

Research Article

LEARNING ENGLISH IN TODAY'S GLOBAL WORLD

A COMPARATIVE STUDY OF AT HOME, ANGLOPHONE, AND LINGUA

FRANCA STUDY ABROAD

Zeynep Köylü *

Department of English, University of Basel, Basel, Switzerland

Nicole Tracy-Ventura 

Department of World Languages, Literatures, and Linguistics, West Virginia University, WV, USA

Abstract

In comparative studies focusing on context of learning, the main contexts under investigation have been study abroad (SA), at-home formal instruction (AH), and domestic immersion (IM). With the global status of English and its burgeoning popularity as a medium of instruction in countries where English only holds the status of a lingua franca, a new SA context has emerged. This study compares the L2 learning of English in this new English as a lingua franca study abroad (ELFSA) context to Anglophone SA and AH in terms of oral and written complexity, accuracy, and fluency gains. Participants' perceptions of contextual differences concerning the amount of language contact, use, development, and their views toward English are also explored qualitatively. Apart from indicating equal development on most CAF measures after a semester, the qualitative findings highlight ELFSA as providing a low-anxiety atmosphere that helps sojourners gain ownership of English. Thus, ELFSA emerges as an appealing study abroad context for future sojourners.

INTRODUCTION

Due to high amounts of exposure to the target language (TL) and increased interaction opportunities, study abroad (SA) has received much attention in second language acquisition (SLA) research for its potential to provide an optimum learning environment

We are grateful to the participants for their collaboration and candor; to Amanda Huensch, Rosamond Mitchell, and Amy Thompson for their invaluable feedback and support throughout the original study; and to the anonymous reviewers for their feedback on previous versions of this article. All remaining errors are our own.

* Corresponding author: E-mail: zeynep.koylu@unibas.ch

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

for second/foreign language (L2) development (Borràs & Llanes, 2019). In the literature focusing on learners of English, the context primarily researched is SA in an Anglophone country (e.g., Australia, England, the United States) where the official or de facto language is English. Yet, with the status of English as a lingua franca (ELF) and the increasing availability of exchange programs particularly in Europe (e.g., the European Region Action Scheme for the Mobility of University Students, or ERASMUS)¹ another sojourn context is becoming more accessible to learners of English. In this context, host universities in non-Anglophone countries provide English medium instruction (EMI) for international sojourners. Several countries in Europe are offering such EMI programs, including Austria, Denmark, Holland, Greece, Italy, Spain, among others. Therefore, when sojourners study in these destinations, they will be in a multilingual environment where they will most likely interact in their L1 (with L1 speaking peers and family, in person or virtually), English (in their EMI courses and with other temporary sojourners, L2 and L1 speakers of English), and the language of the host country (in service encounters, with peers, etc). In this article, we follow Köylü (2016) and refer to this context as English as a lingua franca study abroad (ELFSA). Despite the recent increase in ELFSA programs, SA research has just begun to investigate their language learning potential (Pérez-Vidal & Llanes, 2021). Thus far, the limited research has found ELFSA beneficial for general proficiency, written (Llanes et al., 2016) and oral development (Llanes, 2019; Martín-Rubió & Cots, 2018), and affecting positive changes in learner beliefs toward learning English and studying abroad in the European ELF context (Kaypak & Ortaçtepe, 2014). However, to date, there have been no comparative studies focusing on differences between ELFSA, traditional Anglophone-SA, and at-home (AH) foreign language instruction. It is important to investigate similarities and differences between these contexts to discern potential intercultural, linguistic, and personal gains. Empirical evidence is required when making suggestions to students about the strengths of different learning contexts.

