participants who performed the task in their L1 English experienced nontarget-language interference. Participants were faster and made fewer errors on cognates than noncognates in their L2, but were slower and made more errors on cognates in their L1. Since these cognate effects seem to be modulated by L2 proficiency, as well as by recent exposure, one might expect cognate effects to weaken or disappear when bilinguals are no longer exposed to one of their languages.

1.5.4 Training studies

A few studies have also used training of new words to test for cognate effects. In a paired-associate training study De Groot and Keijzer (2000) trained Dutch students, who were experienced foreign language learners (specifically of L2 English) on associations of Dutch words with newly introduced pseudowords. The authors found that after the training, participants were faster at translating cognates (between Dutch and the trained pseudo-language) than noncognates and they also retained cognates better than noncognates when they were retested one week after the training.

Similarly, Tonzar, Lotto, and Job (2009) found that school children being trained on L2 words acquired (English-German) cognates better than noncognates (as tested directly after

training) and also retained the cognates better for up to a month after training (when the last test was performed). While the participants in De Groot and Keijzer's (2000) study and the children in Tonzar et al.'s study were exposed to the L2 words (or pseudowords) only during a brief training session, in the current study, it will be tested if analogous cognate facilitation effects can be found with adult learners who have acquired L2 words in a naturalistic setting (classroom learning and L2 immersion in France). The above experiments show that bilinguals' two languages interact not only at the phonological but also at the lexical level. The current study examines whether these interactions also affect patterns of language retention versus attrition after the end of an L2 immersion period.

In this chapter, previous literature on L2 attrition was reviewed, which has found attrition, retention, and improvement across time in different phonological and lexical tasks in L2 speakers of various languages. Furthermore, models of L2 speech acquisition were discussed with regards to their relevance to changes across time that may happen after active L2 acquisition has stopped. Finally, previous research on phonological processing of L2 French by native English speakers, as well as the influence of cognate status on L2 lexical processing was discussed.

The remainder of the dissertation is organized as follows: Chapter 2 introduces the current study. Chapter 3 presents two experiments which investigate changes across time in two speech perception tasks: an AX discrimination task (phonological retention versus attrition) and in a lexical decision task (lexical retention versus attrition). Linguistic factors (sound status and cognate status) as well as the extralinguistic factor of continued French study (i.e., whether bilinguals were taking French classes after their return from France or not) are taken into

consideration. Chapter 4 presents two speech production experiments: a word repetition task (phonological retention versus attrition) and a picture naming task (lexical retention versus attrition). The influence of the same linguistic and extralinguistic factors as in the perception experiments on changes across time will be investigated (i.e., sound status, cognate status, continued French). Chapter 5 looks at the individual differences between speakers and their influence on changes across time in the linguistic tasks in more detail. Finally, chapter 5 concludes the dissertation with a general discussion.

Chapter 2: The current study

2.1 Overall design of the study

This section describes the participants and procedures as they pertain to the study as a whole, i.e., to the experiments on speech perception and speech production, as well as the individual differences analyses. For each experiment the methodology specific to the particular experiment will be described in the respective chapter.

Participants

Seventeen native English speakers who were L2 speakers of French participated in the study. All participants were undergraduate students at Northwestern University who had participated in a study-abroad program in France in the term preceding the first experimental session (Time 0). An additional three participants were excluded from the analyses because they did not return for the second experimental session (Time 1), fifteen participants were excluded because they did not return for the third session (Time 2), and two were excluded because their background questionnaires revealed that they had not participated in a study-abroad program immediately preceding the first experimental session, but rather had participated at an earlier time. The remaining seventeen participants were included in the analyses and will be referred to as "bilinguals." Nine of these bilinguals continued taking French language classes upon their return to the United States, while eight bilinguals did not continue with formal French instruction.

In addition to the bilingual participants, there were two control groups consisting of eight native speakers of French and eight native monolingual English speakers. The native French speakers provided the native-speaker baseline for the performance on the French language tasks,

while the native English speakers provided a baseline for the performance of English speakers with no previous experience with French. Since the native French speakers resided in the United States at the time of testing, they all spoke English at varying degrees of proficiency, but all of them reported to be more proficient in French than English (On a scale from 0 to 10, self-rated Listening Proficiency in French: M = 10.0, SD = 0; English: M = 8.0, SD = 0.76; Speaking Proficiency in French: M = 10.0, SD = 0; English: M = 7.5, SD = 1.51). The native French speakers were undergraduate and graduate students as well as lecturers at Northwestern University. The native English speakers were undergraduate students at Northwestern University and were not proficient in any language other than English (self-rated Listening Proficiency in a non-French L2: M = 5.1, SD = 2.23; Speaking Proficiency: M = 3.9, SD = 2.53) and had never studied French in any formal way.

To provide a control for effects of task familiarity, the native English speakers participated in two experimental sessions (Time 0 and Time 1). Because the majority of the native French speakers were in the United States only temporarily (e.g., for an academic semester), they only participated in one experimental session at Time 0.

All participants received financial compensation for their participation ($10 per hour, plus a $10 completion bonus for participants returning for a second and third experimental session).

Procedure

The bilingual participants returned to the laboratory at three times. Once around two months after their return from their study-abroad stay in France (days after bilinguals' return

from France: *M* = 78, *SD* = 14.7; Time 0), a second time approximately five months after their return from France (*M* = 166 days, *SD* = 10.8; Time 1), and a third time about 9 months after their return (*M* = 283, *SD* = 12.3; Time 2). It was not possible to test participants immediately after their return from France. Because of this, no claims can be made here about what happens immediately after the end of an immersion experience in terms of retention versus attrition. Rather, this dissertation will focus on patterns of retention and attrition within the first two to nine months after the end of immersion.

The English native speaking participants also participated in two experimental sessions (Time 0 and Time 1). Just as for the bilinguals Time 0 and Time 1 were about three months apart for the English native speakers (days between Time 0 and Time 1 for bilinguals: *M* = 87, *SD* = 9.6; days between Time 0 and Time 1 for English native speakers: *M* = 88, *SD* = 2.1). The number of days between Time 0 and Time 1 did not differ significantly for bilinguals versus English native speakers (t (36) = -0.38, $p > 0.05$).

At each testing session (Time 0, Time 1, and Time 2), bilingual participants completed two perception tasks (AX discrimination, lexical decision) and four production tasks (word/nonword repetition, picture naming, oral summary of a dialogue, sentence reading) in French. The procedures of these specific tasks are discussed in detail below. Furthermore, information on their language background, motivation, language anxiety levels, phonological processing abilities, as well as executive function was collected as described below.

In addition to the two French perception and production tasks reported in this dissertation (AX discrimination, lexical decision, word/nonword repetition, picture naming), bilingual and

French native participants also completed a conversational speech task (oral summary of an auditory dialogue) and a sentence reading task, the results of which are not reported here. Bilingual participants also completed analogous perception and production tasks (AX discrimination, lexical decision, word/nonword repetition, picture naming, and sentence reading) in English, the results of which are not reported here.

In addition to these linguistic tasks, at each of the three experimental sessions, bilinguals filled out two questionnaires on their language background (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007; and the Study Abroad Language Background Questionnaire[1] designed for this study). Bilinguals also completed an attitude-motivation questionnaire adapted from the mini-Attitude Motivation Test Battery (AMTB; Gardner & McIntyre, 1993; see Appendix C). This questionnaire consisted of twelve out of the fifteen questions from the mini-AMTB (the questions about classroom anxiety and attitudes towards the language class and teacher were left out, since not all bilinguals were taking classes at the time of testing).

Furthermore, at each experimental session the bilinguals completed several tasks designed to assess other cognitive functions outside speech perception and production. This included two tasks from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) measuring their Phonological Memory Quotient (Memory for

---

[1] The three versions of the questionnaire given to the bilingual participants at Time 0, Time 1, and Time 2 can be found in Appendix B.

Digits task) and Phonological Awareness Quotient (Blending Nonwords task) and the Behavior Rating Inventory of Executive Function-A (BRIEF-A).

English monolinguals also completed the LEAP-Q, the two CTOPP tasks, as well as the BRIEF at each of their two experimental sessions. Native French speakers only completed the LEAP-Q. All of these tasks that participants performed in addition to the French perception and production tasks will be described in more detail in Chapter 5 on the influence of individual differences on language retention.

Chapter 3: Retention of L2 speech perception

3.1 Introduction

In the following sections, retention of phonological and lexical skills in speech perception will be investigated. Participants are English-French bilinguals who have just returned from a study-abroad semester in France. The bilinguals' performance is compared to two control groups: native French speakers and native English speakers with no knowledge of French. Bilinguals are tested three times within the first nine months after their return from France to determine to what extent they retained their L2 knowledge. Furthermore, the influence of continued exposure to French (for bilinguals who continue taking French classes upon their return to the United States) on language retention is taken into consideration.

Phonological retention is measured in an AX discrimination task in which participants are asked to determine whether two words or nonwords that are presented to them auditorily are phonologically identical or not. Pairs that are not identical differ in one sound only (e.g., ***d****ommage* versus ***t****ommage*). This task does not require any lexical knowledge, as output phonemes can be retrieved directly from the input (Nozari, Kittredge, Dell, & Schwartz, 2010). This task also aims to determine whether phonological retention affects various types of sounds (consonants versus vowels, contrasts that are similar versus different in the L1) differently.

Lexical retention is measured in an auditory lexical decision task. Participants are asked to determine whether an item played to them is a French word or not. Actual French words are matched with nonwords which are created by changing one sound in the word (e.g., the word *dommage* becomes the nonword *tommage*). This task taps into both the phonological and lexical

processing levels. In addition to examining overall retention of these types of knowledge, this task will also be used to investigate if retention is modulated by the lexical status of a word.

Previous research has suggested that language comprehension seems to be fairly resistant to language attrition, both at the lexical level (e.g., Murtagh, 2003) and at the phonological level (e.g., Weltens, 1989), and that under certain circumstances even improvement can be found despite the absence of further input from a language (e.g., Cohen, 1986; Weltens, 1989). The current study aims to extend previous research by investigating L2 retention versus attrition in speech perception at both the phonological and the lexical level in the same group of participants and across three points in time using a longitudinal design, thus allowing for direct comparison across time and across processing levels.

3.2 Experiment 1: Retention of phonological skills

3.2.1 Introduction

Experiment 1 was designed to explore retention versus attrition of phonological skills in speech perception, namely in an AX discrimination task. Participants were asked to decide whether two French words played to them sounded the same or different. The items in each pair were either two tokens of the same French word, or they were minimal pairs that differed in one critical sound only (a vowel or consonant contrast). Pairs were either word-word or word-nonword pairs. Since this task targets sound perception only, bilinguals, French native speakers, as well as English native speakers (with no knowledge of French) were able to participate. Bilingual participants were tested at three times after their return from their study-abroad program to investigate if their perception of French sounds changed after a reduction (for

bilinguals continuing to take French classes) or absence (for bilinguals not continuing to take French classes) of further exposure to French.

3.2.2 Methods

Participants

All three groups of participants as described above (i.e., the bilinguals, English native speakers, and French native speakers) participated in the AX discrimination task.

Materials

The stimuli in the AX discrimination task consisted of 52 French words and 52 contrast words/nonwords produced by a female native French speaker (see Appendix A for a complete list of the stimuli). The contrast items were created by changing one sound in the word; e.g., *dehors* (= *outside*) was turned into the contrast nonword *tehors*. The stimuli were presented in 52 word-word pairs and 52 word-nonword pairs, where the items in the word-nonword pairs differed by one sound – the critical sound. The word-word pairs consisted of two distinct tokens of the target word. The participants never saw the words in written form, but were only exposed to them auditorily. Note that whether participants were familiar with the word meanings was not directly relevant to this task; they were only asked if they detected a sound difference between the two stimuli in each pair.

The critical sounds consisted of five vowel pairs and two consonant pairs. Each sound pair occurred in two word-nonword pairs so that each of the two sounds once occurred in a real French word and once in a nonword. Example words for each sound pair are shown in Table 1.

Contrast	Example pairs					
Vowels	n of word-word pairs	n of word-nonword pairs	Status	Target word	Transcription	Translation
u – y	y = 4	y – u = 4	word	dur	[dy:ʁ]	hard
	u = 3	u – y = 4	contrast nonword	dour	[du:ʁ]	nonword
œ – y	œ = 4	œ – y = 4	word	sœur	[sœ:ʁ]	sister
	y = 4		contrast word	sûr	[sy:ʁ]	sure
e – ɛ	e = 4	e – ɛ = 4	word	perdre	[pɛʁdʁ]	lose
	ɛ = 4	ɛ – e = 4	contrast nonword	pérdre	[peʁdʁ]	nonword
ɛ – ɛ̃	ɛ = 4	ɛ – ɛ̃ = 4	word	bain	[bɛ̃]	bath
	ɛ̃ = 4	ɛ̃ – ɛ = 4	contrast nonword	baie	[bɛ]	bay
o – õ	o = 4	o – õ = 4	word	peau	[po]	skin
	õ = 4	õ – o = 4	contrast word	pont	[põ]	bridge
Consonants						
d – t	d = 4	d – t = 4	word	dommage	[dɔma:ʒ]	adversity
	t = 4	t – d = 4	contrast nonword	tommage	[tɔma:ʒ]	nonword
n – ɳ	n = 4	n – ɳ = 4	word	jaune	[ʒo:n]	yellow
	ɳ = 4	ɳ – n = 4	contrast nonword	jaugne	[ʒo:ɳ]	nonword

Table 1. Critical sounds and example target words and nonwords for Experiment 1.

Procedure

Participants were seated in front of a computer in a sound-attenuated booth. They were played the critical sounds over headphones at a comfortable level. Visual and auditory stimuli

were presented to the participants and their responses were recorded using Superlab 4 software (Cedrus Corporation, 2006). Instructions for the task were presented in writing on the computer screen and were also read aloud to the participants by the experimenter. Because of the low French proficiency level of some bilingual participants all instructions were given in English and visual prompts during the experiment (asking the participants to respond to each stimulus pair) were presented in English as well.

After having been presented with visual and oral instructions for the experiment, participants completed four practice trials on stimulus pairs that did not appear in the actual experiment. After the practice trials, participants were told to ask any questions they still had about the experiment, and if they had no more questions to press any key on the keyboard to start the experiment.

Each experimental trial started with the appearance of a cross-hair in the center of the computer screen. Simultaneously, participants were played a French stimulus pair (word-word or word-nonword). Immediately after the stimulus pair a prompt ("Same or different?") appeared on the computer screen to which participants responded by clicking the appropriate key on the keyboard. The keyboard resolution was 4 msec. Participants responded by clicking color-coded keys on the computer keyboard. The "4" key on the number pad of the keyboard was covered with a red sticker, and the "+" key with a blue sticker. Participants were asked to click the red key if they thought the speaker on the recording had said the same word twice, and to click the blue key if they thought the speaker had said two different words. Upon click of one of the two keys the experiment automatically advanced to the next trial. Experimental trials were played in random order for each participant. The AX discrimination task lasted about 30 minutes.

Coding

The Superlab software automatically recorded participants' responses ('same' or 'different') and their reaction times. Responses were subsequently labeled as 'correct-same', 'incorrect-same', 'correct-different', and 'incorrect-different', which made it possible to calculate each participant's hit and false-alarm rates for the calculation of participants' d' scores.

d' is a measure that attempts to quantify a person's sensitivity for detecting a signal while taking into account their response bias. In the current AX discrimination experiment, for example, the signal was always the target sound which had to be detected against the 'noise' of the contrast sound. Since the 'different' trials were the trials of interest in this study (i.e., are participants able to detect a difference between the contrasting sounds?) one could have looked at the accuracy rate on the 'different' trials only. However, this would not have accurately represented the participants' performance. People have different criteria of how conservative they are with their responses. If one participant is very conservative and only responds with 'different' if they are absolutely certain the two sounds are different, whereas another participant who is much less conservative will say a sound pair is 'different' even if they are not entirely sure, the latter participant will most likely display a higher accuracy rate for 'different' trials, but they will also have a higher false alarm rate, i.e., they will respond 'different' to more 'same' trials than the first participant. d' takes into account these different response biases.

The d' measure utilized here is calculated by taking the hit rate H (i.e., the proportion of 'different' trials participants called 'different') and the false alarm rate F (i.e., the proportion of 'same' trials participants called 'different'), z-transforming them and subtracting the transformed false alarm rate from the transformed hit rate: $d' = z(H) - z(F)$ (Macmillan & Creelman, 2005)

The reaction times were recorded in milliseconds. Since reaction times between participants and test sessions differed by more than an order of magnitude, they were log-transformed and the log-transformed scores were used in the RT analyses reported below. All trials with reaction times shorter than 5 milliseconds were excluded from the d' and reaction time analyses (Baayen & Milin, 2010).

Orthogonal contrast coding was used to represent categorical variables in the regressions. The advantage of contrast coding over dummy coding (where the levels of an analysis are assigned values of 0 and 1) is that contrast coding allows the researcher to investigate differences between levels (e.g., consonant versus vowel) whereas in dummy coding one can only make single comparisons of each level to a baseline which results in the need to run multiple regressions to look for differences between levels (cf. Davis, 2010).

Analysis

Statistical analyses of the data were performed using maximal linear mixed effects models (lmer; Baayen, 2008) with by-subject random slopes for Time and Sound Status for the d' and RT analyses, and additionally, by-item random slopes for Bilingualism, Time, Continued French, and Sound Status for the RT analyses. Although for this analysis method, it is preferable to also include by-subject and by-item slopes for the interactions of the variables, none of the interactions converged and so had to be dropped from the models. These lmers were performed on the d' scores as well as the log-transformed reaction times (RTs). To determine which main effects and interactions contributed significantly to the model fit (i.e., had a significant effect on d' scores and reaction times), each main effect and each interaction was removed from the fixed effects structure of the maximal model one at a time, and the model lacking the effect of interest

was re-fit. The relative fit of the reduced vs. maximal model was compared using likelihood ratio tests. If the maximal model was improved by the inclusion of the main effect or interaction removed in the reduced model, this was interpreted as the main effect or the interaction making a significant contribution to variance.

For each component of the model (main effect of interaction), the estimate (*Est.*) and standard error (*SE*) of the maximal model are reported, as well as the result of the likelihood ratio test for the corresponding reduced model (X^2(df) with the associated p-value). Analyses were performed on the following independent variables:

- *Bilingualism* (French native speaker, Bilingual, English native speaker)
- *Time* (Time 0, Time 1, Time 2); Time 0 corresponds to the first experimental session that took place about two months after the bilinguals' return from France. Time 0 performance provides the baseline to which any attrition or improvement occurring between then and Time 1 (second experimental session, about 3 months after the first session) and Time 2 (third experimental session, about 7 months after the first session) is compared.
- *Continued French* (Continued, Not Continued); the nine bilinguals who continued taking French classes and the eight bilinguals who were no longer exposed to French after their return from France will be looked at separately in terms of their performance across time.
- *Sound Status* (Vowel, Consonant)
- *PAM Contrast* (Category Goodness, Two Category); category goodness assimilation contrasts are predicted to be more difficult for bilinguals and native English speakers to discriminate than two category assimilation contrasts. Since for French native

speakers both contrast types represent phonemic contrasts, no such difference is expected for them.

3.2.3 Results

Baseline performance at Time 0

d':

There was a significant main effect of Bilingualism for bilinguals versus French native speakers (*Est.* = 1.11, *SE* = 0.35, $X^2(1) = 8.88$, $p < 0.01$) and a marginally significant effect for bilinguals versus English native speakers (*Est.* = -0.54, *SE* = 0.28, $X^2(1) = 3.73$, $p = 0.05$). The French native speakers were better at discriminating French stimulus pairs than were the bilinguals, while the bilinguals in turn were better than the native English speakers. However, the bilinguals who were still taking French classes and those that were no longer exposed to French did not differ significantly at Time 0 ($X^2(1) = 0.19$, $p > 0.05$).

There was a significant main effect of Sound Status (vowel versus consonant; *Est.* = 0.60, *SE* = 0.20, $X^2(1) = 7.99$, $p < 0.01$) with participants being more sensitive to French consonant contrasts than vowel contrasts. In addition, the interaction of Sound Status and Bilingualism was significant for bilinguals versus French native speakers (*Est.* = 1.70, *SE* = 0.48, $X^2(1) = 10.72$, $p < 0.01$). The advantage of discriminating consonant contrasts versus vowel contrasts was bigger for French native speakers than for bilinguals, who in fact treated vowel and consonant contrasts similarly. Furthermore, bilinguals did not differ from English native speakers with regards to their discrimination of vowel versus consonant contrasts; the corresponding interaction failed to reach significance ($X^2(1) = 0.58$, $p > 0.1$; see Figure 1).

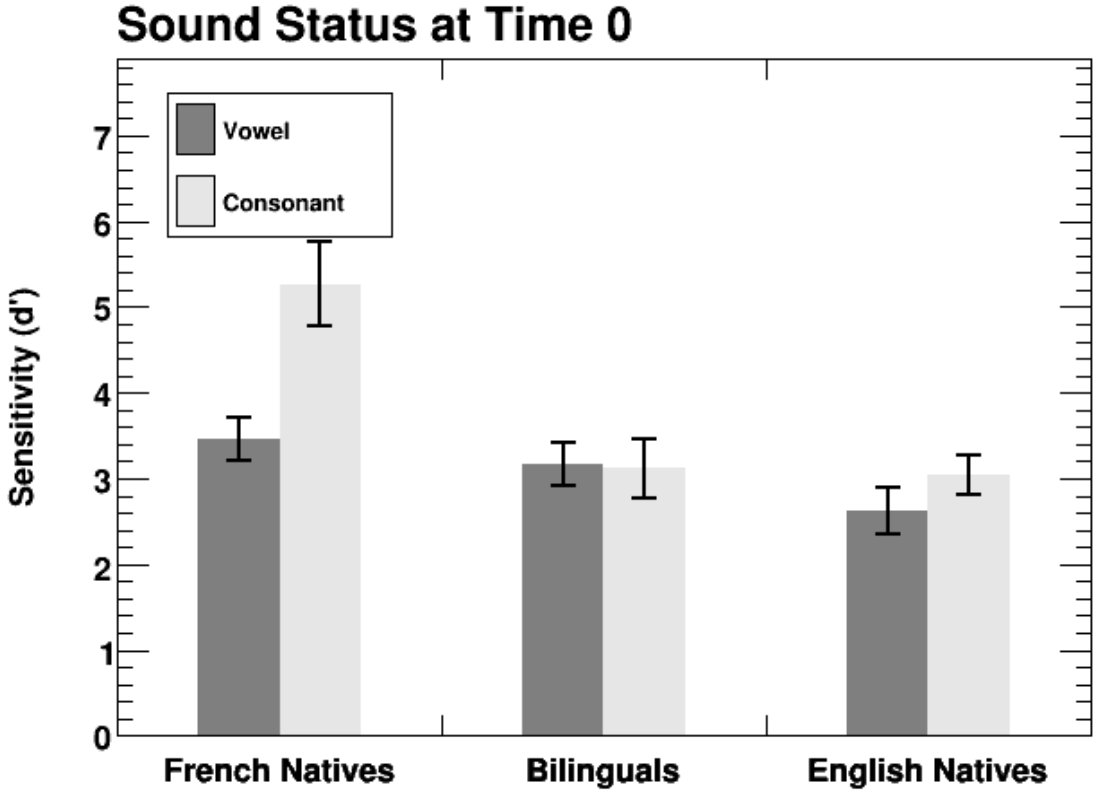

Figure 1. Interaction of Sound Status and Bilingualism. Mean d' for each participant group at Time 0, separated by sound type. Error bars denote standard error.

Finally, there was a significant main effect of PAM Contrast at Time 0 (*Est.* = 0.65, *SE* = 0.22, $X^2(1) = 8.31$, $p < 0.01$). In addition, there was a marginally significant interaction of Bilingualism and PAM Contrast for bilinguals versus native French speakers (*Est.* = 0.73, *SE* = 0.43, $X^2(1) = 2.94$, $p = 0.09$).While French native speakers were more sensitive to category goodness assimilation contrasts (e.g., /u-y/), vs. two category assimilation contrasts (e.g., /t-d/), bilinguals showed no difference in sensitivity between the two contrast types. The interaction of PAM Contrast and Bilingualism for bilinguals versus native English speakers was also

marginally significant (*Est.* = -0.74, *SE* = 0.44, $X^2(1) = 2.85$, $p = 0.09$). Native English speakers showed a nonsignificant trend in the direction opposite that of the French speakers, i.e., a trend towards being more sensitive to two category assimilation contrasts (see Figure 2).

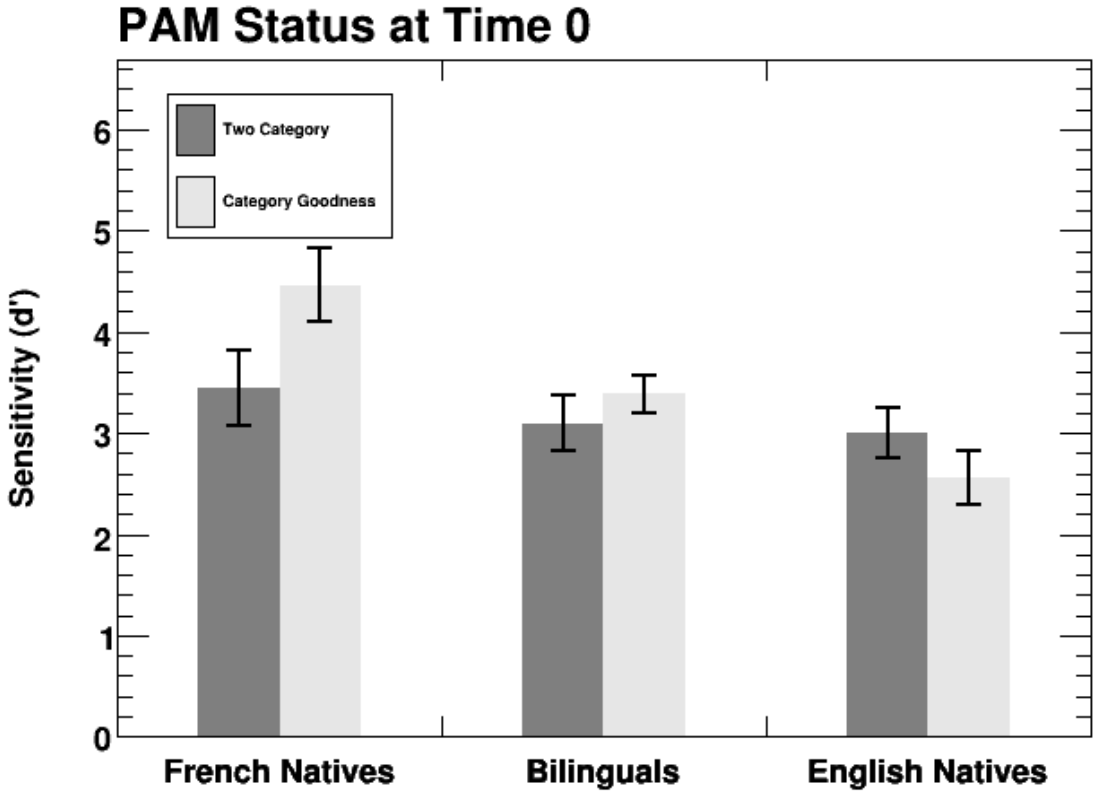

Figure 2. Interaction of PAM Contrast and Bilingualism. Mean d' for each participant group at Time 0, separated by PAM Contrast. Error bars denote standard error.

Reaction time:

There was no main effect of Bilingualism, i.e., the control groups responded equally fast as the bilinguals (bilinguals vs. French natives: $X^2(1) = 1.87$, $p > 0.1$; bilinguals vs. English natives: $X^2(1) = 0.06$, $p > 0.1$).

Furthermore, there were no effects of PAM Contrast for RTs for any of the analyses conducted for Experiment 1 ($X^2(1)$s < 0.02, $p > 0.1$). The distinction between two category assimilation contrasts and category goodness assimilation contrasts did not seem to affect participants' reaction times.

Finally, consonant and vowel reaction times cannot be compared directly, since the vowels and consonants didn't appear in the same places in the stimulus words. All vowels appeared as the second segment in each word (e.g., /u/ in 't**ou**sser'). In contrast, the consonants were either the first segment in the word in the case of /d/ and /t/ (e.g., /d/ in '**d**ommage') or the ultimate or penultimate segment in monosyllabic words, i.e., the third or fourth segment in the word, in the case of /ɳ/ (e.g., /ɳ/ in 'gro**gn**e'). RTs, however, are by necessity influenced by linear position in the word. For example, a discrimination of the initial sounds in *dommage* vs. *tommage* can be performed earlier than a discrimination of the final sounds in *gronne* vs. *grogne*. Therefore, RT data for vowels versus consonants at Time 0 are not reported here.

Changes across the first five months (Time 0 versus Time 1)

French native speakers only participated in one experimental session, while English native speakers participated in two sessions. Time difference analyses for Time 0 versus Time 1 were therefore only performed for the bilinguals and the English native speakers.

a) Practice Effects: Bilinguals versus English native speakers

d':

There was a significant main effect of Bilingualism (*Est.* = -0.54, *SE* = 0.26, $X^2(1) = 4.26$, $p < 0.05$). Bilinguals were better at discriminating French contrasts than English monolinguals (see Figure 3).

However, there was no main effect of Time, i.e., the performance of neither the bilinguals nor the English native speakers changed from Time 0 to Time 1 ($X^2(1) = 0.06$, $p > 0.1$). Neither was there an interaction of Time and Bilingualism ($X^2(1) = 0.38$, $p > 0.1$). Furthermore, there was no main effect of Sound Status ($X^2(1) = 2.11$, $p > 0.1$) nor a significant interaction of Bilingualism and Sound Status ($X^2(1) = 0.49$, $p > 0.1$). Bilinguals and English native speakers were equally sensitive to vowel and consonant contrasts at both times.

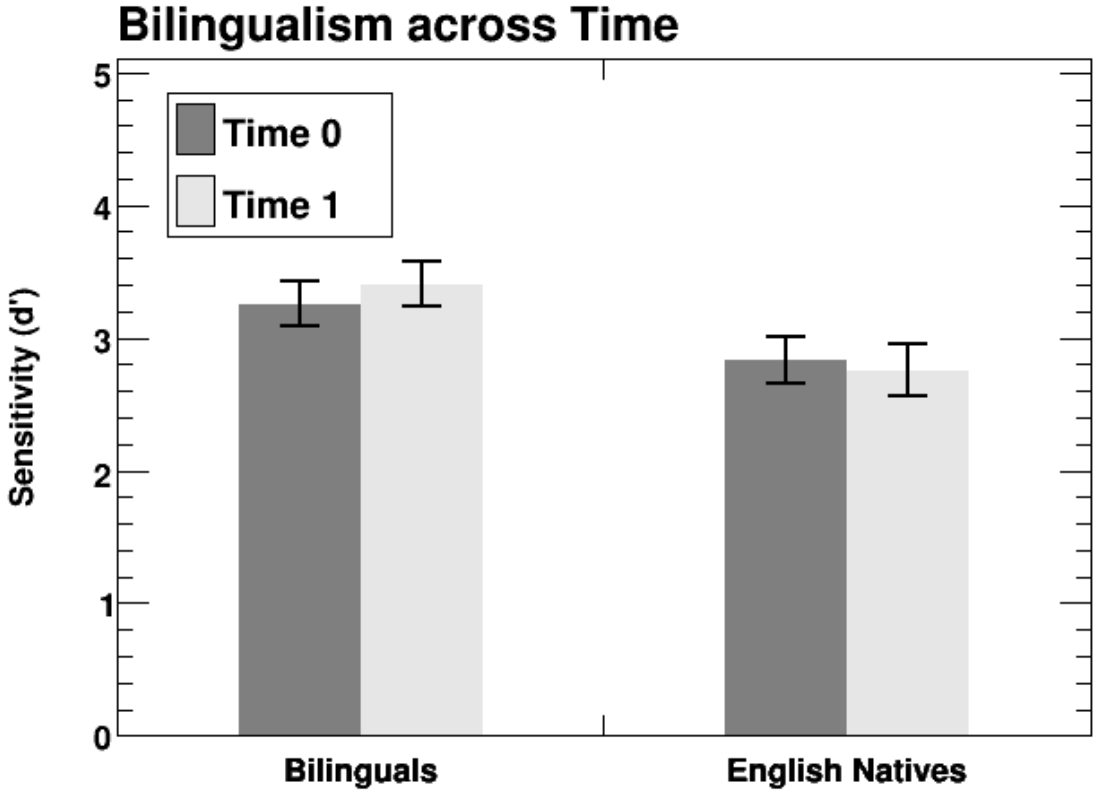

Figure 3. Main effect of Bilingualism. Mean d' for each participant group at Time 0 and Time 1. Error bars denote standard error.

There were no significant main effect of PAM Contrast and no significant interactions of PAM Contrast and Time or Bilingualism ($X^2(1)$s < 2.70, p > 0.1). While a marginally significant interaction of PAM Contrast and Bilingualism was found at Time 0 for bilinguals versus native English speakers, the lack of any effects in the analysis of Time 0 versus Time 1 suggest that the marginal interaction at Time 0 may in fact not represent actual processing differences between the two groups of participants, but may be caused by small, but importantly nonsignificant, trends in opposite directions for the bilinguals and native English speakers at Time 0 that

disappeared when more data was added to the analysis in the form of Time 1 data. This claim is further strengthened by two follow-up analyses looking at each group of participants (bilinguals, native English speakers) separately at Time 0. No significant effect of PAM Contrast was found for either group (Bilinguals: $X^2(1) = 2.55$, $p > 0.1$; native English speakers: $X^2(1) = 0.77$, $p > 0.1$).

Reaction Time:

There was a significant main effect of Time (*Est.* = -0.14, *SE* = 0.07, $X^2(1) = 4.35$, $p < 0.05$), with participants being faster at Time 1 than at Time 0, and a significant interaction of Time and Bilingualism (*Est.* = -0.14, *SE* = 0.05, $X^2(1) = 5.89$, $p < 0.05$). Bilinguals improved their reaction times from Time 0 to Time 1 while the English native speakers did not (see Figure 4). The fact that bilinguals responded faster at Time 1 than at Time 0, whereas English native speakers did not improve their reaction times suggests that the shorter RTs at Time 1 for the bilinguals are not simply due to practice effects (i.e., participants are faster at Time 1 because they are familiar with the task from Time 0). If this were the case, we would expect English native speakers to show the same effect. Rather, this result suggests that the bilinguals' processing of French has improved between Time 0 and Time 1.

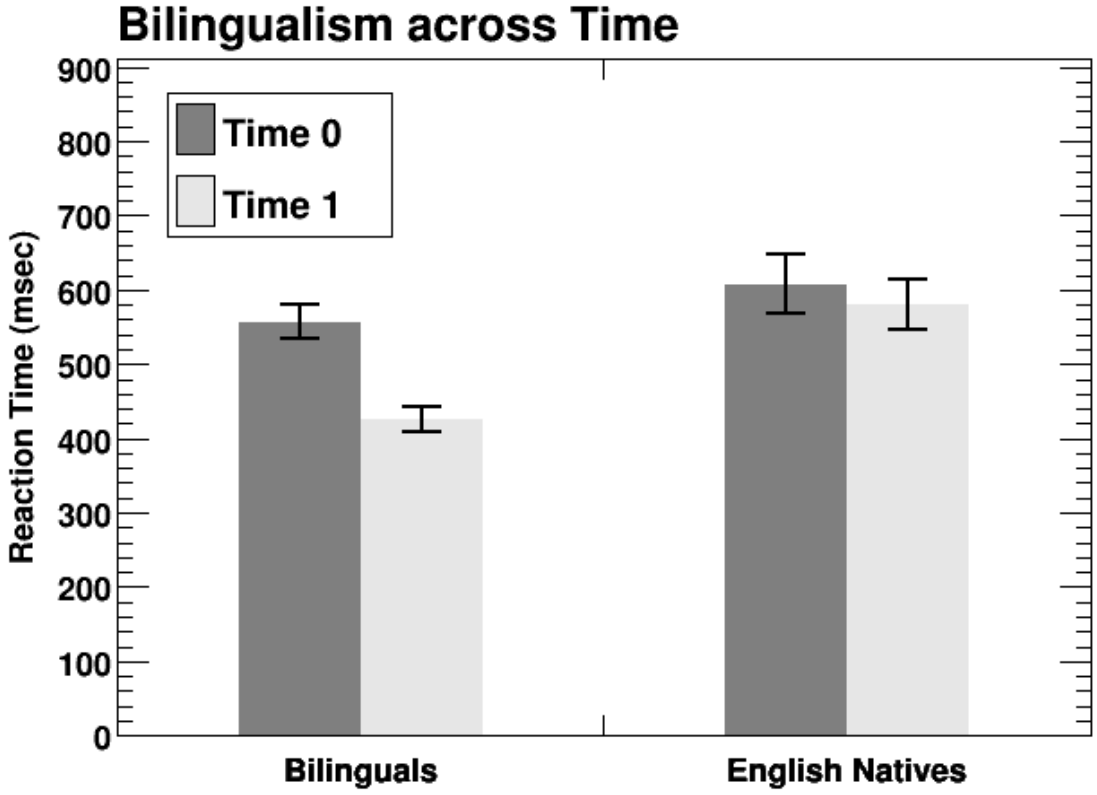

Figure 4. Interaction of Bilingualism and Time. Mean reaction times for each participant group at Time 0 and Time 1. Error bars denote standard error.

There were no significant interactions of Sound Status and Bilingualism ($X^2(1) = 0.43$, $p > 0.1$) and Sound Status, Bilingualism and Time ($X^2(1) = 0.21$, $p > 0.1$), suggesting that the bilinguals and the native English speakers did not treat consonants and vowels differently from each other, and that there was no change across time for either group of participants.

One could argue that the improvement in reaction times across time could simply reflect the fact that some of the bilinguals continued taking French classes after their return. To determine if this is indeed the case, the following sections examine whether there are systematic

differences across participants who did continue taking French classes versus those who did not continue taking French.

b) Effect of Continued French study: Bilinguals who continued studying French versus bilinguals who did not

d':

There was a significant interaction of Continued French and Time (*Est.* = 0.97, *SE* = 0.40, $X^2(1) = 5.74$, $p < 0.05$; see Figure 5) with the bilinguals who continued with French improving their sensitivity to French contrasts from Time 0 to Time 1, while the bilinguals who were no longer exposed to French did not show any changes from Time 0 to Time 1. It is noteworthy, that while it is not surprising that only the bilinguals who were still exposed to French improved in this task, the bilinguals who were no longer exposed to French did not lose their ability to discriminate French sound pairs, either, in the first five months after their return from France. Rather, their discrimination skills showed no appreciable decline despite the absence of input from French.

