Variables affecting the maintenance of L2 proficiency and fluency four years post-study abroad

Amanda Huensch, Nicole Tracy-Ventura, Judith Bridges, and Jhon A. Cuesta Medina
University of South Florida

This study explored the attrition / maintenance of second language (L2) proficiency by examining longitudinally the oral skills of a group of L2 French and L2 Spanish participants ($n = 33$) four years after study abroad, and three years after completing an undergraduate degree in languages. Multiple regressions were conducted to determine the extent to which language contact / use and attained proficiency at the end of study abroad could predict changes in fluency and oral proficiency. Results demonstrated that those variables that improved significantly during study abroad (e.g., speech rate) were maintained four years later. The amount of target language contact / use played a role in maintenance of aspects of fluency such as speech rate and frequency of silent pauses, whereas proficiency attained at the end of study abroad played a role in the use of corrections. Both language contact / use and proficiency attained are important variables in the long-term maintenance of overall proficiency.

Keywords: Spanish, French, fluency, proficiency, longitudinal study, foreign language attrition

1. Introduction

Research on study abroad (SA) has demonstrated that general proficiency and oral fluency are some of the main areas where learners typically make linguistic gains (e.g., Du, 2013; Guntermann, 1995; Hernández, 2010; Huensch & Tracy-Ventura, 2017a, 2017b; Kinginger, 2009, 2011; Magnan & Lafford, 2012; Mora & Valls-Ferrer, 2012; Yang, 2016). However, less is known about what happens in the long term to gains made during SA because research has generally followed a pre-test / post-test design without the addition of delayed post-tests. The small body

https://doi.org/10.1075/sar.17015.hue
Study Abroad Research in Second Language Acquisition and International Education 4:1 (2019), pp. 96–125.
issn 2405-5522 | e-issn 2405-5530

of study abroad literature that has included delayed post-tests has been limited to 12–15 months post-study abroad, a time during which participants continued to receive formal instruction. These studies have focused on aspects such as vocabulary (Pizziconi, 2017), grammar (Howard, 2009; Regan, 2005), and writing (Pérez-Vidal & Juan-Garau, 2009; Sasaki, 2011). Results of this work have demonstrated that gains made while abroad are largely maintained with formal instruction, and this finding seems to hold for some measures of oral fluency as well (Huensch & Tracy-Ventura, 2017b; Llanes, 2012). However, less is known about the long-term evolution of oral skills once participants are no longer receiving formal instruction. As Regan (2005, p. 193) wrote over a decade ago, "if indeed there is evidence that living in the speech community is beneficial, is it also true that these benefits are lasting?" This is an important question not only because it has implications for language teaching and SA program administration, but also because it has the potential to contribute to our understanding of foreign language attrition (Bardovi-Harlig & Stringer, 2010; Schmid & Mehotcheva, 2012) and the individual and contextual variables that influence it.

To investigate the long-term evolution of foreign language proficiency – including the possible outcomes of attrition, maintenance, and/or development – it is preferable to follow the same group of learners longitudinally, to administer a variety of language assessments, and to gather background data on internal and external factors such as language contact and use (Bardovi-Harlig & Stringer, 2010; Schmid & Mehotcheva, 2012). Following a longitudinal design makes it possible to empirically establish the highest proficiency level attained, or 'peak attainment' (Bardovi-Harlig & Stringer, 2010), which is essential when examining long-term retention or attrition. In general, the research suggests that proficiency level attained might predict the amount of attrition / retention (Mehotcheva, 2010; Murtagh, 2003). Of course every language learner is exposed to different amounts and types of input following the instructional period which could explain why learners show considerable variation in maintenance / attrition. However, the amount of reduced exposure (input and time) post-instruction has yet to emerge as a strong predictor of foreign language attrition (Mehotcheva, 2010; Murtagh, 2003; Weltens, 1989).

The current study fills a gap in our understanding of the long-term evolution of foreign language proficiency by examining longitudinally the oral skills of a group of L2 French and L2 Spanish participants who obtained a bachelor's degree in these languages. As part of their degree requirements, they spent their third year of a four-year degree abroad in a Spanish- or French-speaking country. Longitudinal data were collected before, during, and after their year abroad as part of the Languages and Social Networks Abroad Project (LANGSNAP: Mitchell, Tracy-Ventura, & McManus, 2017; http://langsnap.soton.ac.uk). Data from a follow-up

study were collected for the current study three years after LANGSNAP concluded (https://scholarcommons.usp.edu/langsnap). In order to investigate to what extent participants maintained the gains they made during SA and to what extent the variables of proficiency attained and amount of language exposure / use played a role in attrition / maintenance, oral fluency and proficiency data were compared across three time points: pre-sojourn (May 2011), at the end of participants' stay abroad (May 2012), and four years after study abroad (May 2016).

1.1 Effect of study abroad on linguistic gains in the short- and long-term

A substantial body of research has demonstrated the positive and immediate effects of SA on a number of linguistic variables, including the development of proficiency and oral fluency. For example, gains in proficiency have often been measured orally using the ACTFL Oral Proficiency Interview or Simulated Oral Proficiency Interview (e.g., Di Silvio, Diao, & Donovan, 2016; Hernández, 2010; Magnan & Back, 2007; Segalowitz & Freed, 2004), although examples also exist of studies which measured gains in proficiency using other standardized tests (Kinginger, 2008). Research on the development of L2 fluency during SA has generally demonstrated consistent improvement in both short-term (Llanes & Muñoz, 2009) and long-term (Huensch & Tracy-Ventura, 2017b; Serrano, Llanes, & Tragant, 2011) stays as well as across multiple languages such as Chinese (Du, 2013; Kim, Dewey, Baker-Smemoe, Ring, Westover, & Eggett, 2015), English (Mora & Valls-Ferrer, 2012; Serrano et al., 2011), French (Towell, Hawkins, & Bazergui, 1996), Russian (Di Silvio et al., 2016), and Spanish (Huensch & Tracy-Ventura, 2017a; Segalowitz & Freed, 2004).

Fluency, however, is a multidimensional construct and not all aspects have been shown to develop similarly. For example, while it is the case that global / compound measures (e.g., speech rate, mean length of run) typically improve, the development of breakdown and repair fluency is less straightforward. Huensch and Tracy-Ventura (2017b) analyzed the longitudinal development of oral fluency of 24 English L1 university learners of Spanish from LANGSNAP. Data were collected six times over nearly two years: once before, three times during a nine-month stay abroad, and twice after returning home (four months and eight months after living abroad) for their final year of university when they once again received formal instruction in Spanish. Nine measures of utterance fluency were analyzed. Results demonstrated that those measures which showed early and significant improvement while participants were abroad (e.g., mean syllable duration) were those that were best maintained the year after returning home. In contrast, those measures which were slower to develop while abroad (e.g., speech

rate, number of silent pauses) showed signs of attrition within eight months after returning home.

Huensch and Tracy-Ventura (2017b) is one of only a handful of studies which have incorporated delayed post-tests, a trend likely due to second language acquisition researchers' focus on measuring immediate effects of SA on language development. Llanes (2012) included a delayed post-test in her study comparing the linguistic development of children in a short-term SA program ($n = 9$) to that of children enrolled in at-home language classes ($n = 7$) on measures of both oral and written language production. Data were collected three times from both groups: before, immediately after a two-month study abroad program, and one year after returning home. Multiple linguistic variables were measured, including oral fluency, which was operationalized as speech rate (pruned syllables per minute). Results showed that participants from the SA group made greater gains in fluency than the at-home group, and that one year later those gains had been retained. Not all gains made during the stay abroad were maintained, however, after one year. In particular, participants became less accurate in oral production and less fluent in writing although they continued to receive English instruction during that time.

Results of Pérez-Vidal and Juan-Garau's (2009) longitudinal study investigating undergraduate English majors at a university in Barcelona demonstrated a number of statistically significant gains in L2 writing after a three-month SA period that were also maintained 15 months later. The seven participants who completed the delayed post-test had still been receiving formal instruction during the 15-month period, but no significant development occurred during that time. Regan (2005) explored the negation patterns of five Irish university-level learners of L2 French in sociolinguistic interviews conducted before, directly after, and one year after SA. She found a large increase in participants' use of *ne* deletion after their year abroad, showing similarity to French native speakers. The participants continued to receive formal instruction once they returned to Ireland. While it was hypothesized that their rate of *ne* deletion might decrease because they were no longer exposed to more colloquial language, results suggested the opposite. The increase they had made during their year abroad was maintained one year later.

Although there is some diversity in the studies just reviewed (e.g., children vs. adult participants, oral vs. written language, and the particular linguistic features under investigation), the findings suggest that when gains are made during SA, they are largely maintained 12–15 months later. All of the participants in those studies, however, continued to receive formal instruction, which likely contributed to this finding. Whether learners can maintain their L2 abilities after formal instruction has ended has not received much attention. Research focusing on

foreign language attrition comes the closest to addressing this question but, there, too, the research is limited.