Motivated by this gap in the literature, this study investigates, for Turkish L1 learners of English, the contextual influence of traditional Anglophone-SA (England, in this study), ELFSA (non-Anglophone European countries offering EMI),² and intensive-AH instruction (a university in Turkey also providing EMI) on oral and written complexity, accuracy, and fluency (CAF) development in L2 English. This study also compares sojourners' experiences in the three contexts through questionnaire and interview data. A major contribution of this study is the exploration of ELFSA for learning English abroad as little is currently known about learners' experiences in this context in terms of the type and amount of input and interaction. The fact that ELFSA is often more affordable and easily accessible to many learners also makes this context worthy of further exploration (González et al., 2011).

LITERATURE REVIEW

COMPARING L2 LEARNING ACROSS CONTEXTS

In SLA study abroad research, the traditional classification of learning environment has included three main contexts: (1) SA, a temporary stay in a TL-speaking environment varying in length and whether formal instruction is provided; (2) domestic immersion

(IM), which typically involves intensive content and language integrated learning at home, with additional extracurricular activities in the TL; and (3) AH, traditional foreign language instruction (Segalowitz & Freed, 2004).

For L2 learners, SA is considered the best route to linguistic, academic, cultural, and personal development (Pérez-Vidal, 2017). The burgeoning availability of international mobility programs has increased scholarly attention to the contribution of SA to L2 development. SLA research has reported gains in receptive and productive skills, some aspects of oral (McManus et al., 2021) and written production (Llanes & Muñoz, 2013), and pragmatic and intercultural competence (Taguchi, 2015) (see Sanz & Morales-Front, 2018 for recent overviews). However, findings related to linguistic development are disappointingly mixed, and one way that scholars have attempted to understand the effectiveness of SA is to compare it with different learning contexts, either using a within-subjects or between-subjects design.

A few longitudinal studies have compared the same learners' linguistic development first AH, then during SA, and again AH such as the Study Abroad and Language Acquisition project (SALA: Pérez-Vidal, 2014) and the Languages and Social Networks Abroad Project (LANGSNAP: Mitchell et al., 2017). Findings from these projects highlight the benefits of SA for oral and written fluency, and some aspects of accuracy. Improvement in written complexity was found in the SALA project only (Pérez-Vidal & Barquin, 2014). These studies also demonstrate that gains made during SA are generally retained afterward with AH instruction.

Studies adopting a between-subjects design have compared learners in different groups, primarily SA versus AH. Results of these studies also highlight the advantages of SA for oral fluency gains (Llanes & Muñoz, 2013; Llanes & Serrano, 2017; Mora & Valls-Ferrer, 2012; Segalowitz & Freed, 2004; Serrano et al., 2016). For example, Serrano et al. (2011) investigated differences in CAF development between (1) SA in Britain and intensive AH instruction (approximately 25 hours/week) and (2) SA and AH semi-intensive instruction (approximately 10 hours/week). They found that overall gains were similar between the SA and intensive AH groups, yet the SA group significantly improved more on oral fluency and lexical complexity compared to the semi-intensive AH group. Serrano et al. (2016) investigated the case of teenagers (M age = 14.37) in AH ($n = 58$) and 3-week SA ($n = 54$) in terms of grammar, formulaic sequences, and written and oral production. All participants had significant gains on most measures regardless of group membership, while the SA participants significantly outperformed the AH on oral lexical richness. Fewer studies have compared three learning contexts (IM to SA and AH instruction), also yielding important findings (Freed et al., 2004; Segalowitz & Freed, 2004). Freed et al. (2004) found that learners in the French IM group developed more in oral fluency and interacted more in French than the SA group, suggesting that intensive immersion in a CLIL setting would help develop the L2 without sojourning.

Contextual affordances in SA and AH, such as the amount of input, interaction, and opportunities for intercultural development, have also been scrutinized through quantitative and qualitative measures to justify the presence or lack of oral gains. Questionnaires documenting the type and amount of L2 contact as well as the nature of interaction and interlocutors in SA (e.g., Freed et al., 2004; Mitchell et al., 2017) have been used, plus interviews (McGregor, 2021; Tanaka, 2007), to explore differences between contexts.