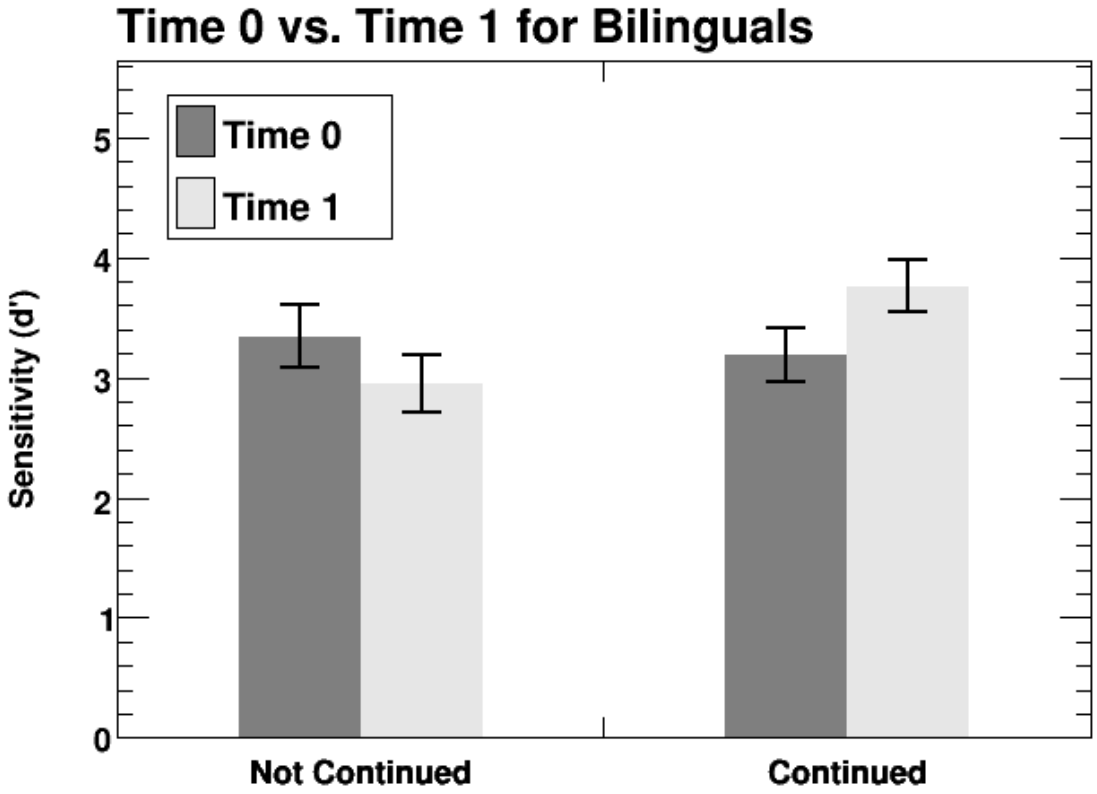

Figure 5. Interaction of Continued French and Time. Mean d' for each participant group at Time 0 and Time 1. Error bars denote standard error.

There were no significant main effect or interactions of Sound Status ($X^2(1)$s < 0.42, p > 0.1), i.e., participants did not differentiate between vowels and consonants. Finally, there were no significant main effect or interactions of PAM Contrast ($X^2(1)$s < 0.22, p > 0.1), either. Bilinguals did not differentiate between two category assimilation contrasts and category goodness assimilation contrasts.

Reaction Time:

There were no significant main effect of Continued French and no significant interactions ($X^2(1)$s < 2.18, $p > 0.1$), suggesting that the improvement of RTs across time described above holds true for all bilinguals, regardless of whether they were still exposed to French or not. In other words, in spite of reduced input – or even no input – from the L2, the bilinguals got faster at discriminating French sound contrasts in the first five months after their return from France.

Changes across the first nine months (Time 0 versus Time 2)

Only bilinguals participated at Time 2. Therefore, no analyses of Bilingualism were performed for Time 0 versus Time 1. Rather, the analyses focused on differences between bilinguals who continued taking French after their study-abroad program and those who did not.

d':

There was a significant main effect of Time (*Est.* = 0.48, *SE* = 0.24, $X^2(1) = 4.12$, $p < 0.05$). The bilinguals were more sensitive to French sound contrasts at Time 2 than at Time 0. There was also a significant main effect of Sound Status (*Est.* = 0.40, *SE* = 0.19, $X^2(1) = 4.40$, $p < 0.05$), indicating that the bilinguals were better at discriminating consonant contrasts than vowel contrasts. The interaction of Sound Status and Time was marginally significant (*Est.* = 0.60, *SE* = 0.34, $X^2(1) = 3.36$, $p = 0.07$), showing that at Time 0, the bilinguals did not discriminate between consonants and vowels, while at Time 2, bilinguals were better at the consonant contrasts than the vowel contrasts (see Figure 6).

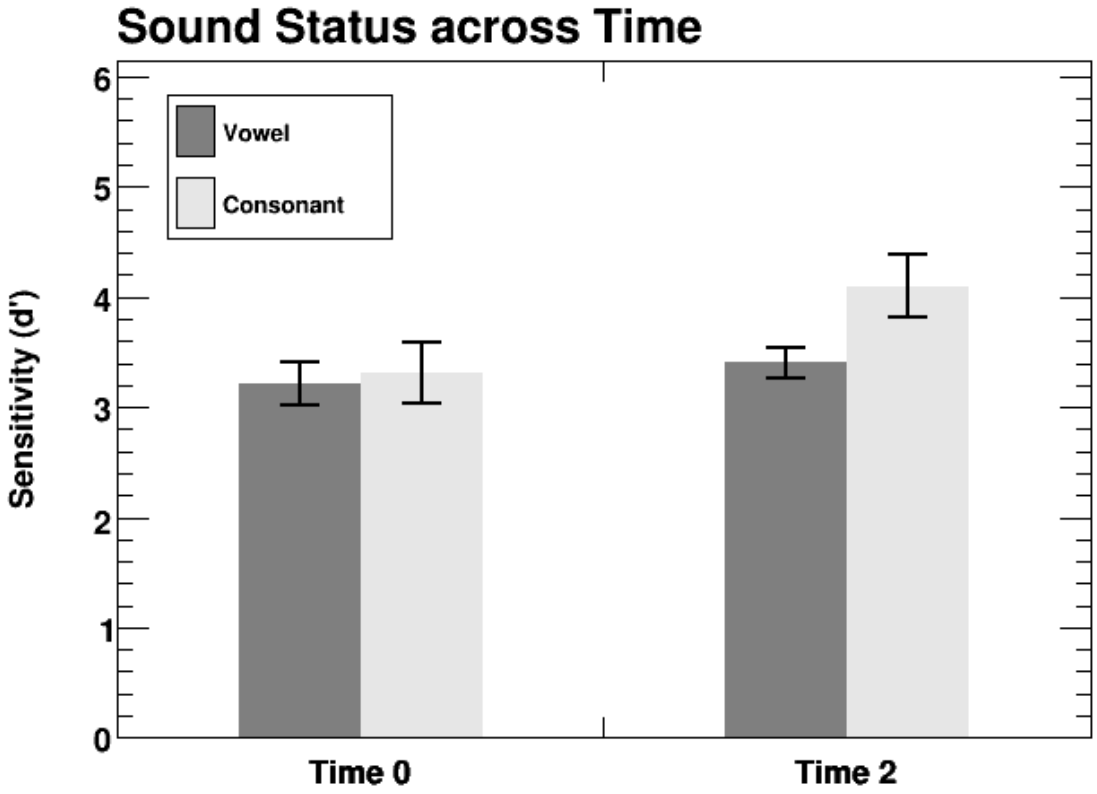

Figure 6. Interaction of Sound Status and Time. Mean d' for bilinguals at Time 0 and Time 2, separated by Cognate Status. Error bars denote standard error.

However, there was no main effect of Continued French ($X^2(1) = 0.41$, $p > 0.1$) and no interaction of Continued French and Time ($X^2(1) = 0.03$, $p > 0.1$), suggesting that both groups of bilinguals performed similarly at both times, regardless of whether they continued taking French classes or not. This suggests that the effects of L2 immersion persist across a fair amount of time (i.e., at least nine months past immersion) even in the absence of continued exposure to the L2.

Finally, there were no significant main effect and no significant interactions of PAM Contrast for Time 0 versus Time 2 ($X^2(1)$s < 0.48, $p > 0.1$). The bilinguals were equally sensitive to both contrast types at both times.

Reaction Time:

There was a significant main effect of Time (*Est.* = -0.35, *SE* = 0.07, $X^2(1) = 15.92$, $p < 0.0001$). Bilinguals were faster overall at Time 2 than at Time 0 (see Figure 7). However, there was no main effect of Continued French ($X^2(1) = 0.01$, $p > 0.1$) and no significant interaction of Continued French and Time ($X^2(1) = 0.01$, $p > 0.1$), i.e., both groups of bilinguals were faster at Time 2 than Time 0, regardless of whether they continued studying French or not. Finally, there were no effects of Sound Status and PAM Contrast for Time 0 versus Time 2 ($X^2(1)$s < 0.88, $p > 0.1$), suggesting that improvement in RT was not modulated by Sound Status or PAM Contrast.

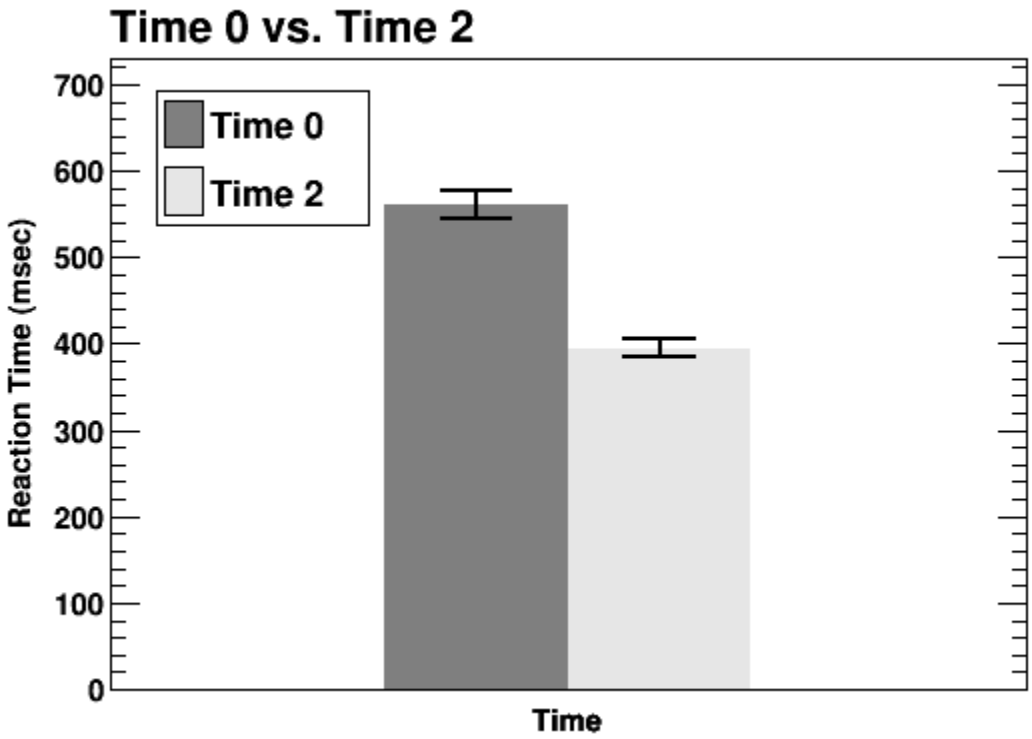

Figure 7. Main effect of Time. Mean reaction times for bilinguals at Time 0 and Time 2. Error bars denote standard error.

3.2.4 Discussion

Experiment 1 tested whether English-French bilinguals who had recently returned from a study-abroad program in France differed from native French speakers in an AX discrimination task in French. Furthermore, it was tested whether any changes in the bilinguals' proficiency occurred across the first five to nine months after the bilinguals' return from France and whether these changes differed for bilinguals who were still taking formal French classes and those who were not taking classes anymore. Finally, it was investigated whether sound status (consonant versus vowel) and contrast type (two category assimilation contrast versus category goodness

assimilation contrast) of certain target sounds in the experimental words influenced attrition, retention, and improvement.

Effects of bilingualism

This experiment set out to examine whether bilinguals who have recently returned to their L1 environment from an L2-immersion program experience attrition of phonological perception skills in the L2. To answer this question a group of L1 English – L2 French speaking bilinguals performed an AX discrimination task of French minimal pairs at three times upon their return home from a 4-month study abroad program in France. Their performance was compared to a native English speaking control group and a native French speaking control group. Furthermore, it was investigated whether L2 retention versus attrition in the ability to discriminate L2 sound contrasts was influenced by sound status of the target sounds (vowel vs. consonant) or by PAM contrast type (two category assimilation contrast vs. category goodness assimilation contrast).

Participants were first tested about two months after the bilinguals' return from France to get a baseline performance level to which the further two tests, which took place three and seven months later, respectively, were compared.

The results from this test session at Time 0 are in line with previous research (Best & Tyler, 2007) in that the bilinguals, as proficient L2 speakers, were better at discriminating L2 sound contrasts than naïve listeners (i.e., the native English speakers). However, even proficient bilinguals may not become completely native-like during an immersion period, as was shown by the fact that the native French speakers were significantly more sensitive to the French sound

contrasts than the bilinguals. While the native French speakers were more sensitive to the French sound contrasts, their response times were no faster than the bilinguals' or the native English speakers.

Changes across time

The main question of interest in this section was whether there was any change in discrimination performance across time for the bilingual participants. The three possibilities here were that the bilinguals would lose some discrimination ability in the L2, that they would retain their skills across time, or that they would improve across time (especially if they were still exposed to French).

While intuitively, one may expect L2 speakers' abilities in L2 discrimination to decline after the end of an immersion period, the bilingual participants in the current study actually improved their AX discrimination skills after their return from France. Specifically, when tested five months after their return (i.e., at Time 1), the bilinguals – but not the monolinguals – had improved their reaction times in the task. They were not more sensitive to the contrasts, but they required less processing time to perform at a similar accuracy level as at Time 0. While one could argue that participants may have been faster at Time 1 because they were already familiar with the task, this seems unlikely to be the reason for the improvement in the bilinguals. We did not see the same improvement in the native English speakers, who had an equivalent amount of experience with the specific experimental task.

The discrimination data also suggest that bilingual participants retained L2 knowledge after their return from France. While the bilinguals' sensitivity to the French sound contrasts had

not improved since Time 0, it, importantly, also had not declined, which suggests that their perceptual skills remained robust across three months with limited or no exposure to French.

Effects of continued exposure to French

As was mentioned earlier, about half of the bilinguals (nine out of seventeen) who participated in this study kept taking formal French classes after the end of their study-abroad program. Because it would be more likely for these bilinguals to show improvement in French, since they still received L2 input – albeit much less than during their stay in France – it is important to compare bilinguals who continued taking French classes to those who were no longer exposed to French. The latter group may be more likely to show attrition, or at least would be less likely to show L2 improvements such as the ones just described. In the next section, therefore differences between the two groups of bilinguals will be discussed.

In terms of sensitivity to French sound contrasts, at Time 0 the two groups of bilinguals performed similarly, but a significant interaction of Continued French and Time was found at Time 1. The bilinguals who continued taking French classes achieved higher d' scores at Time 1 than at Time 0, while the bilinguals who were no longer exposed to French performed similarly at Time 0 and Time 1. However, by Time 2, both bilingual groups showed improvements in sensitivity relative to Time 0.

It is not surprising that the participants who still took French classes improved after their return from France. After all, they were still exposed to French regularly and received further training in the language. Importantly, the amount of exposure they received, which was much less than during the immersion period during their study-abroad program, was enough for their

perceptual skills to improve significantly already during the first five months after their return (Time 1), and to still show improvements after nine months (Time 2), relative to their performance two months after their return (Time 0).

It has to be noted here that even the continuing bilinguals experienced very limited exposure to French between Time 1 and Time 2, since a three-month summer break from college took place during those four months. Therefore, even in the almost complete absence of L2 input (no bilingual was exposed to French for more than three weeks during the summer break) the bilinguals showed improvement by Time 2.

While the non-continuing bilinguals did not improve from Time 0 to Time 1, it is worth noting that at Time 1 they were no less sensitive to French sound contrasts than at Time 0, which we would expect had attrition taken place. In other words, no signs of attrition of L2 perceptual skills were noticeable despite the absence of input from the language. Even more striking, however, is that even the bilinguals who were no longer exposed to French exhibited greater sensitivity to French sound contrasts nine months after the end of the immersion period relative to two months after the end. Similarly to the continuing bilinguals during the college break, the non-continuing bilinguals improved in spite of the absence of L2 input.

The fact that bilinguals who were no longer exposed to French performed similarly at Time 0 and Time 1, but actually improved by Time 2 raises an interesting question about the reason for this 'delay' of an effect. In fact, it seems like there may be a plateau phase after the end of L2 immersion during which perceptual skills remain stable, and which may last anywhere from five to nine months, and after which further improvement may show. Cohen (1986) first

posited such a “settling in process” when he found improvement in a student’s L2 vocabulary after summer recess.

The current study examined reaction times in addition to sensitivity measures. Both groups of bilinguals were faster at discriminating French sound contrasts both at Time 1 and at Time 2 than at Time 0. As discussed in the previous section, the absence of this decrease in RTs in the native English speaking participants suggest that the bilinguals’ performance is not simply a result of familiarity with the task. Rather, it seems to be caused by further L2 processing that had taken place between Time 0 and Times 1 and 2.

Furthermore, note that even the bilinguals who were no longer exposed to French were faster at Time 1 than at Time 0 – even though their sensitivity to the contrasts had not improved. This improvement in performance after L2 immersion suggests that some processing may still take place. This can result in improvements which can be measured earlier for some measures (such as RT), but which take longer to show noticeable improvements for other measures (such as sensitivity).

Researchers often use accuracy and reaction times to study bilingual language processing. Frequently, RTs have been found to be a more sensitive measure of the influence of bilingualism on certain tasks than accuracy, i.e., an effect on RT may be found even in the absence of an effect on accuracy (e.g., Martin-Rhee & Bialystok, 2008, in a bilingual Simon task; Schwartz & Kroll, 2006, in bilingual sentence reading). It makes sense, therefore, that if subtle improvements in L2 processing are going on that they would first show up in participants’ reaction times and only later be measurably different in d'.

In any case, the current data suggest that further perception processing takes place even after the end of an L2 immersion period during the first five to nine months after the end of immersion. Noticeable improvements from this period of further processing may be found earlier if there is continued exposure to L2 input, as in the case of the bilinguals taking French classes. If there is no more input, however, some further processing still seems to take place but show noticeable improvements only after a longer period of time and at different lengths of time for different measures.

Effects of sound status

The analyses of Sound Status (vowel vs. consonant) suggest that native French speakers, bilinguals, and native English speakers are not equally sensitive to vowel versus consonant contrasts. Namely, while all three groups of listeners were about equally sensitive to sound contrasts involving vowel contrasts, the native French speakers were significantly more sensitive to consonant contrasts than both the bilinguals and the native English speakers.

Furthermore, vowels and consonants were affected differently by changes across time. While direct comparisons of reaction times to vowels and consonants are not meaningful because of the occurrence of the two sound types in different places in the target words, some interesting changes across time for vowel vs. consonant RTs were found. First of all, the bilinguals improved their reaction times to consonants but not vowels across the first five months after their return from France. Furthermore, by nine months after the end of the immersion period the bilinguals' sensitivity to consonant contrasts had improved relative to vowel contrasts. It seems, therefore, that the consonant contrasts had actually become more salient for bilinguals within

those first five to nine months after their return from France, despite the limitation, or even absence, of L2 input in the months preceding each experiment.

Differences in the processing of vowels and consonants are a well-known phenomenon. Previous research has shown that listeners have shorter reaction times for consonants than vowels in sound identification tasks (Studdert-Kennedy, Liberman, & Stevens, 1963) and phoneme monitoring tasks (Cutler & Otake, 1994). Differences in vowel versus consonant processing have also been found in ERP (event related potential) studies, where a larger negativity was associated with consonant than vowel processing (Carreiras, Gillon-Dowens, Vergara, & Perea, 2007) and in studies with language-impaired populations (Caramazza, Chialant, Capasso, & Miceli, 2000) where contrasting selective difficulties in producing vowels and consonants have been found in aphasics. Consonant facilitation has, furthermore, been reported by Birdsong (2007), who has found that native-English learners of French are able to become more native-like in their productions of consonants than vowels. The current study extends previous research on consonant facilitation in that the present findings suggest that further processing of French sounds after the end of an immersion period may lead to even greater consonant facilitation over time (at least within the first nine months after the end of an immersion period).

While the current data suggest that the native French speakers and bilinguals experienced some kind of facilitation in consonant processing relative to vowel processing, it would be interesting to know what specifically caused this facilitation with regards to the French consonant and vowel system. One reason for facilitation of consonant processing relative to vowel processing may have been the greater regional variation of vowels compared to

consonants (Hawkins, 1993), which may have made it more difficult for native French speakers to discriminate vowel contrasts than consonant contrasts. For example, while VOT of stop consonants differs somewhat across dialects of French (Caramazza & Yeni-Komshian, 1974), the differences are not so great that a listener would confuse a voiced stop for a voiceless stop. The distinction of open versus closed vowels, on the other hand, which varies regionally, may in some cases change the meaning of a word (e.g., open /ɛ/ as in *tait*, Engl. '(he) keeps silent' vs. closed /e/ as in *thé*, Engl. 'tea'). In addition to regional variation for vowels, vowel quality also varies depending on syllable and consonant context which leads to difficulties in vowel discrimination even in native French speakers (Féry, 2003; Gottfried, 1984). While this regional and context-dependent variability should not directly interfere with the discrimination of the contrast pairs in the current study (e.g., listeners should be able to detect a difference in vowel quality even if a particular realization is different from their native dialect), this variability of vowels in the language may make vowel discrimination generally more difficult for speakers of French.

As for the bilinguals, while they were able to improve their consonant discrimination across the first nine months after their return from France, this general difficulty in vowel discrimination in French may have prohibited them from getting better at discriminating vowel contrasts despite further phonological processing that seemed to have taken place during this time. In fact, since bilinguals and native French speakers performed similarly at Time 0 with regards to vowel discrimination, this suggests that both groups may have already performed at ceiling with regards to discriminating vowel contrasts, and therefore, no further improvement in the bilinguals would have been expected after Time 0.

Effects of contrast type

The current study has also found differences in the processing of different types of sound contrasts (based on contrasts introduced by PAM). In a comparison of two category assimilation contrasts and category goodness assimilation contrasts, varying degrees of sensitivity was found for native French speakers, bilinguals, and native English speakers. Namely, native French speakers were more sensitive to category goodness assimilation contrasts relative to two category assimilation contrasts, whereas the bilinguals were equally sensitive to both, and native English speakers showed no significant difference between the two patterns.

These results are not necessarily what one would expect based on the predictions of the PAM for naïve listeners (native English speakers) and the PAM-L2 for L2 learners (bilinguals). PAM predicts that sound contrasts that map onto similar contrasts in the L1 (i.e., two category assimilation contrasts such as /d-t/) are easy even for naïve listeners to discriminate, whereas contrasts that map onto a single phoneme category in the L1 (e.g., category goodness assimilation contrasts such as /u-y/) are more difficult to discriminate (Best, McRoberts, & Goodell, 2001). While there was a trend for the native English speakers to be more sensitive to two category assimilation contrasts than category goodness assimilation contrasts, this difference was not significant.

PAM-L2 also predicts that as L2 learners become more proficient in the L2, they acquire new phoneme categories for sounds that do not fit into an existing L1 category. This would be the case for one of the two sounds in each category goodness assimilation contrast. Bilinguals may in fact, therefore, be more sensitive to these contrasts than to two category assimilation contrasts after they have reached a certain level of proficiency (the PAM does not specify what

this level would be). One could argue that the bilinguals in the current study may simply not show a difference between the two category types because they have already improved their discrimination of the more difficult contrast (because of the formation of such new phoneme categories). At the same time, they may not yet be proficient enough to be better at discriminating contrasts containing these sounds. However, this would still not explain, why the native English speakers also showed no significant difference between the two contrast types.

Another possibility would be that the contrast types used in the present study were simply easy to discriminate even for naïve listeners and that both the bilinguals and the native English speakers performed at a ceiling level of what would be expected of nonnative listeners. Only the native French speakers were able to perform at a higher level in the discrimination of the more difficult category goodness assimilation contrasts.

With the current data set – which contained only two different two category assimilation contrast types and four category goodness assimilation contrast types – it was not possible to more specifically test the predictions of PAM on the bilinguals versus native English speakers. Future studies may be able to address this with a data set containing a more complete set of PAM contrast types specifically designed to test the predictions of PAM on language retention versus attrition.

3.2.5 Conclusion

The results for d' and RT in this L2 AX discrimination task show that learning that takes place during an L2 immersion period seems to be robust for at least nine months after the end of the immersion period. Rather than losing perceptual skills, the bilinguals in this study even

seemed to have improved their sensitivity to L2 sound contrasts during this period. This was the case even in the absence of continued L2 input. While improvement seems to have taken place even without further L2 exposure, improvement was speeded up by continued formal exposure to the language, even if exposure was much reduced compared to the immersion period (i.e., a few hours per week of class time, compared to several hours a day during the study-abroad stay).

The above results suggest that bilinguals returning from a study-abroad program in France, while further processing takes place, initially may enter a plateau phase where neither improvement nor attrition of L2 sound discrimination skills occur to a noticeable degree, and after five to nine months of further processing the improvement becomes measurable. This plateau phase can be shortened, or avoided altogether, if further exposure to the L2 takes place. Only a few hours of exposure per week seem to be enough to lead to improvements in bilinguals after they complete their study-abroad program.

The current results, furthermore, suggest that sound status (vowel versus consonant) modulates the changes across time, with the bilinguals experiencing consonant facilitation that leads to improvements in consonant processing across time even in the absence of continued exposure to the L2.

AX discrimination does not require participants to have any lexical knowledge about the target language, as has been shown by the native English speakers performing well on this task. In fact, participants could even rely on purely acoustic information to complete this task, and the kind of low-level processing necessary for a discrimination task may be affected by different changes across time than a more processing-heavy task such as lexical decision, where

phonological as well as lexical knowledge are necessary. Experiment 2 tested whether similar patterns of retention and improvement of L2 French perceptual skills can be found if lexical processing is required by participants.

3.3 Experiment 2: Retention of lexical skills

3.3.1 Introduction

While Experiment 1 investigated changes across time in sound perception, in Experiment 2 the focus was on how a reduction, or absence, of input after the bilinguals' study-abroad experience affects their lexical processing skills. In each trial in the lexical decision task participants were presented with a single French word (or nonword) and were asked to determine whether the item was a French word or not. Nonwords differed from real French words in that one sound from a French word was changed. Experiment 2 took place at the same three times as Experiment 1, and the influence of cognate status of target words and the fact whether bilinguals continued taking French classes on the retention of French words was explored. Since this task involved lexical knowledge of French, the native English speakers did not participate in Experiment 2.

3.3.2 Methods

Participants

The same seventeen bilinguals as in Experiment 1 participated in this experiment, as did the eight native French speakers. Since participants who are not familiar with French cannot perform a lexical decision task, the English monolinguals were not included in Experiment 2.

Materials

The stimuli in the lexical decision task consisted of 40 French words and 40 matched nonwords (see Appendix A for complete list of stimuli). There were also 42 word-fillers and 42 nonword-fillers which were not included in the analyses. All nonwords were created by changing one sound in the matched French word. For example, the French word 'panier' (*basket*) was turned into a nonword by exchanging the /p/ sound for a /b/, thus creating the matched nonword 'banier'. 20 of the words were French-English cognates that were similar in sound and meaning across the two languages (e.g., 'bananes' – 'bananas'), and 20 words were French-English noncognates, i.e., the translation equivalent of the French word did not sound similar in English (e.g., 'panier' – 'basket'). The cognates and noncognates were matched for word frequency (cognates $M = 41.50$, $SD = 57.03$; noncognates $M = 52.38$, $SD = 109.37$; $t\ (38) = -0.36$, $p > 0.05$), word-initial segment overlap between matched cognates and noncognates (always at least the first segment, where possible also the second and third; $M = 1.55$, $SD = 0.51$, and syllable length (cognates $M = 1.7$, $SD = 0.66$; noncognates $M = 1.5$, $SD = 0.51$; $t\ (38) = 1.07$, $p > 0.05$).

Words were considered cognates if they, phonologically, shared the same consonant structure in the two languages, with only the vowel being different. Since the vowel inventories in French and English are so different, it was not possible to create a set of cognate stimuli that were matched for consonants and vowels. However, because of the shared consonants, the majority of the cognates were also orthographically identical in the two languages, or only had an additional final 'e' in French (e.g., 'cactus' – 'cactus', 'guitare' – 'guitar'; n = 15). The rest of the words differed in that one letter was different in the two languages ('bananes' – 'bananas', 'tomate' – 'tomato'), the French word had one more letter than the English word ('papier' –

'paper'), two letters were switched when comparing the French and the English word ('lettre' – 'letter'), or the French word had accents on some vowels where the English word did not ('téléphone' – 'telephone'). Note, however, that the participants never saw the stimuli in written form. They were only exposed to them auditorily.

While participants were not explicitly tested about their lexical knowledge of the stimulus words, based on consultation with college-level French instructors as well as vocabulary lists in an elementary French textbook (Valdman, Pons, & Scullen, 2006), it was determined that students who have completed an elementary level French course (i.e., the majority of the bilingual participants in the study) should be familiar with the French stimulus words. Table 2 shows a few example cognates and noncognates.

	Example words				
Lexical Status	Word	Transcription	Translation	Matching nonword	Transcription
Cognate	bananes	[banan]	bananas	pananes	[panan]
	cactus	[kaktys]	cactus	cactousse	[kaktus]
Noncognate	panier	[panje:]	apple	banier	[banje:]
	citron	[sitʁõ:]	lemon	citraux	[sitʁo:]

Table 2. Example words and nonwords for Experiment 2.

Procedure

For the lexical decision task in Experiment 2 participants were seated in front of a computer in a sound-attenuated booth. They were played the target stimuli over headphones at a comfortable level. Visual and auditory stimuli were presented to the participants and their responses were recorded using Superlab 4. Instructions for the task were presented in writing on

the computer screen and were also read aloud to the participants by the experimenter. As in Experiment 1, because of the low French proficiency level of some bilingual participants all instructions were given in English and visual prompts during the experiment (asking the participant to respond to each stimulus) were presented in English as well.

After having been presented with visual and oral instructions for the experiment, participants completed four practice trials on stimuli that did not appear in the actual experiment. After the practice trials, participants were prompted to ask any questions they still had about the experiment, and if they had no more questions to press any key on the keyboard to start the experiment.

Each experimental trial started with the appearance of a cross-hair in the center of the computer screen, and simultaneously participants were played a French word or nonword. Immediately after the stimulus sound a prompt ("Word or not?") appeared on the computer screen to which participants responded by clicking the appropriate key on the keyboard. Participants responded by clicking color-coded keys on the computer keyboard. The "4" key on the number pad of the keyboard was covered with a red sticker, and the "+" key with a blue sticker. Participants were asked to click the blue key if they thought the speaker on the recording had said an actual French word, and to click the red key if the thought the speaker had said an item that is not a real French word. Upon click of one of the two keys the experiment automatically advanced to the next trial. Experimental trials were played in random order for each participant. The lexical decision task lasted about 15 minutes.

Coding

The Superlab software automatically recorded participants' responses ('word' or 'nonword') and their reaction times. Responses were subsequently labeled as 'correct' or 'incorrect' and participants' accuracy rate was calculated from this as the proportion of correct responses out of all responses.

The reaction times were recorded in milliseconds. Reaction times tend to be skewed because there is a definite lower limit but no upper limit, while the logs of reaction times usually resemble normal distributions more closely. The reaction times in the current study were therefore log-transformed and the log-transformed scores were used in the RT-analyses reported below. All reaction times shorter than 5 milliseconds were removed. As in Experiment 1, categorical data were contrast coded.

Analysis

Logistic mixed effects regressions were performed on the accuracy rates and linear mixed effects regressions on the log-transformed RTs (Baayen, 2008), and significance of main effects and interactions was determined through likelihood ratio tests comparing the model fit of the maximal versus reduced models (from which main effects and interactions were removed one by one). The regressions for accuracy and log-RT included by-subject random slopes for Time and Cognate Status and by-item random slopes for Bilingualism, Time, Continued French, and Cognate Status.

For each such model comparison, the estimate (*Est.*) and standard error (*SE*) of the maximal model are reported as well as the chi-square value and degrees of freedom X^2(df), and

the p-value from the chi-squared distribution (*p*) for the model comparison of maximal versus reduced model. Analyses were performed on the following independent variables:

- *Bilingualism* (French native speaker, bilingual, English native speaker)
- *Time* (Time 0, Time 1, Time 2)
- *Continued French* (Continued, Not Continued)
- *Cognate Status* (Cognate, Noncognate)

3.3.3 Results

Baseline performance at Time 0

Accuracy:

There was a significant main effect of Bilingualism (*Est.* = -0.20, *SE* = 0.03, $X^2(1) = 27.52$, $p < 0.0001$), with native French speakers being more accurate at identifying French words than bilinguals, but there was no significant main effect of Cognate Status ($X^2(1) = 0.03$, $p > 0.1$) and no significant interaction of Bilingualism and Cognate Status ($X^2(1) = 0.23$, $p > 0.1$). Neither group of participants treated cognates and noncognates differently (see Figure 8). However, there was a significant main effect of Continued French (*Est.* = 0.09, *SE* = 0.04, $X^2(1) = 5.42$, $p < 0.05$), with the bilinguals who continued taking French classes being more accurate at identifying French words than the bilinguals who were no longer exposed to French.

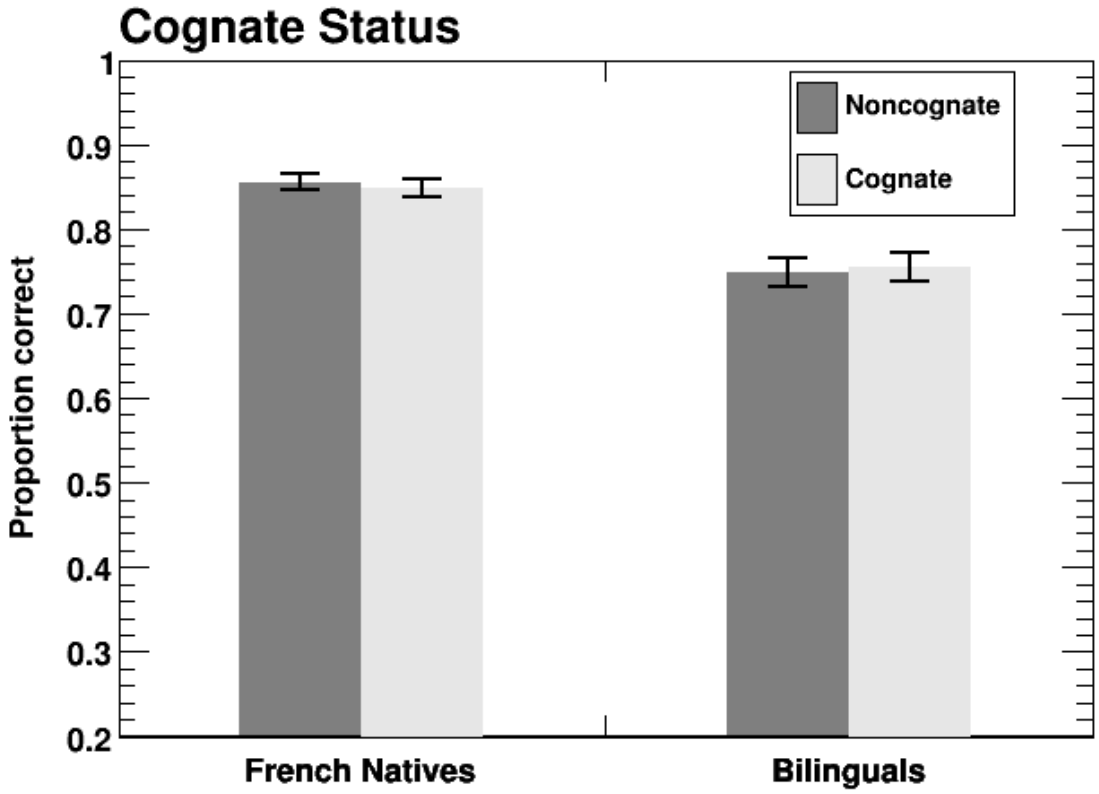

Figure 8. Main effect of Bilingualism. Mean accuracy rate for each participant group at Time 0, separated by Cognate Status. Error bars denote standard error.

Reaction Time:

There were no main effects or interactions of Bilingualism ($X^2(1)$s < 1.57, $p > 0.1$), Continued French ($X^2(1)$s < 0.16, $p > 0.1$), or Cognate Status (*Est.* = -0.11, *SE* = 0.07, $X^2(1)$s < 2.28, $p > 0.1$) at Time 0. Both groups of bilinguals as well as the French native speakers reacted equally fast to stimuli, regardless of whether they were French noncognates or French-English cognates.

Changes across the first five months (Time 0 versus Time 1)

Note that since the native French speakers did not participate in multiple sessions, these analyses were unable to separate out the influence of task practice effects on performance.

Accuracy:

There was a significant main effect of Time (*Est.* = 0.03, *SE* = 0.01, $X^2(1) = 4.34$, $p < 0.05$) suggesting that the bilinguals improved from Time 0 to Time 1. The marginally significant interaction of Continued French and Time (*Est.* = 0.03, *SE* = 0.02, $X^2(1) = 3.04$, $p = 0.08$; see Figure 9) reveals that the bilinguals who continued taking French classes improved from Time 0 to Time 1, while the bilinguals who were no longer exposed to French did not.

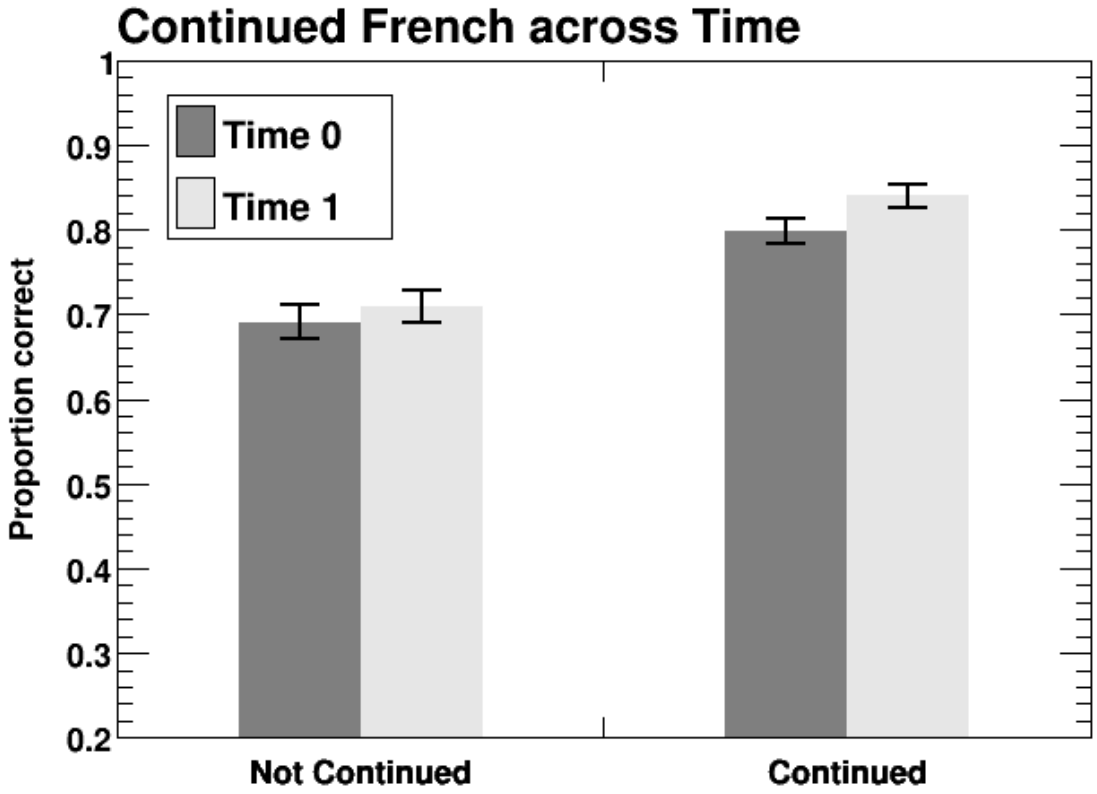

Figure 9. Main effect of Continued French. Mean accuracy rate for each participant group for Time 0 and Time 1. Error bars denote standard error.