1.2 Foreign language attrition

According to Schmid and Mehotcheva (2012), foreign language (FL) attrition refers specifically to the loss of an additional language that was learned via explicit instruction. This definition is used to distinguish between attrition of L2s which are acquired naturalistically as well as first language (L1) attrition. Despite an early interest in investigating L2/FL attrition (e.g., Bahrick, 1984; Weltens, 1989), little is known about the long-term evolution of FL proficiency. While some previous research has focused on FL attrition of French (e.g., Gardner, Lalonde, Moorcroft, & Evers, 1987; Weltens, 1989) and Spanish (e.g., Bahrick, 1984), most studies investigated FL attrition in school-age children (e.g., Cohen, 1986; Weltens, 1989) and/or what happens to students' L2 abilities during summer vacation (e.g., Gardner et al., 1987). Bahrick (1984) was the largest study of Spanish FL attrition by English L1 speakers ($n=773$), but this research followed a cross-sectional design and did not investigate oral language skills. To our knowledge, Mehotcheva (2010) is the only study which focused on FL attrition post-study abroad and included variables related to oral fluency.

Participants in Mehotcheva (2010) included Dutch- and German-speaking university students who had lived abroad in Spain as part of an academic exchange or internship. Cross-sectional data from 51 participants who varied in length of stay abroad (four-twelve months) and length of attrition (1–96 months), as well as longitudinal data from five participants who completed a range of assessments (e.g., a sociolinguistic interview, a C-test, and a picture-naming test) when they returned from their stay abroad and one year later, were analyzed for evidence of FL attrition. The main results of the longitudinal data demonstrated considerable individual variation among the five participants, but in general suggested that the participants were experiencing attrition, particularly in their reduced accessibility to lexical items. Disfluency markers were examined in the sociolinguistic interviews and results demonstrated an increase in filled pauses and repetitions after one year. Due to the larger number of participants in the cross-sectional data, Mehotcheva was able to investigate variables which were predictors of retention using multiple regression. Initial proficiency (measured via self-assessment) was a significant predictor of language retention in the use of filled pauses; those participants who reported having a lower proficiency at the onset of attrition (i.e., after returning from study abroad) produced more filled pauses. In contrast, little support was found for the role of disuse, language contact, or length of exposure during the post-study abroad and instructional period

for the various linguistic variables. Results for the other disfluency markers examined (e.g., repetitions and corrections) were unclear.

Results of this study and others point to interesting questions about the role of internal (e.g., age, motivation, proficiency) and external factors (e.g., amount of language use, length of exposure) in FL attrition and retention years later. In general, the limited research to date suggests that proficiency level attained might predict the amount of attrition / retention, more so than the amount of reduced exposure (input and time) which has not been shown to be a strong predictor of FL attrition (Mehotcheva, 2010; Murtagh, 2003; Weltens, 1989). Based on analysis of Bahrick (1984), Neisser (1984) claimed that there may be a general critical threshold related to L2 proficiency and long-term retention. That is, once learners reach that level, their linguistic knowledge becomes less vulnerable to attrition. Language use may also play a role in FL maintenance. According to Paradis (2007), lack of use is the cause of language attrition. His 'Neurolinguistic Theory of Bilingualism and the Activation Threshold Hypothesis' seeks to explain how the level of activation of linguistic items in memory is influenced by frequency and recency of use (Paradis, 2004). The more often items are activated, the lower the threshold will be to activate them. Findings from studies testing Paradis' claim have been mixed, however, and suggest that other internal factors, such as proficiency in the case of FL attrition, might be more important than language use. Thus, prior research and theorizing motivate the inclusion of both proficiency level attained and language contact / use as relevant variables in research on attrition / maintenance.

1.3 The current study

Longitudinal research focusing on the long-term evolution of foreign language proficiency post-study abroad has primarily been limited to investigating the retention of linguistic gains made one year later. Therefore, little is known about the extent to which gains in various linguistic areas are maintained after participants are no longer receiving formal instruction. Furthermore, it is unclear to what extent variables such as language contact and attained proficiency play a role in language maintenance. Thus, the current study investigates the following research questions:

1. To what extent do speed, breakdown, and repair fluency (in an oral, picture-based narrative) of L2 French and Spanish learners change four years after study abroad?
2. To what extent does the oral proficiency (as measured by an elicited imitation test) of L2 French and Spanish learners change four years after study abroad?

3. To what extent can post-study abroad language exposure and attained proficiency at the end of study abroad predict the utterance fluency and oral proficiency changes of learners four years after study abroad?

2. Methods

2.1 Participants

Participants are a subset of those from LANGSNAP (Mitchell, Tracy-Ventura, & McManus, 2017), a nearly two-year study which tracked 56 French and Spanish undergraduate degree students before, during, and after their required third year abroad.[1] The final data collection wave in that project was in February 2013, during participants' final year of university. In spring 2016, all of the original participants were invited to participate in the current study. A total of 33 (59%) agreed to participate: 15 Spanish L2 learners (twelve females) and 18 French L2 learners (17 females). One participant from each language group had either missing or unusable data. Therefore, a total of 14 Spanish participants and 17 French participants were included in the current analysis. Table 1 summarizes the ages and years studying the L2 of the participants before they studied abroad (May 2011).

Table 1. Participant demographic information at pre-sojourn

	French L2 learners (*n* = 17)	**Spanish L2 learners (*n* = 14)**
Mean Age (*SD*)	19.7 (0.6)	20.5 (1.3)
Years Studying L2 (*SD*)	10.5 (3.0)	6.1 (3.7)

2.2 Project timeline

The complete LANGSNAP timeline is outlined in Table 2, including the dates of data collection waves, where they occurred, and what version of the narrative task was used. The three waves used in the current analysis were the pre-sojourn, in-sojourn 3, and post-sojourn 3. In-sojourn 3 (approximately nine months after participants had been residing abroad) was the last SA data point.[2] Post-sojourn 3

1. Unlike typical US study abroad programs, participants had a choice of placement (exchange students, language teaching assistants, or workplace interns). For further details on placement types, see Mitchell, McManus, and Tracy-Ventura (2015).

2. Two participants' last study abroad data points were from in-sojourn 2; thus, their data from that collection round are used in the current study.

(approximately four years after participants returned home from study abroad) was chosen to explore the long-term effects of SA.

Table 2. Project timeline

Date	May 2011	November 2011	February 2012	May 2012	October 2012	February 2012	May 2016
Location	Home	Abroad	Abroad	Abroad	Home	Home	Varied
Wave	Pre-sojourn	In-sojourn 1	In-sojourn 2	In-sojourn 3	Post-sojourn 1	Post-sojourn 2	Post-sojourn 3
Narrative	Cat	Sisters	Brothers	Cat	Sisters	Brothers	Cat
Proficiency Test	Yes	No	Yes	No	Yes	No	Yes

2.3 Materials and procedures

2.3.1 *Oral production data*

Speech samples were elicited in LANGSNAP using picture-based narratives designed to be maximally similar to each other and administered in sequence approximately one year apart (except for the final data collection wave which was three years later) to avoid task repetition effects (see Table 2). Each story was approximately 15 pages in length and included a limited number of written prompts in the target language. The stories began with images representing typical activities for the characters before turning to images depicting a series of main events. In the current study, data come from The Cat Story (Dominguez, Tracy-Ventura, Arche, Mitchell, & Myles, 2013), which begins by showing the normal daily activities of a little girl and her cat, followed by the events of an atypical day when the cat chases a bird and gets lost.[3] Both characters make unsuccessful attempts to find each other, but eventually they enjoy a happy reunion.

Participants were given a few minutes to familiarize themselves with the story. Then, they were asked to re-tell the story in their own words, and they were able to continue looking at the pictures during their narrations. No time limit was imposed; thus, participants' narrations varied in length. Table 3 provides the

3. Recall that two participants' data come from the in-sojourn 2 data collection wave which used The Brothers Story. To examine narrative task comparability, data from French and Spanish native speakers were analyzed, as reported in Huensch and Tracy-Ventura (2017a), and results indicated no significant differences in fluency across the tasks. Thus, there is evidence that these versions of the narrative task are similar to each other and any learner differences in fluency across the narratives are attributed to time rather than differences in task version.

means, standard deviations, and ranges of recording length at each data collection wave.[4]

Table 3. Means, standard deviations, and ranges for recording length at each wave

	French L2 learners	Spanish L2 learners
Pre-sojourn	5:17 (2:03)	3:54 (1:15)
	2:59–10:29	2:28–6:10
In-sojourn 3	3:22 (1:33)	3:04 (1:03)
	1:48–6:57	1:33–4:53
Post-sojourn 3	3:27 (0:49)	3:28 (1:10)
	1:45–5:22	1:51–4:39

2.3.2 *Proficiency data: Elicited imitation test (EIT)*

French and Spanish versions of the same proficiency test were administered. These tests, based on elicited imitation, have been found to be valid and reliable measures of proficiency (Bowden, 2016; Tracy-Ventura et al., 2014). The Spanish version was created by Ortega (2000) and the French test was designed by Tracy-Ventura, McManus, Norris, and Ortega (2014) for LANGSNAP. Both versions of the EIT consist of 30 statements that gradually increase in syllable length. French items ranged from 7–19 syllables in length, and the Spanish items ranged from 7–17 syllables. Participants listened to and then repeated (after a short pause to minimize parroting) the 30 items as accurately as possible. The EIT was administered via a recording, took approximately ten minutes to complete, and participants' responses were audio-recorded for scoring (see Section 2.4). Both versions of the test are available on the Instruments for Research into Second Languages (IRIS, <www.iris-database.org>) digital database.