For example, Freed et al. (2004) found that the IM students reported more out-of-class engagement in L2 French than the SA and AH groups, as a result of extracurricular activities planned for the IM group. The extra L2 use likely contributed to their greater improvement in oral fluency. This finding highlights an important assumption in study abroad, that of the abundant availability of NS interaction necessary for L2 development (McGregor, 2021). Research has shown that high amounts of input and interaction with NSs during SA is not guaranteed (Tanaka, 2007).

As revealed through interview data, Tanaka (2007) demonstrated how sojourners struggle to access consistent NS interaction during SA, even with homestays, either because the Japanese-L1 participants in this study had lower proficiency in English or the host family members were too busy with their lives and did not engage in conversations with them: “My host family watched television after dinner and I didn’t want to disturb them. I found it was more difficult to make an environment where I could talk to native speakers than I thought.” [NZ17] (Tanaka, 2007, pp. 45–46). Tanaka’s participants, however, did reveal that NNS interaction was a significant source of English practice as they could communicate with Chinese friends quite well with their “shaky English” (p. 46). McGregor (2021) argues that the benefits of L2 peer interaction have been overlooked and undervalued within the SA paradigm. Focusing on pragmatic competence especially in face-threatening situations and using conversation analysis methods, she demonstrates how L2 peer interaction during a short-term SA is a valuable resource for L2 development.

Results suggesting that other intensive learning contexts can promote learning, for example where NNS-NNS interaction is a significant source of TL use, encourage exploration of new contexts such as ELFSA. Because of the growing popularity of EMI programs in non-Anglophone countries, it is now possible to examine the ELFSA context comparative to others. In response to Mitchell’s (2021) and Pérez-Vidal and Llanes’s (2021) call for furthering SA research in the European perspective, the current study is an attempt to understand how sojourners engage with EMI and NNS interaction in a way to conceptualize the nature of learning in an EMI/ELF multilingual setting compared to traditional Anglophone-SA.

ELF DURING STUDY ABROAD

ELF can be defined as “the contact language between people who share neither a common native tongue nor a common culture, and for whom English is the chosen foreign language of communication” (Firth, 1996, p. 240). With the spread of ELF and the “Englishisation of higher education” (Martin-Rubió & Cots, 2018, p. 97), its use and effects upon L2 development have begun to attract attention among SLA scholars (e.g., Llanes, 2019). This new context is also of interest as it provides a multilingual learning environment to sojourners through mobility schemes like ERASMUS. For example, an ERASMUS exchange student in Belgium will likely have contact with Flemish and French, along with EMI at the host university, and ELF with other international students. An ERASMUS exchange is considered indispensable with its “international conviviality” and opportunities to develop intercultural skills and employment capacities afterward (Cairns, 2017, p. 729). Although originally designed to promote European identity through the acquisition of European languages (ERASMUS+, 2021), a major ERASMUS objective is to help promote global citizenship among youth across Europe through

the promotion of “a less nationally oriented uber-European generation” (Cairns, 2017, p. 728).

Thus far, ELFSA has mostly received attention for the study of qualitative variables such as language learner identity construction (Kalocsai, 2014), perceptions and beliefs about language learning in an ELF context (Kaypak & Ortaçtepe, 2014), and awareness of communicative skills and how confidently they are used (Martin-Rubió & Cots, 2018), with research yielding positive results for ERASMUS students. ELFSA was found to contribute to sojourners developing new repertoires of shared ways of speaking and identity affiliation in an ELF-resourced community (Kalocsai, 2014, p. 5) as they negotiated and mediated their use of English with other fellow ELF speakers (Borghetti & Beaven, 2017). Kaypak and Ortaçtepe (2014) similarly demonstrated that ELFSA sojourners changed their focus from accuracy to intelligibility, resulting in more frequent interaction. Borghetti and Beaven (2017) found ELFSA to be a less anxiety-bearing context as sojourners reported feeling less embarrassed, with fewer concerns about their language skills being judged by interlocutors. In return, they reported gaining crucial accommodation, negotiation, and cooperation skills as ELF users as they “shaped ‘ELF’ competence to their own needs” (Kalocsai, 2014, p. 203). Such findings highlight several positive effects of ELFSA, yet little is known about its characteristics in terms of the type and amount of input and interaction, and how those influence L2 gains.