There was no main effect of Cognate Status ($X^2(1) = 0.01$, $p > 0.1$) and no interaction of Continued French and Cognate Status ($X^2(1) = 1.28$, $p > 0.1$). Participants were equally accurate at identifying cognates and noncognates.

Reaction Time:

There was a significant main effect of Time (*Est.* = -0.34, *SE* = 0.05, $X^2(1) = 22.22$, $p < 0.0001$), with bilinguals being faster at Time 1 than at Time 0 (see Figure 10). However, there

was no main effect of Continued French ($X^2(1) = 0.14$, $p > 0.1$), no significant main effect of Cognate Status ($X^2(1) = 2.43$, $p > 0.1$), and no significant interactions ($X^2(1)$s < 0.09, $p > 0.1$).

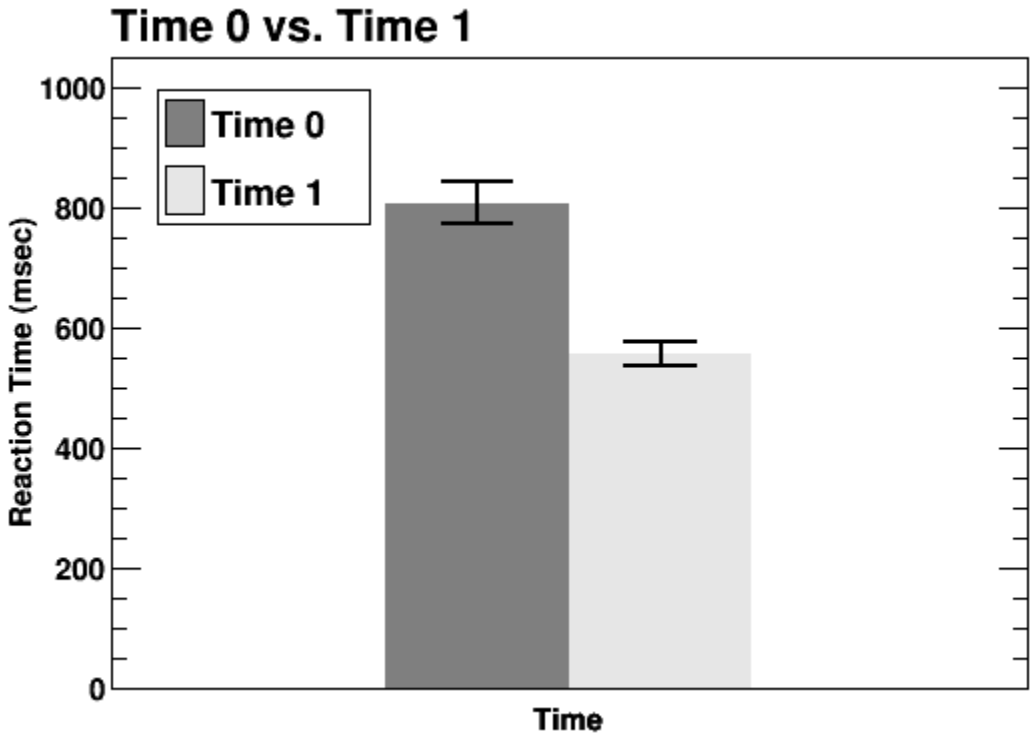

Figure 10. Main effect of Time. Mean reaction time for bilinguals at Time 0 and Time 1. Error bars denote standard error.

Changes across the first nine months (Time 0 versus Time 2)

Accuracy:

There were no significant main effects of Continued French ($X^2(1) = 1.38$, $p > 0.1$), Time ($X^2(1) = 0.01$, $p > 0.1$), or Cognate Status ($X^2(1) = 0.01$, $p > 0.1$), and no significant interactions ($X^2(1) = 1.13$, $p > 0.1$). Both groups of bilinguals treated cognates and noncognates similarly and showed no changes in accuracy from Time 0 to Time 1.

Reaction Time:

There was a significant main effect of Time (*Est.* = -0.32, *SE* = 0.10, $X^2(1) = 8.17$, $p < 0.01$; see Figure 11), with the bilinguals being faster at Time 2 than Time 0.

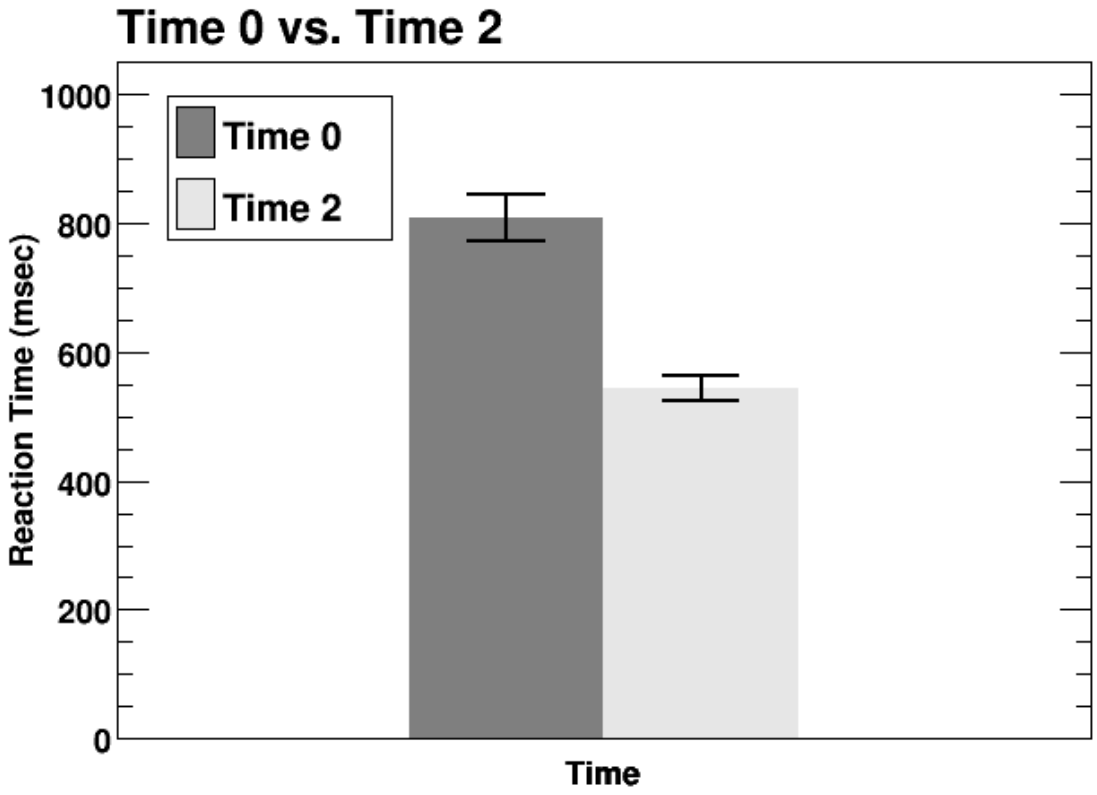

Figure 11. Main effect of Time. Mean reaction time for bilinguals at Time 0 and Time 2. Error bars denote standard error.

However, there was no main effect of Continued French ($X^2(1) = 0.18$, $p > 0.1$) and no significant interaction of Continued French and Time ($X^2(1) = 0.01$, $p > 0.1$), i.e., both groups of bilinguals improved their reaction times to a similar degree from Time 0 to Time 2, regardless of whether they were still exposed to French or not. The interaction of Continued French and

Cognate Status was not significant, either ($X^2(1) = 0.01$, $p > 0.1$). Neither group of bilinguals differentiated between cognates and noncognates.

3.3.4 Discussion

Experiment 2 tested whether English-French bilinguals who had recently returned from a study-abroad program in France differed from native French speakers in a French lexical decision task. Furthermore, it was tested whether changes across time (within the first five to nine months after the bilinguals' return from France) occurred, and whether these changes differed for bilinguals who were still exposed to French upon their return and those who were no longer exposed to their L2. Finally, it was investigated whether cognate status of target words influenced any of the above differences.

Effects of bilingualism

The results from the first test at Time 0 show that the bilinguals were not native-like even after a four-month immersion period in France: the native French speakers were more accurate than the bilinguals at identifying French words. However, the fact that the bilinguals, while less accurate, were not any slower at responding than the native French speakers, suggests that they were indeed quite fluent in their L2.

Effects of continued exposure to French

The bilinguals who still took formal French classes after their return from France were more accurate than the bilinguals who were no longer exposed to French at Time 0. At the time of the first test, the continuing bilinguals had already started their French classes (about four

weeks earlier) and it therefore makes sense that even at Time 0, they were already better at identifying French words than the bilinguals who received no more input in French.

However, the reaction times of the two groups of bilinguals did not differ significantly from each other, suggesting that while the non-continuing bilinguals were less accurate than the continuing bilinguals their relatively lower performance level did not slow them down.

Changes across time

The bilinguals also showed some changes across time. As in Experiment 1, the bilinguals seemed to improve across time: their reaction times were shorter both at Time 1 and Time 2 than at Time 0. Note that in Experiment 1, it was shown that only the bilinguals, but not the native English speakers (who were functional English monolinguals with no knowledge of French), improved their reaction times, suggesting that the improved RTs were indeed a sign of improved L2 processing rather than simply a task familiarity effect.

Not all bilinguals improved in terms of accuracy across the first five to nine months after their return, however. While the bilinguals who continued taking French classes became more accurate between Time 0 and Time 1, the bilinguals who were no longer exposed to French did not. Even though the L2 input was much reduced upon their return to the United States, the continuing bilinguals further improved their accuracy at identifying French words. It seems therefore that regular L2 input, even if it amounts to only a few hours per week, can lead to further improvements in an L2. If there is no more input in the L2, on the other hand, this does not necessarily lead to attrition, as was shown by the non-continuing bilinguals. Rather than

losing their ability to identify French words, their performance was not significantly different at Time 0 and Time 1.

This pattern of results suggests that L2 lexical perceptual knowledge is not subject to attrition in the first five months or so after the end of immersion, even in the absence of further L2 input. Moreover, even limited input in the L2 can lead to significant improvements during this time period.

However, processing does still seem to be sensitive to the amount of input an L2 speaker receives. This is shown by the decline in accuracy exhibited by the continuing bilinguals at Time 2 relative to Time 1. While at Time 1 their accuracy had improved relative to Time 0, at Time 2, their accuracy rate was no longer higher than at Time 0. This suggests that they had lost some of their L2 perceptual skills by Time 2. This can be interpreted as a reaction to even more reduced input after Time 1. The four months between the testing at Time 1 and Time 2 included three months of summer break, during which the bilinguals did not take any formal French classes. None of the bilinguals were exposed to French for more than three weeks total during that break, and the continuing bilinguals only started French classes again about four weeks before the testing at Time 2. It seems, therefore that while a reduction in L2 input immediately after the end of an immersion period does not lead to attrition of perceptual skills, further reduction (or absence) of input a few months down the line (or specifically, between five and nine months after the end of the immersion period), can indeed lead to attrition of lexical decision skills.

The fact that the accuracy rates at Time 0 and Time 2 were not significantly different also suggests, however, that the attrition that seems to have taken place over the bilinguals' summer

break still left them with a proficiency level that was not significantly different from that soon after their return from France.

As for the non-continuing bilinguals, who were not exposed to French at all since their return from France, their lexical decision skills seem to have remained stable across the first nine months after their return. This suggests that if there was attrition in their perceptual skills, it would have been limited to the first two months after the end of the immersion period. No further significant changes in proficiency took place after that amount of time.

Overall, with continued L2 input, there was initial improvement in lexical decision within the first five months after the bilinguals' return, but attrition set in as soon as L2 input ceased. However, a level of proficiency similar to that soon after the end of immersion remained even nine months after the bilinguals' return to their L1 environment. In the case of bilinguals who had not received further input in the L2 at all during the first nine months, neither improvement nor attrition were found across that period of time.

However, the current data do not allow us to address the question of whether the non-continuing bilinguals did indeed experience attrition previous to the test sessions. If present, such attrition would be parallel to the continuing bilinguals' decline over the summer break, which would in turn be expected to have leveled out after about two months (and possibly with continued L2 input after the summer break would improve again). This possibility may present a worthwhile question for future research.

Effects of cognate status

Neither the native French speakers nor the bilinguals exhibited processing differences between cognates and noncognates, both at baseline, and with regards to changes across time (for bilinguals). While bilingual language processing experiments frequently have found cognate effects (facilitation or interference), the lack of a cognate effect in the present study is in line with some previous research.

First of all, the absence of an effect for native French speakers – even though they are indeed French dominant French-English bilinguals – is not surprising. Previous research has shown that bilinguals are much less likely to show a cognate effect in speech perception if they perform a task in their L1 rather than their L2 (e.g., Lemhöfer & Dijkstra, 2004), as was the case with the native French speakers in the current study. Words in the L1 are thought to be accessed faster than L2 words, which makes it more difficult for a cognate that activates the nontarget-L2 to influence processing of an L1 word than vice versa (Dijkstra, 2005).

The absence of effects for the bilinguals, even though they performed the task in their L2, is also in line with previous research on bilingual lexical decision. A number of studies have used repetition priming and its effects on visual lexical decision to investigate processing differences between cognates and noncognates. In repetition priming, nontarget language words are briefly shown (primed) to see if reaction times of their translation equivalents in the target language are shorter following such a prime. If the RTs to these primed items are shorter than those to unprimed items, this suggests co-activation of the nontarget language. Studies with moderately proficient bilinguals (Cristoffanini, Kirsner, & Milech, 1986) have found repetition effects for cognates but not noncognates. However, studies with high-proficiency bilinguals have

found no priming for either cognates or noncognates (Kerkman & De Bot, 1989; as cited in Woutersen, De Bot, and Weltens, 1995), and studies with near-native bilinguals have found priming effects for both cognates and noncognates (Woutersen, Cox, Weltens, & De Bot, 1994). While the majority of repetition priming tasks has been conducted in the visual modality, Woutersen et al. (1995) have found parallel effects in a direct comparison of visual and auditory lexical decision tasks. Crucially, the above studies suggest that a cognate facilitation effect is found only with moderately proficient bilinguals, while highly proficient or near-native bilinguals do not show a distinction between cognate and noncognate primes in lexical decision RTs. In the current study, no differences have been found for cognates versus noncognates in a simple lexical decision task (without priming). The lack of cognate effect was found for RTs as well as accuracy. This suggests that the bilinguals in the current study can be considered highly fluent, and furthermore that the lack of a cognate effect seems to not only affect reaction times but also accuracy times. From the lexical decision paradigm used in the current study, it is not possible to conclude, however, whether there was parallel activation of the nontarget language English that would have boosted the bilinguals' performance for both cognates and noncognates (analogous to the priming effects found for cognates and noncognates in highly proficient bilinguals in Cristoffanini et al.) or whether parallel activation did not play a role in this task and there was no boost for either cognates or noncognates (analogous to the lack or repetition priming found in near-native bilinguals in Kerkman and De Bot, 1989).

3.3.5 Conclusion

The results of the lexical decision task suggest that even in the absence of further L2 input, bilinguals do not necessarily experience lexical attrition in speech perception within the

first nine months after the end of an L2 immersion period. This was shown by the non-continuing bilinguals in the current study, whose accuracy at identifying French words did not change within those nine months. Moreover, even without further input, bilinguals may improve their lexical decision skills, much like the bilinguals in the current study who were faster at Time 1 and Time 2 than at Time 0, regardless of whether they were still exposed to French or not.

Further exposure to French does seem to affect L2 processing, however. The bilinguals who were still taking French classes not only did not show attrition in the first five months, but their accuracy at identifying French words actually improved. However, as soon as they stopped taking classes, their accuracy rates dropped again, so that they were back at the level of performance they were at soon after their return from France. The precise time course of attrition that may have taken place in the four months between the tests at Time 1 and Time 2 would be a worthwhile question to investigate in future research. Finally, cognate status of words does not seem to modulate bilinguals' lexical decision performance across time, likely partly because of the high proficiency level of the bilinguals, and partly because of the properties of the cognates used in this study (which overlapped orthographically *and* phonologically across French and English).

3.4 General discussion

AX discrimination and lexical decision tap into different levels of processing. One can perform a discrimination task even in the absence of lexical knowledge of the target language, or in fact, in the absence of any knowledge of the target language at all. While there are differences in the perceptibility of certain types of sound contrasts in a language a listener is not familiar with (Best et al., 2001; Flege, 1995) any listener can perform the task to some extent even if only

based on purely acoustic discrimination. This is not possible in lexical decision, of course, which requires both phonological and lexical processing of L2 stimuli. This heavier processing load may lead to different patterns of language attrition, retention, and improvement after the end of an L2 immersion period, compared to phonological processing tasks such as AX discrimination. The present data support this difference.

One set of results both tasks have in common is that no attrition, measurable either through lower d' scores, lower accuracy rates, or longer reaction times, was detected in the first five months after the end of the participants' study-abroad experience. This was true for bilinguals who were still exposed to French, and even for those who were no longer exposed to French in any formal way.

However, there were differences in how sound processing (AX discrimination) and lexical processing (lexical decision) were affected by continued exposure to the L2. Bilinguals who were still taking French classes improved their sound processing skills steadily across the first nine months post-immersion (both their Time 1 and Time 2 scores were better than their Time 0 scores). This suggests that further exposure to an L2, even if the amount of exposure is less than during the immersion period, aids in further advancing sound processing skills. This was shown in that the bilinguals achieved higher d' scores and improved RTs at Time 1 and 2 relative to Time 0.

In lexical processing, however, the effects of continued exposure to the L2 seem to be less resistant to attrition. While the continuing bilinguals improved their lexical decision accuracy rate during the first five months after their return from France (while they were taking

French classes), they showed signs of attrition by nine months after the end of their study-abroad program, likely because the four months between Time 1 and Time 2 included three months of college summer break during which the bilinguals exposure to French was extremely limited. Notably, this same hiatus did not lead to attrition of the bilinguals' sound discrimination skills.

While the continuing bilinguals' accuracy rate dropped over the summer break, they further improved their reaction times, suggesting that some L2 processing was still taking place, which led to improved reaction times, but that active L2 input is needed to retain a high level of the lexical knowledge required in a lexical decision task.

Bilinguals who did not continue taking French classes did not improve their sound discrimination skills in the first five months after the end of L2 immersion, but they did improve within nine months post-immersion. Their sound processing changed differently across the first nine months post-immersion compared to the bilinguals who continued taking French. They seemed to have hit a plateau for a few months and then finally advanced again between five and nine months after their return from France. Such a 'settling in process' (Cohen, 1986) has been found in previous research on language attrition. Cohen found improved lexical skills in Spanish L2 students after they had not been exposed to the L2 over summer recess. The circumstances of language study and the timeline of testing was different in Cohen's study and the current study. It should be noted that Cohen's students were only exposed to Spanish in the classroom while living in an L1 environment, and the period between the end L2 exposure and testing was only the duration of summer recess (approximately 2 months). However, it is certainly possible that similar effects would be found in immersion students and over a longer period of time.

As for lexical decision, the bilinguals who did not continue taking French classes showed no changes in accuracy across the first nine months after the end of the immersion period, but just as the continuing bilinguals they improved their reaction times. This suggests that continued lexical processing is going on even in the absence of continued exposure.

In sum, continued exposure seems to lead to improvements in sound processing, but even in the absence of continued exposure students catch up with those continuing within nine months after exposure. Lexical processing, on the other hand seems to rely more on continuous input, in the absence of which no improvements seem to be made, and in addition if input ceases, signs of attrition can soon be found.

Chapter 4: Retention in L2 speech production

4.1 Introduction

In the following sections, retention versus attrition of phonological and lexical skills in speech production will be investigated. The same participants (bilinguals and native French and English speaking control groups) as in the in the perception experiments participated in the two production experiments. Bilinguals were tested three times within the first nine months after their return from France to determine if they exhibit any form of attrition or whether they have been able to retain their L2 production skills. However, only the results from the test at Time 0 (two months after the bilinguals' return from France) and Time 2 (nine months after their return) are analyzed here. If effects of time since return are present, this analysis should allow us to detect them; any effects of attrition/retention should be strongest at this longest length of time. While the perception experiments in the current study, in line with previous research (e.g., Cohen, 1986; Murtagh, 2003; Weltens, 1989) have shown that perceptual skills such as those required in an AX discrimination task or a lexical decision task are fairly robust within the first nine months after the end of an L2-immersion program, previous research has also shown that speech production seems to be more susceptible to attrition (e.g., Bahrick, 1984; Tomiyama, 2000).

The current study aims to extend previous research by investigating L2 retention versus attrition in speech production at both the phonological and the lexical level in the same group of participants across two points in time in a longitudinal design, allowing for direct comparison across time and across processing levels.

Phonological retention is measured in a word repetition task in which participants were asked to repeat French words and nonwords presented to them in a recording of productions by a

native French speaker. Previous research has shown that while lexical processing may be involved in word repetition tasks it is not necessary to perform the task, i.e., output phonology can be retrieved directly from input phonology without having to retrieve phonemes from the lexicon (Nozari et al., 2010). The task also aimed to determine whether patterns of attrition versus retention are influenced by different types of sounds (consonants versus vowels, phonemes that are similar versus different in the L1 and the L2).

Lexical retention was measured using a picture naming task. Participants were asked to name a series of pictures in French. The pictures showed either French-English cognates or French noncognates. As lexical processes are required to relate pictures and sounds (cf. Nozari et al., 2010), this task tapped into both the phonological and lexical processing levels and it will be investigated whether patterns of attrition versus retention are modulated by the lexical status of a word.

4.2 Experiment 3: Retention of phonological skills

4.2.1 Introduction

Experiment 3 was designed to explore retention versus attrition of phonological skills in speech production, which was assessed through a word repetition task. Participants listened to a recording of a French speaker saying French words and nonwords and were asked to repeat each word or nonword after the recording. Only the word productions will be analyzed here. Each target word contains a critical vowel or consonant sound that was submitted to acoustic analyses. All three groups of speakers (bilinguals, French native speakers, and English native speakers) participated in this task, but only the productions of the bilinguals and the French native speakers

were analyzed for this study. To maximize the effect any changes across time might have on the participants' productions, only Time 2 productions were compared to Time 0 productions.

4.2.2 Methods

Participants

All three groups of participants as described above participated in the word repetition task, but only the productions by the bilinguals and the French native speakers were analyzed for this study. In addition, in order to maximize the likelihood of finding attrition only the productions of bilinguals who were no longer exposed to French were analyzed.

Materials

The stimuli in the repetition task consisted of 20 French words (and 20 nonwords, which are not included in the present analysis). Each word contained one of five critical sounds, which was either a vowel (/u/, /y/, /œ/) or a consonant (/t/, /d/). Each critical sound was also either a sound that exists both in French and English, i.e., a 'similar' sound (/u/, /t/, /d/), or a sound that exists in French but has no close equivalent in English, i.e., a 'new' sound (/y/, /œ/; cf. Flege, 1987). In the case of the stop consonants (/d/ and /t/), the critical sound was always the initial sound of the word, and in the case of the vowels (/u/, /y/, /œ/), the critical sound was always the second sound in the word, immediately following a single initial consonant. All of the words in the repetition task had already appeared in the AX discrimination task. The participants never saw the words in written form, but were exposed to them auditorily only. The complete list of critical sounds and target words containing each sound are shown in Table 3 below.

Critical sound	Sound type	Cross-language status	Target words	Transcription	Translation
u	vowel	similar	bouche	[buːʃ]	mouth
			douleur	[dulœːʁ]	pain
			poulet	[puleː]	chicken
			tousser	[tuseː]	to cough
y	vowel	new	dur	[dyːʁ]	hard
			fûtes	[fyːt]	(you) were
			pur	[pyːʁ]	pure
			zut	[zyːt]	darn
œ	vowel	new	neuve	[nœːv]	new
			peur	[pœːʁ]	fear
			soeur	[sœːʁ]	sister
			veuve	[vœːv]	widow
t	consonant	similar	tante	[tãːt]	aunt
			taureau	[toʁoː]	bull
			tellement	[tɛlmãː]	such
			tige	[tiːʒ]	stem (noun)
d	consonant	similar	dehors	[dəɔːʁ]	outside
			dites	[diːt]	(you) say
			dommage	[domaːʒ]	harm (noun)
			donc	[dõːk]	therefore

Table 3. Critical sounds and target words from Experiment 3.

Procedure

Participants were seated in front of a computer in a sound-attenuated booth. They were played the target words over headphones at a comfortable level. Visual and auditory stimuli were presented to the participants and their responses were recorded using Superlab 4 software. Instructions for the task were presented in writing on the computer screen, and they were also read aloud by the experimenter. As in Experiments 1 and 2, instructions were given in English so that participants with a low proficiency in French would be able to perform the task correctly.

After having been presented with visual and spoken instructions, participants completed four practice trials on stimulus words that were not included in the actual experiment. After the practice trials, participants were prompted to ask any questions, and when they had no more questions, they pressed a key on the keyboard to start the experiment.

Once the experiment started, the screen remained blank and participants were played one stimulus word at a time. Participants then repeated the stimulus into the microphone. Participants were asked to respond as fast and as accurately as possible. After the participants had spoken the word, they then pressed the space bar to move on to the next trial. The experimental trials were played to the participants in random order and interspersed with 100 filler trials. Fillers consisted of 50 French words and 50 nonwords that did not contain any of the critical sounds in the same word positions as the experimental trials. Each stimulus and filler item was played to the participants twice. The repetition task lasted about 20 minutes.

Coding

All sound recordings were annotated to a text grid using Praat version 5.2.11 (Boersma & Weenink, 2011) and the EasyAlign tool (Goldman, 2008). Target words and segments were aligned and labeled in multiple steps by the author and a second coder. In a first step, a text grid with two tiers was created using the EasyAlign plug-in for Praat, which was written to help with labeling text grids for French sound recordings. The plug-in can handle individual French words or entire sentences and can return a text grid with tiers for sentences, words, phonetic transcription of the words, syllables and phonemes. It aligns the items in the tiers in two steps in between which hand-correction of the text grid tiers is necessary. The aligner read in each sound file and a text file containing all the recorded words in the order they appeared in the recording

(the list was adapted from the Superlab output) and returned a text grid with one tier in which the words from the sound file were labeled with the items from the Superlab output list (the 'label' tier), and a second tier in which they were labeled phonetically (based on the database built into EasyAlign rather than the actual productions), which was called the 'phono' tier.

After this step the text grids were manually checked for accuracy and corrected if necessary (e.g., if the aligner missed a word it was hand segmented and labeled in the text grid). Next, EasyAlign was run again and in this second step returned a text grid with an additional two tiers: one with the words segmented and labeled orthographically (the 'words' tier, which was in fact identical to the 'label' tier), and a tier in which the individual sounds in each word were segmented and labeled (the 'sounds' tier).

While the EasyAlign plug-in was accurate at labeling French sound segments, it was not able to reliably align the boundaries of the sounds in the text grid. This was especially true for the recordings by the bilinguals, since their pronunciations were not always native-like. Therefore, substantial hand-corrections of all text grids were necessary before they could be used to extract formant and voice onset time information.

When the text grids had been labeled satisfactorily the voice onset time for each target consonant as well as the formant values for F1 and F2 for each target vowel were extracted and saved to a txt-file for further analysis. The formant measurements were made automatically with an LPC formant tracking algorithm in Praat. F1 and F2 values that differed by more than 200 Hz from the native French speakers' mean for that category were hand checked and corrected if necessary (cf. Smiljanic & Bradlow, 2005).

F1 and F2 measurements were taken at the midpoint of each vowel. Since F1 values correspond to vowel height (e.g., the high vowel /i/ has a lower F1 value than the low vowel /a/), and F2 values correspond to vowel frontness (e.g., the front vowel /i/ has a higher F2 value than the back vowel /u/). Since French vowels differ in quality from English vowels (and some French vowels do not exist in English at all), F1 and F2 provide good measures to compare the French vowel productions by the French native speakers and the bilinguals.

Voice onset times (VOT) for stop consonants were measured from the release of the stop until the first glottal pulse of the first periodic wave of the vowel. In the case of pre-voicing, the duration was measured from the first glottal pulse of the periodic wave until the onset of the burst of the stop consonant. Pre-voicing was labeled as such and entered with a negative value into the analysis spreadsheet. VOT provides a useful crosslinguistic measure, because French stop consonants tend to have shorter VOTs than English stop consonants. Finally, the categorical variables Bilingualism and Time were contrast coded.

Interrater reliability was assessed by having both coders analyze 100 tokens of VOTs for /d/ and /t/ across five bilingual participants. The coders had a high rate of agreement. The average absolute deviation across coders was 1.6 msec and 95% of deviations were less than 3 msec.

Analysis

Linear mixed effects regressions (lmer) with Female as a fixed effect and by-subject random slopes for Time and by-item random slopes for Bilingualism and Time were performed on the F1 and F2 values for the critical sounds in the case of vowels, and the VOT in the case of

consonants. Female was included as a fixed effect to control for the fact that the number of female participants was not equal across the bilinguals and the native-French speakers (6 females among 8 bilinguals and 5 females among 8 native French speakers). As in the perception experiments, each main effect (i.e., Bilingualism, Time) was removed from the maximal model and the model lacking the effect of interest was re-fit. The relative fit of the reduced vs. maximal model was compared using likelihood ratio tests. If the maximal model was improved by the inclusion of the main effect, this was interpreted as the main effect making a significant contribution to variance. For each component of the model, the estimate (*Est.*) and standard error (*SE*) of the maximal model are reported, as well as the result of the likelihood ratio test for the corresponding reduced model (X^2(df) with the associated *p*-value). Analyses were performed on the following independent variables:

- *Bilingualism* (French native speaker, Bilingual)
- *Time* (Time 0, Time 2); as in the perception experiments, the Time 0 tests took place about 2 months after the bilinguals' return from France, and Time 2 was about 7 months later (nine months after the bilinguals' return)

Productions that fell more than 3 standard deviations outside the mean across participants for each sound were identified as outliers and excluded from the analyses. This led to the exclusion of 2.5% of the data (8 words). Furthermore, one native French speaker's productions of /y/ and /œ/ lay entirely outside of 2 standard deviations of the overall mean for all native French speakers. This speaker possibly spoke with a different regional accent than the other speakers (all native French speaking participants' were from France, but their regional dialects

were not elicited in the background questionnaires). All of this participant's vowel productions were consequently excluded from the analyses.

4.2.3 Results

Baseline performance at Time 0

Voice onset time:

There was no significant main effect of Bilingualism for /d/ and /t/ ($X^2(1)$s < 0.25, p > 0.1; see Figure 12), suggesting that the bilinguals' productions were in fact not significantly different from the native French speakers' productions. In terms of VOT, bilinguals were native-like for these 'similar' sounds. Table 4 shows the mean VOTs for bilinguals and native French speakers.

Sound	French speakers		Bilinguals	
	M (msec)	*SD*	*M* (msec)	*SD*
/d/	-104	6.1	-108	8.9
/t/	54	4.6	59	4.7

Table 4. Mean VOTs for target consonants in Experiment 3.

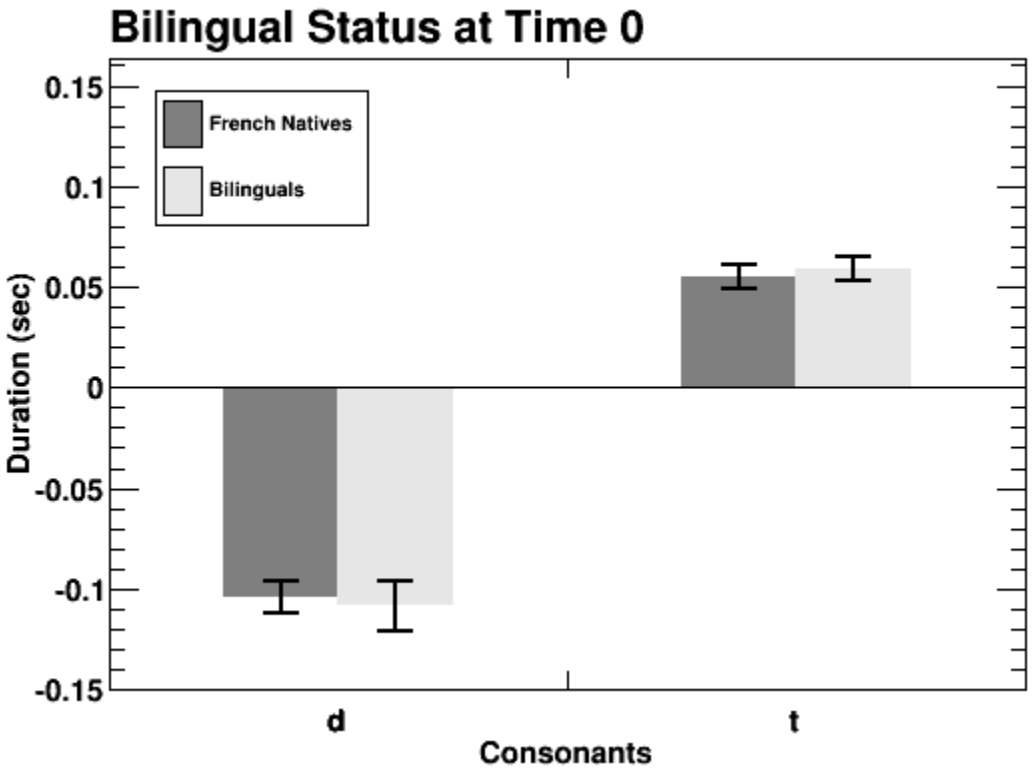

Figure 12. No significant main effect of Bilingualism for /d/ and /t/. Mean VOT for each consonant at Time 0, separated by participant group. Error bars denote standard error.

F1:

There were no significant main effects of Bilingualism for any of the three vowels (/u/, /y/, and /œ/; $X^2(1)$s < 2.70, $p > 0.1$), suggesting that the bilinguals were able to produce the F1s of the vowels in a native-like fashion. Table 5 shows F1 and F2 frequencies for bilinguals and native French speakers.

Sound	French speakers				Bilinguals			
	F1		F2		F1		F2	
	M (Hz)	*SD*	*M* (Hz)	*SD*	*M* (Hz)	*SD*	*M* (Hz)	*SD*
/u/	341	109	1206	351	357	82	1366*	402
/y/	330	66	1887	218	377	57	1826	294
/œ/	526	97	1557	181	554	67	1595	174

Table 5. F1 and F2 frequencies for target vowels in Experiment 3.

Note. An asterisk in the Bilinguals column indicates a significant difference from the native French speakers' formant frequency (i.e., not native-like production).

F2:

There was a significant main effect of Bilingualism for /u/ (*Est.* = 248.53, *SE* = 99.74, $X^2(1) = 5.50$, $p < 0.05$), but no main effects for /œ/ and /y/ ($X^2(1)$s < 0.67, $p > 0.1$). This suggests that bilinguals were able to produce the F2 for /œ/ and /y/ in a native-like fashion, but they were not able to do so for /u/. The bilinguals' F2 values were significantly higher than the native French speakers, i.e., their /u/ sounds were fronted relative to the native French speakers' /u/. Since the prototypical English /u/ has a higher F2 value than the prototypical French /u/ (cf. Flege, 1987), these results suggest that the bilinguals' French /u/ was assimilated to the English /u/ category resulting in the more English-like F2 value compared to the native French speakers'.

The vowel space diagrams in Figure 13a and Figure 13b show that the bilinguals' and native French speakers' productions for /u/ differ along the F2-axis. Figure 13a shows the mean for each vowel for each participant, and Figure 13b shows the means for F1 and F2 for each group.

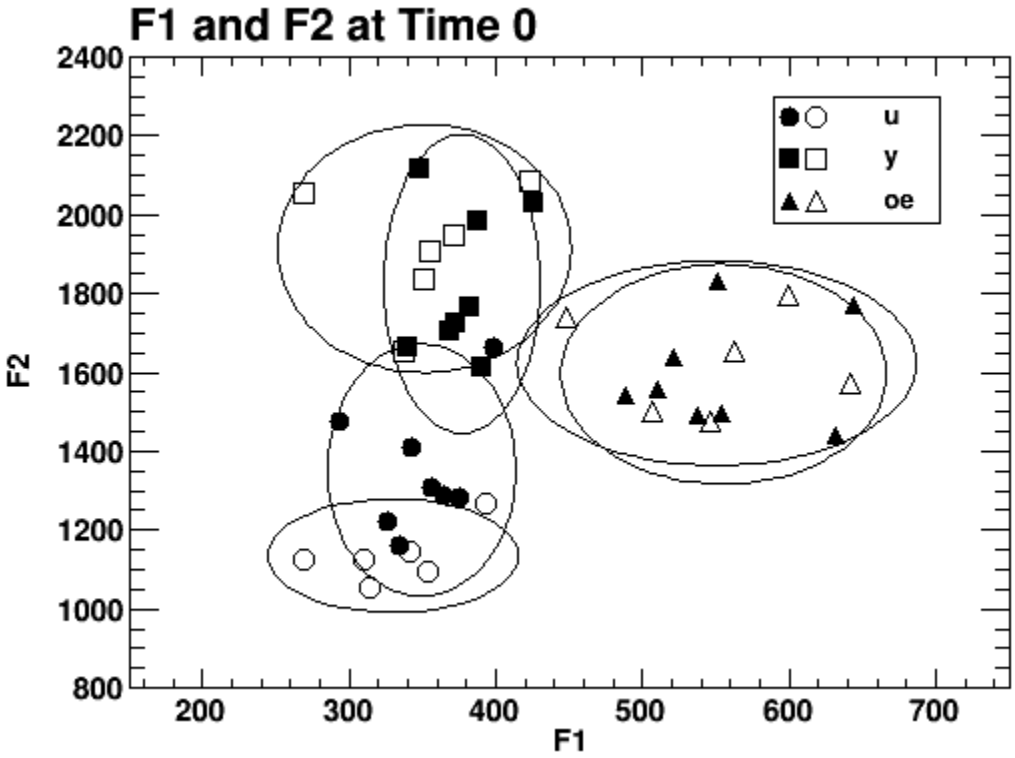

Figure 13a. Significant main effect of Bilingualism on F2 for /u/ (participant means shown). Ellipses represent the area covering 2 standard deviations from the means of F1 and F2 for each group of participants. Filled symbols show the bilinguals' productions, open symbols show the native French speakers' productions.