2.3.3 *Language contact data*

The post-sojourn 3 data collection wave included several tasks designed to gather information about the language exposure (contact and use) that participants had with their L2 in the three years that had passed since they graduated from university. Tasks used to collect language exposure data included two interviews (one in English and one in their L2) and two online questionnaires. The interviews

4. With respect to the means at the pre-sojourn, there were two French participants whose narrations lasted 9.5 and 10.5 minutes (the remaining 29 participants' narrations were between 2.5 and 6.5 minutes). These two longer narrations resulted in the difference between the French and Spanish averages (without these two narrations, the average French narration time is four minutes 39 seconds). Nevertheless, the fluency measures reported in the analysis are normalized, such that they are not dependent on the length of sample.

each lasted approximately 10–20 minutes and included questions related to participants' travel or residence abroad since graduation; their self-reported language abilities; the ways (if any) in which they had used their L2; their professional or academic activities; their social networks and the languages they used with the most important / frequent contacts in their daily lives; and, generally, what their lives were currently like.

The first questionnaire was a background questionnaire that contained some content overlap with the interviews and was designed to similarly collect information about the participants' exposure to the L2 since graduation. The questionnaire was adapted from Mehotcheva (2010). The second questionnaire was the Language Engagement Questionnaire (LEQ, McManus, Mitchell, & Tracy-Ventura, 2014). The LEQ (available on IRIS) was designed to collect information about the frequency and type of contact and use the learners had with the different languages in their repertoire (e.g., L1, L2, L3, L4). The questionnaire begins by eliciting from the participants any languages they use on a regular basis. Then, for each language, the participants indicated how frequently they use that language (on a six-point scale) for a variety of different activities (e.g., watching television, listening to music, reading magazines, writing texts, teaching a class).

2.4 Data coding and analysis

Data coding for the narrative recordings followed the same procedure used for previous investigations of fluency in LANGSNAP: narrative data were transcribed in CLAN following CHAT conventions (MacWhinney, 2000) and checked by at least two members of the research team. Instances of filled pauses (e.g., 'uh', 'um'), repetitions (i.e., repetition of a word or phrase), and corrections (i.e., modification of a word or phrase) were annotated in the transcripts using the symbols, &, [/], and [//], respectively, to allow for automatic counting in CLAN. Any written prompts included in the narratives were coded as such and omitted from the analysis. Silent pauses (longer than 250 ms) were coded in Praat (Boersma & Weenik, 2015) using the 'Annotate To TextGrid (silences)…' command and manually checked by a member of the research team. Syllables were counted manually based on the transcripts by a member of the research team, with a subset counted by a second member of the research team. Interrater reliability for syllable counts, calculated using Cronbach's alpha because the data were continuous, was high, at .99 for both the French and Spanish data. Five fluency variables representing aspects of speed, breakdown, and repair fluency (Skehan, 2003; Tavakoli & Skehan, 2005) were used in the current study. The variables were chosen based on previous research including Llanes (2012) and Mehotcheva (2010). Speed fluency was represented by speech rate, which was

calculated by dividing the total number of syllables by the total speaking time in seconds (including pauses). Breakdown fluency was represented by the number of filled pauses per 100 words and the number of silent pauses per 100 words. Repair fluency was represented by the number of repetitions per 100 words and the number of corrections per 100 words.

The EIT scoring is based on the five-point rubric (0–4) described in Ortega (2000), giving a maximum possible score of 120 points. Each item received one of the following possible scores: 4 (exact repetition), 3 (meaning preserved, but small changes in grammar), 2 (slight departures in meaning), 1 (important content missing), and 0 (little to no original content).

Regarding language exposure, interview and questionnaire data were triangulated to categorize participants as having 'intense' or 'limited' exposure depending on the consistency and intensity of their contact with the L2 since graduation. Participants coded as intense exposure were typically those that had additional experience living abroad or used their L2 often for work. For example, four participants were living abroad at the time of the post-sojourn 3 data collection wave; five others had additional overseas experiences during their three years after graduation (i.e., they lived abroad in a French- or Spanish-speaking country for nine-twelve months); several had L2-speaking partners; and others used the L2 extensively at work. In all, twelve of the participants (six French, six Spanish) were categorized as intense exposure. Those categorized as limited exposure were mostly working in their home countries in a job that required little to no L2 use. These participants' LEQ responses either did not include the L2 as a language used on a regular basis or demonstrated practically no engagement with the L2. Any travel in an L2-speaking country was typically limited to several days / weeks for holidays. In all, 19 of the participants (eleven French, eight Spanish) were categorized as limited exposure.

3. Results

Research question one investigated the change in fluency and oral proficiency over time of L2 French and Spanish learners. Because the participants in the current study are a smaller subset of LANGSNAP participants, we felt it was important to test whether the same results would be found over time with this particular subgroup. Therefore, separate mixed within / between repeated-measures ANOVAs (one for each independent variable) with time as within-subjects variable (pre-sojourn, in-sojourn 3, post-sojourn 3) and group as between-subjects variable (French, Spanish) were conducted. In cases where a main effect of time and no interaction between time and group were found, planned comparisons (with Bonferroni adjustments) were conducted among the three time

points. Table 4 provides the means and standard deviations of the five fluency variables for each group at each time point and Figures 1a–e present the box plots. Some of the assumptions of the ANOVA test were violated (e.g., normally distributed data, presence of outliers; see Table 4 and Figures 1a–e); however, given that the ANOVA is relatively robust to violations of normality and that results are interpreted using Cohen's *d*, the ANOVA results are reported with a Greenhouse-Geisser correction (Larson-Hall, 2016). Effect sizes of the ANOVA tests are reported using partial eta squared and of the planned comparisons using *d* (with 95% confidence intervals [CI]). The *d* scores are interpreted following field-specific recommendations for within-subjects comparisons: small = 0.60, medium = 1.00, large = 1.40 (Plonsky & Oswald, 2014).

Regarding speech rate (number of syllables / total time), results of the repeated-measures ANOVA indicated a significant main effect for time $F(1.60, 46.38) = 66.61$, $p < .001$, $\eta_p^2 = .697$, but no main effect for group $F(1, 29) = 0.52$, $p = .478$, $\eta_p^2 = .018$, nor an interaction between group and time $F1.60, 46.38) = 0.31$, $p = .687$, $\eta_p^2 = .011$. Planned comparisons indicated a significant difference and large effect between the pre-sojourn and in-sojourn 3 ($p < .001$, $d = 1.60$, 95% CI [1.01, 2.15]), and a significant difference and medium effect between the pre-sojourn and post-sojourn 3 ($p < .001$, $d = 1.24$, 95% CI [0.68, 1.77]). There was no significant difference (and a negligible effect) between in-sojourn 3 and post-sojourn 3 ($p = .435$, $d = -0.16$, 95% CI [−0.66, 0.34]). These results suggest that for both groups of learners, speech rate improved significantly during study abroad and those improvements were maintained four years after returning home, and three years after graduating from university.

Regarding the number of silent pauses, results of the repeated-measures ANOVA indicated a main effect of time, $F(1.68, 48.65) = 61.68$, $p = .000$, $\eta_p^2 = .680$, and a main effect of group, $F(1, 29) = 4.37$, $p = .045$, $\eta_p^2 = .131$, but no interaction between group and time, $F(1.68, 48.65) = 0.32$, $p = .691$, $\eta_p^2 = .011$. Planned comparisons indicated significant differences and medium effects between the pre-sojourn and in-sojourn 3 ($p < .001$, $d = -1.15$, 95% CI [−1.67, −0.60]), and the pre-sojourn and post-sojourn 3 ($p = .000$, $d = -1.01$, 95% CI [−1.53, −0.47]). There was no significant difference (and a negligible effect) between in-sojourn 3 and post-sojourn 3 ($p = 1.000$, $d = 0.07$, 95% CI [−0.43, 0.56]). Similar to the findings for speech rate, these results suggest that for both groups of learners, silent pauses decreased significantly during study abroad and those improvements were maintained four years later.