To date, few studies have investigated L2 users’ linguistic development after ELFSA. Llanes et al. (2016), Llanes (2019), and Martin-Rubió and Cots (2018) investigated the case of university-level Spanish/Catalan bilinguals learning English in ELFSA (e.g., Belgium, Denmark, Germany, Holland, and Italy) over a semester. Llanes et al. (2016) indicated significant gains in general proficiency, measured by the Oxford Placement Test (OPT), and written lexical complexity. Considering oral skills, Llanes (2019) reported significant development in proficiency (OPT), speech rate, and lexical complexity after sojourning in a Nordic or Mediterranean country. No significant changes in accuracy were detected, suggesting that sojourners might have prioritized fluency at the expense of accuracy. Martin-Rubió and Cots (2018) reported gains in oral fluency and accuracy through descriptive statistics (no statistical tests were computed). They also reported a positive relationship between sojourners’ increasing self-confidence and oral proficiency on account that they were in a low-anxiety ELF context away from “a native-speakerist discourse based on a deficit-model of the foreign language learner” (p. 110). Hence, the type of interlocutors in sojourn contexts could be a variable influencing linguistic development while abroad. Thus far, the limited research on language learning during ELFSA shows promising results in both oral and written skills as measured by CAF. Whether it is as beneficial as traditional SA for the development of these skills, however, is in need of empirical research.

RESEARCH QUESTIONS

Based on the review of the literature presented, this study investigates the following research questions:

1. To what extent does the type of context (SA, ELFSA, intensive AH) have an effect on the oral and written performance of English as measured by fluency, accuracy, and syntactic and lexical complexity development over time?

2. To what extent do the amount and types of L2 contact differ across contexts?
3. How do learners' lived experiences in these different learning contexts compare in terms of language contact, L2 development, and their views toward English?

METHODOLOGY

CONTEXTS

The two sojourn contexts compared in this study were provided by the ERASMUS mobility scheme for Turkish students studying at several public or private universities: (1) SA, in this case England, and (2) ELFSA, which in the current study included 10 different EU countries (described in the next section). In both sojourn contexts, participants took 9–12 hours of EMI classes at their host institution, as per the exchange program requirements. However, the amount of coursework varied among participants. Some reported frequent assignments and others only minimal coursework. Attendance was required for all sojourners.

The AH context investigated was an EMI program at a public university in Turkey, offering 18 hours of content classes (the entire curriculum) per week. Participants were third-year degree-program (BA) students majoring in American culture and literature. Typical activities included listening to lectures, giving presentations, reading, and writing essays. All exams were also administered in English. Due to the high amount of contact hours, we operationalize this context as intensive AH instruction, following Serrano et al. (2011).

PARTICIPANTS

Participants included 47 Turkish L1 learners of English from various public and private universities in Turkey (36 females, 11 males, *M* age = 22). Only data gathered from those who completed all the instruments were analyzed.

Sojourner participants included those in the SA (*n* = 8) and ELFSA (*n* = 24) groups who spent their 16-week semester abroad as ERASMUS students (29 undergraduates, 3 graduates), were majoring in a variety of EMI programs, and had no previous sojourn experience (see Table 1). They participated in ERASMUS to continue their studies at a European institution as exchange students. Their major motivation was to experience living and studying in Europe. All courses abroad were content courses; they did not take specific language courses. Per ERASMUS requirements, they documented intermediate proficiency in English through institutional tests administered at their home universities.