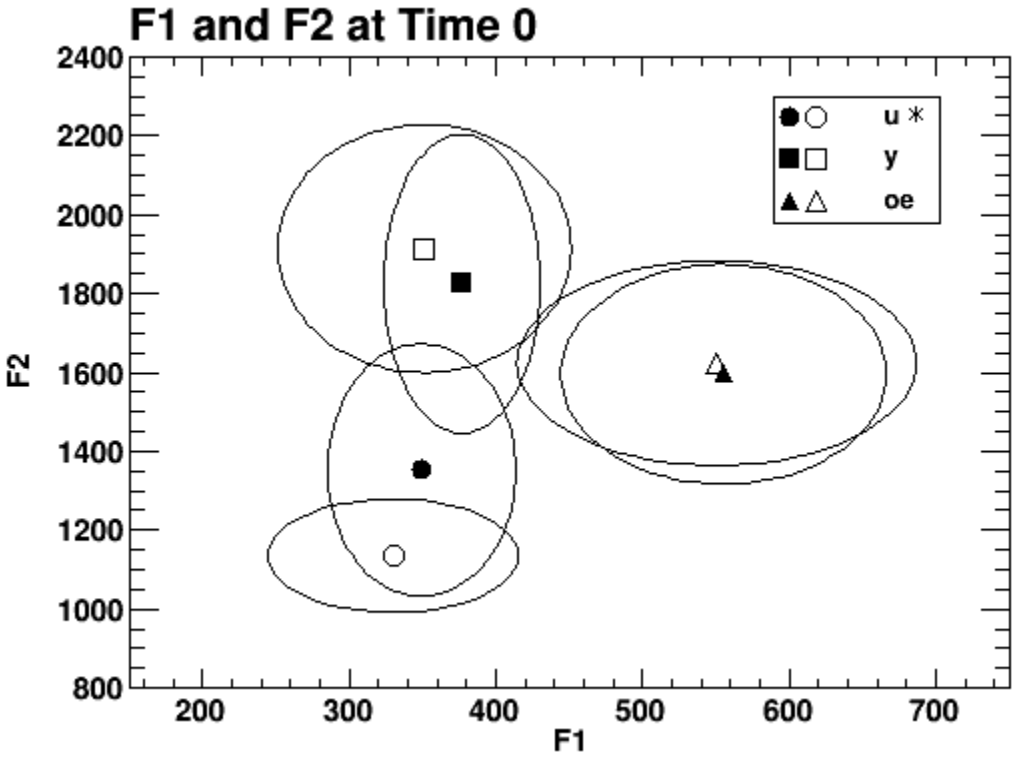

Figure 13b. Significant main effect of Bilingualism on F2 for /u/ (group means shown). Ellipses represent the area covering 2 standard deviations from the means of F1 and F2 for each group of participants. Filled symbols show the bilinguals' productions, open symbols show the native French speakers' productions.

Note. An asterisk indicates a significant difference between the native French speakers' and the bilinguals' formant frequency (i.e., not native-like production).

Changes across the first nine months (Time 0 versus Time 2)

Acoustic variables:

Only the bilinguals' productions were analyzed for changes across time, since the native French participants only participated in the testing at Time 0. There were no significant main

effects of Time for VOT, F1, or F2 for any of the sounds ($X^2(1)$s < 21.20, p > 0.1), suggesting that the bilinguals' productions remained stable across the first nine months after their return from France. Figures 14a and 14b show that the F1 and F2 values were no different for the bilinguals at Time 0 and at Time 2. Figure 14a shows the average formant values for vowel for each bilingual. Figure 14b shows the average formant values for the whole group of bilinguals at Time 0 and Time 2.

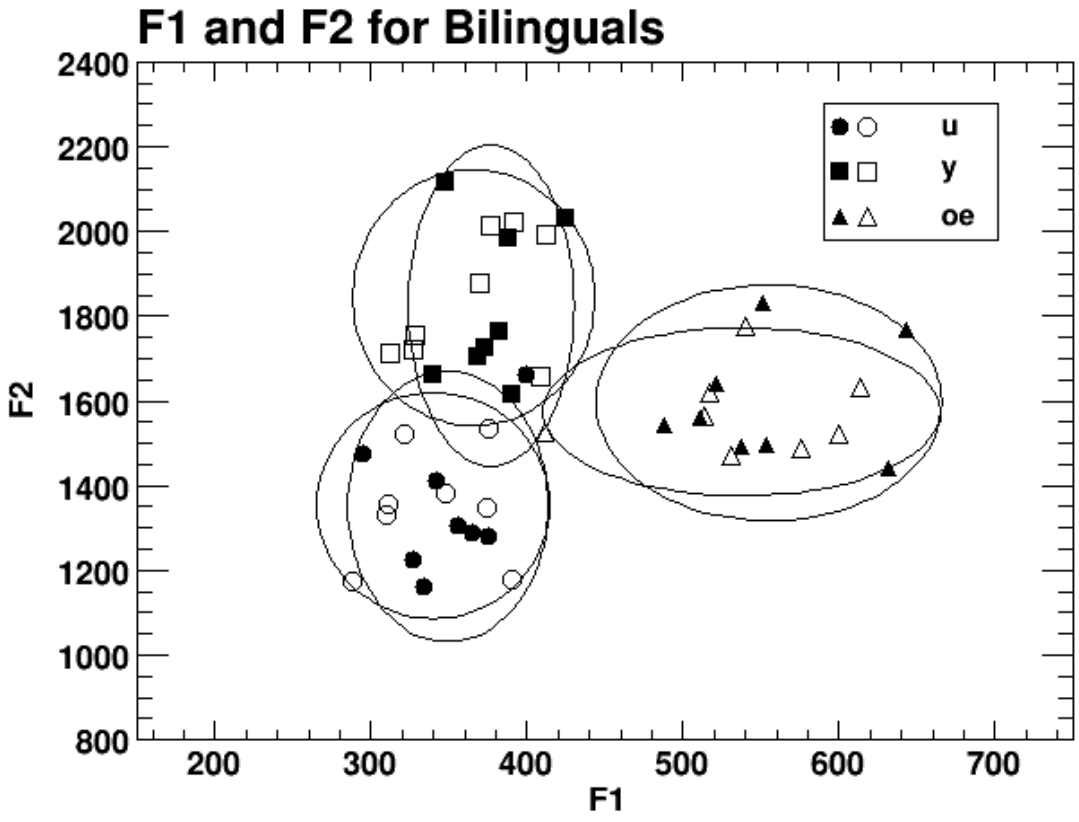

Figure 14a. No significant main effect of Time on F1 and F2 (participant means shown). Ellipses represent the area covering 2 standard deviations from the means of F1 and F2 at each time. Filled symbols show the bilinguals' productions at Time 0, open symbols show their productions at Time 2.

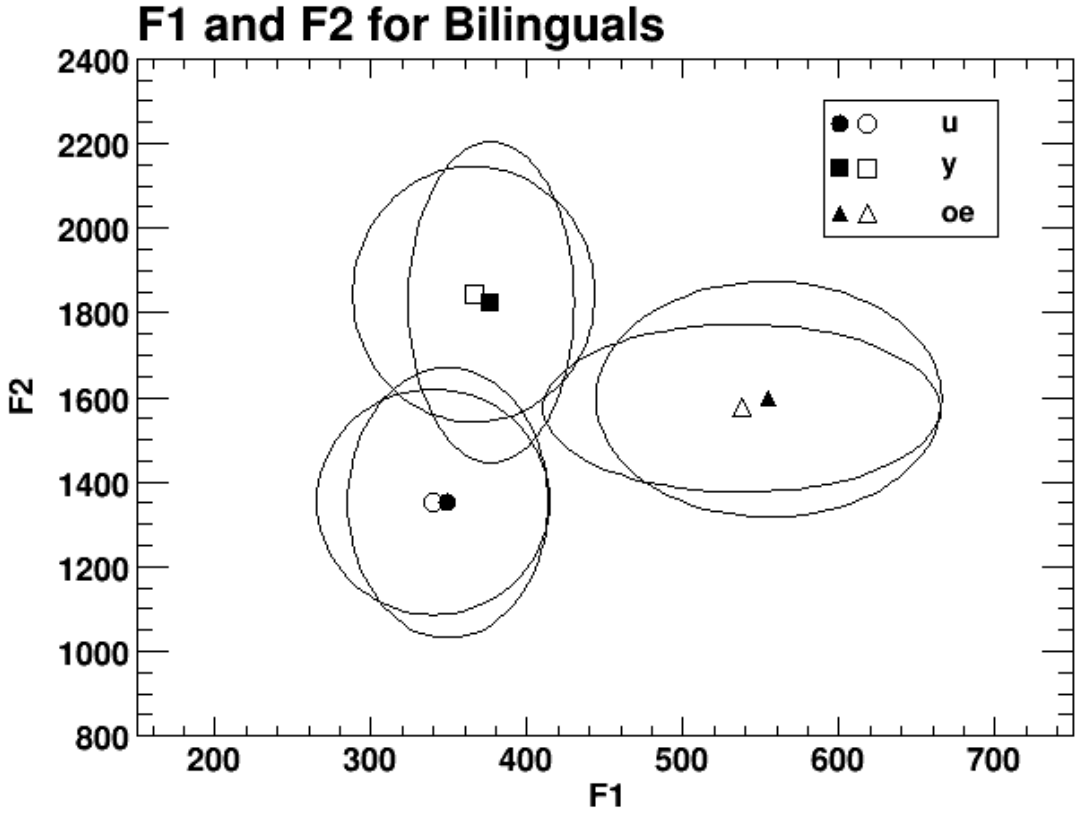

Figure 14b. No significant main effect of Time on F1 and F2 (group means shown). Ellipses represent the area covering 2 standard deviations from the means of F1 and F2 at each time. Filled symbols show the bilinguals' productions at Time 0, open symbols show their productions at Time 2.

4.2.4 Discussion

Experiment 3 examined through a word repetition task whether bilinguals who had recently returned from a study-abroad stay in France differed from native French speakers in their productions of French words, and crucially, also whether the bilinguals' productions changed within the first nine months after their return. Acoustic measures were taken from a set

of target sounds. Namely, for vowels F1 and F2 were measured, and for consonants voice onset time was measured. Similarity of sounds and sound contrast pairs across languages was also taken into consideration in the interpretation of the results.

No changes across time were found, i.e., the bilinguals neither showed a significant amount of attrition, nor improved significantly within the first nine months after their return from France. This suggests that the skill level in pronunciation of L2 sounds acquired during a semester in an L2 environment is robust against attrition for at least those nine months. This is in line with previous research on L2 phonological attrition versus retention in speech production. Murtagh (2003) found that English-native learners of Irish, who studied Irish in a secondary school setting did not show signs of attrition eighteen months after the end of Irish training in a word reading task targeted at pronunciation accuracy of fourteen different phonemes.

As was discussed for speech perception above, the lack of attrition during those first nine months after the end of the L2 immersion period are compatible with Cohen's (1986) idea of a settling-in processing during periods of non-use. However, this would also imply that further improvements would appear after the end of this process during which learners consolidate information and unlearn incorrect patterns. This was not the case in the current data. It is possible that further improvements would indeed be found at a later point in time. However, Cohen found improvements (at a lexical level) already after a period of three months. Of course, the participants in Cohen's study did not study their L2 Spanish in an immersion environment, and so it is possible that the time frame for such a settling-in phase and improvements following the phase would be different for speakers who have been immersed in an L2 environment. Future research, looking at the development in speech production beyond nine months post-immersion

would be necessary to see whether further improvements are indeed possible after nine months have elapsed since the end of an immersion period.

While there were no changes across time in the bilinguals' speech production, some interesting differences between the bilinguals' and the native French speakers' productions at Time 0 were found. The bilinguals' differed in their nativelikeness of the three acoustic measures.

For F1 no distinction between similar and new vowels was found, since all vowels were produced with native-like F1 frequencies. The lack of a difference in F1 for similar versus new vowels is in line with Flege's explanation that variability along the F1 continuum does not seem to be a main contributor to nativelikeness of high vowels like /y/ and /u/ (he does not mention /œ/), and that, in fact, little variability is expected by different speaker groups. In line with the current research, English-French bilinguals in Flege (1987) were also native-like in their F1 frequencies for French vowels.

As for F2, a similar/new distinction is made. The SLM would predict that L2 learners are better at producing new sounds than similar sounds, which is indeed what was found for F2. This suggests that the bilinguals, as advanced learners of French, have formed separate phoneme categories in French for the new sounds /y/ and /œ/, which allowed them to produce them in a native-like fashion. Since /u/ in French, however, is similar to /u/ in English, no new category needed to be formed, but rather, the bilinguals may have interpreted the French /u/ as an exemplar of the English /u/ category, which prevented them from acquiring a native-like pronunciation for this sound.

Specifically, the bilinguals' F2 frequencies for /y/ and /œ/ were not significantly different from the native French speakers' F2. The bilinguals F2 for /u/, however, was significantly higher – and therefore more English-like – than the native French speakers'. These results are in line with previous research by Flege (1987) who found no significant differences in F2 for /y/ in native English learners of French at various proficiency levels compared to monolingual French speakers. Just like the current study, Flege also found that F2 frequencies for /u/, however, were significantly higher for native English learners of French compared to monolingual French speakers. The results of the current study, therefore, confirm the predictions of the SLM and suggest that the bilinguals in the current study have indeed formed separate phoneme categories for the French vowels /y/ and /œ/, which have no close equivalents in English, whereas the bilinguals seem to have assimilated the French /u/ to the English /u/ and, therefore, not produced them in a native-like fashion. For these bilinguals, new sounds were in fact acquired more native-like than similar sounds.

While the results for F2 for the vowels nicely fit into the SLM framework, the VOT results are not what the SLM would predict. The bilinguals' VOTs were native-like for the similar sounds /d/ and /t/. Flege (1987) found that native English learners of French differed from monolingual French speakers in their VOTs for /t/. Specifically, the English speakers' VOTs were longer – and thus more English-like – than the French monolinguals'. Flege also found, however, that fluent L2 speakers were able to produce VOTs in a much more native-like fashion than less proficient L2 speakers. In fact, Flege tested a group of Americans living in France whose VOTs were not different from the native French control group's. This suggests that the bilinguals in the current study, who had recently spent four months immersed in French in

France, had reached a high enough L2 proficiency to have acquired even those similar sounds to a native-like degree.

The above results are also in line with Birdsong (2007), who showed that native English late learners of French were indeed able to acquire VOTs in French consonants at a native-like level, even if those consonants were similar to the L1 English. Nine out of 22 native-English late learners of French were considered native-like for VOT (for /p/, /t/, /k/) in French.

French-English bilingual studies that compared VOTs across languages or speaker groups have generally focused on the voiceless stops /p/, /t/, and /k/. There is, therefore, no comparable data of French productions by native English bilinguals for the /d/ in the current study. However, based on Ryalls, Larouche, and Giroux (2003), the productions of the native French speakers and the bilinguals appear to lie within the range of native French speaker VOT values for /d/ (Ryalls et al. found average VOTs of -91 msecs in native French speakers, compared to the -108 msecs for the bilinguals and the -104 msecs for the native French speakers in the current study). In fact, the speakers in the current study prevoiced the stops even more than the speakers in Ryalls et al., i.e., they were more French-like and less English-like. One reason for this difference may be that the speakers in Ryalls et al. were from Montreal, and thus spoke Canadian French, whereas the participants in the current study were either French (the native French speakers came from various regions in France) or had studied abroad in France (all bilinguals had spent their semester abroad in Paris). In fact, Caramazza and Yeni-Komshian (1974) found that Canadian French VOTs tend to be longer than European French VOTs for voiceless stops produced by native speakers. It is possible that since not all native French speakers spoke Parisian French, some differences among native French speakers and between native French speakers and

bilinguals may have been due to dialectal differences. A suggestion for future research would, therefore, be to use a more homogeneous group of native speakers that speak a regional variant identical to the variant that the bilinguals were exposed to while studying abroad. Such a dialect difference, however, does not pose a major problem for the current study, since the important point was that the bilinguals were able to produce the voiced stops in a native-like fashion regardless of regional variant – and if anything, the bilinguals would have been expected to be more different from the native speakers, if the native speakers' dialects were not all identical to the variant the bilinguals had acquired. Furthermore, while this was not predicted by the SLM, Flege (1987) himself, as well as Birdsong (2007) found that fluent L2 speakers are able to produce even VOTs of consonants that are similar in the L1 and the L2 in a native-like fashion. While previous research has only shown this in the VOTs of voiceless stops, the current study extends this finding to the voiced stop /d/.

The studies cited above show that native-like VOT productions are in fact possible for fluent bilinguals, and that dialectal differences between speakers may have influenced the similarity versus difference between speakers in different studies. However, there is one major difference between the current study and previous bilingual production studies, which for the most part have found bilinguals' VOTs to be *not* native-like, and which may be the main contributor to the nativelikeness of the bilinguals' VOTs in the current study. Previous VOT production studies have used word naming paradigms, where participants named written words out loud (e.g., Fowler et al., 2008), or other tasks where linguistic stimuli were presented in written form and had to be pronounced (e.g., Sancier & Fowler, 1997). In the current task, however, participants repeated French words that were presented to them *auditorily*. Previous

research in English using such a word repetition, or 'shadowing', paradigm has found that speakers' VOTs were significantly closer to the VOTs in stimulus words they were asked to repeat than in productions of the same words that the speakers named out loud when they were presented to them in written form. This was true both for the repetition of English words (Goldinger, 1998; Shockley, Sabadini, & Fowler, 2004) and for nonwords (Goldinger, 1998). These studies show that speakers are able to immediately incorporate into their productions perceptual details from auditory stimuli (Goldinger, 1998). While bilingual speech production studies have not generally used auditory word repetition paradigms, it stands to reason that bilinguals would also benefit from perceptual details in the input in their L2 productions, which would lead to more native-like productions than if the stimuli were presented in written form (as is the case in previous literature on monolingual versus bilingual VOTs).

But why were the bilinguals in the current study able to become native-like for the VOTs and not the F2 values for /u/? One reason that was mentioned for sound perception above was the greater variability in vowels than consonants across regional variants (Hawkins, 1993), which may make the native-like acquisition of vowels more difficult in general. In addition, it may be the case that sound duration (such as VOT) may be an easier cue to sound identification for native English speakers than spectral cues such as F1 and F2. Gottfried and Beddor (1988) came to this conclusion for their L1 English – L2 French participants in a French vowel identification task. The bilingual participants were more likely to use duration as a cue to vowel identity than did French monolinguals. The authors explained this as a result of the greater prominence of duration in vowel identity in English than French. While Gottfried and Beddor analyzed only vowels and not consonants, it could still be the case that bilinguals are in general better at

reaching native-like mastery of an ‘easy’ cue like sound duration (vowel duration or VOT) than the more difficult spectral cues. With the limited set of vowels and consonants used in this experiment, it is not possible, however, to definitively state how different cues may influence the nativelikeness of different sound types (such as similar versus new and vowels versus consonants) to varying degrees. In future research a bigger and well-balanced set of vowels and consonants that are new or similar sounds for L2 learners could shed further light on the question of how L2 sounds are acquired.

The results from the current study provide further evidence that L2 learners can indeed achieve native-like proficiency in the production of new and similar sounds to varying degrees for VOT, F1, and F2.

4.2.5 Conclusion

In a word repetition task English-French bilinguals differed from native French speakers on different cues to sound identity. The bilinguals were native-like in their VOTs for /d/ and /t/. This suggests that after an L2 immersion experience, bilinguals can produce native-like VOTs even for sounds that are similar in their two languages. In terms of formant frequencies, while F1 does not seem to pose much difficulty for bilinguals (they were able to produce native-like F1 frequencies for all vowels), the bilinguals’ F2 frequencies were native-like for the new vowels /y/ and /œ/, but not for the similar vowel /u/. These results indicate that bilinguals seem to be affected by similarity of sounds across languages, but that this similarity seems to affect the various acoustic cues to sound identity differently.

Furthermore, the bilinguals did not show any changes in their sound productions across the first nine months after the end of the L2 immersion period. The pronunciation skills bilinguals acquire during a study-abroad stay seem to be robust against attrition for at least the first nine months after the bilinguals' return to an L1 environment.

4.3 Experiment 4: Lexical Retention

4.3.1 Introduction

Experiment 4 set out to explore retention versus attrition of lexical skills in speech production, which was assessed through a picture naming task. Participants were asked to name pictures of common items in French. Each target picture showed either a French-English cognate word or a French noncognate. Only the bilinguals and the French native speakers participated in this task. To maximize the effect any changes across time might have on the participants' productions, only Time 2 productions were compared to Time 0 productions.

4.3.2 Methods

Participants

The same participants as in Experiment 3 took part in Experiment 4, i.e., the eight bilinguals who were no longer exposed to French and who participated at Time 2 as well as the eight native French speakers. Since the picture naming task required lexical knowledge, the native English speakers did not participate in this experiment.

Materials

The stimuli in the picture naming task consisted of 40 pictures that depicted 20 French noncognates and 20 French-English cognates. The pictures were black and white line drawings

with gray shadings. The cognates and noncognates were matched for word frequency (cognates *M* = 41.50, *SD* = 57.03; noncognates *M* = 52.38, *SD* = 109.37), word-initial segment overlap between matched cognates and noncognates (always at least the first segment, where possible also the second and third; *M* = 1.55 of phonemes shared across first three positions, *SD* = 0.51) and syllable length (cognates *M* = 1.7, *SD* = 0.66; noncognates *M* = 1.5, *SD* = 0.51). The forty words in the picture naming task had already appeared in the lexical decision task (either as target items or fillers). The experiment also included ten pictures showing false friends, but these were not included in the analyses described below. A few example cognates and noncognates are shown in Table 6. The complete list of target items can be found in Appendix A.

	Example words	
Lexical Status	Word	Translation
Cognate	bananes	bananas
	cactus	cactus
Noncognate	tête	head
	gants	gloves

Table 6. Example target words for Experiment 4.

Procedure

Participants were seated in front of a computer in a sound-attenuated booth. The pictures were presented to them using Superlab 4 software. Instructions for the task were presented in writing on the computer screen, and they were also read aloud by the experimenter. As in the previous experiments, instructions were given in English so that participants with a low proficiency in French would be able to perform the task correctly. After having been presented with visual and spoken instructions, participants completed four practice trials with stimulus

pictures that were not included in the actual experiment. After the practice trials, participants were asked to ask any questions, and if they had no more questions, they pressed a key on the keyboard to start the experiment.

Each trial started with a cross-hair in the center of the screen which remained there for 500 ms, at which time it was replaced by the target picture. Participants were instructed to say the name of the picture as quickly and accurately as possible into the microphone in front of them. They were instructed to say the word in isolation, i.e., without a determiner, to avoid hesitation sounds like "um" and to say "Je ne sais pas," or "I don't know," if they didn't recognize a picture or did not know its French name. The experiment was self paced, i.e., participants clicked the space bar after they had named a picture to move on to the next trial. The experimental trials were presented to the participants in random order and each picture was shown to the participants twice. The picture naming task lasted about 10 minutes.

Coding

Superlab provided the picture onset times for each trial. Reaction times were calculated by subtracting the picture onset time for each trial from the time at the onset of the sound wave for each word production. Furthermore, participants' responses were hand-labeled as correct or incorrect by listening to the sound recordings. A response was considered to be correct when no more than one sound was missing from a word, added to it, or uttered in the wrong place. Participants' accuracy rate was calculated as the proportion of correct responses divided by the total number of responses.

The reaction times were recorded in milliseconds. Since reaction times between participants and test sessions differed by more than an order of magnitude, they were log-transformed and the log-transformed scores were used in the RT-analyses reported below. All trials with reaction times shorter than 5 milliseconds were removed. RTs were only analyzed for correct trials. As in the previous experiments, categorical data were contrasts coded.

Analysis

Logistic regressions with Female as a fixed effect and by-subject random slopes for Time and by-item random slopes for Bilingualism and Time were performed on the accuracy rates and linear mixed effects regressions on the reaction times. As in the previous experiments, maximal and reduced models were compared using likelihood ratio tests. For each component of the model, the estimate (*Est.*) and standard error (*SE*) of the maximal model are reported, as well as the result of the likelihood ratio test for the corresponding reduced model (X^2(df) with the associated *p*-value). Analyses were performed on the following independent variables:

- *Bilingualism* (French native speaker, Bilingual)
- *Time* (Time 0, Time 2)

4.3.3 Results

Baseline performance at Time 0

Accuracy:

At Time 0, there were significant main effects of Bilingualism (*Est.* = -0.24, *SE* = 0.07, $X^2(1) = 10.81$, $p < 0.01$) with French native speakers being more accurate overall than bilinguals, and Cognate Status (*Est.* = 0.17, *SE* = 0.07, $X^2(1) = 5.37$, $p < 0.05$) with participants being more

accurate at naming cognates than noncognates. The main effect of Repetition was marginally significant (*Est.* = 0.03, *SE* = 0.02, $X^2(1) = 3.03$, $p = 0.08$), with participants being more accurate at Repetition 2 than Repetition 1. Furthermore, the interaction of Bilingualism and Cognate Status was significant (*Est.* = 0.35, *SE* = 0.11, $X^2(1) = 9.75$, $p < 0.01$; see Figure 15), revealing that French native speakers were equally accurate at naming cognates and noncognates, while bilinguals were more accurate at cognates than noncognates. Finally, the interaction of Bilingualism and Repetition was marginally significant (*Est.* = 0.06, *SE* = 0.04, $X^2(1) = 2.95$, $p = 0.09$), with bilinguals showing a bigger improvement from Repetition 1 to Repetition 2 than the native French speakers.

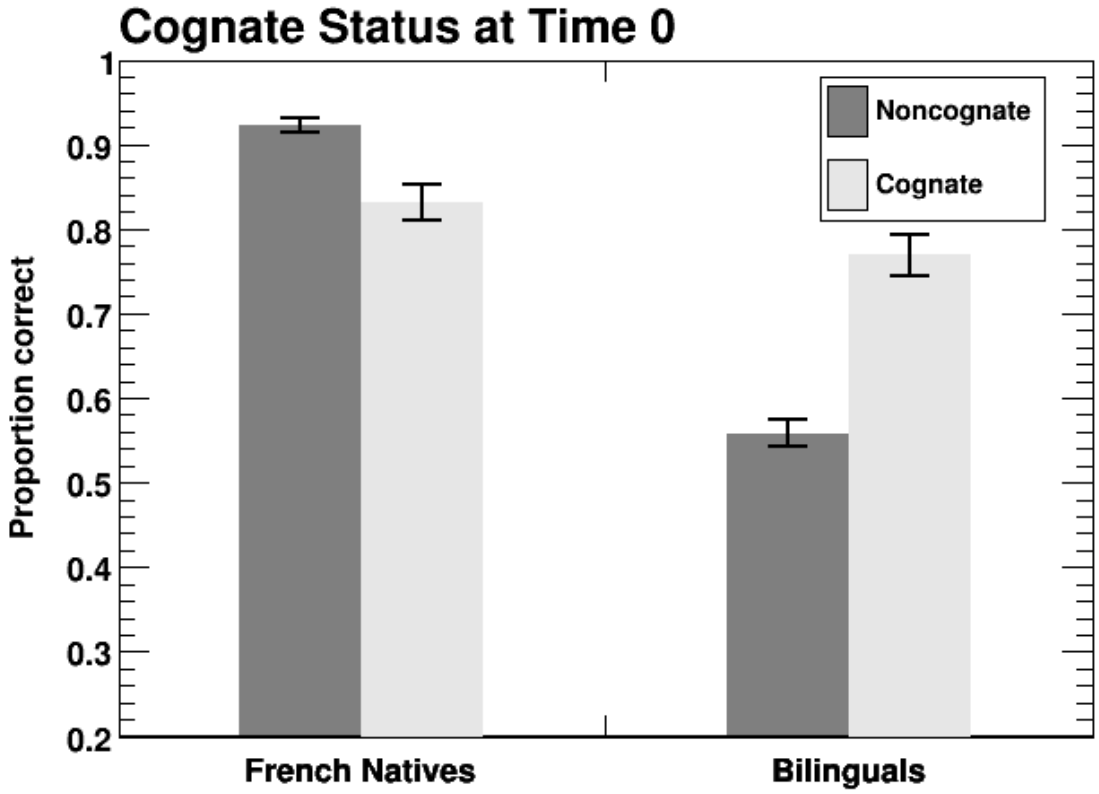

Figure 15. Significant interaction of Bilingualism and Cognate Status. Mean accuracy rate for each participant group at Time 0, separated by Cognate Status. Error bars denote standard error.

Reaction time:

There was a marginally significant main effect of Cognate Status (*Est.* = -0.07, *SE* = 0.04, $X^2(1) = 3.53$, $p = 0.06$), with participants being faster at identifying cognates than noncognates, and the interaction of Cognate Status and Bilingualism was significant (*Est.* = -0.12, *SE* = 0.05, $X^2(1) = 6.62$, $p < 0.05$; see Figure 16), revealing that the main effect was driven by the bilinguals being faster at identifying cognates than noncognates, whereas no such difference was found for

the native French speakers. The main effect of Repetition was also significant (*Est.* = -0.28, *SE* = 0.04, $X^2(1) = 25.19$, $p < 0.0001$), with participants being faster at Repetition 2 than Repetition 1.

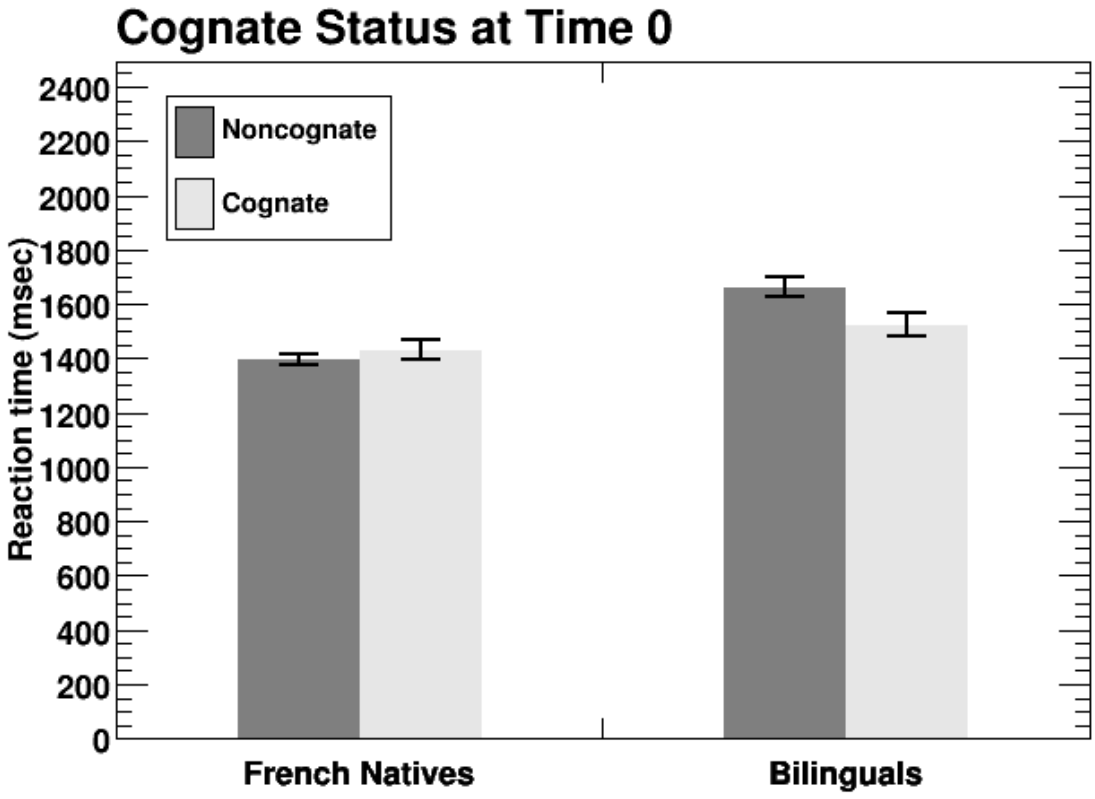

Figure 16. Significant interaction of Bilingualism and Cognate Status. Mean reaction times for each participant group at Time 0, separated by Cognate Status. Error bars denote standard error.

Changes across the first nine months (Time 0 versus Time 2)

Accuracy:

There was a significant main effect of Time (*Est.* = 0.07, *SE* = 0.02, $X^2(1) = 6.18$, $p < 0.05$), with bilinguals being more accurate at Time 2 than Time 0. Furthermore, there was a significant main effect of Cognate Status (*Est.* = 0.37, *SE* = 0.10, $X^2(1) = 12.76$, $p < 0.001$; see

Figure 17), with bilinguals being more accurate at naming cognates than noncognates, and a significant main effect of Repetition (*Est.* = 0.05, *SE* = 0.02, $X^2(1) = 5.39$, $p < 0.05$), with bilinguals being more accurate at Repetition 2 than Repetition 1. No significant interactions were found ($X^2(1)$s < 0.97, $p > 0.1$).

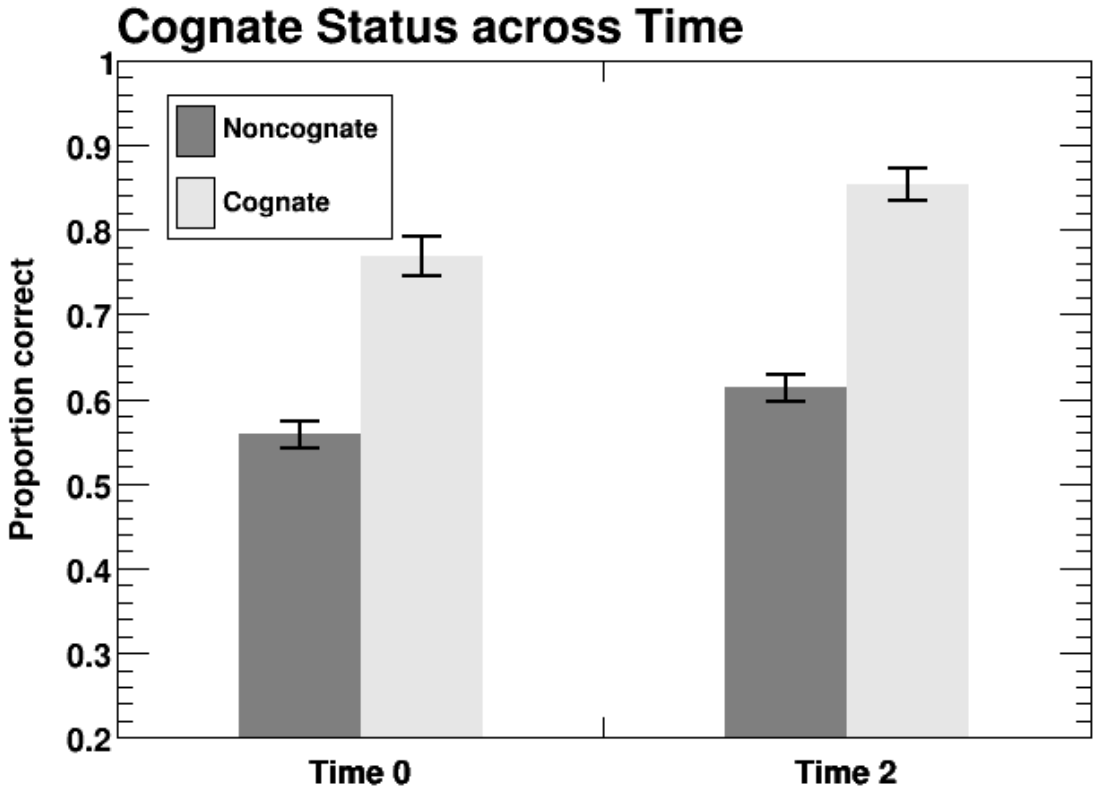

Figure 17. Significant main effects of Time and Cognate Status. Mean accuracy rate for bilinguals at Time 0 and Time 1, separated by Cognate Status. Error bars denote standard error.

Reaction time:

There was a significant main effect of Time (*Est.* = -0.14, *SE* = 0.04, $X^2(1) = 10.11$, $p <$ 0.01), with bilinguals being faster at Time 2 than Time 0, and significant main effect of Cognate

Status (*Est.* = -0.11, *SE* = 0.04, $X^2(1) = 5.73$, $p < 0.05$; see Figure 18), with bilinguals being faster at naming cognates than noncognates, and Repetition (*Est.* = -0.28, *SE* = 0.03, $X^2(1) = 20.87$, $p < 0.0001$), with bilinguals being faster at Repetition 2 than Repetition 1. However, there were no significant interactions ($X^2(1)$s < 0.20, $p > 0.1$).

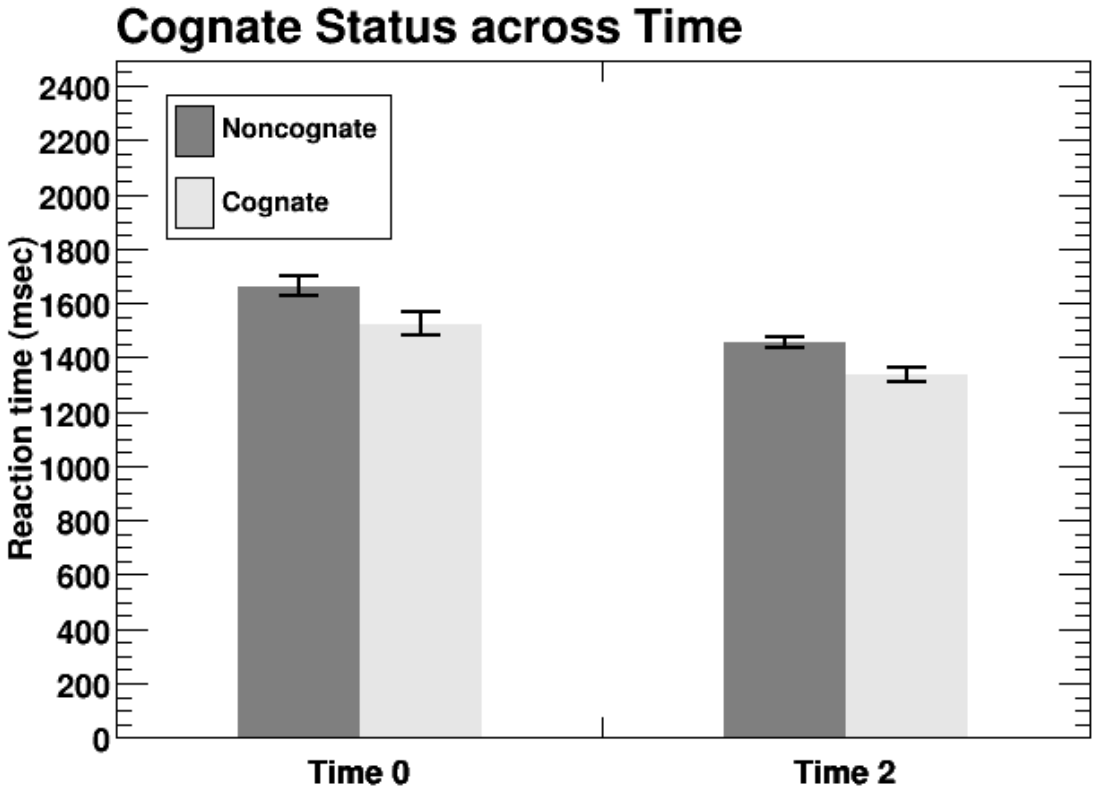

Figure 18. Significant main effects of Time and Cognate Status. Mean reaction times for bilinguals at Time 0 and Time 1, separated by Cognate Status. Error bars denote standard error.

4.3.4 Discussion

English-French bilinguals who had recently returned from an L2-immersion period in France participated in a French picture naming task in order to determine whether their

performance (accuracy and reaction times) differed from that of native French speakers, and most importantly, whether the bilinguals' productions changed across the first nine months after their return from France. Furthermore, it was tested whether cognate status of the target words influenced the bilinguals' performance.

Effects of bilingualism

A significant main effect of Bilingualism showed that the bilinguals were not native-like in their picture naming performance. The native French speakers were significantly more accurate than the bilinguals at naming pictures in French. This is in line with previous research that has shown bilingual disadvantages in various tasks such as picture naming, lexical decision, and eye fixation times (e.g., Gollan, Slattery, Goldenberg, Van Assche, Duyck, & Rayner, 2011). It is noteworthy, however, that while the bilinguals were less accurate than the native French speakers, they were not significantly slower at this task. In other words, even though the task may have been harder for them, this did not slow them down. Since the reaction times were only measured for correctly named items, this is not just a sign of a speed-accuracy trade-off, but rather suggests that when bilinguals were familiar with a French word they were able to process it equally fast as the native French speakers.

Changes across time

As in the previous experiments, rather than showing attrition within the first nine months after their return to the United States, the bilinguals actually improved across time. Their accuracy rates increased from Time 0 to Time 2, and their RTs became shorter. This was in spite of the fact that the bilinguals whose picture naming data was analyzed were no longer exposed to French. These results are in line with some previous studies on L2 attrition that have found

retention (Bahrick, 1984; Murtagh, 2003) and in the case of high-proficiency L2 speakers even improvement in lexical production (Murtagh, 2003) over a period of two or more years (up to fifty years in the case of Bahrick).