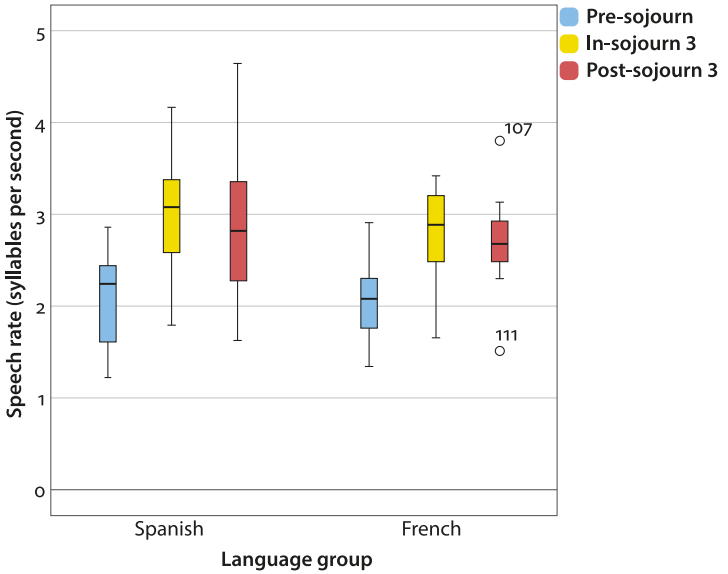
Regarding the number of filled pauses, results of the repeated-measures ANOVA indicated a main effect of time $F(1.63, 47.36) = 34.36$, $p = .000$, $\eta_p^2 = .542$, and an interaction between group and time $F(1.63, 47.36) = 8.83$, $p = .001$, $\eta_p^2 = .233$, but no main effect of group $F(1, 29) = 0.04$, $p = .843$, $\eta_p^2 = .001$. Given the group

Table 4. Means (and standard deviations) of the fluency variables for each group separated by time

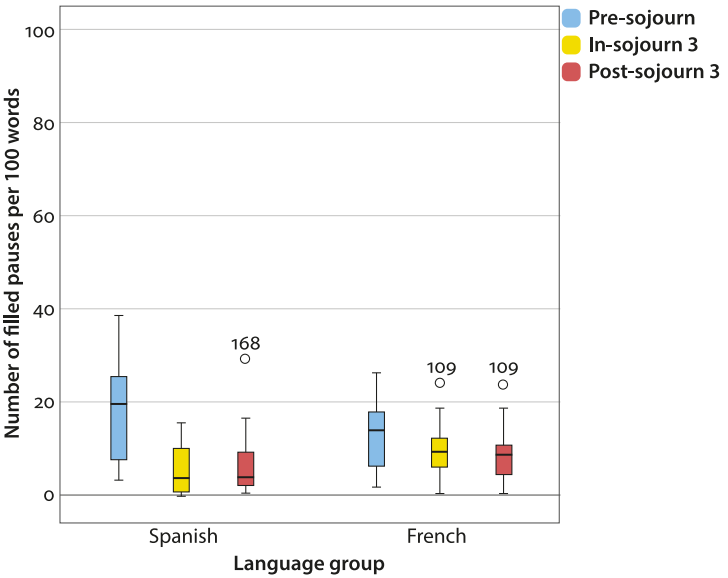
	French			Spanish		
	Pre-sojourn	In-sojourn 3	Post-sojourn 3	Pre-sojourn	In-sojourn 3	Post-sojourn 3
Speech rate	2.03	2.80	2.71	2.10	2.98	2.87
(syllables/ second)	(0.42)	(0.48)	(0.46)	(0.53)	(0.64)	(0.89)
Filled Pauses	13.29	9.37	9.14	18.10	5.19	7.10 +
per 100 words	(7.76)	(6.23)	(6.26)	(10.42)	(5.25)	(7.89)
Silent Pauses	35.61	21.67	21.48	44.84	28.79	30.69
per 100 words	(10.11)	(7.44) +	(6.87) +	(18.78)	(12.16) +	(16.08)
Repetitions	3.62	3.80	3.49	4.32	3.74	2.89
per 100 words	(1.98)	(2.27)	(2.42)	(2.76)	(2.04)	(1.94)
Corrections	3.18	1.89	2.25	3.94	3.86	3.61
per 100 words	(2.18) +	(1.27)	(1.49) +	(1.51)	(1.77)	(1.80)

+ Shapiro-Wilk test $p < .05$

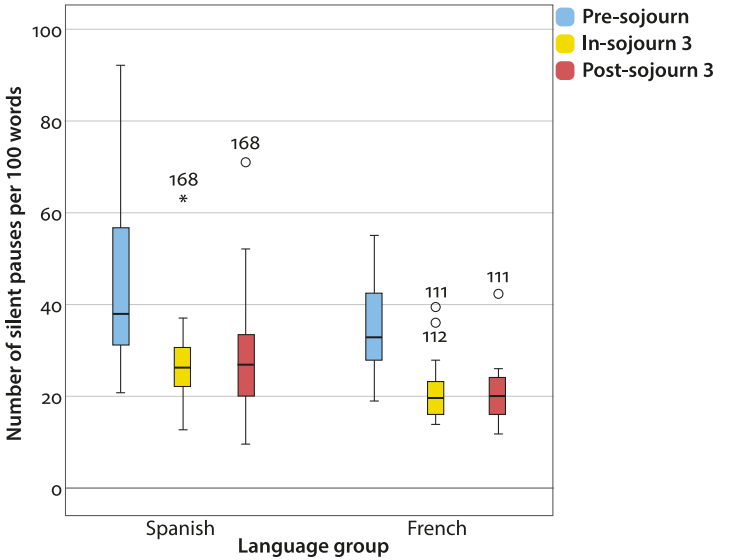
by time interaction, a separate repeated-measures ANOVA was conducted for each group with time (pre-sojourn, in-sojourn 3, post-sojourn 3) as the within-subjects variable. The ANOVA for the Spanish group indicated a main effect of time, $F(1.66, 21.55) = 30.67$, $p < .001$, $\eta_p^2 = .702$, and planned comparisons indicated a significant difference and large effect between the pre-sojourn and in-sojourn 3 ($p < .001$, $d = -1.56$, 95% CI [−2.36, −0.68]), and a significant difference and medium effect between the pre-sojourn and post-sojourn 3 ($p < .001$, $d = -1.19$, 95% CI [−1.96, −0.36]). There was no significant difference (and a negligible effect) between in-sojourn 3 and post-sojourn 3 ($p = .523$, $d = 0.28$, 95% CI [−0.47, 1.02]). The ANOVA for the French group also indicated a main effect of time, $F(1.58, 25.33) = 5.50$, $p = .015$, $\eta_p^2 = .256$; however, planned comparisons (with Bonferroni adjustments) only approached significance (with small effects; however CIs passed through zero) for the comparison of pre-sojourn and in-sojourn 3 ($p = .060$, $d = -0.56$, 95% CI [−1.23, 0.14]), and pre-sojourn and post-sojourn 3 ($p = .065$, $d = -0.59$, 95% CI [−1.26, 0.11]); the difference between in-sojourn 3 and post-sojourn 3 was not significant and the effect size was close to zero ($p = 1.000$, $d = -0.04$, 95% CI [−0.64, 0.71]). These results suggest that the Spanish group


Pre-sojourn
In-sojourn 3
Post-sojourn 3
Speech rate (syllables per second)
0
1
2
3
4
5
107
111
Spanish
French
Language group

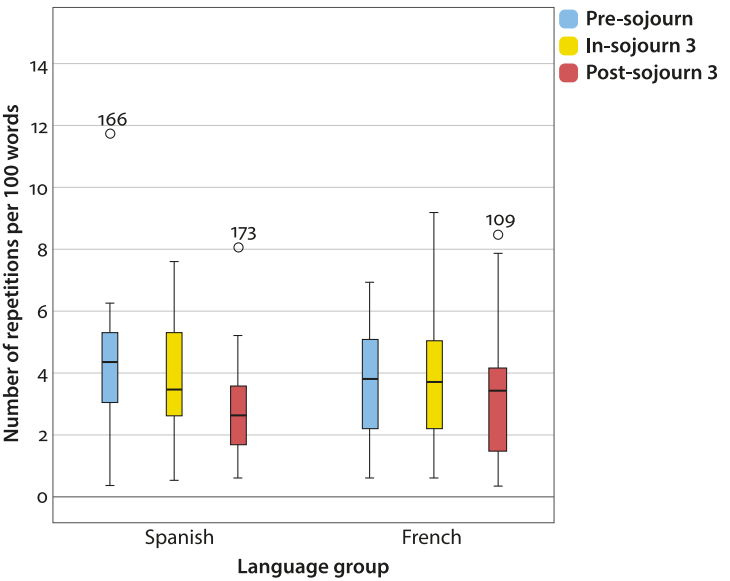
a.


Pre-sojourn
In-sojourn 3
Post-sojourn 3
Number of filled pauses per 100 words
0
20
40
60
80
100
168
109
109
Spanish
French
Language group

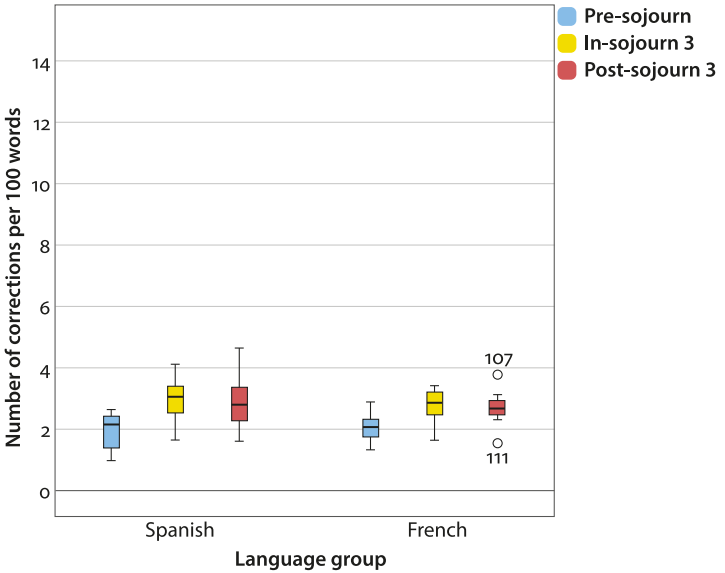
b.