SA participants were studying in different universities located in the South, Midlands, or West of England, whereas the ELFSA participants were enrolled in different universities across Europe (see Table 1). The lower number of participants in the SA group is due to fewer universities having mutual learning agreements³ with British institutions. Also, many Turkish sojourners preferred to undertake their sojourn semester in countries with lower cost of living rates (Turkish National Agency, 2015).

Sojourners were not allowed to work in the host country per ERASMUS requirements. Considering living arrangements, family-stay was the least preferred (*n* = 1 in the SA group), with the most popular type being single or shared dorm-stays/student housing

TABLE 1. Participants' demographics

| Group | Destination | Majors | Predeparture/ Initial Proficiency |
|----------------|---|--|---|
| SA (N = 8) | England | Business Administration, Engineering, Graphic Design, Language Teaching | B2 |
| ELFSA (N = 24) | Austria (2), Czechia (1), Denmark (1), Finland (2), Germany (4), Greece (1), Holland (5), Italy (5), Poland (2), Portugal (1) | Archeology, Business Administration, Engineering, Law, Psychology | B1, B2 |
| AH (N = 15) | Turkey | American Culture and Literature | B1, B2 |

($n = 6$ single dorm room, $n = 25$ shared dorm room with two or more students). None of the participants reported sharing rooms with compatriots or English NSs.

AH participants ($n = 15$) were all third-year language majors, studying American culture and literature (ACL) who attended intensive weekly EMI classes (four core and two elective classes) in American poetry, novel, drama, and geography/cultural landscape with frequent assignments. The rationale behind including this group was to find AH participants with relatively similar amounts of TL exposure and output opportunities in their AH context.

INSTRUMENTS

A series of instruments were administered to assess oral and written development in English. Additionally, a questionnaire on language contact was administered and interviews were conducted with a subset of the participants ($n = 9$). Instruments and measures are explained below.

Oral performance task

Oral production was elicited by a 1-minute TOEFL-style speaking prompt: *What would you like to do in your free time and why?* at pre- and postsojourn. This task was selected as it was familiar to the participants and the prompt requires no technical vocabulary. Prior to this task, participants took part in a short biographical interview also in English as a warm-up activity. After receiving the prompt, 10 seconds preparation time was provided.

Written performance task

A 15-minute computerized composition writing task was administered at pre- and posttest. In response to the prompt, "your past, present, and future expectations," participants produced a paragraph of at least seven lines in a standard word processor. Time-on-task was recorded for each participant as some completed the task before the 15-minute limit. The writing prompt, first used by Llanes and Muñoz (2013), was chosen because it did not require specialized knowledge or domain/topic-specific vocabulary.

Elicited imitation task

A 30-item elicited imitation task (EIT) in English (Ortega et al., 1999) was also administered to assess proficiency. EIT is an oral production test requiring participants to repeat sentence stimuli as accurately as possible, and stimuli range from 7 to 19 syllables in this particular EIT. Each item is scored using a rubric ranging from 0–4, allowing a maximum of 120 points. This instrument is presented only to explain the preliminary steps taken to select the proper statistical tests, as predeparture proficiency was considered a potential covariant. Results regarding pre- and postsojourn proficiency are discussed in Köylü (2021).

LIQ and interviews

To examine the type and frequency of activities conducted in English throughout the semester, a series of online Language Interaction Questionnaires (LIQ: a modified version of the Language Engagement Questionnaire designed by McManus et al., 2014) were completed by all participants once every fourth week over the 16-week semester. LIQ (Cronbach's $\alpha = .90$) data analyzed for this study included 13 6-point Likert scale items (11 items on the type and frequency of activities engaged in English and 2 items on type of interlocutor, NS or NNS—e.g., *How frequently did you interact with a NS throughout the semester?*)⁴ and 1 item measuring the total hours of self-reported L2 contact (Total $n = 14$ items). The 11 Likert-scale items focused on activities in English including listening, reading, writing, speaking, and internet use either for leisure or academic purposes (e.g., *How frequently do you read something in English for academic purposes?*). The full instrument is available on IRIS (<https://www.iris-database.org/iris/app/home/detail?id=york:938936>).