While Bahrick also found retention of productive skills across time, his results, at first, seem to be contradictory to the current results in terms of the time-course of attrition versus retention. Bahrick suggests that there is an initial period of three to six years after the end of L2 exposure during which some lexical attrition may take place and after which proficiency levels remain stable. Since in the current experiment, however, only a time period of nine months post-immersion was tested, this would coincide with the period of attrition in Bahrick's study. There are a few differences between Bahrick's study design and the current study that make a direct comparison of results impossible.

First of all, the shortest interval between the end of L2 exposure and testing in Bahrick's study was one year. Therefore, Bahrick cannot make any claims about what happens in terms of attrition versus retention in the first twelve months after the end of L2 exposure. It is possible, that further processing may have happened in Bahrick's participants during those early months, and that attrition only set in later (i.e., at some later point during the first three to six years post-exposure, which was the time frame during which Bahrick found attrition).

The time patterns of attrition versus retention may also differ in the two studies because the participants in the current study were immersed in an L2 environment for four months prior to the experiments, whereas in Bahrick's case no L2 immersion had taken place, i.e., the

participants were exposed to the L2 in classroom settings only. The amount of exposure may also affect the patterns of attrition versus retention.

Finally, Bahrick tested written production, whereas the current study tested oral production. While both written and oral tasks require lexical access, the different production modalities may be affected differently by attrition. In conclusion, Bahrick's results and the current results are not incompatible, but future research would have to investigate attrition versus retention in speech production in the same individuals starting soon after the end of exposure and lasting several years into the period of non-use to see if further improvement takes place, if and when attrition finally sets in and if proficiency starts to remain stable at a certain point in time.

Murtagh (2003) tested lexical attrition versus retention in a speech production task eighteen months after the end of L2 exposure. She found retention of lexical production skills in low to moderate proficiency speakers (skill level at baseline) and improvement in high proficiency speakers. Like Bahrick, Murtagh cannot make any claims about patterns of attrition versus retention in the first year after the end of exposure, but her data are in line with the results of the current study. In fact, Murtagh's results suggest that further improvements in lexical production could be expected for up to eighteen months (Murtagh did not test participants after more than eighteen months post-exposure) after the end of immersion. Again, future research testing speech production in the same individuals at different points in time throughout the first eighteen months or more after the end of L2 immersion could test whether there are indeed further improvements beyond nine months post-immersion.

Effects of cognate status

The current study also investigated whether the bilinguals treated cognates differently from noncognates in picture naming, and whether cognate status of the target words influenced patterns of attrition versus retention. First of all, the bilinguals were more accurate and had shorter RTs for cognates than noncognates, whereas the native French speakers showed no such difference.

The fact that the native French speakers, who are French dominant bilinguals, did not show a cognate effect is in line with previous research which has shown that bilinguals are less likely to show a cognate effect when naming pictures in their L1 rather than their L2 (Costa, Caramazza, & Sebastián-Gallés, 2000). It has been argued that words in the L2 are less strongly activated than the corresponding L1 words (Kroll & Stewart, 1994), which makes it less likely for cognates to exert an effect on L1 lexical processing, where target language words are more strongly activated than in L2 processing where they are less strongly activated.

Having argued that the native French speakers were less likely to show a cognate facilitation effect because they performed the picture naming task in their L1, the same reasons apply that make cognate facilitation likely for the bilinguals who performed the task in the L2. And it was in fact the case that the bilinguals in the current study were more accurate and faster at naming cognates in French than noncognates. This is in line with a growing number of bilingual speech production studies that have found cognate facilitation in picture naming (Costa et al., 2000), fewer tip-of-the tongue states for cognates than noncognates (Gollan & Acenas, 2004), and word naming (Jared, Friesen, & Haigh, 2008).

Cognate status did not influence the changes that took place across time in the bilinguals' productions. The bilinguals improved their naming accuracy and speed for cognates as well as noncognates to a similar degree. While the bilinguals did not show an additional cognate facilitation effect which would have led them to improve more for cognates than noncognates, they did still experience cognate facilitation at Time 2 in that they were still more accurate and faster at naming cognates than noncognates nine months after the end of L2 immersion. This suggests that cognate facilitation is a robust effect that persists even in the absence of L2 input – at least given that further L2 processing is still taking place, as was the case for the bilinguals in the current study.

4.3.5 Conclusion

In a picture naming task it was shown that while bilinguals are not quite native-like even after a four-month immersion period, they seem to improve in accuracy and speed during the first nine months after the end of the immersion period. This is the case even in the absence of further L2 input. Furthermore, the bilinguals in this study experienced a cognate facilitation effect in that they were faster and more accurate at naming French-English cognates than French noncognates. These results are in line with previous research in L2 attrition in speech production and with research on bilingual speech production more generally.

4.4 General Discussion

When comparing bilinguals to native speakers of the L2, and when looking at changes across time in bilingual speakers, one would expect differences between word repetition and picture naming, since the processing levels involved in the completion of each task are not the same. Specifically, when naming a picture, a set of lexical candidates for that picture

(supposedly for most pictures there are several options for what to call the item in the picture) are activated, and the most appropriate target item has to be selected from them. Once a target item is selected, first, phonological encoding, and subsequently, phonetic encoding of that item takes place (Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991). The stages are not necessarily discrete, i.e., some feedback between the different processing levels is assumed to take place (Costa, Roelstraete, & Hartsuiker, 2006). Finally, the word is articulated.

In a word repetition task, the input is not a nonlinguistic item that requires the speaker to search for lexical candidates, but rather, the target word is already presented to the speaker in auditory form. The first step in processing is, therefore, an auditory analysis of the item and its recognition. Recognition can include recognition of the item as a lexical item in the case of a real word that is familiar to the listener, or it can simply refer to the recognition of the phonological string in the case of a nonword or a word that is not familiar to the listener (Klein, Watkins, Zatorre, & Milner, 2006). The following steps would be identical to the ones described for picture naming, i.e., phonological and phonetic encoding and articulation. Crucially, in word repetition, unlike in picture naming, a speaker does not have to be familiar with a target word's meaning to complete the task (Nozari et al., 2010).

In addition to the different processing levels involved, the variables of interest were different in the two tasks, as well. In word repetition, acoustic analyses of various cues to sound identity (duration, formant frequencies) were performed, i.e., the bilinguals' performance at a fine-grained phonetic level was measured. In picture naming, on the other hand, accuracy and reaction times were measured, which do not take the phonetic level into consideration at all.

The fact that different processing levels were involved in the two tasks and that the variables of interest were different, made it likely that the bilinguals' performance compared to native French speakers and changes in performance across time would not be parallel for word repetition and picture naming. This was, indeed, the case.

In word repetition, the bilinguals were native-like in their productions of some sounds for particular cues. In the case of the stop consonants /d/ and /t/, for which only VOT was measured, the bilinguals were completely native-like, whereas in the case of vowels, for which F1 and F2 were measured, the bilinguals were not completely native-like for all of them. The fact that the bilinguals were native-like for some sounds, however, suggest, that it is indeed possible for fluent bilinguals to achieve native-like pronunciation after an L2 immersion experience (at least from what can be judged from the cues that were included in the present analysis).

In picture naming, the bilinguals were not completely native-like in their performance. While they were not significantly slower than the native French speakers in their responses, they were less accurate. Selection of a lexical target item, which was not necessary in word repetition, therefore, seemed to be more difficult for the bilinguals than the phonetic processes involved in production of the target words and the actual articulation at the phonetic level. On the other hand, lexical processing improved across time for the bilinguals, while their pronunciation accuracy did not change across time. It seems, therefore, that after the end of an L2 immersion period further processing at the lexical level may take place, whereas phonetic processing skills remain stable across the first nine months post-immersion.

In future studies, it would be interesting to investigate the influence of the different processing levels on performance across time in more detail. For example, in the current study, acoustic measurements were only taken for the items in the repetition task, but not for the items in the picture naming task. It may be the case, however, that the bilinguals' pronunciation accuracy is not the same if they have to also access the lexical processing level in order to produce target words. In other words, it is possible that bilinguals are less native-like in their pronunciation if the task requires a higher processing load when also having to make a lexical selection rather than just repeating a sound string (which may be meaningful to the speaker or not).

Chapter 5: The influence of individual differences between speakers on improvements across time

5.1 Introduction

The previous two chapters described the performance of a group of English-French bilinguals who had recently returned from a study-abroad stay in France in a number of French speech perception and production tasks. For three of the tasks, namely AX discrimination, lexical decision, and picture naming, the bilinguals showed improvements in their performance across the first nine months after the end of the L2 immersion period. In the previous chapters, the bilinguals were treated as a fairly homogeneous group. The only difference between bilingual participants that was taken into consideration in the analyses so far was, whether the bilinguals continued taking formal French classes after their return from France or, alternatively, whether they were no longer exposed to French in any formal way.

As with any group of people, however, other differences between the bilinguals existed as well, including their level of L2 proficiency, their use of French during and after their stay in France, their attitudes towards learning French, and differences in cognitive skills. In the present section, these individual differences as well as any influence they may have on the improvements that have been found in the bilinguals' performance across time will be investigated in more detail.

Since improvements across time were found for some of the experiments in the previous two chapters, it will be informative to look at previous research in L2 acquisition to decide which individual differences between learners may have an impact on improvements across time, even

though only some of the bilinguals in the present study continued taking French classes after their return from France.

This section continues with a brief presentation of individual differences that have been found to influence L2 acquisition. Next, an overview of individual differences investigated in L2 attrition studies will be given. Finally, an analysis of individual differences between the participants in the current study and their impact on performance changes across time will be presented and discussed.

5.2 Individual differences in L2 acquisition

5.2.1 Age of acquisition

The age at the onset of L2 acquisition has frequently been looked at as a determining factor in the ultimately attained proficiency level in L2 speakers. In fact, a critical period (or sometimes less strictly formulated, a sensitive period) has been suggested in which a language has to be acquired if a learner wants to attain native-like proficiency (Johnson & Newport, 1989; Oyama, 1976). While the term critical period implies that native-like attainment is impossible after the end of this period, a less strict view is that high proficiency levels can still be attained even by older learners (cf. Birdsong, 2007, who found native-like pronunciation of vowels and consonants by late L2 learners of French), but that there seems to be a linear trend of progressively lower ultimate attainment with higher age of acquisition (e.g., Oyama, 1976; Flege, Yeni-Komshian, & Liu, 1999b). These claims have been made especially with regard to grammar (Johnson & Newport, 1989) and phonological processing (e.g., Flege et al., 1999b). A critical period has not generally been associated with lexical acquisition. Even late learners are

normally able to acquire L2 vocabulary successfully (although word learning difficulties have been found in children who received cochlear implants after early childhood; Lederberg & Spencer, 2005). For lexical processing, the age of acquisition of individual words seems to matter more than the age of acquisition of the L2 in general (Morrison, Ellis, & Quinlan, 1992).

It should also be noted that in most studies, age of acquisition refers to the age of immigration to an L2 country, rather than the age at which learners were first exposed to the L2 in their home country. Since in the current study, the age of immigration to France was uniform across all bilinguals (i.e., they were all 20 or 21 years old at the start of the immersion period), it was decided to instead investigate whether the age at which participants first were exposed to French (whether in a classroom setting or otherwise) influenced their performance on the linguistic tasks in the current study.

5.2.2 Amount of language use

Previous research has found that bilinguals in an L2 environment tend to have better accents in the L2 and tend to be better at L2 speech perception if they use their L2 more and their L1 less. Guion, Flege, and Loftin (2000) found more native-like L2 vowel productions in a group of Quichua-Spanish bilinguals who seldom used Quichua, compared to those bilinguals who used their L1 often. Flege, Mackay, and Meador (1999a) reported better English vowel discrimination in Italian-English bilinguals who rarely used their L1 Italian compared to bilinguals who used their L1 often.

Because the bilinguals in the current study lived in an L1 environment at the time of testing, their L1 use was naturally higher than that of most immigrants in an L2 country, but

differences in amount of L2 use may still have been an important factor in language retention and improvement.

5.2.3 Cognitive skills

Previous research on second language acquisition has found that certain general cognitive skills are predictive of second language performance (especially L2 vocabulary acquisition). Two such skills are phonological awareness and working memory. Phonological awareness refers to the ability to recognize and manipulate sounds and syllables (cf. Stahl & Murray, 1994), phonological working memory refers to the ability to keep auditory information in short-term memory and make it available for active manipulation to complete a task, such as repeating a string of digits (cf. Becker & Morris, 1999). Masoura and Gathercole (1999) tested Greek primary-school children who were studying English on their phonological awareness (using a nonword repetition task) and their L2 vocabulary knowledge in a translation task. The authors found that the English vocabulary measure was significantly correlated with the children's nonword repetition performance, i.e., children who were better at repeating nonwords (indicating a higher degree of phonological awareness) were also able to translate a greater number of words from Greek into English and vice versa.

Paradis (2011) found significant correlations of children's phonological working memory (assessed in a digit span test) and phonological awareness (tested in a nonword repetition task) with vocabulary size and morphological knowledge in the children's L2 English. Miyake (1998), furthermore, found a significant correlation of working memory span and sensitivity to L2 word order in Japanese adults studying English.

5.2.4 Executive function

Previous research has suggested that there is a link between certain functions of the executive system in the brain and bilingualism. These functions, which are mainly performed in the frontal regions of the brain, are mostly higher-order cognitive functions including planning, decision making, and inhibition (Gazzaniga, Ivry, & Mangun, 2002). A number of studies have found higher inhibitory control in bilinguals compared to monolinguals (e.g., Bialystok, Craik, & Luk, 2008; Blumenfeld & Marian, 2010). One account of this effect is that it derives from the need for bilinguals to continuously suppress one of their languages while using the other. This practice with suppression of a language generalizes, enhancing bilinguals' ability to inhibit irrelevant information in non-linguistic tasks. Consequently, bilinguals perform above the level of monolinguals on tasks testing inhibitory control such as the Simon task, which requires the inhibition of attention to a specific cue (Bialystok, Craik, Klein, & Viswanathan, 2004) or the Stroop task, which requires the inhibition of a habitual response (Martin-Rhee & Bialystok, 2006).

Recently, linguistic studies have used the Behavior Rating Inventory of Executive Function (BRIEF-A; Roth, Isquith, & Gioia, 2005) to test for correlations between executive functioning and language skills in clinical populations such as hearing impaired individuals (e.g., Pisoni, Conway, Kronenberger, Henning & Anaya, 2010) or individuals with specific language impairment (e.g., Hughes, 2006). Elevated BRIEF scores, compared to healthy adults, have been found in hearing impaired individuals and individuals with specific language impairment. While the BRIEF has not been previously used specifically to investigate the role of executive functioning in bilingual language processing, it measures a range of executive functions (rather

than just an individual function, such as inhibition) whose potential relationship with L2 proficiency and language retention versus attrition can be tested.

5.2.5 Motivation and attitudes towards the L2

Research on motivation in L2 acquisition is frequently based on Gardner's (1985) framework of motivation and attitudes towards an L2, which can be measured through his Attitude Motivation Test Battery (AMTB). Although this was originally devised to measure the attitudes of Canadian learners of French, it has been adapted to fit various languages and learning contexts since its inception.

Gardner identified five concepts related to motivation and attitudes towards the L2 that are hypothesized to influence L2 acquisition. *Integrativeness* is described as an openness to identify with the L2 community. The concept is related to motivation because if an individual wants to actively participate in another culture, a certain degree of knowledge of that culture's language is necessary. A second concept that Gardner identified comprises *attitudes towards the learning situation*. These include an L2 learner's perception of their language class and language teacher. Thirdly, *motivation*, in Gardner's terms, can be measured by the effort a learner puts into studying the L2, by the extent to which the learner wants to achieve a high proficiency in the language, and how much pleasure the learners gains from the language learning experience itself. Gardner's fourth concept is *instrumental orientation* which relates to pragmatic reasons for L2 acquisition (e.g., because the learner thinks the language is necessary for job success). Finally, *language anxiety* concerns apprehensions a learner may have in the language classroom or in other situations where the language is used.

A multitude of studies have used the AMTB or variations thereof and have found an influence of one or more of the abovementioned concepts on L2 acquisition (e.g., Bernaus & Gardner, 2008, for English; Muchnick & Wolfe, 1982, for Spanish; Murtagh, 2003, for Irish).

5.3 Individual differences in L2 attrition, retention, and improvement

Some of the variables presented in the previous section have already been used in L2 attrition studies which investigated the influence of these individual differences on retention versus attrition. The most frequently investigated variables are briefly summarized below.

5.3.1 Initial L2 proficiency

L2 proficiency is often the focus of L2 acquisition studies in that they investigate which factors lead to higher ultimate proficiency levels. In L2 attrition studies, on the other hand, initial proficiency, i.e., the proficiency level at the baseline test to which later performances are compared, is usually investigated as an important factor affecting patterns of attrition versus retention.

Initial L2 proficiency was found to be an indicator for the amount of attrition L2 speakers experienced across time in Bahrick (1984), where learners who had achieved a higher proficiency level during their training were able to retain vocabulary knowledge for longer than those who had only achieved a lower proficiency level (as measured by the years of L2 training they had received). Weltens (1989) found that learners who had studied French for six years in secondary school performed at a higher level at various linguistic tasks four years after the end of L2 instruction than learners who had studied French for four years. Finally, Murtagh (2003) also found that learners who had studied Irish in an immersion program in secondary school were

more likely to improve their spoken language skills than learners who had studied Irish in a less intensive program.

5.3.2 Amount of language use

While previous studies on L2 attrition have frequently tested participants with little to no exposure to the L2 since the end of the L2 instructional period (Bahrick, 1984; Weltens, 1989), L2 learners are often still exposed to the L2 in some form even after formal instruction has ended. Murtagh (2003), therefore, specifically investigated effects of continued exposure on L2 retention and attrition. She found that the amount of use of Irish since the end of secondary school was significantly correlated with the performance level in spoken Irish and with overall Irish proficiency eighteen months after the end of Irish instruction.

5.3.3 Motivation and attitudes towards the L2

Weltens (1989) assessed attitude and motivation through questionnaire items asking about participants' opinions of their language classes, their attitudes towards French and their estimation of the usefulness of French in their future. Weltens only found a weak relationship between these attitudinal factors and language retention, but he also claimed that his instruments to measure motivation and attitudes (i.e., only three question in a questionnaire) may not have been adequate to determine such a relationship.

Murtagh (2003) used a version of the AMTB (Gardner, 1985) to measure motivational strength and attitudes towards Irish in her participants, and she found that motivational strength (i.e., the amount of effort learners were willing to invest into L2 acquisition), the desire to learn Irish, and language class anxiety were significantly correlated with overall proficiency, and

speaking and listening skills in particular, eighteen months after the end of formal instruction. Students who were highly motivated and had showed a desire to learn Irish (at the time of the baseline test) were more likely to retain their skills, whereas students who experienced a lot of anxiety in the language classroom were less likely to retain their skills.

5.4 The current study

In this part of the study, twenty variables that represent individual differences between the participants in this study will be looked at in more detail in terms of their influence on the attrition and the improvement across time that has been found in the perception and production tasks. The influence of these individual differences on the improvements found in AX discrimination from Time 0 to Time 2, in lexical decision from Time 0 to Time 1, and in picture naming from Time 0 to Time 2 will be investigated.

5.4.1 Methods

Design

In addition to the linguistic tasks described in the previous two sections, bilinguals filled out a number of questionnaires and performed a few cognitive tasks at each of the three testing sessions. From the data collected through these means, twenty variables were selected that, based on findings from previous research on L2 acquisition and L2 attrition, were likely to have had an influence on the bilinguals' performance across time. In order to reduce the number of variables to a number that is manageable for a linear mixed effects regression, the parameters were combined into sets of highly correlated variables (i.e., factors) through an exploratory factor analysis. These factors were included as independent variables in regression models to determine

their influence on the changes across time found for d' in the AX discrimination task, and for the changes found for accuracy in the lexical decision and picture naming tasks.

Procedure and variables

At each of the three experimental sessions, bilinguals filled out two questionnaires on their language background (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007; and the Study Abroad Questionnaire, SAQ, designed for this study – see Appendix B for the SAQ). The LEAP-Q assessed the bilinguals' L2 acquisition history, their proficiency levels for listening, speaking, reading and writing at Time 0, Time 1, and Time 2, the amount of L2 use after their study-abroad experience, knowledge of other languages, and the amount of time spent in an L2 environment. The SAQ adapted some of the questions from the LEAP-Q (which are aimed at proficiency and exposure at the time of testing) to inquire about proficiency levels and amount of use during the study-abroad program and at the end of the program (i.e., before any attrition would have set in). Furthermore, the SAQ assessed the duration and type of formal instruction received (e.g., high school versus college level French classes), travel to French-speaking countries, and the amount and type of exposure to French since the end of the study-abroad program. From these questionnaires, the following individual differences variables were included in the analysis:

- Variables related to the amount of French training received/proficiency level attained *before* the start of the study-abroad program: age of acquisition in years, number of years of French-study before the study-abroad stay, number of years of college level French-instruction;

- Variables related to the amount of L2 use *during* and *after* the study-abroad stay: percent of French use (vs. other languages) during the stay in France, percent of French use at the time of the baseline test (Time 0), amount of French use (listening) at Time 1 (on a scale from 0 = none to 10 = always), amount of French use (speaking) at Time 1;

- Variables related to the initial proficiency level (all on a scale from 0 = none to 10 = perfect): self-reported proficiency in L2 listening right before the study-abroad stay, self-reported proficiency in L2 speaking right before the start of the study-abroad stay, proficiency in listening at the end of the study-abroad stay, proficiency in speaking at the end of the study-abroad stay, proficiency in listening at Time 0 (two months after the end of the study-abroad stay), proficiency in speaking at Time 0;

Bilinguals also completed the mini-Attitude Motivation Test Battery (AMTB; Gardner & McIntyre, 1993), which included 11 questions that targeted four of the five attitudinal-motivational concepts described above (integrativeness, motivation, instrumental orientation, language anxiety). Classroom anxiety was not included in the current study, since not all participants were taking French classes, and thus no meaningful comparisons of this variable across participants could have been made. Each of the questions was rated by participants on a scale from 1 to 7 (depending on how much they identified with a given statement) and an average score for each of the four concepts was calculated from participants' ratings on the individual questions. The variables from the AMTB included in the analyses were:

- Integrativeness
- Motivational strength

- Instrumental orientation
- Language anxiety

Since previous research has shown an influence of certain cognitive factors on L2 acquisition (e.g., Masoura & Gathercole, 1999), at each experimental session the bilinguals also completed two cognitive tasks from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) measuring their Phonological Memory Quotient (Memory for Digits task), and their Phonological Awareness Quotient (Blending Nonwords task). In the Memory for Digits task, participants were played a recording of twenty strings of digits that became progressively longer (ranging from two to eight digits per string), and they were asked to repeat each string back to the experimenter. If a participant repeated all twenty strings correctly, they received a score of 20 on this task. The raw scores were normalized by age (based on the CTOPP's normalization scale) and the resulting phonological memory quotient was entered into the analyses. In the Blending Nonwords task, participants were played a recording in which a speaker said individual sounds or sound clusters (ranging from two to eight sounds/clusters) that participants had to string together and pronounce the nonword formed by these sounds (e.g., participants would hear [ʃ – ɔ – b – oʊ] and they would say [ʃɔboʊ]). Participants received a score of 18 if they pronounced all eighteen words correctly. The scores were age-normalized and the resulting phonological awareness quotient was entered into the analyses. The cognitive variables included in the analyses were:

- Working memory span (i.e., phonological memory quotient)
- Phonological awareness (i.e., phonological awareness quotient)

Finally, bilinguals also completed the Behavior Rating Inventory of Executive Function-A (BRIEF-A; Roth, Isquith, & Gioia, 2005) questionnaire each time. The BRIEF-A is a self-report rating scale of a number of executive functions that can be administered to adults over the age of 18. Participants rate their own behavior in 75 real-world situations as applying to them never, sometimes or often. The inventory comprises nine clinical subscales that get combined into three composite scores: the subscales for inhibition, shifting, emotional control, and self-monitoring make up the composite score for the Behavioral Regulation Index (BRI), and the subscales for Initiation, Working Memory, Planning and Organization, Task Monitoring, and Organization of Materials make up the composite score for the Metacognition Index (MI). The BRI and MI scores in turn get combined into the Global Executive Composite (GEC). Higher scores are indicative of problematic behavior. The inventory is age-normed for 18-90 year old participants and scores are considered elevated if they are above the mean of 50, and they are considered to lie in the clinical range if they are above 65. The variable included in the analyses is:

- Executive function (i.e., BRIEF Global Executive Composite)

The variables included in the analyses are listed in Table 7, and descriptive statistics for each variable are presented.

Variable	Code	Mean	SD	Min	Max
Age of acquisition	AOA	13.60	4.24	1	20
Years of French before study-abroad	Yrs_Before	6.02	3.56	0	15
Years of French in college	Yrs_College	1.28	0.90	0	3
Percent French usage during study-abroad	%_During	54.30	20.41	15	90
Percent French usage at Time 0	%_T0	6.63	5.52	0	25
Use at Time 0 (Listening)[1]	Exp_Listen_T0	1.40	0.89	0	3
Use at Time 0 (Speaking)[1]	Exp_Speak_T0	1.37	0.93	0	3
Listening proficiency before study-abroad[2]	Prof_Listen_Bef	4.60	2.43	0	9
Speaking proficiency before study-abroad[2]	Prof_Speak_Bef	4.13	2.10	0	8
Listening proficiency at end of study-abroad[2]	Prof_Listen_End	7.33	2.09	2	10
Speaking proficiency at end of study-abroad[2]	Prof_Speak_End	6.67	1.99	2	9
Listening proficiency at Time 0[2]	Prof_Listen_T0	6.53	2.06	1	9
Speaking proficiency at Time 0[2]	Prof_Speak_T0	5.77	2.18	1	9
Working memory span (CTOPP)	WorkMem	13.17	1.93	8	16
Phonological awareness (CTOPP)	PhonAware	13.90	1.63	10	17
Executive Function (BRIEF GEC)	BRIEF_GEC	52.13	7.39	37	66
Integrativeness (AMTB)[3]	Integrative	5.57	0.86	3.7	7
Motivation (AMTB)[3]	Motivation	5.38	1.38	1	7
Instrumental orientation (AMTB)[3]	Instrumental	3.50	1.87	1	7
Language anxiety (AMTB)[3]	Anxiety	4.37	1.45	2	7

Table 7. Descriptive statistics of individual differences variables.

[1] Self-reported rating on a scale from 0 (never) to 10 (always).

[2] Self-reported rating on a scale from 0 (none) to 10 (perfect).

[3] Self-reported rating on a scale from 1 to 7.

5.4.2 Correlations among variables

Since several of the variables referred to similar concepts (e.g., those that are related to L2 proficiency and use, or to motivational and attitudinal factors), a high degree of collinearity among the variables was expected. A correlation matrix (Pearson *r*) among the mean values of the variables was calculated and is presented in Table 8.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	AoA																			
2	Yrs_Before	**-.87**																		
3	Yrs_College	**-.74**	**.90**																	
4	%_During	-.18	.33	**.39**																
5	%_T0	**-.45**	**.45**	**.43**	.26															
6	Exp_Listen_T1	-.20	.22	.32	.15	**.49**														
7	Exp_Speak_T1	-.20	.24	.32	.23	**.44**	**.86**													
8	Prof_Listen_Bef	**-.69**	**.76**	**.66**	.26	**.43**	.12	.23												
9	Prof_Speak_Bef	**-.58**	**.72**	**.62**	.17	**.41**	.24	.33	**.88**											
10	Prof_Listen_End	**-.66**	**.77**	**.69**	**.40**	**.46**	.19	.32	**.81**	**.69**										
11	Prof_Speak_End	**-.62**	**.79**	**.70**	**.38**	**.49**	.19	.28	**.85**	**.81**	**.94**									
12	Prof_Listen_T0	**-.63**	**.77**	**.73**	**.43**	**.44**	.30	**.40**	**.84**	**.74**	**.90**	**.88**								
13	Prof_Speak_T0	**-.52**	**.72**	**.69**	**.46**	**.54**	.33	**.43**	**.75**	**.74**	**.87**	**.90**	**.91**							
14	WorkMem	-.05	-.07	-.12	-.17	-.27	-.25	-.12	.18	.14	.01	-.01	.10	-.03						
15	PhonAware	.08	.12	.14	.35	-.20	-.14	-.01	.16	.14	.09	.14	.21	.21	**.37**					
16	BRIEF_GEC	.27	-.15	-.16	-.12	**-.37**	-.29	-.16	-.32	-.24	-.25	-.24	-.27	-.19	.12	.00				
17	Integrative	.09	-.16	.04	.16	-.13	-.17	-.21	-.22	**-.40**	-.10	-.15	-.16	-.10	.01	.14	.04			
18	Motivation	-.15	.22	.25	**.51**	-.11	-.16	-.17	.20	.11	.16	.20	.29	.22	.34	**.48**	-.13	.33		
19	Instrumental	-.23	**.40**	**.54**	.34	.11	.06	-.05	**.38**	**.38**	.29	**.41**	**.41**	**.38**	.14	.30	-.20	.24	**.60**	
20	Anxiety	-.04	.11	.00	-.35	.05	-.20	-.04	.04	.02	.06	.08	-.06	-.11	-.07	-.29	.18	-.34	**-.45**	-.22

Table 8. Correlations among individual differences variables.

Note. Correlations equal to or greater than .36 are significant at the $p < 0.05$ level (t (28) = 2.05; two-tailed).[2]

[2] With a *p*-level of 0.05 across multiple pairwise comparisons these correlations are not conservative, i.e., there is a risk of Type I (false positive) errors occurring across the large number of correlations. Rather than correcting this by adjusting the *p*-level for the correlations (and thus risking an increase in Type II – false negative – errors) a more sophisticated analytic technique, i.e., factor analysis, is introduced in the next section to investigate dependencies between variables.

The correlation matrix confirms the high degree of collinearity among the variables. This pattern is similar to the collinearity among individual differences variables reported by Murtagh (2003) in her L2 attrition study with learners of Irish. Specifically, in the current study, all variables related to French speaking and listening proficiency were significantly correlated with each other (rs > 0.68, p < 0.05). High speaking and listening proficiency (based on self-reports) before the study-abroad experience seems to be indicative of high speaking and listening proficiency at the end of the study-abroad program and the proficiency two months after the end of the study-abroad stay (at Time 0). The proficiency variables were also significantly correlated with most of the variables representing instruction and use of French before the study-abroad experience. Speaking and listening proficiency before and at the end of the study-abroad program as well as at Time 0 were negatively correlated with age of acquisition (rs < -0.51, p < 0.05), with participants who started studying French at a younger age achieving higher proficiency levels at these three times. Proficiency at these three times was also correlated with the years of French-study before the study-abroad stay and with the years of college French, in that bilinguals with more years of French and more years of French in college also achieved higher proficiency levels at these three times (rs > 0.61, p < 0.05). Furthermore, proficiency levels were correlated with the percent of French use during the study-abroad program and at Time 0 (rs > 0.37, p < 0.05). Only the speaking and listening proficiency levels before the study-abroad program were not significantly correlated with the amount of French use abroad. In other words, low proficiency speakers used French just as much while they were in France as the higher proficiency speakers. Finally, listening and speaking proficiency at Time 0 were significantly correlated with the amount of French use (listening and speaking) at Time 1 (rs >

0.39, $p < 0.05$). Bilinguals who were at a high proficiency level at Time 0 tended to use French more frequently at Time 1 than those bilinguals who were at a lower proficiency level at Time 0.

Most of the variables dealing with French use and instruction before and during the study-abroad stay and at Time 0 were also significantly correlated with each other. The bilinguals who had started studying French at a younger age had generally also studied French for more years and had taken more years of college French than bilinguals who had started studying French at a later age (rs > -0.74, $p < 0.05$). Years of French study overall were correlated with years of French study in college ($r = 0.90$, $p < 0.05$). Furthermore, bilinguals with a younger age of acquisition and more years of French study (in college and overall) tended to use French more at Time 0 than the bilinguals with less L2 experience (rs > |0.42|, $p < 0.05$). Finally, the bilinguals with more years of college French tended to use French more during their study-abroad stay ($r = 0.43$, $p < 0.05$).

The variables related to years of French study (college and overall) were also correlated with an instrumental orientation to French as an L2 (rs > 0.39, $p < 0.05$), i.e., bilinguals with more years of French study tended to think of French as more useful for their future (e.g., they deemed it important for their future career). Finally, the students who used French more during their stay in France also tended to be the more motivated students ($r = 0.51$, $p < 0.05$).

The amount of French use (listening and speaking) at Time 1, i.e., about five months after the end of the study-abroad program was significantly correlated with the amount of use at Time 0 (rs > 0.43, $p < 0.05$). Furthermore, the amount of listening and speaking at Time 1 were, not surprisingly, also correlated with each other ($r = 0.86$, $p < 0.05$).

In terms of motivation and attitudes, instrumental orientation (i.e., seeing an importance of French for one's future) was significantly correlated with speaking and listening proficiency before the study-abroad program, speaking proficiency at the end of the program, and speaking and listening proficiency at Time 0 ($rs > 0.37$, $p < 0.05$). The high proficiency bilinguals tended to be the ones who saw future importance for their French abilities. Interestingly, there was a negative correlation between speaking proficiency before the study-abroad stay and integrativeness ($r = 0.40$, $p < 0.05$). The bilinguals with low speaking proficiency at the onset of the study-abroad program tended to be the ones who were the most interested in the French culture at Time 0. Furthermore, motivation was positively correlated with instrumental orientation ($r = 0.60$, $p < 0.05$) and negatively correlated with language anxiety ($r = -0.45$, $p < 0.05$). Highly motivated bilinguals tended to see French as more useful for their future and tended to be less anxious about using the language than less motivated bilinguals.

Motivation was also correlated with phonological awareness ($r = 0.48$, $p < 0.05$). Highly motivated bilinguals tended to be the ones with high phonological awareness. This could be because students with high phonological awareness are more likely to be interested in studying foreign languages, or alternatively, putting effort into becoming fluent in another language (as is the case with highly motivated students) may improve one's phonological awareness. Phonological awareness was also correlated with working memory span ($r = 0.37$, $p < 0.05$). Bilinguals with high phonological awareness tended to also have high working memory spans, which is not that surprising given that both of these cognitive tasks require one to keep sounds or digits in memory. Finally, executive function was correlated with the percent of French use at Time 0 ($r = 0.37$, $p < 0.05$). Bilinguals who used French more at Time 0 tended to be the ones

with higher general executive composite scores (indicating a trend towards 'problematic' behavior). Why the executive function measure was only correlated with French use at Time 0 cannot clearly be explained from the current data.

5.4.3 Factor analysis

Clearly these individual difference measures are highly interrelated. They may therefore be accounted for by appealing to a much smaller number of dimensions. To achieve this, an exploratory factor analysis with orthogonal (varimax) rotation was conducted on twenty-one variables. The use of an orthogonal rotation is important if the factors are entered as independent variables into a linear mixed regression analysis, since the regression assumes that variables are orthogonal. The standard procedure of including eigenvalues[3] equal to or greater than 1.00 was used to determine the number of factors (Vogt, 2007). Three factors were extracted and factor loadings (i.e., correlations of each input variable with a factor) were reported. The variables were included in the factor with the highest factor loading (primary loading). Variables may also have secondary loadings, which are the second highest loadings that are also significant. Loadings are generally not considered significant if they are below .40 (Vogt, 2007). The three factors that were extracted by the factor analysis accounted for 58.7% of the total variance. The rotated factor matrix and the factor loadings of each variable are presented in Table 9.

---

[3] An eigenvalue is a measure related to the amount of variance explained by a factor. In the current study, since there are 21 initial variables, all eigenvalues add up to 21. A factor's eigenvalue divided by 21 is equal to the percent of variance explained by the factor. A factor with an eigenvalue under 1 cannot explain any more variance than a single input variable and therefore does not help to summarize the data (cf. Vogt, 2007).

	Factor 1	Factor 2	Factor 3
Variables	Proficiency	Exposure	Attitude
Age of acquisition	**-.68**	-.13	
Years of French before study-abroad	**.82**	.15	.10
Years of French in college	**.72**	.27	.20
Percent French usage during study-abroad	.34	.16	**.51**
Percent French usage at Time 0	**.46**	.44	-.15
Continued French	.22	**.56**	
Use at Time 1 (Listening)		**.98**	
Use at Time 1 (Speaking)	.23	**.84**	-.11
Listening proficiency before study-abroad	**.89**		
Speaking proficiency before study-abroad	**.80**	.16	
Listening proficiency at end of study-abroad	**.95**	.10	
Speaking proficiency at end of study-abroad	**.96**	.10	
Listening proficiency at Time 0	**.90**	.23	.20
Speaking proficiency at Time 0	**.80**	.26	.15
Working memory span (CTOPP)		-.24	.29
Phonological awareness (CTOPP)	.13	-.12	**.52**
Executive Function (BRIEF GEC)	-.21	-.27	-.12
Integrativeness (AMTB)	-.18	-.13	**.40**
Motivation (AMTB)	.16	-.12	**.90**
Instrumental orientation (AMTB)	.35		**.60**
Language anxiety (AMTB)	.12	-.24	**-.56**

Table 9. Factor loadings for variables.

Note. Significant primary loadings are in bold.

The first factor received loadings (.48 to .96) from variables related to length of L2 study before the immersion period and speaking and listening proficiency at various times during and after the immersion period. Based on the variables that make up Factor 1, it is best summarized as *Proficiency*. The Proficiency factor accounted for 34.2% of the total variance.

The second factor received loadings (.56 to .98) from the variables L2 use (listening) at Time 1, L2 use (speaking) at Time 1 and Continued French. The factor received a secondary loading (.44) from the variable percent French usage at Time 0. While this variable was included in the Proficiency factor, it was also fairly highly correlated with the second factor. This makes sense, since the variable also refers to the use of French after the end of the study-abroad program. Since all three variables deal with exposure to the L2 after the end of the immersion period, the factor was called *Exposure*. The Exposure factor accounted for 13.0% of the total variance.

Six variables (.40 to .90) were loaded onto the third factor, which accounted for 11.6% of the total variance. The variables included in this factor deal with motivation and attitudes (integrativeness, motivation, instrumental orientation, language anxiety), the cognitive function of phonological awareness, and the percentage of French usage during the stay in France. The loadings were somewhat lower than of the variables loaded onto the first factor, indicating that they were less highly correlated with each other and the factor as a whole. It is not surprising that the four motivational/attitudinal measures from the AMTB were grouped together. With regards to the other two variables that were loaded onto the second factor, it seemed to be the case that highly motivated L2 learners sought to use the language more when in an L2 environment and seemed to also be able to work with difficult sound patterns (as in the blending nonwords task). Motivation to be a fluent speaker of another language may very well be related to putting effort into repeating unfamiliar sound sequences (which is necessary in the early stages of L2 acquisition). Since the items holding together this factor seem to be the ones related to motivation and attitudes towards the L2, this factor was termed *Attitude*.

Two variables, namely working memory span and executive function, did not receive high loadings for any factor (i.e., were not highly correlated with any factor). This is in line with the correlation analyses presented in the previous section where working memory span only significantly correlated with phonological awareness (r = .37) and executive function was only significantly correlated with the percent of French usage at Time 0 (r = -.37). These two variables were therefore included as separate variables in the regression analyses.