Pre-sojourn
In-sojourn 3
Post-sojourn 3
Number of silent pauses per 100 words
100
80
60
40
20
0
168
168
111
112
111
Spanish
French
Language group

c.


Pre-sojourn
In-sojourn 3
Post-sojourn 3
Number of repetitions per 100 words
14
12
10
8
6
4
2
0
166
173
109
Spanish
French
Language group

d.



e.

Figure 1. Box plots of the five fluency variables at each data collection wave for each group[5]

produced fewer filled pauses at the end of study abroad and those gains were maintained four years later, whereas the French group's decrease in the use of filled pauses after study abroad did not appear to be statistically meaningful.

Regarding the number of repetitions, the results of the repeated-measures ANOVA indicated no main effect of time, $F(1.94, 56.34) = 1.57$, $p = .218$, $\eta_p^2 = .051$, no main effect of group, $F(1, 29) = 0.00$, $p = .986$, $\eta_p^2 = .000$, and no interaction between group and time, $F(1.94, 56.34) = 1.03$, $p = .363$, $\eta_p^2 = .034$. Unlike the results for speech rate and silent pauses, these results suggest that for both groups of learners, the number of repetitions did not change significantly during study abroad or four years later.

Regarding the number of corrections, the results of the repeated-measures ANOVA indicated a main effect of group, $F(1, 29) = 8.23$, $p = .008$, $\eta_p^2 = .221$, but no main effect of time, $F(1.84, 53.41) = 2.64$, $p = .085$, $\eta_p^2 = .083$, and no interaction between group and time, $F(1.84, 53.41) = 1.65$, $p = .204$, $\eta_p^2 = .054$. These results suggest that, similar to the findings for repetitions, the number of corrections did not change significantly during study abroad or four years later.

---

5. The numbers appearing in the box plots (e.g., 166) are the participant numbers and were included because the data are publicly available. Thus, it is possible to track participants across studies.

Finally, the results of the learners' proficiency scores are reported; however, recall that the EIT was not administered at in-sojourn 3, thus results are from the post-sojourn 1 data collection wave. Table 5 provides the means and standard deviations of the EIT scores separated by group and Figure 2 displays the box plots.

Table 5. Descriptive statistics for the elicited imitation test

		Mean	Standard deviation
French ($n = 17$)	Pre-sojourn	65.71	14.22
	Post-sojourn 1	89.76	9.97
	Post-sojourn 3	96.18	9.63
Spanish ($n = 14$)	Pre-sojourn	80.64	11.55
	Post-sojourn 1	104.57	8.54
	Post-sojourn 3	101.00	12.17

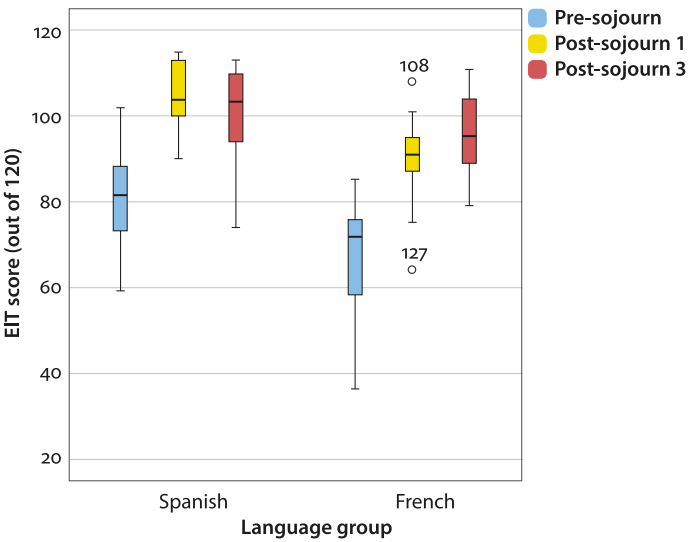


Figure 2. Box plots of EIT scores at each data collection wave for each group

Results of the repeated-measures ANOVA indicated a significant main effect for time $F(1.77, 51.31) = 115.62$, $p = .000$, $\eta_p^2 = .799$ and group $F(1, 29) = 11.41$, $p = .002$, $\eta_p^2 = .282$, and an interaction between group and time $F(1.77, 51.31) = 4.77$, $p = .016$, $\eta_p^2 = .141$. Given the significant interaction between time and group, separate repeated-measures ANOVAs were conducted for each group with time as the within-subjects variable. Both tests indicated a main effect of time: $F(1.79, 23.33) = 53.33$, $p < .001$, $\eta_p^2 = .804$ for the Spanish group, and $F(1.75, 27.94) = 70.04$, $p < .001$, $\eta_p^2 = .814$ for the French group. Planned comparisons

for the Spanish group indicated significant differences and large effects between the pre-sojourn and post-sojourn 1 ($p < .001$, $d = 2.36$, 95% CI [1.33, 3.24]), and the pre-sojourn and post-sojourn 3 ($p < .001$, $d = 1.72$, 95% CI [0.81, 2.53]), but no significant difference between post-sojourn 1 and post-sojourn 3 ($p = .346$, $d = -0.34$, 95% CI [−1.08, 0.42]. Planned comparisons for the French group indicated significant differences and large effects between the pre-sojourn and post-sojourn 1 ($p < .001$, $d = 1.96$, 95% CI [1.10, 2.73]) and the pre-sojourn and post-sojourn 3 ($p < .001$, $d = 2.51$, 95% CI [1.56, 3.34]), and a significant difference between post-sojourn 1 and post-sojourn 3, however, with a small effect whose CIs passed through zero ($p = .027$, $d = 0.65$, 95% CI [−0.05, 1.33]. These suggest that oral proficiency improved during study abroad and those improvements were maintained in the long term.

The second research question explored the extent to which post-study abroad language exposure and proficiency could predict changes in fluency and proficiency from in-sojourn 3 to post-sojourn 3. Given the relatively small number of participants, it was necessary to limit the number of independent or explanatory variables to those two that seem most likely to contribute to language development, maintenance, or attrition over time (Larson-Hall, 2016). Our two predictor variables were motivated by previous research and theories of attrition: (1) language exposure (as a variable taking into account the consistency and intensity of L2 contact and use between returning home from study abroad and post-sojourn 3) and (2) oral proficiency (as measured by the EIT at post-sojourn 1). Language exposure was coded as a binary variable (0/1), based on participants being classified as limited exposure ($n = 19$) or intense exposure ($n = 12$).

In addition, a correlation analysis was conducted using the change scores for the five fluency variables and the EIT, and the two explanatory variables (see Table 6). Based on this preliminary step, a decision was made to limit the regression analyses to those predictor variables (i.e., change scores) where significant correlations were found with at least one of the explanatory variables: speech rate change, silent pause change, corrections change, and proficiency change.

The assumptions for multiple regression were checked and met for each analysis. When checking for influential points in the analyses using studentized deleted residuals and leverage and Cook's Distance values, no individual showed up as problematic for all three tests. Therefore, we proceeded with the analyses. Table 7 provides the means and standard deviations for each of the dependent variables.

The results of the multiple regressions are provided in Table 8. Figures 3a–d are scatterplots of each dependent variable plotted against EIT scores with those categorized as limited exposure represented by the black dots and those categorized as intense exposure represented by the white dots. This visual information is helpful for interpreting the results of the regressions. For speech rate,

Table 6. Correlations between the language variables and exposure and proficiency (EIT)

	Exposure	Speech rate change	Silent pause change	Filled pause change	Repetitions change	Corrections change	Proficiency change	SA end proficiency
Exposure	1	.470 **	−.427 *	−0.229	0.003	−0.038	.381 *	−0.088
Speech rate change		1	−.686 **	0.017	−0.061	0.064	.201	−0.160
Silent pause change			1	0.313	−0.136	0.151	−.291	0.115
Filled pause change				1	0.020	0.168	−.338	0.238
Repetitions change					1	0.195	.355	−0.004
Corrections change						1	.319	−0.503 **
Proficiency change							1	−0.502 **
SA end proficiency								1

* $p < .05$ ** $p < .01$

Table 7. Means and standard deviations of the multiple regression dependent variables