5.4.4 Regressions on individual difference factors and variables

Statistical analyses of the data were performed on the data sets for which significant changes across time were found in Experiments 1 to 4. Namely, the bilinguals had shown improvements from Time 0 to Time 2 for AX discrimination, from Time 0 to Time 2 for picture naming, and from Time 0 to Time 1 for lexical decision.

For AX discrimination, a separate linear mixed effects regression was run for each factor/individual differences variable with d' as the dependent variable. Given the relatively small number of participants, it was not possible to run multiple regressions that included all of the factors simultaneously. In these regressions, the independent factors include Time plus a single factor (Proficiency, Attitude) or variable (residualized Exposure measure, Working Memory, Executive Function). The regressions also included by-subject and by-item random slopes for Time and the single factor/variable included in the regression. For lexical decision and picture naming accuracy, corresponding logistic regressions were run. A significant interaction of a factor with time is interpreted as the factor influencing the change across time in d' or accuracy. As in the previous experiments, significance of main effects and interactions was determined through likelihood ratio tests.

The Exposure factor consisted of the three variables Continued French, Exposure to French (listening) at Time 1, and Exposure to French (speaking) at Time 1. However, Continued French (i.e., the fact whether bilinguals continued taking French classes after their return to the U.S.) is a partial subset of Exposure (since overall exposure to listening and speaking in French also includes exposure received in French class). To control for this correlation, Exposure was residualized against Continued French and the resulting residualized exposure measure (henceforth *Exposure*) was used as the independent variable in the regression analysis, rather than the Exposure factor. This measure can explain the effect of exposure over and above the already assumed effect of Continued French.

For each component of the model (main effect or interaction), the estimate (*Est.*) and standard error (*SE*) of the maximal model are reported, as well as the result of the likelihood ratio test for the corresponding reduced model (X^2(df) with the associated *p*-value).

The independent variables in each analysis were Time, Factor 1 (Proficiency), Factor 2 (Attitude), the residualized Exposure measure, and Working Memory and Executive Function. The main interest of the analyses was the interaction of each factor with Time. If a significant interaction was found between Time and a factor or variable, this was interpreted as the factor or variable adding significantly to the variance, i.e., having a significant effect on the changes across time.

5.5.5 Results

Phonological changes in speech perception: AX discrimination

For d' at Time 0 versus Time 2, there was a significant main effect of Proficiency (*Est.* = 0.009, *SE* = 0.003, $X^2(1) = 5.43$, $p < 0.05$), suggesting that bilinguals who had a higher initial proficiency level were overall more sensitive to the French sound contrasts in the AX discrimination task. There were no significant main effects of Attitude ($X^2(1) = 0.39$, $p > 0.1$) and Exposure ($X^2(1) = 0.92$, $p > 0.1$), suggesting that having a positive attitude towards the L2 culture and having more exposure to the L2 in the first five months after the end of the immersion period does not necessarily lead to overall higher sensitivity to L2 sound contrasts across the first nine months after the end of immersion. However, the interaction of Time and Attitude (*Est.* = 0.007, *SE* = 0.004, $X^2(1) = 2.99$, $p = 0.08$) was marginally significant, suggesting that bilinguals who were motivated and had a positive attitude towards the L2 culture improved more from Time 0 to Time 2. Figure 19 shows a scatter plot of the change in d' from Time 0 to Time 2 and Attitude with a least-squares regression line fit to the data.

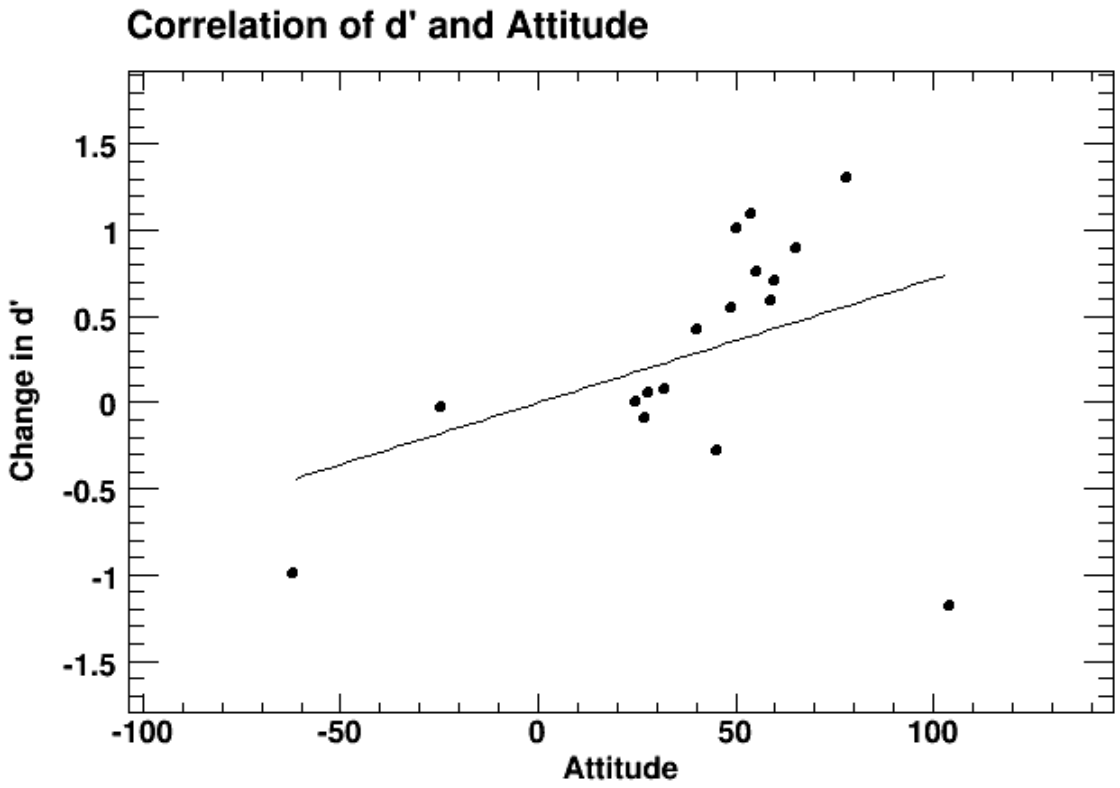

Figure 19. Significant correlation of change in d' and Attitude. Mean change in d' from Time 0 to Time 2 correlated with Attitude with least-squares regression line fit to the data.

There was also a significant interaction of Time and Exposure (*Est.* = 0.55, *SE* = 0.07, $X^2(1) = 14.38$, $p < 0.001$; see Figure 20), with bilinguals who had more exposure to French between Time 0 and Time 1 improving more by Time 2 than those who had less exposure to French.

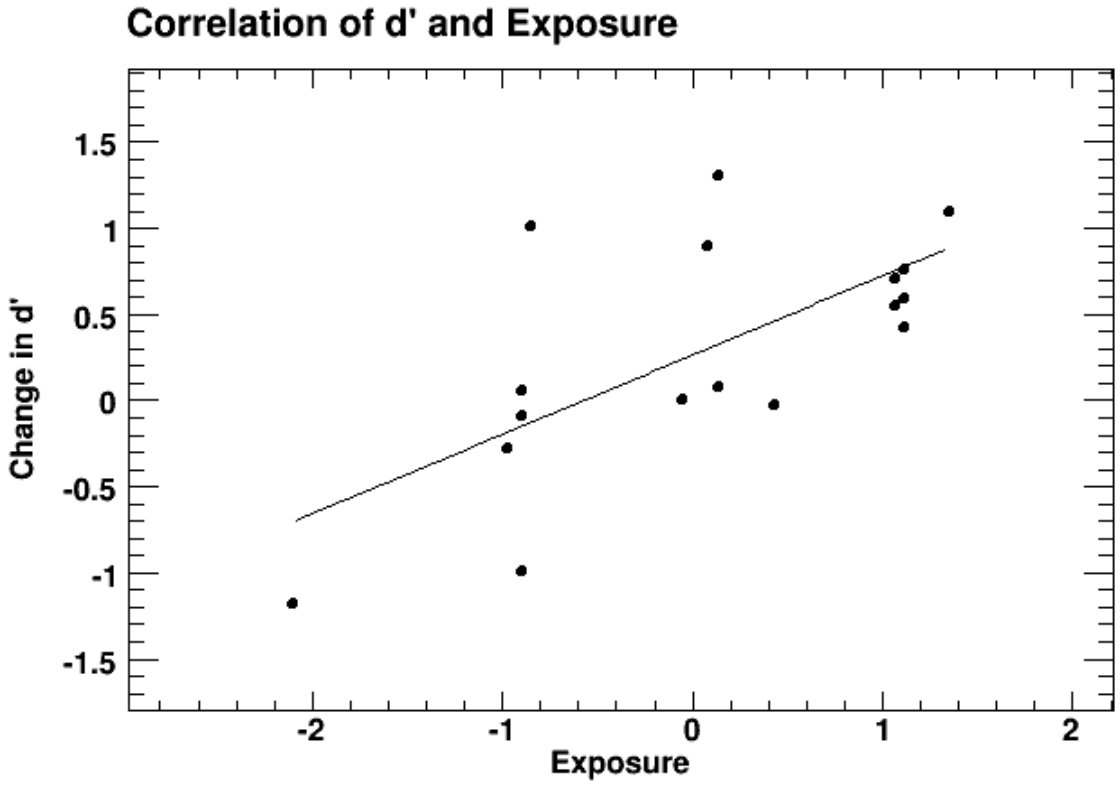

Figure 20. Significant correlation of change in d' and Exposure. Mean change in d' from Time 0 to Time 2 correlated with Exposure with least-squares regression line fit to the data.

The interaction of Time and Proficiency was not significant, however ($X^2(1) = 0.17$, $p > 0.1$), suggesting that improvement across time is likely in an AX discrimination task regardless of initial proficiency.

Finally, there were no significant main effects or interactions of Working Memory ($X^2(1)$s < 0.44, $p > 0.1$) or Executive Function ($X^2(1)$s < 0.14, $p > 0.1$), suggesting that differences in these cognitive abilities do not influence performance in an AX discrimination task in an L2 in the first nine months after the end of an immersion period.

Lexical changes in speech perception: lexical decision

There was no significant main effect of Exposure ($X^2(1) = 0.24$, $p > 0.1$), suggesting that bilinguals' exposure to an L2 after the end of an immersion period does not influence their overall performance on a lexical decision task. The main effects of Proficiency (*Est.* = 0.002, *SE* = 0.0007, $X^2(1) = 10.65$, $p < 0.01$) and Attitude (*Est.* = 0.002, *SE* = 0.0005, $X^2(1) = 5.44$, $p < 0.05$), on the other hand, were significant, suggesting that high proficiency bilinguals and those who have a positive attitude towards the L2 culture may be better at identifying French words than low proficiency bilinguals and those who have a less positive attitude.

The interaction of Time and Proficiency was not significant ($X^2(1) = 1.65$, $p > 0.1$), suggesting that bilinguals with low initial proficiency and bilinguals with high initial proficiency improved at a similar rate. The interaction of Time and Attitude, however, was marginally significant (*Est.* = 0.0006, *SE* = 0.0003, $X^2(1) = 3.78$, $p = 0.05$; see Figure 21), suggesting that bilinguals who were highly motivated and had a positive attitude towards the L2 improved their L2 lexical decision skills more than the less motivated bilinguals.

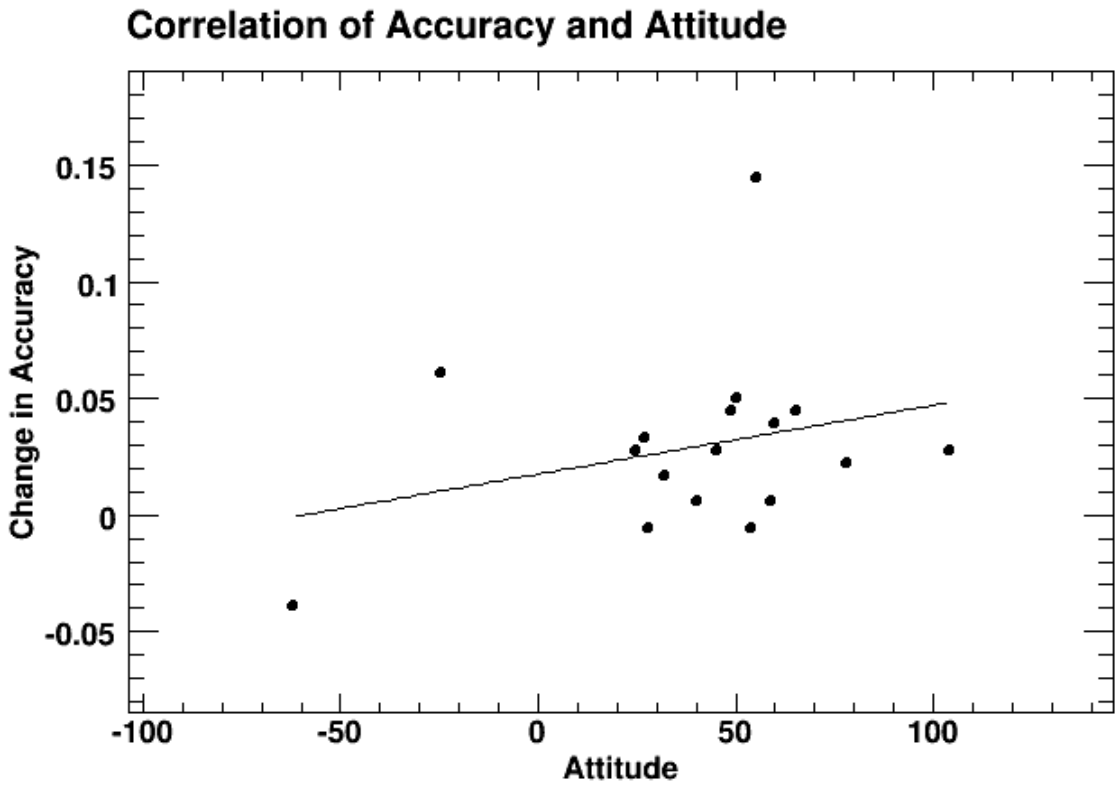

Figure 21. Significant correlation of change in accuracy and Attitude. Mean change in accuracy from Time 0 to Time 1 correlated with Attitude with least-squares regression line fit to the data.

The interaction of Time and Exposure was not significant ($X^2(1) = 1.93$, $p > 0.1$), indicating that bilinguals who had more exposure to French in the first five months after their return from France did not improve their lexical decision skills any more than bilinguals with less exposure to French.

Finally, there were no significant main effects or interactions of Working Memory ($X^2(1)$s < 0.11, $p > 0.1$) or Executive Function ($X^2(1)$s < 0.64, $p > 0.1$), suggesting that differences in these cognitive abilities do not influence performance in lexical decision task in an L2 in the first nine months after the end of an immersion period.

Lexical changes in speech production: picture naming

In order to maximize the likelihood of finding changes across time, changes across time for Time 2 versus Time 0 were investigated and only participants who did not continue taking French classes were included in the analyses.

There was a marginally significant main effect of Proficiency (*Est.* = 0.002, *SE* = 0.001, $X^2(1) = 3.47$, $p = 0.06$) and a significant main effect of Exposure (*Est.* = 0.11, *SE* = 0.05, $X^2(1) = 3.90$, $p < 0.05$), but not of Attitude ($X^2(1) = 1.90$, $p > 0.1$), suggesting that those bilinguals with higher initial proficiency and those with more exposure in the first five months after their return from the study-abroad program were overall better able to name pictures in French within the first nine months after their return home. However, whether they were highly motivated or not did not seem to influence their overall performance.

When looking at the changes across time, on the other hand, the bilinguals who were highly motivated and had a positive attitude towards the French culture improved more than the bilinguals who were less motivated, as was demonstrated by the marginally significant interaction of Time and Attitude (*Est.* = 0.0007, *SE* = 0.0004, $X^2(1) = 2.74$, $p = 0.09$; see Figure 22).

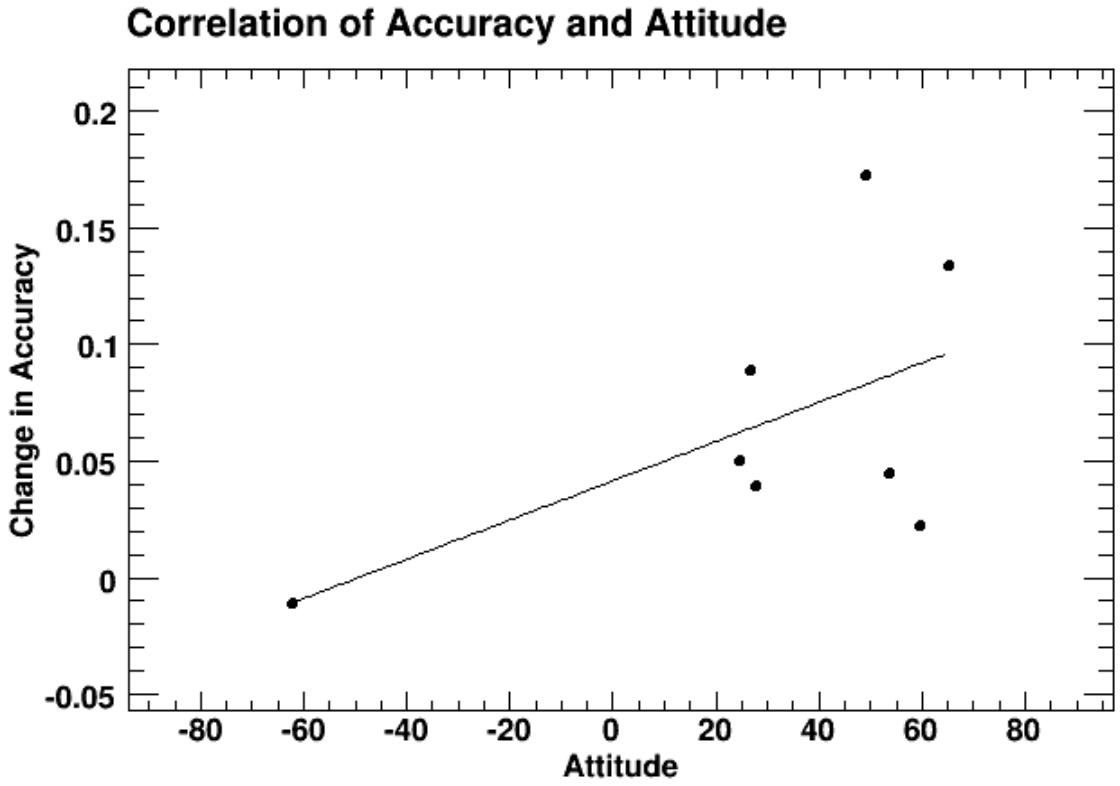

Figure 22. Significant correlation of change in accuracy and Attitude. Mean change in accuracy from Time 0 to Time 2 correlated with Attitude with least-squares regression line fit to the data.

While higher initial proficiency and more exposure at Time 1 led to better overall performance in the picture naming task, these two factors did not influence the bilinguals' changes across time, as was suggested by the nonsignificant interactions of Time and Proficiency ($X^2(1) = 0.01$, $p > 0.1$) and Time and Exposure ($X^2(1) = 0.02$, $p > 0.1$).

Furthermore, the main effect and interaction of Working Memory were not significant ($X^2(1)$s < 0.42, $p > 0.1$), suggesting that working memory span does not seem to influence performance in an L2 picture naming task in the first nine months after the end of an L2 immersion period.

Finally, the main effect of Executive Function (*Est.* = -0.01, *SE* = 0.006, $X^2(1) = 6.48$, $p < 0.05$) and the interaction of Executive Function and Time (*Est.* = -0.004, *SE* = 0.003, $X^2(1) = 7.09$, $p < 0.01$; see Figure 23) were significant, suggesting that bilinguals with higher Executive Function scores (i.e., leaning towards the 'problematic' end of the scale) were overall less accurate at picture naming in French, and in addition, improved less across time.

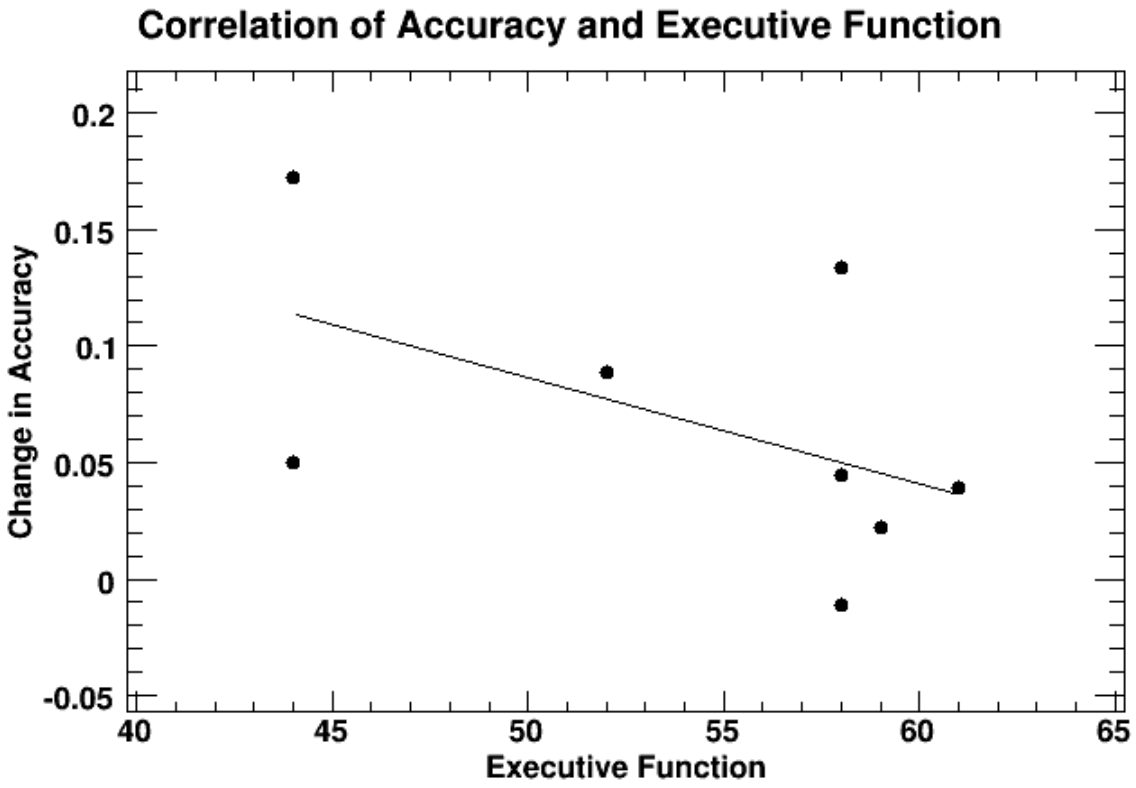

Figure 23. Significant correlation of change in accuracy and Executive Function. Mean change in accuracy from Time 0 to Time 2 correlated with Executive Function with least-squares regression line fit to the data.

5.4.3 Discussion

Individual differences variables

In the previous sections, it was investigated how bilinguals who had recently returned from a semester abroad performed in L2 phonological and lexical perception and production tasks across the first nine months after their return home. Since it is clear from previous research that a multitude of individual differences between language users influence their performance on L2 tasks (e.g., Flege, Yeni-Komshian, & Liu, 1999b; Gardner, 1985; Guion, Flege, & Loftin, 2000), and that some of these have been found to influence the retention of the L2 after the end of formal language instruction (Bahrick, 1984; Murtagh, 2003; Weltens, 1989), the present chapter investigated differences between the bilingual participants in this study in detail.

Information on individual differences between the bilinguals pertaining to age of acquisition, amount of L2 use, L2 proficiency, cognitive skills, executive function, motivation to acquire the L2, and language anxiety were collected from the participants in this study, and a correlation analysis showed that many of the variables were correlated with each other. A few groups of correlated variables emerged, which revealed that the different variables pertaining to the amount of language use before the study-abroad program were correlated with each other. It is not too surprising that these variables were not independent of each other, since, naturally, speakers who have started studying a language at an earlier age are more likely to have studied the language for longer than those who started later and would consequently achieve a higher proficiency in the language. Previous research has also shown that age of acquisition as well as amount of use influence the level of ultimate attainment in an L2 (e.g., Piske, MacKay, & Flege, 2001). Furthermore, language proficiency at different points in time (i.e., before and after the

study-abroad program) were correlated with each other, suggesting that the bilinguals who were more proficient in French before the study-abroad program were still the most proficient in speaking and listening at the end of the program and after two months back in the United States.

Furthermore, the results showed high correlation between variables related to the bilinguals' motivation to learn French as an L2, their attitudes towards the French culture, the anxiety they experience when using French, and L2 proficiency. This suggests that the more L2 learners found knowledge of the L2 useful for their future, and the more they were personally interested in the L2 culture, the more motivated they were to acquire the language and the less anxiety they experienced when using the language. Furthermore, this seemed to help them achieve a higher level of proficiency in the L2. Again, it is not surprising that being motivated to learn a language and not experiencing a great deal of anxiety when speaking it would lead to higher proficiency levels in the language.

Cognitive skills such as working memory span and phonological awareness have been shown in previous research to influence L2 acquisition. For example, children with larger working memory spans and higher phonological awareness tend to be better at acquiring new words in an L2 than children with shorter working memory spans and lower phonological awareness. In the current study, however, working memory span and phonological awareness were not significantly correlated with proficiency, i.e., the bilinguals with the largest working memory span and the highest phonological awareness quotient were not necessarily the most proficient L2 speakers (based on their self-reports of proficiency in listening and speaking). Two things to bear in mind here are that the variables pertaining to cognitive skills have been found to be most influential on vocabulary acquisition in an L2 (Masoura & Gathercole, 1999; Paradis,

2011), whereas 'speaking and listening skills' includes more than just vocabulary knowledge. For overall speaking and listening proficiency these two cognitive skills may, therefore, not be equally advantageous. Furthermore, Masoura and Gathercole (1999) also point out that the advantage seems to be greatest in early vocabulary learning, since more advanced learners receive additional help from the large body of vocabulary already in their lexicon, which makes good working memory and phonological awareness less necessary to be successful in acquiring new words. Therefore, it may also be the case that the bilinguals in the current study were too advanced learners of French to show an influence of working memory span and phonological awareness on their L2 performance.

Finally, executive function, as measured by the Global Executive Composite (GEC) taken from the Behavior Rating Inventory of Executive Function (BRIEF-A), was not correlated with proficiency, suggesting that overall executive function does not seem to influence L2 acquisition, even though individual measures (such as inhibition of irrelevant information) have been shown to be correlated with fluent bilingualism (e.g., Blumenfeld & Marian, 2010). However, as will be discussed in the section on the regression analyses, executive function was found to influence performance in one of the behavioral tasks, namely the picture naming task. One more interesting fact is that the average GEC score of the bilinguals in the current study was quite high (M = 52.13, SD = 7.39), with scores of 50 and above being considered elevated and indicative of problematic behavior. In other words, the bilingual participants in general seemed to have high GEC scores relative to the population as a whole. These results cannot easily be explained. One could speculate that students who choose to study abroad, or who choose to study French as a foreign language, are more likely to exhibit problems with behavior related to

executive function, or it could simply be the case that the population of college students from which these individuals were drawn tend to have elevated executive function scores. The fact that the GEC scores of the native English speakers in the current study (who were not fluent in any other languages) were similar to those of the bilinguals (M = 53.13, SD = 14.72; $t(36) = 0.27$, $p > 0.1$) suggests that the latter is the more likely scenario. With the data at hand, however, it is not possible to come to a conclusion about the meaning of the BRIEF scores with regards to the groups of participants. Future research would have to systematically compare different groups of participants (e.g., college students who choose to study abroad versus those that do not) to see if there are any patterns of elevated executive function scores related to L2 acquisition.

Factor analysis

To get a clearer picture of how the variables pertaining to individual differences between bilinguals are related to each other, an exploratory factor analysis was conducted on the twenty-one variables described above. Three factors were extracted that were labeled Proficiency, Attitude, and Exposure. All individual variables except working memory span and executive function (BRIEF GEC score) were assigned to one of the factors. Because working memory span and executive function received low loadings for all of the factors (which means they were not highly correlated with any of them) they were entered into the regression analyses as individual variables.

The remaining variables had high factor loadings with three factors, which suggests that the individual variables made up aspects of the concepts proficiency, attitude, and exposure, and that proficiency, attitude and exposure were the most predictive variables for overall L2 performance and change across time.

Regression analyses

Whether it is indeed the case that Proficiency, Attitude, Exposure, Working Memory, and Executive Function influence bilingual performance on L2 perception and production tasks, and most importantly, whether they have an influence on improvements across time that happen after the end of an L2 immersion period, was tested in a series of linear mixed effects regression analyses (lmers) and logistic regressions which were conducted on the changes across time in d' for AX discrimination at Time 0 versus Time 2, and on the changes in accuracy for lexical decision at Time 0 versus Time 1, and for picture naming at Time 0 versus Time 2.

The regression analyses revealed that Proficiency, Exposure and Attitude did indeed influence the bilinguals' performance in the three perception and production experiments, both in terms of overall performance and for improvement across time. Working Memory did not influence the bilinguals' performance overall or across time, while Executive Function influenced the bilinguals overall performance and the changes across time for picture naming.

a) Overall performance

Proficiency influenced how well the bilinguals were able to discriminate L2 sound contrasts, identify French words, and name pictures in French. It is not surprising that variables like age of acquisition, years of French study before the immersion program, and proficiency in speaking and listening before and after the immersion program influenced how well the bilinguals performed these L2 tasks. Naturally, higher proficiency speakers of a language should perform better on L2 listening and speaking tasks, and the current results confirm that the variables that received high loadings for the Proficiency factor indeed had this positive effect on L2 performance.

When it comes to L2 perception, as in the AX discrimination task and the lexical decision task, the fact whether bilinguals were still regularly exposed to French did not influence performance levels. This was not true for the production task, i.e., picture naming, however. Bilinguals with more exposure to French (listening and speaking) in the first five months after their return from France performed at an overall higher level than participants with less exposure to French. Note that Continued French (i.e., whether bilinguals kept taking French classes or not) was residualized out of the Exposure factor extracted from the factor analysis. Therefore, this measure represents any exposure to listening and speaking that the bilinguals received over and above French classes. While the results from Experiment 2 (lexical decision) suggested that bilinguals who continued taking French classes were overall better at identifying French words at Time 0 and Time 1 than the bilinguals who did not continue taking French classes, the same was not true for this Exposure measure. This difference could be interpreted such that overall more exposure to spoken French does not provide enough of an advantage to be more accurate at identifying French words, while taking French classes does. Students may get something out of their French classes that helps them perform better on a lexical decision task compared to bilinguals who are not taking French classes. This could, for example, be the additional exposure to reading and writing, receiving instruction from a teacher, or the learning environment in the classroom. The above results are, however, consistent with Experiment 1 (AX discrimination) for which no effect of Continued French was found for Time 0 versus Time 2. The regressions on the factors suggest that continued exposure to listening and speaking outside of the classroom is not crucial either for improvement in the perception of French. The bilinguals were able to

discriminate sound contrasts and identify French words quite accurately, regardless of whether they were still exposed to French regularly or not.

Finally, the significant main effect of Attitude found for the lexical decision task suggests that being highly motivated to acquire an L2, being interested in the L2 culture and not experiencing a lot of anxiety when using the L2, are all beneficial for lexical decision. While all bilinguals, highly motivated or not, performed at a high level in the AX discrimination task, a positive attitude towards French boosted the bilinguals' performance in the lexical decision task. In other words, having a personal interest in the L2 and identifying with the culture more seems to aid lexical processing in the language, but does not seem to be necessary for performance at a high level in sound discrimination.

The above results are in line with previous research on L2 acquisition, which has found benefits of L2 exposure (or the resulting lesser use of the L1) and of motivation and attitudes towards the L2 culture on L2 proficiency (e.g., Bernaus & Gardner, 2008; Guion, Flege, & Loftin, 2000), as well as correlations between bilingualism/L2 proficiency and executive function measures (Masoura & Gathercole, 1999; Paradis, 2011). While previous research has found effects of these factors on L2 performance in learners still in the learning process, or in bilinguals living in an L2 environment, the current results suggest that these factors influence L2 performance even up to nine months after the end of immersion to the language.

b) Improvements across time

Effects of proficiency

While more proficient speakers showed overall better performance on the three linguistic tasks, Proficiency did not seem to be an important factor for improvements across time. In all three tasks, high and low proficiency speakers improved to similar degrees across the first five to nine months after the end of the immersion period.

This is not parallel to previous findings in L2 attrition studies: Bahrick (1984) found that learners who had achieved a higher proficiency level during their training were able to retain vocabulary knowledge for longer than those who had only achieved a lower proficiency level. Similarly, Weltens (1989) found that learners who had studied French for six years in secondary school performed at a higher level at various linguistic tasks four years after the end of L2 instruction than learners who had studied French for four years. Finally, Murtagh (2003) also found that learners who had studied Irish in an immersion program in secondary school were more likely to retain their spoken language skills than learners who had studied Irish in a less intensive program.

There are three main differences between these three studies and the current study, however. Firstly, the time points at which L2 performance was measured were much earlier in the current study than in the previous studies. Bahrick did not test any learners earlier than one year after the end of L2 instruction (up to 50 years after the end of instruction), Murtagh tested her learners 18 months after the end of instruction, and Weltens tested his participants four years after the end of L2 instruction. It is, therefore, possible that proficiency may not be as influential

on L2 performance in the first few months after the end of immersion or instruction in the L2, but may later become a factor in retaining a higher level of the L2.

Secondly, the previous three studies looked at L2 learners who were living in an L1 environment and were exposed to the L2 purely in classroom settings, whereas the current study investigated L2 performance in bilinguals who had been immersed in the L2 for four months prior to the experiments. It may, therefore, be that the time course of retention versus attrition/improvement is different for language immersion versus classroom learning and that influences of proficiency on this time course show up at different points in time. In other words, highly proficient L2 learners who were exposed to their L2 in a classroom setting only may improve more than bilinguals at a lower proficiency level immediately after the end of instruction. In contrast, after an immersion period, both high and low proficiency speakers may improve at first, with only highly proficient speakers continuing to improve later; alternatively, proficiency may not play a role at all in improvement after a period of L2 immersion. Future studies would have to compare classroom learners to immersion learners throughout the first year after the end of instruction/immersion and beyond to test for these possibilities.

Thirdly, the definition of 'proficiency'[4] is not the same in the three studies. In Bahrick and Weltens, proficiency was equivalent to years of training and in Murtagh, additionally,

---

[4] Note that Bahrick, Weltens, and Murtagh did not actually call this measure "proficiency", but rather referred to it as "level of training" (Bahrick), "years of training" (Weltens), or "instructional background" (Murtagh), but these were equated with "proficiency" in the current study, since they were the closest variables to the Proficiency factor .

intensity level of training (immersion school versus regular secondary-school L2 courses). In the current study, however, the Proficiency factor was made up of age of acquisition, years of French training prior to immersion (college and overall L2 training), and self-reported proficiency in listening and speaking before immersion, at the end of the immersion period, and two months after the end of immersion. None of the previous studies had included the learners' own estimate of proficiency at various points in time into their proficiency variable, which may indeed not correspond exactly to the level of training the students had received. Had the learners' subjective proficiency judgments been taken into consideration, results in the previous studies may have been different. (Note, however, that the fact that self-reported proficiency and years of training were highly correlated in the present study suggests that these two measures do indeed affect L2 retention similarly.)

Effects of attitude

The current study also found that Attitude (comprised of degree of motivation to study the L2, attitudes towards the L2 culture, deemed importance of the L2 for one's future, anxiety when using the L2, and phonological awareness) significantly influenced the amount of improvement found in phonological and lexical perception and production tasks in L2 French. Bilinguals with a more positive Attitude score (i.e., those bilinguals who were highly motivated, less anxious about using French, deemed it more important for their future, were more interested in the French culture, and had a higher phonological awareness quotient) improved more in their L2 perception and production skills across the first five to nine months post-immersion than the bilinguals with a less favorable attitude towards the L2.

While attitude and motivation have been shown to influence L2 acquisition (e.g., Bernaus & Gardner, 2008; Gardner & McIntyre, 1993), the current study showed that the same factors influence improvement across time after the end of an immersion period. This was true for perception skills at the phonological (AX discrimination) and lexical level (lexical decision) and in production at the lexical level (picture naming). Murtagh (2003) previously investigated the influence of motivational factors on language retention. Similar to the current results, she found that learners who were highly motivated and had shown a desire to learn the L2 were more likely to retain their skills, whereas students who experienced a lot of anxiety in the language classroom were less likely to retain their skills. While Murtagh tested for attrition 18 months after the end of language instruction in secondary-school students (in listening, speaking, and global proficiency level tasks), the current study extends these results by demonstrating that similar patterns of the influence of L2 attitudes on retention are found in former immersion students, that attitude already influences changes in performance across the first five to nine months after the end of immersion, and that a positive attitude towards the L2 not only is beneficial for the retention of the L2, but also makes improvement in the L2 after the end of immersion more likely.

Effects of exposure

The regressions suggest that Exposure influenced improvement across time in sound processing (AX discrimination) but not in lexical processing (lexical decision, picture naming). At first glance, this is not in line with the results of Experiment 1 (AX discrimination) and Experiment 2 (lexical decision) in the current study. Experiment 1 found that all bilinguals, i.e., those that continued taking French classes and those that did not, improved to a similar degree

from Time 0 to Time 2. However, the factor regression suggests that the amount of exposure the bilinguals received between Time 0 and Time 1 over and above their French classes did significantly influence how much they improved by Time 2. Of course, while both measures dealt with the amount of exposure they did not probe exactly the same thing. The residualized Exposure measure represented effects of exposure over and above the exposure provided by Continued French. The Exposure measure was, however, limited to exposure to spoken French, while the Continued French variable, while being limited to formal instruction in a classroom environment, included not just listening and speaking, but also reading and writing, formal instruction etc. Therefore, Continued French covered a broader spectrum of L2-usage. The different results obtained for Continued French across time versus Exposure across time may therefore mean that for sound processing it is mainly the amount of exposure specifically to spoken French (through listening and speaking) that helps a learner improve after the end of an immersion period, regardless of whether the learner is taking French classes. Therefore, bilinguals who were overall more exposed to spoken French improved more from Time 0 to Time 2 in their sound discrimination abilities than those who were less exposed overall, whereas bilinguals who were taking French classes improved to a similar degree to those that did not receive formal instruction after the end of the immersion period.

The results of the factor regressions further suggest that the opposite pattern is true for lexical processing. The results for the lexical decision task in Experiment 2 revealed that the bilinguals who continued taking French classes improved more from Time 0 to Time 1 in identifying French words than those who did not take classes. However, the bilinguals who were more exposed to listening and speaking (represented by the residualized Exposure measure)

improved at the same rate from Time 0 to Time 1 as the bilinguals who were less exposed to spoken French. The data here, therefore, suggest that for lexical decision, language instruction in a classroom setting, exposure to reading and writing and other benefits that may come from taking L2 classes boost the improvement in lexical processing skills after the end of an immersion period. In the case of lexical processing, those added benefits of classroom instruction may, therefore, be more important to improvement across time than the overall amount of exposure to L2 speaking and listening.

Effects of executive function

Executive function did not influence performance across time in L2 perception, but it was found that bilinguals with lower executive function scores (i.e., farther from the 'problematic behavior' end of the scale) improved more across time in picture naming than the bilinguals with higher executive function scores. This is in line with previous research, which has found that executive functioning is correlated with linguistic skills. For example, bilinguals have been found to be better at inhibiting irrelevant linguistic or nonlinguistic information than monolinguals (e.g., Blumenfeld & Marian, 2011), and in clinical studies it has been found that adults who were born deaf and who received cochlear implants in late childhood or later exhibited certain language difficulties and also tended to have BRIEF scores closer to the problematic-behavior end of the scale than hearing adults (Pisoni, Conway, Kronenberger, Henning, & Anaya, 2010). Why executive function seemed to have influenced speech production but not speech perception is not immediately clear. It may be that production is more affected by these differences than perception simply because it is a more complex task (i.e., it involves perception *and* production). Future studies that are designed specifically to investigate the

influence of executive functioning on language perception versus production would have to be undertaken to find the actual reasons for this difference.