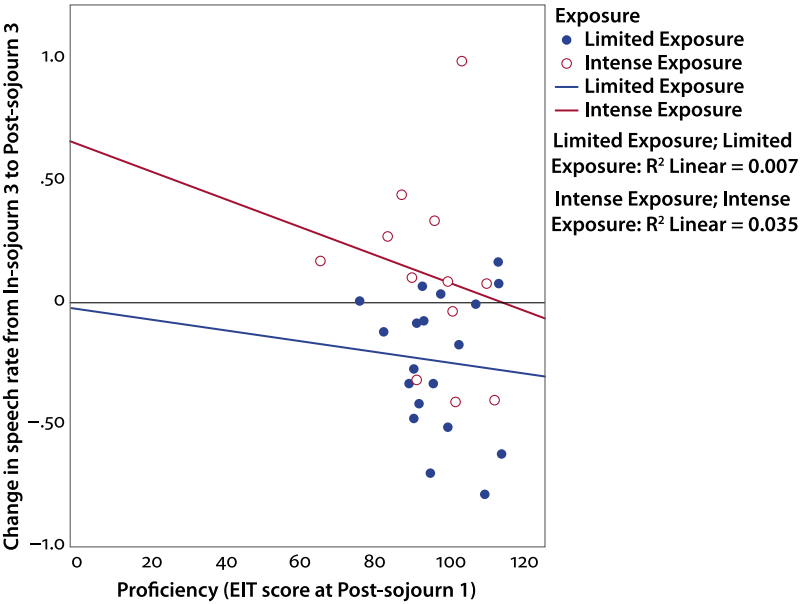
	Mean	Standard deviation
Speech rate change	−0.10	0.37
Silent pause change	0.75	6.46
Corrections change	0.09	1.17
Proficiency change	1.90	9.73

results indicated a statistically significant model, with an R^2 value of .18. Only language exposure ($p = .010$) contributed significantly to prediction of change in speech rate. As shown in Figure 3a, more participants coded as intense exposure (white dots) increased or maintained their speech rate compared to those coded as limited exposure (black dots). For silent pauses, results indicated a model approaching statistical significance, with an R^2 value of .13, and again language exposure ($p = .021$) was the only variable which contributed significantly to prediction of change. As shown in Figure 3b, more of the participants coded as limited exposure increased their use of silent pauses, which is a sign of disfluency. For neither speech rate nor silent pauses was study abroad end proficiency a significant predictor of change, $p = .619$ and $p = .650$, respectively.

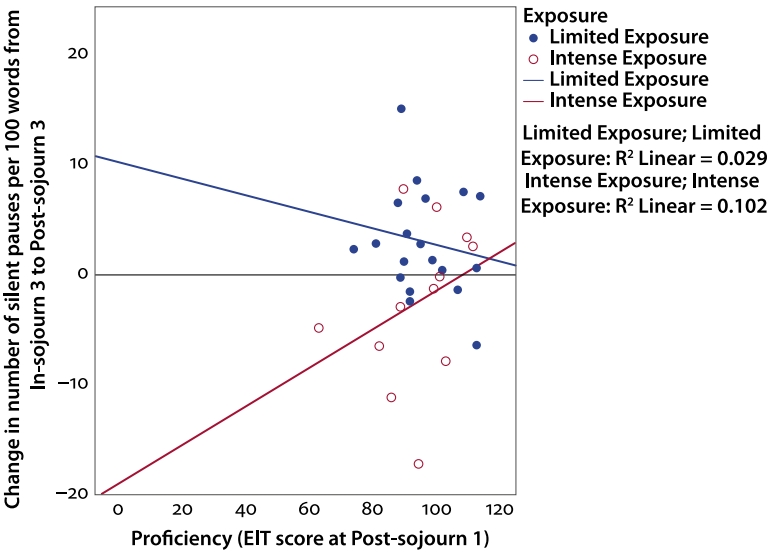
For corrections, results indicated a statistically significant model, with an R^2 value of .21, but in this case SA end proficiency ($p = .004$) was the only variable that contributed significantly to prediction of change. Those who scored higher on the EIT at the end of their stay abroad tended to use fewer corrections four years later regardless of exposure. For oral proficiency, results indicated a statistically significant model, with an R^2 value of .32. Both SA end proficiency and exposure contributed significantly to the model, but SA end proficiency ($p = .004$) contributed more to explaining the variance in proficiency change than exposure ($p = .032$). As shown in Figure 3d, the higher the SA end proficiency (further to the right on the x-axis), the less one's proficiency changed (dots closer to the 0 line) regardless of exposure post-graduation (both white and black dots). This result suggests that when higher proficiency levels are attained, they are more likely to stabilize and are more resistant to change whether or not language contact is intense or limited. At lower proficiencies, however, language contact is more likely to contribute to maintenance of proficiency. As shown in Figure 3d, those participants with a lower SA end proficiency, but who had intense exposure, were able to retain or improve their proficiency gains four years later. Taken together, these results suggest that both SA end proficiency and language exposure are important predictor variables in maintenance / attrition of oral fluency and proficiency over time; however, the relative importance of each variable appears to depend on what linguistic aspect is being measured.

Table 8. Regression results

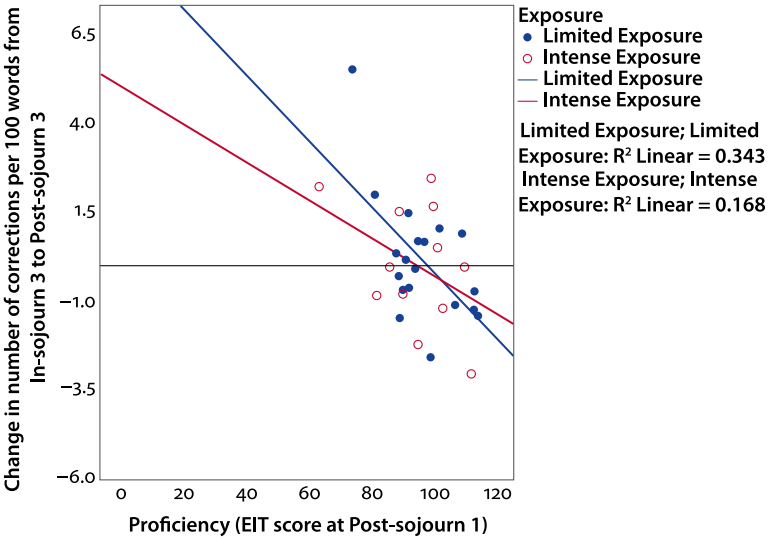
		B	SE_b	β	*F*	*p*	R^2
Speech rate change	Intercept	.127	.507		$F(2, 28) = 4.30$	.024	.18
	Exposure	.342	.123	.459*		.010	
	Proficiency	−.004	.005	−.120		.475	
Silent pauses change	Intercept	−1.240	9.157		$F(2, 28) = 3.25$	.054	.13
	Exposure	−5.478	2.229	−.420*		.020	
	Proficiency	.043	.093	.078		.650	
Corrections change	Intercept	7.29	2.32		$F(2, 28) = 4.91$	.015	.21
	Exposure	−.284	.564	−.082		.618	
	Proficiency	−.074	.024	−.510*		.004	
Proficiency change	Intercept	36.68	12.19		$F(2, 28) = 8.10$	.002	.32
	Exposure	6.89	2.97	.340*		.032	
	Proficiency	−.387	.12	−.472*		.004	


Change in speech rate from In-sojourn 3 to Post-sojourn 3
1.0
.50
0
−.50
−1.0
0
20
40
60
80
100
120
Proficiency (EIT score at Post-sojourn 1)
Exposure
Limited Exposure
Intense Exposure
Limited Exposure
Intense Exposure
Limited Exposure; Limited Exposure: R² Linear = 0.007
Intense Exposure; Intense Exposure: R² Linear = 0.035

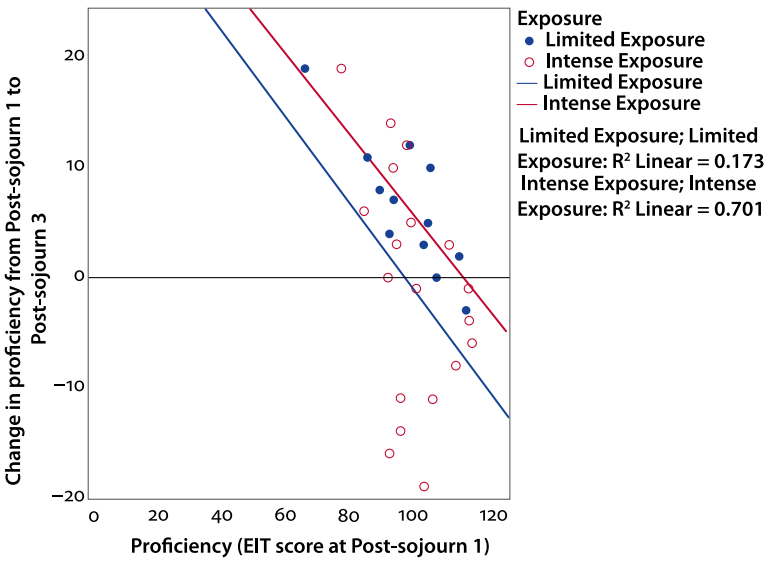
a.


Change in number of silent pauses per 100 words from In-sojourn 3 to Post-sojourn 3
20
10
0
−10
−20
0
20
40
60
80
100
120
Proficiency (EIT score at Post-sojourn 1)
Exposure
Limited Exposure
Intense Exposure
Limited Exposure
Intense Exposure
Limited Exposure; Limited Exposure: R² Linear = 0.029
Intense Exposure; Intense Exposure: R² Linear = 0.102

b.



c.



d.

Figure 3. Scatterplots of change variables and proficiency

4. Discussion

The purpose of this study was to begin to investigate the long-term evolution of oral fluency and proficiency in a group of L2 French and L2 Spanish participants who, as part of their undergraduate degree requirements as languages majors, spent an academic year abroad. After returning home, they spent another year taking courses before graduating with their degree. Data for the current study were collected approximately four years after they returned from their year abroad and three years after they were no longer receiving formal instruction. Oral language data were collected via a picture-based narrative and a proficiency test based on elicited imitation. These data were used to investigate whether changes in fluency and proficiency after study abroad could be predicted by the amount of L2 contact since graduation and proficiency attained at the end of their year abroad.

Research question 1 investigated to what extent the fluency and oral proficiency of L2 French and Spanish learners changed four years after study abroad. To examine this question, we compared results of five measures of utterance fluency (speech rate, number of silent pauses per 100 words, number of filled pauses per 100 words, number of repetitions per 100 words, and number of corrections per 100 words) at three different data collection waves: pre-sojourn (May 2011), the end of their sojourn (in-sojourn 3, May 2012), and most recently (post-sojourn 3, summer 2016). Results of a proficiency test administered pre-sojourn, post-sojourn 1, and post-sojourn 3 were also compared. Overall, results of group averages demonstrated that in general when the two language groups showed improvement on a measure from pre-sojourn to in-sojourn 3, that gain was maintained at post-sojourn 3. In particular, the average increase in speech rate and decrease in silent pause frequency evidenced by the learners during their time abroad were maintained four years after study abroad. Previous research that has included delayed post-tests (e.g., Howard, 2009; Llanes, 2011; Regan, 2005) has demonstrated that gains made during SA are maintained in the short term (12–15 months post-study abroad) for learners continuing to receive formal instruction. In our previous work we also found that the original group of LANGSNAP participants maintained some of their gains the year immediately following study abroad, while they continued to receive instruction. When the follow-up 2016 data were collected, the participants were no longer receiving formal language instruction, yet the group means were not significantly different from the end of study abroad (four years earlier) on the measures of speech rate and silent pause frequency. Thus, it does not appear to be the case that formal instruction contributes to the findings of overall maintenance for the groups on these two measures. Based on the predictions of the 'Critical Threshold Hypothesis' (Neisser, 1984) and

the 'Activation Threshold Hypothesis' (Paradis, 2004), other variables such as proficiency level attained and continued language use (in more informal contexts) after graduation could predict retention over time. As a result, these two variables were the focus of research question 2.

The findings related to filled pauses demonstrated a significant decrease in the frequency of filled pauses for the Spanish group between the pre-test and the end of study abroad that was retained four years later. In contrast, the French group did not demonstrate the same decrease as a result of study abroad. One possible explanation of this finding could relate to the levels at pre-sojourn; the Spanish group on average produced 18 filled pauses per 100 words ($SD = 10.42$) at the pre-test whereas the French group produced 13 per 100 words ($SD = 7.76$). By in-sojourn 3, the Spanish group decreased to five filled pauses per 100 words ($SD = 5.25$) whereas the French group only decreased to nine per 100 words ($SD = 6.23$). Cross-linguistic differences could also help explain this finding. For example, Huensch and Tracy-Ventura (2017a) compared native speakers of English, French, and Spanish completing the same task and results indicated that Spanish native speakers produced fewer filled pauses per second compared to French native speakers ($d = 1.57$). The number of participants in each group was relatively small; thus, future research is needed to corroborate this finding. The longitudinal data from Mehotcheva (2010) demonstrated an increased use of filled pauses and repetitions one year after study abroad, which differs from the current study. However, her longitudinal data were only based on five participants.

No change over time was found for both groups on the measures of number of repetitions and corrections. Therefore, in general, the frequency of repetitions and corrections did not change (see also Huensch & Tracy-Ventura, 2017a), but it is possible that the distribution of what was repeated or corrected (e.g., grammar, lexis, phonology, message) may have changed. For example, Kormos (2000) demonstrated that native speakers and advanced learners more frequently produced repairs which modified linguistic appropriateness / content whereas pre-intermediate learners more frequently produced repairs which targeted linguistic errors. Similarly, using stimulated recall, Kahng (2014) demonstrated that when asked to comment on instances of pausing and hesitation in their speech, advanced learners of English more often commented on the content of their message whereas lower proficiency learners more often commented on issues pertaining to grammar and lexis. Thus, while the quantity of repairs did not change over time for participants in the current study, it is possible that there were qualitative differences in what was repeated or corrected. Additional analyses that take into consideration where within the utterance the repetition or correction occurs could also be enlightening in this regard (see Schmid & Fägersten, 2010).

Regarding oral proficiency, results demonstrated that the Spanish group maintained the gains they made during study abroad four years later. Perhaps surprisingly, the French group improved significantly from post-sojourn 1 to post-sojourn 3, although the effect size was small and the confidence intervals passed through zero suggesting that this finding is not statistically meaningful. Nevertheless, we inspected individual change scores to interpret this finding. Participants coded as intense exposure showed improvements at post-sojourn 3, and on average they improved 10.3 points. In comparison, there was a lot of individual variation in the group coded as limited exposure with some whose scores decreased by more than ten points, some whose scores did not change at all, and some whose scores improved by ten or more points. Their average change score was 4.2. Individual change scores for the Spanish group appeared to follow a different pattern: the change scores of participants coded as intense exposure ranged from −3 to 10 with an average improvement of 6.2. Each participant in the limited exposure group showed a decrease in proficiency ranging from −1 to −19 with an average decrease of −8.3.

Research question 3 investigated to what extent the variables of language exposure post-instruction and SA end proficiency could predict changes in oral fluency and proficiency four years later. Results of the current study suggest that both are important variables in predicting maintenance / attrition but that the relative importance of each one depends on the linguistic variable that is measured. For example, the maintenance of variables related to speed and breakdown fluency (e.g., speech rate and silent pauses) appears to be influenced by language exposure only. In other words, the more language exposure, the lower the chance of attrition in these areas, and in some cases the higher the chance of development. This result provides some support for the 'Activation Threshold Hypothesis' (Paradis, 2007) which claims that frequency and recency of access of individual items contributes to lowering of the activation threshold. The practical effect is that those items become easily accessible and that attrition "is a result of lack of long term stimulation" (Paradis, 2007, p. 125). Paradis also allows that motivation as well as an advanced level of proficiency may have a protective effect. However, SA end proficiency was not a significant predictor of change in speech rate and silent pauses, though it was for number of corrections, a measure that has been characterized as reflecting "the extent to which the learner is oriented towards accuracy" (Ellis & Barkhuizen, 2005, pp. 149–150). Thus, it may be the case that the 'Critical Threshold Hypothesis' (Neisser, 1984) is more applicable to maintenance of accuracy, as well as to lexical access (see Mehotcheva, 2010), compared to the maintenance of fluency which is likely related to continued use. Our results demonstrating that exposure is an important variable in the retention of certain aspects of oral fluency differ from previous research on foreign language attrition.

For example, Mehotcheva (2010) found that SA end proficiency was a significant predictor of language maintenance post-study abroad but that language exposure was not. However, she only measured disfluency variables (e.g., filled pauses, corrections, repetitions) and not other fluency variables such as speech rate and frequency of silent pauses.

Interestingly, the change in EIT scores was the only measure where both variables emerged as significant predictors, but SA end proficiency contributed more to explaining the variance in scores. This finding might suggest that for the maintenance of general proficiency, it is most important how proficient one was at the end of study abroad, but that sustained language exposure has the potential to contribute to continued development over time in those cases where higher proficiency levels were not reached. Participants in the current study were of more advanced proficiency compared to previous research. Thus, whether both variables would appear important in language retention of participants who only reached intermediate proficiency, for example, is a question for future research.

5. Conclusion

The current study was a preliminary investigation into the long-term effects of study abroad on language retention four years later. Some limitations of the study should be acknowledged. First, the current study included oral data from a picture-based narrative. It is also important to look at less controlled oral tasks such as semi-structured interviews. Another potential limitation is the operationalization of language exposure as a binary variable. Despite it being a useful solution for conducting the statistical tests, we acknowledge that it potentially oversimplifies a complex reality. Regarding fluency variables, future investigations should include information about pause duration / location, as well as additional linguistic variables such as accuracy, complexity, and lexis. Finally, given the relatively small number of participants, it was not possible to include more potentially explanatory variables in the regression analysis. For example, it is possible that individual differences in identity, motivation, or personality might play a role as well. Qualitative research would complement our findings by providing more details about the individual participants and their experiences and motivations for maintaining their second languages in the long term.

These limitations notwithstanding, results of the current study demonstrate that, for these particular participants, fluency and oral proficiency were maintained four years after study abroad and three years after no longer receiving formal instruction. The amount of contact and use with the target language played a role in maintenance of aspects of fluency such as speech rate and frequency of

silent pauses, whereas proficiency attained at the end of study abroad played a role in the use of corrections. Both language contact / use and proficiency attained were important variables in the maintenance of overall proficiency. More research is needed to corroborate these findings so that students of foreign languages and those who go abroad can be provided with empirically-based information about the likely evolution of their language skills post-instruction and those variables that are important for long-term retention.