No effects of working memory

While previous research has found that a higher working memory span may lead to improved vocabulary acquisition in an L2 (Masoura & Gathercole, 1999; Paradis, 2011), the current study suggests that working memory does not seem to influence the performance of proficient bilinguals who were immersed in the L2 for a few months previous to testing. This is, in fact, in line with Masoura and Gathercole (1999). They reported that working memory had the biggest influence on L2 vocabulary in beginning learners. Advanced learners have a large existing vocabulary that helps them in the acquisition of new words. Therefore, they do not have to rely on working memory as much when acquiring new words.

6.5 Conclusion

The previous chapters have shown that bilinguals who have recently returned from a semester in France can improve their phonological and lexical processing skills in the first five to nine months after their return home. In the current chapter it was shown that four main factors related to individual differences between L2 learners modulate this improvement across time. While initial proficiency does not seem to influence improvement across time, with high and low proficiency bilinguals improving at comparable rates, the bilinguals' attitude towards the L2 and L2 culture, as well as their exposure to the L2 in the first five months after the end of the immersion period seem to influence how much improvement they experience. Attitude was correlated with improvement for all three tasks, suggesting that highly motivated students who tend to identify more with the L2 culture, feel less anxious when using the L2, and have high

phonological awareness are more likely to improve over time. While phonological awareness may not be something L2 learners can influence, language programs offering study abroad programs and students participating in such programs will benefit from knowing that they can further improve their language skills (phonological and lexical, perceptive and productive) after the end of a study-abroad program if they feel like the language they are studying will be useful for them in the future and if they show a personal interest in the L2 culture. This will motivate them to put more effort into language improvement and may make them feel less anxious about using the language, which all ultimately may lead to more improvement in the L2, compared to less motivated students. This is true even for learners who do not have the possibility to take more language classes. Language programs should foster students' personal identification with the target language culture, since this seems to aid L2 improvement in study-abroad returnees.

Exposure (to listening and speaking in French) also seems to increase improvement across time, but only for phonological tasks, such as L2 sound discrimination, and not necessarily for tasks that require lexical knowledge, such as lexical decision or picture naming. The regression analyses in the previous sections showed that taking French classes upon one's return home may increase improvement across time for lexical tasks. The current results, furthermore, suggest that overall more exposure to spoken French can increase improvements across time. By nine months after the end of the immersion period all bilinguals had improved in the AX discrimination task, whether they were taking classes or not – but the bilinguals who received the most exposure to French in the form of listening and speaking were the ones who had improved the most.

In a lexical task, the amount of exposure to listening and speaking in French may be less crucial. Improvement across time was similar for bilinguals with lots of exposure to spoken French and for those with little exposure. However, taking classes seemed to provide an advantage. Bilinguals taking French classes after their return home improved their lexical decision skills more than the bilinguals who did not take classes. It is not clear, what the specific benefits for the lexical tasks in the current study were that were provided by exposure to French in a classroom setting versus general exposure to listening and speaking in French. It would be worthwhile, however, to address this question in future research.

Chapter 6: Conclusions

In this thesis, it was investigated whether English-French bilinguals who have recently returned from a study-abroad stay in France differ from native French speakers in their perception and production of French sounds and words, and more importantly, whether any changes in the bilinguals' proficiency become evident within the first nine months after their return to the United States. A number of linguistic and extralinguistic variables were analyzed with regard to their influence on changes across time in four linguistic tasks. The results of these analyses are summarized below. The theoretical and practical implications of the results are then discussed, and directions for future research are suggested.

6.1 Summary of results

6.1.1 Lexical and phonological attrition, retention, and improvement in speech perception

In an AX discrimination task and a lexical decision task it was shown that even after four months of L2 immersion bilinguals are not likely to be completely native-like in their phonological and lexical L2 perception. Native French speakers were more sensitive to French sound contrasts and were better at identifying French words than the bilinguals when tested soon after the bilinguals' return from France. While the bilinguals were less accurate, they were equally fast as the native French speakers, suggesting that they were quite fluent in French.

Furthermore, the bilinguals improved across time in both tasks with regards to reaction times: they were faster at discriminating French sound contrasts and at identifying French words five months, and again nine months after the end of the study-abroad program compared to two months after the end of the program. Importantly, the native English speakers did not get faster

from the first to the second test session, suggesting that the bilinguals did indeed get faster because of additional L2 processing that had taken place in the meantime, and that this was not simply an effect of task familiarity (in which case the native English speakers should also have improved their reaction times).

The phonological and the lexical tasks were differentially affected by changes across the first five to nine months after the bilinguals' return from France. While the bilinguals' reaction times improved for all tasks, sensitivity to sound contrasts and accuracy in identifying French words did not improve to the same degree. In AX discrimination, bilinguals who continued taking French classes were more accurate both at Time 1 (five months after the end of immersion) and at Time 2 (nine months after the end of immersion) relative to Time 0. Bilinguals who did not continue taking French classes did not improve until Time 2. While they did not lose their sound discrimination skills, it took them longer to show improvement. This could be attributed to a settling-in phase (cf. Cohen, 1986). During such a settling-in phase, L2 processing may have taken place even though the effects of it did not become evident until Time 2. In lexical decision, on the other hand, the bilinguals who continued taking French improved from Time 0 to Time 1, but they seemed to have lost some of their knowledge again over the summer break between Time 1 and Time 2. At Time 2, their accuracy was back at the same level as at Time 0. The bilinguals who did not continue taking French classes, however, showed no changes across time in accuracy at all. These results suggest that improvement in sound processing may occur even in the absence of further input; in contrast, lexical processing may rely more on continued input.

Some linguistic factors also influenced the changes across time. Specifically, in AX discrimination the native French speakers were overall better than the bilinguals at consonant contrasts whereas they were not significantly different from the bilinguals for vowel contrasts. Furthermore, the bilinguals improved more on the consonant contrasts than the vowel contrasts. As for sound contrast types, the native French speakers were more sensitive to category goodness assimilation contrasts (compared to two category assimilation contrasts), while the bilinguals did not make a distinction. Generally, nonnative listeners tend to be better at the two category assimilation contrasts (Best & Tyler, 2007), but the bilinguals in the current study may have been too high proficiency to show this nonnative perception effect.

Cognate status did not show any effects in the current study, i.e., there was no difference between native French speakers and bilinguals in their identification of cognates versus noncognates in the lexical decision task. Furthermore, the changes across time found for the bilinguals did not seem to be modulated by cognate status of the target words. Previous research (Woutersen et al., 1995) has suggested that cognate effects may disappear in lexical decision once bilinguals have reached a high proficiency level in the L2. Considering that the bilinguals in the current study had recently spent four months in France where they were immersed in the L2, it seem feasible that the bilinguals may in fact have achieved proficiency levels that were too high to be affected by the lexical status of the target words.

With the above results suggesting that changes across time affect lexical and phonological comprehension tasks differentially, it should be noted here that there is in fact an alternative explanation to the lexical versus phonological processing difference. In addition to the two tasks involving different levels of processing, the AX discrimination task and the lexical

decision task also differ in the amount of metalinguistic awareness necessary to complete the task. While both tasks potentially involve metalinguistic processing, it is only crucial for the lexical decision task. Lexical decision (i.e., deciding whether a string of sounds makes up a real word or not) requires the listener to make a metalinguistic judgment about the word status of the stimulus. For AX discrimination, metalinguistic processing may play a role as well, e.g., participants may make their judgments based on whether they heard a word-nonword pair, or whether they heard two identical French words versus two contrasting French words, i.e., taking lexical status into consideration along with the acoustic information. In contrast to lexical decision, however, metalinguistic knowledge is not necessary to complete the task. This is evident from the fact that naïve listeners, such as the native English speakers in the current study, are able to complete this task without great difficulty, based on just the sound differences (listeners can, in fact, complete this task by relying purely on acoustic discrimination of sounds). Therefore, it is possible that the differences across time that were found for AX discrimination and lexical decision were not solely based on lexical versus phonological processing, but may also be due to differences in metalinguistic processing. Future studies could try and tease apart these differences by using tasks that do not require metalinguistic judgments. For example, instead of a lexical decision task, a visual world paradigm could be used, where participants are played a French word and are asked to identify the picture that shows that word. Of course such a task would introduce other problems, such as the fact that lexical and phonological processing are involved in the same task and that participants have to understand the target language.

6.1.2 Lexical and phonological retention, and improvement in speech production

A word repetition task showed that English-French bilinguals who had recently returned to the United States from France were able to pronounce some French sounds more native-like than others. First, the bilinguals were native-like in their productions of the voice onset times of the consonants /d/ and /t/. The SLM, however, would predict that L2 learners would have difficulties with such 'similar' sounds that also exist in the L1. This may reflect a ceiling effect. The bilinguals in this study may have simply acquired excellent processing of consonants in perception (as shown above) as well as production. For vowels, the predictions of the SLM were confirmed. The bilinguals were able to produce native-like F2 values for the new vowels /y/ and /œ/, but not for the similar vowel /u/. However, the bilinguals' productions did not change across the first nine months after the end of the immersion period, neither for similar nor for new sounds. This suggests that pronunciation of L2 sounds remains fairly stable across the first months after the end of immersion – even in the absence of further input from the L2.

While the bilinguals retained their L2 production skills from Time 0 to Time 2 for word repetition, they even improved across the same time span in picture naming. The bilinguals were more accurate and faster nine months after the end of immersion relative to two months post-immersion. This was the case even though none of the bilinguals continued taking French classes.

Furthermore, the bilinguals in the current study seemed to experience cognate facilitation. While the native French speakers did not differentiate between French-English cognates and French noncognates, the bilinguals were more accurate and faster at naming cognates than noncognates. This is in line with previous research that has found a bigger cognate advantage for

participants who perform a task in their L2 than those performing the task in their L1 (Kroll & Stewart, 1994). However, cognate status did not influence the bilinguals' improvement across time. The bilinguals improved their naming accuracy and speed for cognates and noncognates to a similar degree.

6.1.3 Individual differences affecting improvements across time

Twenty-one variables related to the bilinguals' language acquisition history, proficiency level, amount of exposure to French, motivation, attitudes towards the L2 culture, language anxiety, working memory span, phonological awareness and executive function were investigated with regards to their influence on the improvements across time that were found for some of the linguistic tasks. Many of the variables were highly interrelated, but it was possible to extract a few main factors that seemed to have driven the bilinguals' improvement within the first five to nine months after their return from France. These factors were L2 proficiency, exposure to the L2 (specifically in speaking and listening), motivation and attitudes towards the L2 culture, and executive function.

Specifically, higher proficiency bilinguals were more accurate in the AX discrimination, lexical decision and picture naming tasks, but their improvement across time was comparable to that of lower proficiency bilinguals. In other words, achieving a high proficiency in an L2 during a study-abroad program seems to be neither advantageous nor disadvantageous when it comes to improving further after the end of L2 immersion. This was true for speech perception and speech production, and in phonological as well as lexical tasks.

Exposure to French in speaking and listening after the bilinguals' return home affected their improvements differently for different tasks. While bilinguals with greater exposure to French were overall better at naming pictures in French in the first nine months after their return, the amount of exposure did not influence their improvement across time. Exposure did, however, influence their improvements in the AX discrimination task. This suggests that exposure to spoken French is beneficial for improvements in sound discrimination, but may not be enough to lead to improvements in lexical tasks.

High levels of motivation, low language anxiety and a positive attitude towards the L2 culture were found to aid bilinguals in a lexical decision task, and furthermore, they seemed to lead to greater improvements across time for phonological and lexical tasks in speech perception and production. This is in line with previous research on L2 acquisition (Gardner, 1985), which has shown that highly motivated students tend to acquire an L2 more successfully, and with previous studies on language attrition (Murtagh, 2003), which have shown that a positive attitude towards the L2 culture can reduce or prevent attrition after the end of language instruction.

Finally, while working memory did not seem to influence improvements across the first five to nine months after the end of L2 immersion, executive function was found to influence the patterns of improvement found for picture naming. Specifically, bilinguals with higher executive functioning (i.e., the bilinguals who received a lower score on the BRIEF inventory, where higher scores represent a tendency towards problematic behavior) improved more across time than bilinguals with lower executive function abilities. Previous research has found correlations between executive function and bilingualism, in that fluent bilinguals tend to be better at tasks requiring high levels of executive functioning (Blumenfeld & Marian, 2010). The current study

extends this research by suggesting that high levels of executive functioning can also aid bilinguals in improving their language skills across time.

6.1.4 Changes across time in perception versus production

Research on L1 attrition has frequently reported that language comprehension is more resistant to attrition than language production (Ammerlaan, 1996; Hulsen, 2000), and some L2 attrition studies have come to the same conclusion (e.g., Tomiyama, 2000). Other studies have not found clear differences between changes across time for perception versus production (Murtagh & van der Slik, 2004). In the current study, improvements were found for perception *and* production tasks. The only task where no improvement was found was word repetition. However, since the variables of interest in the different task varied (e.g., sound status in AX discrimination, lexical status in lexical decision and picture naming, and acoustic measures in word repetition) direct comparisons across the four tasks are for the most part impossible. The fact that word repetition was the only task for which no improvement was found could be understood as a sign that this task may be more vulnerable to attrition than the other tasks, which would be in line with research that has found more attrition in production tasks. The other production task (picture naming), on the other hand, showed improvements just as the perceptual tasks did, so it could also be the case that acoustic measure require continued input for improvement to take place. The current results, therefore, suggest that perception versus production seem to be affected by changes across time differently depending on the processing levels required. This may account for some of the contrasts with previous results. The production task in Tomiyama's study (where attrition was found) was a verbal fluency task. The processing requirements of this task are not identical to picture naming and it may, therefore, be affected

differently by changes across time. As for the L1 attrition studies, one could expect L1 and L2 attrition to show different patterns simply because of the language status. Furthermore, the L1 attrition studies that have found attrition in production while there was retention in perception tested bilinguals who had not been exposed to their L1 for a number of years. It is certainly possible while the bilinguals in the current study experienced improvements in the first nine months after the end of immersion, attrition may still set in at a later time.

6.2 Theoretical implications

The results of the current study are relevant for two sets of theoretical implications. The first pertains to models of L2 speech acquisition (i.e., the SLM and PAM) and how they can be extended to include language retention. The second set of implications addresses the need for an integrated model of L2 attrition, i.e., one that includes both linguistic and extralinguistic variables and their effects on language attrition, retention, and improvement.

6.2.1 Models of L2 speech acquisition

The Speech Learning Model (SLM; e.g., Flege, 1987) and the Perceptual Assimilation Model (PAM; e.g., Best, 1995) have been widely used in L2 acquisition research to explain how the perception and production of different sounds and sound contrasts varies in difficulty for L2 speakers, depending on the similarity of the sound or contrast to existing L1 sounds or sound contrasts. The SLM states that sounds that are similar in the L1 and L2 are less likely to be acquired in a native-like fashion than sounds that are 'new' sounds in the L2, i.e., those that do not have a close equivalent in a speaker's L1. The PAM, on the other hand, claims that sound *contrasts* that also exist in the L1 are easier to discriminate by naïve listeners as well as L2 learners than L2 contrasts that map onto a single phoneme category in the L1. The current study

suggests that some of these claims are not only true for naïve listeners and learners who are in the process of actively acquiring a language, but even for L2 speakers who do not receive any exposure to the L2 anymore, i.e., those that have 'stopped acquiring' the language. First, however, the PAM predictions were not confirmed in the current study. The bilinguals did not treat different types of sound contrasts differently (i.e., they were not more sensitive to sound contrasts that were similar to existing L1 sound contrasts as the PAM would have predicted). In contrast, the predictions of the SLM, i.e., L2 learners acquire new sounds in a more native-like fashion than similar sounds, were shown to be partially true for the bilinguals' productions (namely for the vowels). In sum, the above results suggest that the SLM can also be applied to proficient L2 speakers who have stopped actively acquiring their L2, whereas no conclusions can be drawn about the PAM predictions with regards to this group of speakers.

6.2.2 Informing an integrated model of L2 attrition

Most previous studies on lexical and phonological L2 attrition have focused either on linguistic factors to language attrition versus retention (e.g., sound similarity across languages in Dugas, 2000), and have restricted themselves to investigating either language production (Gonzalez Moncada, 1995) or perception (Weltens, 1989). Other studies focused on extralinguistic influences on retention and attrition, only testing language ability in a single task, which does not allow one to draw conclusions about patterns of language change more generally (Gardner, Lalonde, Moorcroft, & Evers, 1987). One notable exception is Murtagh (2003) who investigated a variety of extralinguistic influences on retention of L2 Irish which she tested in a variety of perception and production tasks. However, a comprehensive study including well-controlled linguistic tasks that can address a number of linguistic factors for L2 retention (such

as perception *and* production, phonological *and* lexical processing) as well as a variety of extralinguistic factors, such as factors related to proficiency, exposure, attitudes and cognitive skills was still missing. Köpke and Schmid (2004) and Köpke (2007) have expressed the need for a multi-componential model of language attrition, claiming that only the combination of linguistic and extralinguistic factors tested in a variety of tasks would be able to provide a (more) complete picture of language attrition versus retention. Köpke (2007) claims that no single variable, or even cluster of variables, can explain language attrition. Rather, depending on a speaker's situation, different variables may be weighted differently and consequently lead to quantitative and possibly also qualitative differences in changes across time. However, which factors are the most influential on language attrition/retention is not clear and depends on the speaker's situation. While the model proposed by Köpke is aimed at L1 attrition research, a multi-componential view of language attrition is equally relevant for L2 attrition.

Köpke's proposal focuses on the integration of biological and cognitive aspects that may play a role in language attrition, as well as speaker-external factors. Among the biological factors discussed are activation thresholds. Köpke's discussion of these with regards to language attrition is based on Paradis' (1993) Activation Threshold Hypothesis, which defines activation thresholds as the activation levels of an L1 and an L2 dependent on the amount of exposure a speaker has to each language. The more a language is used, the more easily it is activated, and on the reverse side, the less a language is used, the higher is its activation threshold, i.e., the more difficult it is to activate the language. This is true for the language as a whole, and for individual aspects of the language (e.g., lexical access is easier for words that are more frequently encountered than for words a speaker has less exposure to).

The predictions of the Activation Threshold Hypothesis for the current study would be that the retrieval of the L2 should become more difficult across time for bilinguals (since they are less exposed to French upon their return to the USA), and this should be the case especially for bilinguals with little to no exposure to French upon their return. Furthermore, words that are more frequently encountered should be more easily retrieved than less frequently encountered words. Since the Activation Threshold Hypothesis is generally applied to lexical processing, it makes the most sense to compare the hypothesis' predictions to the results of the lexical tasks in the current study (lexical decision and picture naming). The results of the current study are partially in line with the predictions made by the Activation Threshold Hypothesis. While it was not the case that the L2 performance of the bilinguals decreased overall across the first nine months after their return from France, it was indeed the case that in lexical decision, where the two groups of bilinguals were compared, bilinguals who still received more exposure to French performed better than the bilinguals with less exposure to French. This would indicate a lower activation threshold for French in the bilinguals with more exposure. At the lexical level, while it was not possible to test the actual exposure of the bilinguals to individual words, it is generally assumed that bilinguals are overall more exposed to cognates than noncognates (since they are exposed to them in both languages; cf. Gollan & Acenas, 2004). In the current study, bilinguals did indeed show a cognate facilitation effect in speech production (picture naming), suggesting that the activation threshold for cognates may have been lower than for noncognates.

Another biological factor discussed by Köpke is inhibition. In an inhibition account, a language that is used more frequently needs to be suppressed more if a speaker wants to use the language he or she is exposed to less frequently than vice versa (e.g., Bialystok et al., 2008). The

inhibition account predicts that in speakers who never use a language (i.e., the bilinguals who are no longer exposed to French in the current study) the language is completely dormant, whereas most of the available resources when using that language have to be spent on inhibiting the more frequently used language (i.e., English in the current study). This may result in processing difficulties in the L2 that can show up as longer reaction times or lower accuracy rates in the L2 across time. For bilinguals who are still in frequent contact with their L2, an inhibition account would predict easier access to the L2 compared to the bilinguals without exposure to the L2. Again, the results of the current study are partially in line with these predictions: while bilinguals did not show any signs of having to inhibit the L1 more across time (i.e., they did not slow down and did not show decreased accuracy rates across time in French), it was the case that bilinguals with more exposure to French overall performed better than the bilinguals with less exposure, suggesting that the overall amount of inhibition necessary for their L1 English was less than the inhibition necessary for the bilinguals with less exposure to French.

Köpke's multi-componential model for language attrition also includes cognitive factors, such as memory and language aptitude. Memory here includes the influence that both long term memory and working memory may have on language attrition. With regards to long term memory, Köpke cites Ullman's (2001) distinction of declarative versus procedural memory structures. Declarative memory is said to involve facts and events, whereas procedural memory involves cognitive and motor skills and habits. With regards to language, the declarative/procedural model claims that the lexicon is stored in declarative memory, whereas grammar, if acquired early (as is the case in the L1 and early-acquired L2) is stored in procedural memory, but if acquired late is stored in declarative memory. Furthermore, declarative structures

are more vulnerable to attrition. Therefore, the model predicts that the lexicon is likely to be affected by attrition in all bilinguals, whereas grammar is more likely to be affected in late bilinguals (for whom grammar structures are stored in declarative memory) than for early bilinguals (for whom grammar is stored in procedural memory). Since the current study did not investigate L2 grammar, no evidence for or against this difference between lexicon and grammar can be made. However, the fact that changes across time were found for L2 vocabulary, suggests that this may indeed have been the case because the L2 vocabulary was stored in declarative memory, which is more vulnerable to changes across time.

Furthermore, Köpke suggests that working memory may also affect language attrition in that it plays a role in tasks that require high levels of cognitive control. If there is reduced access to one of a bilingual's languages, this leads to higher cognitive demands. Bilinguals with better working memory skills may, therefore, show fewer processing difficulties when using a language they are no longer exposed to. The current study, however, did not find significant effects of working memory span on changes across time, i.e., bilinguals with differing working memory spans improved similarly across the first nine months after the end of their study-abroad program. The results, thus, suggest that working memory may not be a factor that needs to be included in a multi-componential model of language attrition.

Another cognitive function included in Köpke's multi-componential model is language aptitude. Language aptitude is an individual's talent for acquiring a foreign language and has been found to correlate with successful language learning (Abrahamsson & Hyltenstam, 2005; as cited in Köpke, 2007). Greater language learning aptitude could also lead to an increased ability to retain the acquired skills and knowledge, i.e., it may be the case that less attrition would be

seen in individuals with high language aptitude (regardless of their proficiency level in the L2). The current study investigated language aptitude in the form of a phonological awareness task. A factor analysis grouped the phonological awareness variable together with the motivation and L2 attitude variables into the Attitude factor. A higher Attitude score (including higher phonological awareness) was correlated with bigger improvements across time. This suggests two things with regards to Köpke's model: first, language aptitude (as measured by the phonological awareness task) may indeed have an influence on patterns of language attrition (or changes across time more generally), but second, as the factor analysis showed, language aptitude may not be a variable that contributes to language changes independently, but rather may have an influence as a component of a more general Attitude factor that also includes an individual's motivation to learn the L2, their attitudes towards the L2 culture, and language anxiety.

Finally, Köpke recommends that 'external' factors, such as language use, the cultural context of an immigrant (or L2 user), and attitudes towards the languages should be included in a multi-componential model of language attrition. As discussed in the previous paragraphs, the amount of exposure to the L2, did indeed play a role for the improvement the bilinguals in the current study experienced across time. The bilinguals with more exposure to French improved more than the bilinguals who received little or no exposure to French upon their return home.

Cultural context, in Köpke's terms refers to whether immigrants in an L2 country are isolated from other L1 speakers, or whether they participate actively in the immigrant community. Köpke predicts that immigrants with no contact to other immigrants would experience more attrition in terms of performance (i.e., their L1 would be more difficult to access), but their L1 would show less evidence of L2 intrusions than the language of immigrants

who are regularly exposed to the L1 of other members of the immigrant community. While this factor seems less relevant for L2 speakers residing in their L1 country, some parallels can be drawn. For example, the bilinguals in the current study may have a lot of exposure to French, but the exposure may come mainly from other L2 learners, alternatively they may be exposed to French by native speakers residing in the USA, or they may be mainly exposed to French through contacts with French speakers in France, and French language media. These different forms of French input may lead to different patterns of improvement versus attrition in the L2 learners. However, in the current study, the analyzed data did not include the type of French input the bilinguals received, and therefore, no conclusions based on the influence of type of input can be drawn.

As was discussed in the paragraph on language aptitude, the bilinguals' attitudes towards studying French and the French culture was shown to influence the improvements experienced across time. This is in line with Köpke's prediction that a more positive attitude towards the L2 and the L2 culture may make retention of the L2 more likely.

While Köpke (2007) focused on nonlinguistic factors that need to be included in a multi-componential model, Köpke and Schmid (2004) also included the necessity of different linguistic factors to form a complete model of language attrition. The current study aimed to collect data on as many linguistic and extralinguistic factors as possible and to analyze the interplay of these factors and the resulting patterns of L2 attrition, retention, and improvement. While the current study did not attempt to formulate a multi-componential model of language attrition, the results of this study, with the multitude of aspects of language retention investigated, can certainly inform such a model. Specifically, with respect to the situation of university students who have

participated in a study-abroad program a number of variables and clusters of variables (i.e., factors) were isolated that seemed to have the biggest influence on retention and improvement across time (e.g., L2 exposure and attitude). Schmid and Dusseldorp (2010) pointed out that the few studies that have taken a multi-componential approach to the study of language change established composite factors based on intuitive and theoretical grounds. The current study, therefore, has brought the study of language attrition/retention a step further by analyzing empirically validated factor clusters (i.e., the extralinguistic factors in the current analysis were extracted from a list of variables through a factor analysis, and not simply combined based on the intuition of the author).

Furthermore, the current study did not only focus on the different factors' influence on quantitative differences in changes across time, but also investigated qualitative differences (i.e., some factors influenced changes in performance on lexical tasks, whereas other were more influential on phonological tasks). The results from the current study can therefore be used as a guideline for how to weight linguistic and extralinguistic (i.e., individual differences) influences on L2 retention and improvement within the framework of a multi-componential model.

Finally, since the current study interpreted results also with reference to L2 speech acquisition models (SLM and PAM), the results allow for the incorporation of the predictions of these models into a multi-componential model of language attrition/retention. In other words, this would allow the model to make predictions on qualitative differences in retention or improvement based on the similarity of sounds across languages. Language attrition studies have only recently begun to investigate changes at the phonological level, but without information on how changes across time affect the phonological processing level, a model of language attrition

and retention cannot be complete. In sum, with its extensive investigation of phonological retention and improvement, the current study is able to inform a multi-componential model on which cross-linguistic phonological factors seem to influence L2 retention the most after the end of an L2 immersion period. Specifically, the results from the current study suggest that it is important to integrate both linguistic and nonlinguistic factors into a model of language attrition, since no single type of factor alone can explain changes in language proficiency across time, but it is the interplay of linguistic, cognitive, biological, and social factors that shape an individual's language abilities and that determine if and how these abilities change cross time.

6.3 Practical implications

In a more practical way, the current study can inform language teaching and learning practices. First of all, language instructors and students will be happy to hear that contrary to general intuition and learners' self-reports, proficiency in the L2 (in both perception and production) may not diminish as much as the learners themselves may feel is the case. The participants in the current study reported diminished skills in speaking and listening while their performance in the perception and production tasks showed that they had retained the proficiency level from Time 0 (around two months after the end of the study-abroad program) for up to nine months post-immersion and had even improved in some tasks. Some previous L2 attrition studies have found the same surprising discrepancy between participants' self-reports of significant language attrition while linguistic tasks showed retention (Murtagh & van der Slik, 2004) or even improvement (Weltens, 1989).

The current study also showed that there are certain factors that can influence whether there is improvement, or that may change the time course of attrition, retention, and

improvement. One interesting finding was that the bilinguals' attitude towards the L2 culture, combined with high motivational levels, low language use anxiety, and a sense of usefulness of the language in the learner's mind was correlated with bigger improvements across time, regardless of proficiency level and amount of exposure to the language. Language instructors and curriculum designers should take this into consideration. If language courses raise a personal interest in and identification with the target culture in L2 learners, keep learners motivated throughout the learning process, and create a classroom atmosphere that decreases the learners' language use anxiety, this would greatly benefit not just the learners' language acquisition (as shown in previous studies such as Gardner, 1985). Such courses would also help learners retain their language skills (and even improve their skills) once they no longer receive formal instruction in the language.

Furthermore, exposure to listening and speaking also seems to lead to improvement across time for phonological tasks (such as L2 sound discrimination). However, tasks that require lexical knowledge, such as lexical decision or picture naming may not benefit from increased exposure to spoken French. The results of the current study suggest that study-abroad participants who want to further improve their sound processing skills (such as those required in an AX discrimination task) are well advised to expose themselves as much as possible to French in the form of listening and speaking (either in a classroom setting or in a more informal context). Although by nine months after the end of the immersion period all bilinguals in the current study had improved in the AX discrimination task (whether they were taking classes or not), the bilinguals who received the most exposure to French in the form of listening and speaking were the ones who had improved the most.

Study-abroad returnees who want to improve their vocabulary skills, however, may not profit to the same degree from simple exposure to French in the form of listening and speaking. Improvement across time was similar for bilinguals with lots of exposure to spoken French and for those with little exposure. For lexical processing, however, taking classes seemed to provide an advantage. Bilinguals taking French classes after their return home improved their lexical decision skills more than the bilinguals who did not take classes.

6.4 Directions for future research

In the current study, a wide variety of linguistic and nonlinguistic variables were examined with regard to their effects on attrition, retention, and improvement in L2 speakers of French who had recently returned home from a study-abroad program in France. By testing all these different variables, the study laid important groundwork for future research that could concentrate on a subset of the variables and/or tasks and examine the effects of these variables in more detail. For example, since for the AX discrimination task in the current study words were matched for sound status (vowel versus consonant) and PAM contrast type (two category assimilation contrast and category goodness assimilation contrast), only a small set of words could be created for each variable. The same was true for sound status and sound similarity (similar versus new sounds) in the word repetition task. Future studies could provide a more detailed picture of the effects of these variables by focusing on just one of the variables and creating a bigger set of target items for it. The results of the current study, which were certainly able to at least show some trends – such as the differences found between vowel and consonant processing both in speech perception and production, or the lack of an effect for the different PAM contrast types – can guide future studies in the choice of variables.

The SLM and PAM-L2 predict that as a learner's proficiency in the L2 increases and L2 sounds are either assimilated to existing L1 phoneme categories or new categories are created, L1 phonemes may change as well. For example, the SLM states that if a *similar* L2 sound is assimilated to an L1 sound, neither sound is likely to be completely native-like (compared to monolingual speakers' phonemes), but rather both the L1 and the L2 sound fall somewhere in between the respective monolingual L1 and L2 sounds. To give an example, if a native English speaker acquires French /t/ (which has a shorter VOT in French than English), the resulting English VOT is likely to be shorter (i.e., more French-like) than a monolingual English speaker's VOT, and the French VOT is likely to be longer (i.e., more English-like) than a monolingual French speaker's VOT. The opposite is predicted for *new* sounds. If an English speaker acquires French /y/, which perceptually is closest to English /u/, the resulting French /y/ is likely to be more extreme (i.e., more fronted) than a monolingual French speaker's /y/ to be as distinctive as possible from English /u/. At the same time the bilingual's English /u/ is likely to be more extreme in the opposite direction (i.e., farther back) to be distinctive from the newly acquired French /y/ category. Since in the current study, only L2 perception and production was analyzed, it was not possible to address these predictions. In future studies, therefore, it would be worthwhile to also analyze L1 data from bilinguals and track the changes in both languages to see if it is indeed the case that the two phoneme categories in the two languages behave this way, and most importantly, if they change when input from the L2 ends.

Because of the scope of the current study, which covered many different variables, there were limits with regards to the comparability of the individual tasks. For example, in speech production, the analysis of the word repetition task was limited to sound status but did not

include lexical status. The opposite was true for picture naming, for which no sound status information was collected. It would be informative, however, to test whether the effects of sound status found in the repetition task would also appear in the picture naming task. It is possible that effects of lexical status versus sound status are different for tasks that require lexical processing than for those that do not. Lexical versus phonological tasks have not usually been compared directly in L2 attrition studies but a well-controlled study directly comparing effects on a lexical versus a phonological tasks would not only provide further insights into language attrition, but would also be informative for models of bilingual speech processing in general in that it would expand our knowledge of sound and lexical processing in the absence of L2 input. The current study took a first step in that direction.

While the current study was limited to single word and sound perception and production, it would be interesting to examine changes across time for different linguistic aspects such as intonation and word stress, or in different types of tasks, such as in conversational speech – which would furthermore make an analysis of intonation, stress patterns, and speaking fluency possible. To date, no L2 attrition studies have focused on sound processing in conversational speech, although some studies have looked at speaking fluency. For example, Murtagh (2003) investigated L2 fluency in former students of Irish, and she found retention in some tasks (giving an opinion, describing objects) and improvement in one fluency task (telling a story) across the first eighteen months after the end of L2 instruction. Tomiyama (1999) also tested speaking fluency in an eight-year-old native Japanese speaking child, whose L2 English changed across the first 33 months after his return from the United States to Japan. The child's fluency started to decrease around 8 months after the return home. Tomiyama also points out, however, that it is

difficult to disentangle fluency from lexical access, since the loss in fluency was primarily caused by lexical retrieval difficulties. These results suggest that general L2 fluency seems to follow similar patterns of change across time as phonology and lexicon in the current study (i.e., retention and improvement in fluency in Murtagh, and loss of fluency not until 8 months post-immersion in Tomiyama).

The current study also did not investigate syntax and morphology. The three speaking fluency tasks in Murtagh (2003) mentioned above were also analyzed for grammatical accuracy, and the same results were found as for fluency (i.e., retention or improvement). Similarly, Tomiyama (1999) found retention in morphology and syntax in the first 19 months, and only between 20 and 33 months after the child's return to Japan did attrition in morphology and syntax become noticeable. Some previous studies (Moorcroft & Gardner, 1987; Weltens, 1989) compared attrition in grammar and vocabulary and have found that in low-proficiency bilinguals, grammar seems to be affected more by attrition than the lexicon, whereas the opposite tends to be true in high-proficiency bilinguals. This indicates that for the majority of the bilinguals in the current study, who are fairly high proficiency speakers of French, no attrition in grammar would be expected, since they did not show attrition even in the lexicon, which should show signs of attrition first.

Finally, pragmatics is another processing level not investigated in the current study. However, Murtagh (2003) investigated pragmatic skills (in terms of appropriateness of L2 responses) in two tasks that required participants to give instructions or advice and to react appropriately in role play situations. Similarly to the previously mentioned skills, Murtagh found retention of pragmatic skills in the first eighteen months after the end of L2 instruction. It can

therefore be assumed that similar results would have been found with the participants in the current study.

The results from the studies presented above compared to the current study suggest that lexical and phonological processing seem to show similar patterns of changes across time as other processing levels (such as morphology, syntax, and pragmatics). However, since results cannot be compared directly across studies due to differences in participant groups and experimental methodology, it remains up to future research to provide a comprehensive picture of changes in language proficiency across time including all processing levels, i.e., compare changes in syntax, semantics, pragmatics, lexicon, phonology, and phonetics in the same group of participants.

As was pointed out in previous chapters, another difficulty when trying to compare the results of different L2 attrition studies is the fact that they looked for attrition and retention at varying time frames – from as short as two months in Cohen (1986) to as long as fifty years in Bahrick (1984). In the future, it would therefore be desirable to test the same group of participants at various times starting immediately after the end of language instruction or of a study-abroad program up to several years later. Only this way can the long- and short-term patterns of attrition versus retention and improvement really be compared. One possibility for this would be to try to test as many of the participants in the current study again throughout the next few years to see if and when the bilinguals cease to show improvements, and whether different processing levels (lexical versus phonological) and different skills (perception versus production) show a different time course with regards to attrition, retention and improvement in the longer term. Ideally, however, research that tries to explain patterns of language change

across a longer time period would first test bilinguals immediately after the end of an immersion program (or classroom instruction) rather than starting to test them two months after the end of immersion or instruction. Likewise, it would be beneficial to test a monolingual control group at the same times as the bilinguals. In the current study, the monolingual control groups were only available at one or two test sessions, whereas some of the predictions of the current study could have been more powerful if control group data had been available for accuracy, sensitivity and reaction times for all three participant groups at all three times. As was pointed out earlier, the fact that only bilinguals improved their reaction times from Time 0 to Time 1, while the native English speakers did not get any faster, suggests that this improvement is not due to a simple task familiarity effect. While the current study assumed that this pattern extends to the other tasks and to testing at Time 2, future studies should try to include monolingual control groups for all tasks and at all testing sessions to confirm that the patterns indeed hold true across tasks and across time. Furthermore, to control for potential task familiarity effects, a third control group could be included: a group of bilinguals that only gets tested at the initial test and again at the final test session, but not in between (i.e., at the equivalent of Time 0 and Time 2, but not Time 1 in the current study). If this group showed the same improvements at Time 2 as the bilingual group tested all three times, this would provide further evidence that the changes across time indeed were caused by L2 processing changes rather than task familiarity. Future studies may want to include such a bilingual control group.

With regards to participant groups it should also be noted here, that selection bias may have played a role, too, in the current study. In other words, it is possible that only the most motivated bilinguals, or those who were most interested in keeping up their French decided to

participate in the study, and most importantly to return at Time 1 and Time 2 (35 bilinguals participated at Time 0, of which only 17 returned for both Time 1 and Time 2 and were analyzed in this study). While motivation certainly may have played a role, and the results found in this study may therefore be mainly representative for such motivated L2 learners, it was in fact the case that the majority of participants returned at Time 1 (namely, 32 out of the 35 bilinguals who had participated at Time 1), and furthermore, that the drop-out rate from Time 1 to Time 2 was high because Time 2 took place in the following school year, and a number of participants had graduated and moved away before the testing at Time 2.

Finally, to return to the difficulty in comparing results across studies, another issue worth mentioning is that some previous studies tested classroom L2 learners (e.g., Bahrick, 1984; Cohen, 1986; Gardner et al., 1987; Murtagh & van der Slik, 2004; Weltens, 1989), whereas other studies focused on learners who had been exposed to the L2 in an immersion situation in an L2-speaking country (e.g., Cohen, 1989; Dugas, 2000; Mehotcheva, 2010; Sancier & Fowler, 1997; Tomiyama, 2000). It would be worthwhile for future studies to compare the time course of attrition versus retention and improvement in L2 speakers who studied their L2 in a classroom environment versus those who studied in the country where the L2 is spoken. If one assumes that an L2 is acquired more deeply in an immersion situation, it is possible that signs of attrition would show up sooner in classroom-acquired L2 than in immersion-acquired language. However, at this point this is only speculation and it is up to future research to test this hypothesis.

6.5 Conclusions

The current study showed that contrary to what one might expect, English-French bilinguals who had recently returned from a study-abroad program in France, did not show any signs of language attrition, but instead either retained, or even improved their phonological and lexical speech perception and production skills across the first five to nine months after their return to the United States.

Furthermore, certain linguistic factors, such as sound status (vowel versus consonant, similar versus new sounds), but not lexical status (cognate versus noncognate) influenced the changes across time. A number of individual differences between speakers, including continued French study, proficiency level, amount of exposure to French since the return home, attitudes towards the French culture, as well as executive functioning influenced the pattern of attrition versus retention versus improvement in the bilinguals.

The results from the current study suggest that certain aspects of L2 speech acquisition models also apply to situations in which learners no longer receive any input in the L2, and the results of this comprehensive analysis of factors in L2 retention can inform future multi-componential models of language attrition. Finally, the study can also inform language teachers and curriculum designers on which characteristics of language learners they should focus on strengthening to help the learners achieve long-term retention of their L2 skills.

References

Ammerlaan, T. (1996). You get a bit wobbly... Exploring bilingual lexical retrieval characteristics in the context of first language attrition. Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen, Netherlands.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R.* Cambridge, U. K.: Cambridge University Press.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*, 12-28.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113*, 1-29.

Becker, J. T., & Morris, R. G. (1999). Working memory(s). *Brain and Cognition, 41*, 1-8.

Bernaus, M., & Gardner, R. C. (2008). Teacher motivation strategies, student perceptions, student motivation, and English achievement. *The Modern Language Journal, 92*, 387-401.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Timonium, MD: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775-794.

Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 859-873.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19,* 290-303.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 99-116). Amsterdam: John Benjamins.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition, 118,* 245-257.

Boersma, P., & Weenink, D. (2011). Praat: doing phonetics by computer (Version 5.2.11) [Computer program]. Retrieved from http://www.praat.org/

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research* (pp. 279-304). Timonium, MD: York Press.

Bohn, O.-S. & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11,* 303–328.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature, 403*, 428-430.

Caramazza, A., & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics, 2*, 239-245.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America, 54*, 421-428.

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2008). Are vowels and consonants processed differently? Event-related potential evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience, 21*, 275-288.

Cedrus Corporation (2006). *Stimulus Presentation Software SuperLab 4.0*. Pedro, CA: Cedrus Corporation.

Cohen, A. (1986). Forgetting foreign-language vocabulary. In B. Weltens, K. de Bot, & T. van Els (Eds.), *Language attrition in progress* (pp. 143-158). Dordrecht, Netherlands: Foris.

Cohen, A. D. (1989). Attrition in the productive lexicon of two Portuguese third language speakers. *Studies in Second Language Acquisition, 11*, 135-149.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1283-1296.

Costa, A., Roelstraete, B., & Hartsuiker, R., J. (2006).The lexical bias effect in bilingual speech production: evidence for feedback between lexical and sublexical levels across languages. *Psychonomic Bulletin and Review, 13*, 972-977.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: the status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology, 38A*, 367-393.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language, 33*, 824-844.

Davis, M. J. (2010). Contrast coding in multiple regression analysis: Strengths, weaknesses, and utility of popular coding structures. *Journal of Data Science, 8*, 61-73.

De Groot A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign language vocabulary learning and forgetting. *Language Learning, 50*, 1-56.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representations of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90-123.

De Leeuw, E. (2008). When your native language sounds foreign: A phonetic investigation into first language attrition. Unpublished doctoral dissertation, Queen Margaret University, Edinburgh, U. K.

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389-407). Oxford, U.K.: Oxford University Press.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496-518.

Dugas, L .G. (2000). Attrition of pronunciation accuracy among advanced American learners of French. Dissertation Abstracts International A 60/11. (UMI No. 304505283)

Féry, C. (2003). Markedness, faithfulness, vowel quality and syllable structure in French. *Journal of French Language Studies, 13*, 1-33.

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*, 47-65.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research* (pp. 233-272). Timonium, MD: York.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America, 76*, 708-721.

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

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

Flege, J. E., Yeni-Komshian, G., & Liu, H. (1999b). Age constraints on second language acquisition. *Journal of Memory and Language, 41,* 78-104.

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*, 649-663.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London: Edward Arnold.

Gardner, R. C., Lalonde, R. N., Moorcroft, R., & Evers, F. T. (1987). Second language attrition: The role of motivation and use. *Journal of Language and Social Psychology, 6*, 29-47.

Gardner, R. C., & MacIntyre, P. D. (1993). On the measurement of affective variables in second language learning. *Language Learning, 43,* 157–194.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002*). Cognitive neuroscience. The biology of the mind* (2nd ed.). New York: Norton.

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

Goldman, J.-P. "EasyAlign: a semi-automatic phonetic alignment tool under Praat", http://latlcui.unige.ch/phonetique, 2008.

Gollan, T., & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 246-269.

Gollan, T., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and non-cognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1122-1139.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General, 140*, 186-209.

Gonzalez-Moncada, A. (1995). Language attrition and retention among non-native speaker ESL teachers. Dissertation Abstracts International A 57/03. (UMI No. 304205419)

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

Gottfried, T. L., & Beddor, P. S. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech, 31*, 51-75.

Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes, 13*, 601-623.

Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics, 28*, 27-42.

Hawkins, R. (1993). Regional variation in France. In C. Sanders (Ed.), *French today. Language in its social context* (pp. 55-84). Cambridge, U.K.: Cambridge University Press.

Hughes, D. M. (2006). Parent and self-ratings of executive function in adolescents with language impairments and typically developing peers. Dissertation Abstracts International B 67/05. (UMI No. 3216890)

Hulsen, M. (2000). Language loss and language processing. Three generations of Dutch migrants in New Zealand. Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen, Netherlands.

Institute of International Education. (2009, May). *Expanding study abroad capacity at American colleges and universities* (IIE Study Abroad White Paper Series No. 6). New York: Blumenthal, P. & Gutierrez, R. (Eds.).

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America, 97*, 553-562.

Jared, D., Friesen, D., & Haigh, C. (2008, October). *Cross-language phonological activation in bilingual word naming.* Paper presented at the International Conference on Models of Interaction in Bilinguals, Bangor, Wales.

Jared, D. & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44*, 2-31.

Jared, D. & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition, 5*, 225-239.

Jiménez Jiménez, A. F. (2003). Linguistic and psychological dimensions of second language attrition during and after a study abroad experience. Dissertation Abstracts International A 64/09. (UMI No. 304251099)

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Klein, D., Watkins, K. E., Zatorre, R. J., & Milner, B. (2006). Word and nonword repetition in bilingual subjects: A PET study. *Human Brain Mapping, 27*, 153-161.

Köpke, B. (2007). Language attrition at the crossroads of brain, mind, and society. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition. Theoretical perspectives* (pp. 9-37). Amsterdam: John Benjamins.

Köpke, B., & Schmid, M. S. (2004). Language attrition. The next phase. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues* (pp. 1-43). Amsterdam: John Benjamins.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149-174.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet" effect for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50,* 93-107.

Lederberg, A. R., & Spencer, P. E. (2005). Critical periods in the acquisition of lexical skills. In P. Fletcher, & J. F. Miller (Eds.), *Developmental theory and language disorders* (pp. 121-145). Amsterdam: John Benjamins.

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19,* 585-611.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122-142.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America, 125,* 1138-1152.

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

Macmillan, N. A., & Creelman, C. D. (2005). *Signal detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940-967.

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition, 11*, 81-93.

Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology, 4*, 383-388.

Mehotcheva, T. H. (2010). After the fiesta is over: Foreign language attrition of Spanish in Dutch and German Erasmus students. Unpublished doctoral dissertation, Universitat Pompeu Fabra, Barcelona, Spain.

Miyake, A. (1998). Individual differences in second language proficiency: The role of working memory. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339 –364). Mahwah, NJ: Lawrence Erlbaum Associates.

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory and Cognition, 20*, 705-714.

Muchnick, A. G., & Wolfe, D. E. (1982). Attitudes and motivation of American students of Spanish. *Canadian Modern Language Review, 38*, 262-281.

Murtagh, L. (2003). Retention and attrition of Irish as a second language. A longitudinal study of general and communicative proficiency in Irish among second level school leavers and

the influence of instructional background, language use and attitude/motivation variables. Unpublished doctoral dissertation, Rijksuniversiteit Groningen, Groningen, Netherlands.

Murtagh, L., & van der Slik, F. (2004). Retention of Irish skills: A longitudinal study of a school-acquired second language. *International Journal of Bilingualism, 8*, 279-302.

Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language, 63*, 541-559.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*, 261-283.

Paradis, J. (2011). Individual differences in child English second language acquisition. Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*, 213-237.

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of 'interference' in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics, 9*, 133-145.

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

Pisoni, D. B., Conway, C. M., Kronenberger, W., Henning, S., & Anaya, E. (2010). Executive function, cognitive control, and sequence learning in deaf children with cochlear implants. In M. Marschark, & P. E. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (pp. 439-457). Oxford University Press.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/ - /ð/ perception. Evidence of a new developmental pattern. *Journal of the Acoustical Society of America, 109*, 2190-2201.

Roth, R.M., Isquith, P.K., & Gioia, G.A. (2005). *BRIEF-A: Behavior rating inventory of executive function - adult version.* Lutz, FL: Psychological Assessment Resources.

Ryalls, J., Larouche, A., & Giroux, F. (2003). Acoustic comparison of CV syllables in French-speaking children with normal hearing, moderate-to-severe and profound hearing impairment. *Journal of Multilingual Communication Disorders, 1*, 99-114.

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

Schmid, M. S., & Dusseldorp, E. (2010). Quantitative analyses in a multivariate study of language attrition: The impact of extralinguistic factors. *Second Language Research, 26*, 125-160.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*, 197-212.

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

Smiljanic, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America, 118*, 1677-1688.

Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 221-234.

Studdert-Kennedy, M., Liberman, A. M., & Stevens, K. N. (1963). Reaction time to synthetic stop consonants and vowels at phoneme centers and at phoneme boundaries. *Journal of the Acoustical Society of America, 35*, 1900-1900.

Sundara, M. & Polka, L. (2008). Discrimination of coronal stops by bilingual adults. The timing and nature of language interaction. *Cognition, 106,* 234-258.

Sundara, M., Polka, L., & Baum, S. (2006). Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition, 9,* 97-114.

Sundara, M., Polka, L., & Genesee, F. (2006). Language-experience facilitates discrimination of /d/ - /ð/ in monolingual and bilingual acquisition of English. *Cognition, 100*, 369-388.

Sunderman, G., & Kroll, J. F. (2009). When study abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics, 30*, 79-99.

Tomiyama, M. (2000). Child second language attrition: A longitudinal case study. *Applied Linguistics, 21*, 304-322.

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning, 59,* 623-646.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second languages: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105-122.

Valdman, A., Pons, C., & Scullen, M. E. (2006). *Chez nous. Branché sur le monde francophone* (3rd ed.). Upper Saddle River, NJ: Pearson.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review, 9*, 780-789.

Vogt, W. P. (2007). *Quantitative research methods for professionals.* Boston, MA: Pearson.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-Ed.

Weltens, H. H. (1989). *The attrition of French as a foreign language*. Dordrecht, Netherlands: Foris.

Woutersen, M., Cox, A., Weltens, B., & De Bot, K. (1994). Lexical aspects of standard-dialect bilingualism. *Applied Psycholinguistics, 15*, 447-473.

Woutersen, M., De Bot, K., & Weltens, B. (1995). The bilingual lexicon: Modality effects in processing. *Journal of Psycholinguistic Research, 24*, 289-298.

Appendix A: Experimental materials

Table 10 shows the target words and contrast items from Experiment 1 (AX discrimination).

	Target word	Transcription	Translation	Contrast item	Transcription	Translation	Sound contrast
1	bouche	[buːʃ]	mouth	bûche	[byːʃ]	log	u – y
2	douleur	[dulœːʁ]	pain	duleur	[dylœːʁ]	nonword	u – y
3	poulet	[pule:]	chicken	pulet	[pyle:]	nonword	u – y
4	tousser	[tuse:]	to cough	tusser	[tyse:]	nonword	u – y
5	dur	[dyːʁ]	hard	dour	[duːʁ]	nonword	y – u
6	fûtes	[fyːt]	(you) were	foot	[fuːt]	football	y – u
7	pur	[pyːʁ]	pure	pour	[puːʁ]	for	y – u
8	zut	[zyːt]	darn	zoute	[zuːt]	nonword	y – u
9	neuve	[nœːv]	new	nuve	[nyːv]	nonword	œ – y
10	peur	[pœːʁ]	fear	pur	[pyːʁ]	pure	œ – y
11	sœur	[sœːʁ]	sister	sûr	[syːʁ]	sure	œ – y
12	veuve	[vœːv]	widow	vuve	[vyːv]	nonword	œ – y
13	cercle	[sɛʁkl]	circle	cércle	[seʁkl]	nonword	ɛ – e
14	messe	[mɛs]	mass	méçe	[mes]	nonword	ɛ – e
15	naisse	[nɛs]	bore	néçe	[nes]	nonword	ɛ – e
16	perdre	[pɛʁdʁ]	to lose	pérdre	[peʁdʁ]	nonword	ɛ – e
17	dé	[de:]	die	dais	[dɛ:]	canopy	e – ɛ
18	fée	[fe:]	fairy	fait	[fɛ:]	made	e – ɛ
19	pénombre	[penõːbʁ]	twilight	painombre	[pɛnõːbʁ]	nonword	e – ɛ
20	ténèbres	[tenɛːbʁ]	darkness	tainaibre	[tɛnɛːbʁ]	nonword	e – ɛ
21	bain	[bɛ̃:]	bath	baie	[bɛ]	bay	ɛ̃ – ɛ
22	feinte	[fɛ̃:t]	trick	fête	[fɛt]	party	ɛ̃ – ɛ
23	sainte	[sɛ̃:t]	saint	sept	[sɛt]	seven	ɛ̃ – ɛ
24	teint	[tɛ̃:]	complexion	tait	[tɛ]	is silent	ɛ̃ – ɛ

25	mèche	[mɛʃ]	wick	manche	[mɛ̃:ʃ]	sleeve	ɛ – ɛ̃
26	naît	[nɛ]	(she) bears	nain	[nɛ̃:]	gnome	ɛ – ɛ̃
27	paix	[pɛ]	peace	pain	[pɛ̃:]	bread	ɛ – ɛ̃
28	sèche	[sɛʃ]	dry	senche	[sɛ̃ʃ]	nonword	ɛ – ɛ̃
29	faute	[fo:t]	mistake	fonte	[fõ:t]	cast	o – õ
30	peau	[po:]	skin	pont	[põ:]	bridge	o – õ
31	saute	[so:t]	change	sonte	[sõ:t]	nonword	o – õ
32	taux	[to:]	quota	ton	[tõ:]	your	o – õ
33	fonder	[fõ:de]	to found	foder	[fo:de]	nonword	õ – o
34	mon	[mõ:]	my	mot	[mo:]	word	õ – o
35	non	[nõ:]	no	naux	[no:]	nonword	õ – o
36	sonder	[sõ:de]	investigate	soder	[so:de]	nonword	õ – o
37	dehors	[dəɔ:ʁ]	outside	tehors	[təɔ:ʁ]	nonword	d – t
38	dites	[di:t]	(you) say	tite	[ti:t]	nonword	d – t
39	dommage	[doma:ʒ]	harm	tommage	[toma:ʒ]	nonword	d – t
40	donc	[dõ:k]	therefore	tonc	[tõ:k]	nonword	d – t
41	tante	[tã:t]	aunt	dente	[dã:t]	nonword	t – d
42	taureau	[toʁo:]	bull	daureau	[doʁo:]	nonword	t – d
43	tellement	[tɛlmã:]	such	dellement	[dɛlmã:]	nonword	t – d
44	tige	[ti:ʒ]	stem	dige	[di:ʒ]	nonword	t – d
45	jaune	[ʒo:n]	yellow	jaugne	[ʒo:ɳ]	nonword	n – ɳ
46	poitrine	[pwatʁin]	chest	poitrigne	[pwatʁiɳ]	nonword	n – ɳ
47	prenne	[pʁɛn]	(he) take	prègne	[pʁɛɳ]	nonword	n – ɳ
48	soupçonne	[supson]	(he) suspects	soupçogne	[supsoɳ]	nonword	n – ɳ
49	besogne	[bəsoɳ]	task	besonne	[bəson]	nonword	ɳ – n
50	craigne	[kʁɛɳ]	(he) fear	crainne	[kʁɛn]	nonword	ɳ – n
51	grogne	[gʁoɳ]	resentment	gronne	[gʁon]	nonword	ɳ – n
52	souligne	[suliɳ]	(I) underline	soulinne	[sulin]	nonword	ɳ – n

Table 10. Target items from Experiment 1 (AX discrimination).

Table 11 shows the target words and nonwords for Experiment 2 (lexical decision). The target words (but not nonwords) were also used as target words in Experiment 4 (picture naming).

	Target word	Frequency	Translation	Lexical status	Matched nonword
1	bananes	2.74	bananas	cognate	pananes
2	blouse	14.42	blouse	cognate	bluse
3	bus	7.16	bus	cognate	bousse
4	cactus	1.52	cactus	cognate	cactousse
5	carotte	2.48	carrot	cognate	cavote
6	cigarettes	19.71	cigarettes	cognate	cigavettes
7	dentiste	3.23	dentist	cognate	dontiste
8	guitare	5.20	guitar	cognate	guitave
9	lampe	42.61	lamp	cognate	lambe
10	lettre	98.48	letter	cognate	lentre
11	papier	100.23	paper	cognate	bapier
12	pianiste	2.87	pianist	cognate	bianiste
13	prince	55.42	prince	cognate	brince
14	statue	18.48	statue	cognate	statou
15	table	200.58	table	cognate	dable
16	téléphone	60.19	telephone	cognate	déléphone
17	tomate	3.52	tomato	cognate	domate
18	touriste	6.03	tourist	cognate	turiste
19	train	161.55	train	cognate	traie
20	tube	23.52	tube	cognate	dube
21	bateau	41.94	boat	noncognate	banteau
22	barbe	29.71	beard	noncognate	parbe
23	boîte	58.77	box	noncognate	poîte
24	cahier	15.48	notebook	noncognate	canhier
25	canard	9.58	duck	noncognate	canave
26	citron	8.06	lemon	noncognate	citraux
27	diable	33.58	devil	noncognate	tiable
28	gants	16.77	gloves	noncognate	gaugne

29	langue	105.42	tongue	noncognate	logue
30	lait	44.90	milk	noncognate	lain
31	panier	16.32	basket	noncognate	banier
32	poitrine	54.10	chest	noncognate	poitrigne
33	prêtre	19.97	priest	noncognate	prentre
34	stylo	5.68	pen	noncognate	stylon
35	tapis	39.94	rug	noncognate	dapis
36	théière	1.65	teapot	noncognate	taithiere
37	toit	33.03	roof	noncognate	toint
38	tondeuse	0.77	lawnmower	noncognate	tonduse
39	tête	504.74	head	noncognate	téte
40	tuyau	7.19	hose	noncognate	tuyon

Table 11. Target items for Experiment 2 (lexical decision; words and nonwords) and Experiment 4 (picture naming; target words only).

Appendix B: Study Abroad Language Background Questionnaires

Questionnaire 1 (Time 0):

Study Abroad / Language Background Questionnaire 1

Date	Last Name	First Name	Study ID #

1. When did you move to France?

Month/Day/Year	

2. When did you return to the U.S.?

Month/Day/Year	

3. Total duration of stay in France (months & weeks):

4. Did you visit any other countries while abroad? *(Circle)*

Yes No

If yes, please list:

Country				
Duration of stay				
Languages spoken				

5. Had you ever visited a French speaking country before your recent study-abroad stay? *(Circle)*

Yes No

If yes, please list:

Country				
Duration of stay				
Date (Year/Month) of stay				
Did you speak French there? *(Circle)*	No Only Some	No Only Some	No Only Some	No Only Some

6. Please list what percentage of the time you were on average exposed to French and other languages during your stay abroad *(Your percentages should add up to 100%)*:

List language here					
List percentage here					

7. Did you take classes in French during your stay abroad? *(Circle)*

Yes No

If yes, please list:

Language Classes				
Level				
Hours/week in class				
Hours/week spent on homework, preparation etc.				
Duration of class in weeks/months				
On a scale from 1 (no problem at all/too easy) to 7 (couldn't understand anything), how difficult was this class for you? *(Circle)*	1 5 2 6 3 7 4	1 5 2 6 3 7 4	1 5 2 6 3 7 4	1 5 2 6 3 7 4

Content Classes				
Topic				
Hours/week in class				
Hours/week spent on homework, preparation etc. *in French*				
Duration of class in weeks/months				
Was this a class for French students or specifically for international students?	French International	French International	French International	French International
On a scale from 1 (too easy) to 7 (couldn't understand anything), how difficult was this class for you *language-wise*? *(Circle)*	1 5 2 6 3 7 4	1 5 2 6 3 7 4	1 5 2 6 3 7 4	1 5 2 6 3 7 4

8. Did you take classes in English/other languages while abroad? *(Circle)*

Yes No

If yes, please list:

Topic				
Language				
Hours/week in class				
Hours/week spent on homework, preparation etc.				
Duration of class in weeks/months				

9. Did you know French before you went on your study abroad? *(Circle)*

Yes No Minimal

If yes/minimal, please list how long you had learned French (years/semesters):

In elementary school or earlier	Junior High	High school	College

Other educational institution	Specify institution	At what age?

Outside of a school setting (e.g., with family, friends, self-study)	Please specify	At what age?

10. On a scale from 0-10, how would you rate your proficiency in speaking, reading, writing, and understanding spoken French BEFORE you went abroad? *0 = none, 10 = perfect. (Circle).*

Speaking	Reading	Writing	Understanding spoken language
0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10

11. On a scale from 0-10, how would you rate your proficiency in speaking, reading, writing, and understanding spoken French AT THE END of your stay abroad? *0 = none, 10 = perfect. (Circle).*

Speaking	Reading	Writing	Understanding spoken language
0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10

12. How would you rate your writing proficiency in French NOW? *(Circle)*

Writing	0 1 2 3 4 5 6 7 8 9 10

13. Are you currently taking a French class? *(Circle)*

Yes No, but would like to. No, no interest.

If yes, please specify what class(es) you are taking (name and/or number):

14. Do you plan to take more French classes in the future? *(Circle)*

Yes No

15. How often do you *currently* and *on average* **hear** spoken French? *(Circle)*

Never Every day Every week Every month

16. Where do you most often hear French? *(Circle all that apply)*

Home School Work Friends Radio TV Internet Other (specify):

17. Who speaks to you in French? *(Circle all that apply)*

Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

18. How often do you *currently* and *on average* **speak** French? *(Circle)*
Never Every day Every week Every month

19. Where do you speak French? *(Circle all that apply)*
Home School Work Friends Phone Internet Other (specify):

20. With whom do you speak French? *(Circle all that apply)*
Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

21. How often do you currently and on average **read** French? *(Circle)*
Never Every day Every week Every month

22. What do you read in French? *(Circle all that apply)*
Newspapers Magazines Literature Textbooks Letters/e-mail Websites
Other (specify):

23. Where do you read French? *(Circle all that apply)*
Home School Work At friends' homes Library Internet Other (specify):

24. How often do you currently and on average **write** in French? *(Circle)*
Never Every day Every week Every month

25. Who do you write to in French? (Circle all that apply)
Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

26. What do you write in French? *(Circle all that apply)*
Informal letters/e-mails Formal letters Essays/reports for school Chat/instant messaging
Poetry/literature Short notes (to friends, yourself…) Filling out forms Other (specify):

27. Where do you write in French? *(Circle all that apply)*
Home School Work Library Internet Other (specify):

28. Are you still in regular contact with French speakers in France? *(Circle)*
Yes No
- If yes, how often do you keep in touch via the following methods? *(Circle)*

E-mail: Never Every day Every week Every month Sporadically
Phone: Never Every day Every week Every month Sporadically
Letters: : Never Every day Every week Every month Sporadically
Chat/instant messaging: : Never Every day Every week Every month Sporadically
Visits: : Never Every day Every week Every month Sporadically
Other (specify): Never Every day Every week Every month Sporadically

- What language is the contact primarily in? *(Circle)*
French English French and English equally Other languages

29. Are you in contact with native French speakers in the Chicago area? *(Circle)*
Yes No

If yes, how regular is the contact? *(Circle)*

Every day Every week Every month Sporadically

What language do you primarily use? *(Circle)*

French English French and English equally Other languages

30. Are you in contact with other learners of French, with whom you speak French? *(Circle)*

Yes No

If yes, how frequent is your contact? *(Circle)*

Every day Every week Every month Sporadically

31. When speaking with other people who understand both English and French, do you ever mix the two languages in the same conversation? *(Circle)*

Yes, often Yes, but rarely No

32. *During your stay in France*, how often did you do the following? *(Circle)*

- Conversing with native speakers:
- at home with host family or friends
  - Never Every day Every week Every month Sporadically
- at school/work
  - Never Every day Every week Every month Sporadically
- in the course of daily routine (shopping, in transit etc.)
  - Never Every day Every week Every month Sporadically
- on the phone
  - Never Every day Every week Every month Sporadically

- Listening to French radio programs
  Never Every day Every week Every month Sporadically

- Watching French TV programs
  Never Every day Every week Every month Sporadically

- Seeing French movies
  Never Every day Every week Every month Sporadically

- Listening to French vocal music
  Never Every day Every week Every month Sporadically

- Reading newspapers/magazines
  Never Every day Every week Every month Sporadically

- Reading novels/literature
  Never Every day Every week Every month Sporadically

- Reading professional journals/articles
  Never Every day Every week Every month Sporadically

- Translating French written texts into English
  Never Every day Every week Every month Sporadically

- Interpreting spoken French into English
  Never Every day Every week Every month Sporadically

Questionnaire 2 (Time 1):

Study Abroad / Language Background Questionnaire 2

Date	Last Name	First Name	Study ID #

Date of last experiment	
French	
English	

1. Do you plan to be at NU/in Evanston in the Fall quarter 2010? *(Circle)*
Yes No

2. Please list what percentage of the time you are currently and on average exposed to English, French, and other languages *(Your percentages should add up to 100%)*:

Language:	English	French	Other *(specify)*:	Other *(specify)*:
Percentage:				

3. When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you. *(Your percentages should add up to 100%).*

Language:	English	French	Other *(specify)*:	Other *(specify)*:
Percentage:				

4. When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percent of total time. *(Your percentages should add up to 100%).*

Language:	English	French	Other *(specify)*:	Other *(specify)*:
Percentage:				

5. Please name the cultures with which you identify. On a scale from zero to ten *(0 = no identification, 10 = complete identification)*, please rate the extent to which you identify with each culture. (Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc).

List cultures here:			
Extent of identification (Circle).	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10

6. On a scale from 0-10, how would you rate your proficiency in understanding spoken French NOW?
0 = none, 10 = perfect. (Circle).

Understanding	0 1 2 3 4 5 6 7 8 9 10

7. How would you rate your speaking proficiency French NOW?

Speaking	0 1 2 3 4 5 6 7 8 9 10

8. How would you rate your reading proficiency French NOW?

Reading	0 1 2 3 4 5 6 7 8 9 10

9. How would you rate your writing proficiency French NOW?

Writing	0 1 2 3 4 5 6 7 8 9 10

10. In your perception, how much of a foreign accent do you have in French? *0 = none, 10 = pervasive. (Circle).*

Accent	0 1 2 3 4 5 6 7 8 9 10

11. Please rate how frequently others identify you as a non-native speaker based on your accent in French. *0 = never, 10 = always. (Circle).*

Non-native accent	0 1 2 3 4 5 6 7 8 9 10

12. Please rate to what extent you are currently exposed to French in the following contexts. *0 = never, 10 = always. (Circle).*

Interacting with friends	0 1 2 3 4 5 6 7 8 9 10	Listening to radio/music	0 1 2 3 4 5 6 7 8 9 10
Interacting with family	0 1 2 3 4 5 6 7 8 9 10	Reading	0 1 2 3 4 5 6 7 8 9 10
Watching TV	0 1 2 3 4 5 6 7 8 9 10	Language-lab, self-instruction	0 1 2 3 4 5 6 7 8 9 10

13. Have you taken any French classes since the last experiment? *(Circle)*

Yes No

If yes, please list:

Topic/class number									
Hours/week in class									
Hours/week spent on homework, preparation etc.									
Duration of class (list quarter, or number of weeks)									
On a scale from 1 (no problem at all/too easy) to 7 (couldn't understand anything), how difficult was this class for you? *(Circle)*	1 2 3 4	5 6 7	1 2 3 4	5 6 7	1 2 3 4	5 6 7	1 2 3 4	5 6 7

14. Do you plan to take more French classes in the future? *(Circle)*

Yes No

15. Have you visited any French speaking countries since the last experiment?

Yes No

If yes, please list:

Country				
Duration of stay				
Date (year/month) of stay				

Answer 15 – 15 c only if you answered "yes" to 15!

15a. Please list what percentage of the time you were, on average, exposed to French and other languages during this recent visit to a French speaking country *(Your percentages should add up to 100%)*:

List language here				
List percentage here				

15 b. Did you take classes in French during your recent visit?

Yes No

If yes, please list:

Type of class				
Hours/week in class				
Hours/week spent on homework, preparation etc.				
Duration of class in weeks				

15c. Do you feel like your French improved during your recent visit? *(Circle)*

No Some A lot

16. How often do you *currently* and *on average* **hear** spoken French? *(Circle)*

Never Every day Every week Every month

17. Where do you most often hear French? *(Circle all that apply)*

Home School Work Friends Radio TV Internet Other (specify):

18. Who speaks to you in French? *(Circle all that apply)*

Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

19. How often do you *currently* and *on average* **speak** French? *(Circle)*

Never Every day Every week Every month

20. Where do you speak French? *(Circle all that apply)*

Home School Work Friends Phone Internet Other (specify):

21. With whom do you speak French? *(Circle all that apply)*

Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

22. How often do you currently and on average **read** French? *(Circle)*

Never Every day Every week Every month

23. What do you read in French? *(Circle all that apply)*

Newspapers Magazines Literature Textbooks Letters/e-mail Websites
Other (specify):

24. Where do you read French? *(Circle all that apply)*
Home School Work At friends' homes Library Internet Other (specify):

25. How often do you currently and on average **write** in French? *(Circle)*
Never Every day Every week Every month

26. Who do you write to in French? (Circle all that apply)
Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

27. What do you write in French? *(Circle all that apply)*
Informal letters/e-mails Formal letters Essays/reports for school Chat/instant messaging
Poetry/literature Short notes (to friends, yourself…) Filling out forms Other (specify):

28. Where do you write in French? *(Circle all that apply)*
Home School Work Library Internet Other (specify):

29. Are you still in regular contact with French speakers in France? *(Circle)*
Yes No
- If yes, how often do you keep in touch via the following methods? *(Circle)*

E-mail: Never Every day Every week Every month Sporadically
Phone: Never Every day Every week Every month Sporadically
Letters: : Never Every day Every week Every month Sporadically
Chat/instant messaging: : Never Every day Every week Every month Sporadically
Visits: : Never Every day Every week Every month Sporadically
Other (specify): Never Every day Every week Every month Sporadically

- What language is the contact primarily in? *(Circle)*
French English French and English equally Other languages

30. Are you in contact with native French speakers in the Chicago area? *(Circle)*
Yes No
If yes, how regular is the contact? *(Circle)*
Every day Every week Every month Sporadically
What language do you primarily use? *(Circle)*
French English French and English equally Other languages

31. Are you in contact with other learners of French, with whom you speak French? *(Circle)*
Yes No
If yes, how frequent is your contact? *(Circle)*
Every day Every week Every month Sporadically

32. When speaking with other people who understand both English and French, do you ever mix the two languages in the same conversation? *(Circle)*
Yes, often Yes, but rarely No

Questionnaire 3 (Time 2):

Study Abroad / Language Background Questionnaire 3

Date	Last Name	First Name	Study ID #

1. Please list what percentage of the time you are currently and on average exposed to English, French, and other languages *(Your percentages should add up to 100%)*:

Language:	English	French	Other *(specify)*:	Other *(specify)*:
Percentage:				

2. When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you. *(Your percentages should add up to 100%).*

Language:	English	French	Other *(specify)*:	Other *(specify)*:
Percentage:				

3. When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percent of total time. *(Your percentages should add up to 100%).*

Language:	English	French	Other *(specify)*:	Other *(specify)*:
Percentage:				

4. Please name the cultures with which you identify. On a scale from zero to ten *(0 = no identification, 10 = complete identification)*, please rate the extent to which you identify with each culture. (Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc).

List cultures here:			
Extent of identification (Circle).	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10

5. On a scale from 0-10, how would you rate your proficiency in understanding spoken French NOW?
0 = none, 10 = perfect. (Circle).

Understanding	0 1 2 3 4 5 6 7 8 9 10

6. How would you rate your speaking proficiency French NOW?

Speaking	0 1 2 3 4 5 6 7 8 9 10

7. How would you rate your reading proficiency French NOW?

Reading	0 1 2 3 4 5 6 7 8 9 10

8. How would you rate your writing proficiency French NOW?

Writing	0 1 2 3 4 5 6 7 8 9 10

9. In your perception, how much of a foreign accent do you have in French? *0 = none, 10 = pervasive. (Circle).*

Accent	0 1 2 3 4 5 6 7 8 9 10

10. Please rate how frequently others identify you as a non-native speaker based on your accent in French. *0 = never, 10 = always. (Circle).*

Non-native accent	0 1 2 3 4 5 6 7 8 9 10

11. Please rate to what extent you are currently exposed to French in the following contexts. *0 = never, 10 = always. (Circle).*

Interacting with friends	0 1 2 3 4 5 6 7 8 9 10	Listening to radio/music	0 1 2 3 4 5 6 7 8 9 10
Interacting with family	0 1 2 3 4 5 6 7 8 9 10	Reading	0 1 2 3 4 5 6 7 8 9 10
Watching TV	0 1 2 3 4 5 6 7 8 9 10	Language-lab, self-instruction	0 1 2 3 4 5 6 7 8 9 10

12. Have you taken any French classes since the last experiment? *(Circle)*

Yes No

If yes, please list:

Topic/class number								
Hours/week in class								
Hours/week spent on homework, preparation etc.								
Duration of class (list quarter, or number of weeks)								
On a scale from 1 (no problem at all/too easy) to 7 (couldn't understand anything), how difficult was this class for you? *(Circle)*	1 2 3 4	5 6 7	1 2 3 4	5 6 7	1 2 3 4	5 6 7	1 2 3 4	5 6 7

13. Where did you take these classes *(Circle all that apply)*

In the U.S. In a French-speaking country Somewhere else

14. Are you taking any French classes this quarter? *(Circle)*

Yes No

If yes, please list:

Topic/class number				
Hours/week in class				
Hours/week spent on homework, preparation etc.				
Duration of class (list quarter, or number of weeks)				

On a scale from 1 (no problem at all/too easy) to 7 (couldn't understand anything), how difficult was this class for you? *(Circle)*	1 2 3 4	5 6 7	1 2 3 4	5 6 7	1 2 3 4	5 6 7	1 2 3 4	5 6 7

15. Do you plan to take more French classes in the future? *(Circle)*

Yes No

16. Did you visit any French speaking countries over the summer?

Yes No

If yes, please list:

Country				
Duration of stay				
Date of stay				

Answer 16 – 16 c only if you answered "yes" to 16!

16a. Please list what percentage of the time you were, on average, exposed to French and other languages during this recent visit to a French speaking country *(Your percentages should add up to 100%)*:

List language here				
List percentage here				

16 b. Did you take classes in French during your recent visit?

Yes No

If yes, please list (no need to list them again if you listed them in question 12):

Type of class				
Hours/week in class				
Hours/week spent on homework, preparation etc.				
Duration of class in weeks				

16c. Do you feel like your French improved during your recent visit? *(Circle)*

No Some A lot

17. How often do you *currently* and *on average* **hear** spoken French? *(Circle)*

Never Every day Every week Every month

18. Where do you most often hear French? *(Circle all that apply)*

Home School Work Friends Radio TV Internet Other (specify):

19. Who speaks to you in French? *(Circle all that apply)*

Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

20. How often do you *currently* and *on average* **speak** French? *(Circle)*

Never Every day Every week Every month

21. Where do you speak French? *(Circle all that apply)*
Home School Work Friends Phone Internet Other (specify):

21. With whom do you speak French? *(Circle all that apply)*
Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

23. How often do you currently and on average **read** French? *(Circle)*
Never Every day Every week Every month

24. What do you read in French? *(Circle all that apply)*
Newspapers Magazines Literature Textbooks Letters/e-mail Websites
Other (specify):

25. Where do you read French? *(Circle all that apply)*
Home School Work At friends' homes Library Internet Other (specify):

26. How often do you currently and on average **write** in French? *(Circle)*
Never Every day Every week Every month

27. Who do you write to in French? (Circle all that apply)
Family Friends in the U.S. Friends in France Teachers Classmates Other (specify):

28. What do you write in French? *(Circle all that apply)*
Informal letters/e-mails Formal letters Essays/reports for school Chat/instant messaging
Poetry/literature Short notes (to friends, yourself…) Filling out forms Other (specify):

29. Where do you write in French? *(Circle all that apply)*
Home School Work Library Internet Other (specify):

30. Are you still in regular contact with French speakers in France? *(Circle)*
Yes No
- If yes, how often do you keep in touch via the following methods? *(Circle)*

E-mail: Never Every day Every week Every month Sporadically
Phone: Never Every day Every week Every month Sporadically
Letters: : Never Every day Every week Every month Sporadically
Chat/instant messaging: : Never Every day Every week Every month Sporadically
Visits: : Never Every day Every week Every month Sporadically
Other (specify): Never Every day Every week Every month Sporadically

- What language is the contact primarily in? *(Circle)*
French English French and English equally Other languages

31. Are you in contact with native French speakers in the Chicago area? *(Circle)*
Yes No
If yes, how regular is the contact? *(Circle)*
Every day Every week Every month Sporadically

What language do you primarily use? *(Circle)*

French English French and English equally Other languages

32. Are you in contact with other learners of French, with whom you speak French? *(Circle)*

Yes No

If yes, how frequent is your contact? *(Circle)*

Every day Every week Every month Sporadically

33. When speaking with other people who understand both English and French, do you ever mix the two languages in the same conversation? *(Circle)*

Yes, often Yes, but rarely No

Appendix C: Adapted Mini-AMTB questionnaire

Language Questionnaire

DIRECTIONS: Please indicate your opinion after each statement by putting an X in the place that best describes the extent to which you believe the statement applies to you.

1. If I were to rate my feelings about learning French in order to interact with people from France/other

   French speaking countries, I would say it is:

   Weak____:____:____:____:____:____:____Strong

2. If I were to rate my interest in languages other than French or English, I would say that it is:

   Very Low____:____:____:____:____:____:____Very High

3. If I were to rate my attitude toward the French, I would say that it is:

   Unfavorable____:____:____:____:____:____:____Favorable

4. If I were to rate how hard I work at learning French, I would characterize it as:

   Very Little____:____:____:____:____:____:____Very Much

5. If I were to rate my desire to learn French, I would say that it is:

   Very Low____:____:____:____:____:____:____Very High

6. If I were to rate my attitude toward learning French, I would say that it is:

   Unfavorable____:____:____:____:____:____:____Favorable

7. If I were to rate how important it is for me to learn French for employment, I would say that it is:

Very Low____:____:____:____:____:____:____Very High

8. If I were to rate my anxiety when speaking French, I would rate myself as:

Very Calm____:____:____:____:____:____:____Very Nervous

If you are currently taking a French class, please, also answer the following questions:

1. If I were to rate my attitude toward my second language instructor, I would say that it is:

Unfavorable____:____:____:____:____:____:____Favorable

2. If I were to rate my attitude toward my second language course, I would say that it is:

Unfavorable____:____:____:____:____:____:____Favorable

3. If I were to rate my anxiety in my French class, I would rate myself as:

Very Calm____:____:____:____:____:____:____Very Nervous