Acknowledgements

LANGSNAP was funded by the ESRC (award number RES-062-23-2996). LANGSNAP 3.0 was funded by a *Language Learning* Small Research Grant, USF World Faculty Mobility Grant, and a USF New Researcher Grant. We are grateful to our LANGSNAP colleagues, the participants, and various research assistants for their contribution to this study.

References

Bahrick, B. P. (1984). Fifty years of second language attrition: Implications for programmatic research. *The Modern Language Journal*, 68, 105–118. https://doi.org/10.1111/j.1540-4781.1984.tb01551.x

Bardovi-Harlig, K., & Stringer, D. (2010). Variables in second language attrition: Advancing the state of the art. *Studies in Second Language Acquisition*, 32, 1–45. https://doi.org/10.1017/S0272263109990246

Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer (Version 5.4.18) [Computer program]. Retrieved from <http://www.praat.org/>

Bowden, H. W. (2016). Assessing second language oral proficiency for research. *Studies in Second Language Acquisition*, 38(4), 647–675. https://doi.org/10.1017/S0272263115000443

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

Di Silvio, F., Diao, A., & Donovan, W. (2016). The development of L2 fluency during study abroad: A cross-language study. *Modern Language Journal*, 100, 610–624. https://doi.org/10.1111/modl.12343

Dominguez, L., Tracy-Ventura, N., Arche, M., Mitchell, R., & Myles, F. (2013). The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition*, 16, 558–577. https://doi.org/10.1017/S1366728912000363

Du, H. (2013). The development of Chinese fluency during study abroad in China. *The Modern Language Journal*, 97, 131–143. https://doi.org/10.1111/j.1540-4781.2013.01434.x

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

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. https://doi.org/10.1177/0261927X8700600102

Guntermann, G. (1995). The Peace Corps experience: Language learning in training and in the field. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 149–170). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.9.10gun

Hernández, T. A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *The Modern Language Journal*, 94, 600–617. https://doi.org/10.1111/j.1540-4781.2010.01053.x

Howard, M. (2009). Short- versus long-term effects of naturalistic exposure on the advanced learner's L2 development: A case-study. In E. Labeau & F. Myles (Eds.), *The advanced learner variety: The case of French* (pp. 93–123). Oxford: Peter Lang.

Huensch, A., & Tracy-Ventura, N. (2017a). L2 utterance fluency development before, during, and after residence abroad: A multidimensional investigation. *The Modern Language Journal*, 101, 275–293. https://doi.org/10.1111/modl.12395

Huensch, A., & Tracy-Ventura, N. (2017b). Understanding L2 fluency behavior: The effects of individual differences in L1 fluency, cross-linguistic differences, and proficiency over time. *Applied Psycholinguistics*, 38, 755–785. https://doi.org/10.1017/S0142716416000424

Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, 64, 809–854. https://doi.org/10.1111/lang.12084

Kim, J., Dewey, D. P., Baker-Smemoe, W., Ring, S., Westover, A., & Eggett, D. L. (2015). L2 development during study abroad in China. *System*, 55, 123–133. https://doi.org/10.1016/j.system.2015.10.005

Kinginger, C. (2008). Language learning in study abroad: Case studies of Americans in France. *The Modern Language Journal*, 92(S1), 1–124. https://doi.org/10.1111/j.1540-4781.2008.00821.x

Kinginger, C. (2009). *Language learning and study abroad: A critical reading of research*. Houndmills: Palgrave Macmillan. https://doi.org/10.1057/9780230240766

Kinginger, C. (2011). Enhancing language learning in study abroad. *Annual Review of Applied Linguistics*, 31, 58–73. https://doi.org/10.1017/S0267190511000031

Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, 50, 343–384. https://doi.org/10.1111/0023-8333.00120

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R*. New York, NY: Routledge.

Llanes, À. (2012). The short- and long-term effects of a short study abroad experience: The case of children. *System*, 40, 179–190. https://doi.org/10.1016/j.system.2012.05.003

Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37, 353–365. https://doi.org/10.1016/j.system.2009.03.001

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Magnan, S. S., & Back, M. (2007). Social interaction and linguistic gain during study abroad. *Foreign Language Annals*, 40, 43–61. https://doi.org/10.1111/j.1944-9720.2007.tb02853.x

Magnan, S. S., & Lafford, B. A. (2012). Learning through immersion during study abroad. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 525–540). Abingdon, NJ: Routledge.

McManus, K., Mitchell, R., & Tracy-Ventura, N. (2014). Understanding insertion and integration in a study abroad context: The case of English-speaking sojourners in France. *Revue Française de Linguistique Appliquée*, 14, 97–116.

Mehotcheva, T. H. (2010). After the fiesta is over: Foreign language attrition of Spanish in Dutch and German Erasmus students. Unpublished doctoral dissertation, University of Groningen, The Netherlands.

Mitchell, R., McManus, K., & Tracy-Ventura, N. (2015). Placement type and language learning during residence abroad. In R. Mitchell, N. Tracy-Ventura, & K. McManus (Eds.), *Social interaction, identity and language learning during residence abroad* (pp. 115–137). EUROSLA Monographs series.

Mitchell, R., Tracy-Ventura, N., & McManus, K. (2017). *Anglophone students abroad: Identity, social relationships and language learning*. New York, NY: Routledge. https://doi.org/10.4324/9781315194851

Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46, 610–641.

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, The Netherlands.

Neisser, U. (1984). Interpreting Harry Bahrick's discovery: What confers immunity against forgetting? *Journal of Experimental Psychology: General*, 113, 32–35. https://doi.org/10.1037/0096-3445.113.1.32

Ortega, L. (2000). Understanding syntactic complexity: The measurement of change in the syntax of instructed L2 Spanish learners. Unpublished doctoral dissertation, University of Hawai'i, Honolulu, HI.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.18

Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 121–133). Amsterdam: John Benjamins. https://doi.org/10.1075/sibil.33.09par

Pérez-Vidal, C., & Juan-Garau, M. (2009). The effect of Study Abroad (SA) on written performance. *Eurosla Yearbook*, 9, 269–295. https://doi.org/10.1075/eurosla.9.13per

Pizziconi, B. (2017). Japanese vocabulary development in and beyond study abroad: The timing of the year abroad in a language degree curriculum. *The Language Learning Journal*, 45, 133–152. https://doi.org/10.1080/09571736.2013.786119

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912. https://doi.org/10.1111/lang.12079

Regan, V. (2005). From community back to classroom: What variation analysis can tell us about context of acquisition. In J.-M. Dewaele, (Ed.), *Focus on French as a foreign language: Multidisciplinary approaches* (pp. 191–209). Clevedon: Multilingual Matters.

Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45, 81–105. https://doi.org/10.5054/tq.2011.240861

Schmid, M. S., & Fägersten, K. B. (2010). Disfluency markers in L1 attrition. *Language Learning*, 60, 753–791. https://doi.org/10.1111/j.1467-9922.2010.00575.x

Schmid, M., & Mehotcheva, T. (2012). Foreign language attrition. *Dutch Journal of Applied Linguistics*, 1, 102–124. https://doi.org/10.1075/dujal.1.1.08sch

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173–199. https://doi.org/10.1017/S0272263104262027

Serrano, R., Llanes, À., & Tragant, E. (2011). Analyzing the effect of context of second language learning: domestic intensive and semi-intensive courses vs. study abroad in Europe. *System*, 39, 133–143. https://doi.org/10.1016/j.system.2011.05.002

Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36, 1–14. https://doi.org/10.1017/S026144480200188X

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–276). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.15tav

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84–119. https://doi.org/10.1093/applin/17.1.84

Tracy-Ventura, N., McManus, K., Norris, J. M., & Ortega, L. (2014). Repeat as much as you can: Elicited imitation as a measure of oral proficiency in L2 French. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 143–166). Bristol: Multilingual Matters. https://doi.org/10.21832/9781783092291-011

Weltens, B. (1989). *The attrition of French as a foreign language*. Dordrecht: Foris. https://doi.org/10.1515/9783111395937

Yang, J.-S. (2016). The effectiveness of study abroad on second language learning: A metaanalysis. *Canadian Modern Language Review*, 72, 66–94. https://doi.org/10.3138/cmlr.2344

Authors' addresses

Amanda Huensch
University of South Florida
4202 E. Fowler Ave., CPR 107
Tampa, FL 33620
United States of America
huensch@usf.edu

Judith Bridges
University of South Florida
4202 E. Fowler Ave., CPR 107
Tampa, FL 33620
United States of America
jcbridges@mail.usf.edu

Nicole Tracy-Ventura
University of South Florida
4202 E. Fowler Ave., CPR 107
Tampa, FL 33620
United States of America
nkt@usf.edu

Jhon A. Cuesta Medina
University of South Florida
4202 E. Fowler Ave., CPR 107
Tampa, FL 33620
United States of America
jhon@mail.usf.edu