To explore individual participants' experiences and triangulate the quantitative findings, two semistructured interviews were conducted in Turkish with a subsample of the participants, once at the end of the first month and once upon program completion. According to the LIQ results, three participants from each group ($n = 9$) representing either low, medium, or high amounts of contact were purposely selected. Interviewees were asked about amount and type of L2 contact, interlocutors, merits and demerits of their contexts, major difficulties, their social life, and suggestions for future sojourners. Interview data were first translated into English and later back-translated into Turkish to reconcile any differences in meaning.

Procedures

The home universities' international offices for student mobility helped recruit the sojourner participants through a flyer explaining the project. Only volunteers spending the following semester abroad as ERASMUS students (with no previous SA experience) were selected as participants. An intact class of third-year ACL in Turkey were presented the project, and those who volunteered comprised the AH group.

Oral and written tasks were completed twice, once at the onset of the study and again after the participants completed their semester (either at home or abroad), along with the EIT (see Köylü, 2021, for those results).⁵ To minimize task effects, participants randomly

selected the task sequence; no systematic task sequence selection was noticed. A corpus of 97':29" of spoken and 18,201 words of written data was compiled and quantitatively analyzed. Data were first transcribed and coded following CHAT format and then analyzed using CLAN (MacWhinney, 2000) to calculate measures of syntactic and lexical complexity, accuracy, and fluency (CAF).

Measures

Fluency: written fluency was operationalized as the total number of words per total time (W/T). Following Skehan (2009) oral fluency was operationalized through utterance fluency broken down into the subdimensions of (1) speed, (2) breakdown, and (3) repair fluency:

- Speed fluency/speech rate (words/time in seconds): the total number of words excluding words in disfluent production (i.e., pruned speech) divided by total production time in seconds, W/T.
- Breakdown fluency (pauses/time in milliseconds): total duration of silent and filled pauses longer than .250 milliseconds divided by the total time expressed in seconds, P/T.
- Repair fluency (disfluency/time in seconds): the total number of disfluencies determined by the number of repetitions, retraces, and reformulations divided by total time expressed in seconds and multiplied by 60, D/T.

Syntactic complexity: oral data were segmented into analysis of speech unit (ASU) (Foster et al., 2000) and written data segmented into T-units (TU). The total number of finite and nonfinite clauses per ASU and per TU was calculated (CL/ASU for oral; CL/TU for written). For lexical complexity, D values were obtained by the VocD program in CLAN (MacWhinney, 2000) in both modes of production.

Accuracy: both lexical and grammatical errors were coded in the transcript. The total number of errors per ASU and per TU was computed (Err/ASU and ERR/TU). To ensure reliability, another Turkish-L1 researcher fluent in English segmented and coded the data. Disagreements were reviewed and revised after discussing and reaching consensus. Intercoder reliability was measured through the percentage of agreement, reaching 92% as mean reliability percentage for all measures.

Statistical analyses

An *a priori* power analysis was conducted to estimate minimum sample size requirements with the alpha level set to .05 (two-tailed), power level to .80, and effect size to .25 for the three context groups involved using G*Power (Faul et al., 2007). The sample size ($n = 47$) ensured statistical power surpassing the minimum sample size (42) suggested. Given the unequal distribution of participants in each context, we employed the bootstrapping method (1,000 samples) for robust statistics despite the sufficient total sample size. For all tests of mean comparison, we reported bootstrapped 95% confidence intervals (CI) to indicate the levels of measurement precision.

CAF measures were used as dependent variables in the statistical tests to discern inter- and intragroup development over time. Predeparture proficiency scores from the EIT were tested for initial differences among the groups. Using a one-way ANOVA, no differences were found ($F(1,31) = 1.285, p = .266, \text{partial } \eta^2 = .04$), showing no requirements to control for

initial proficiency as a covariant. Therefore, a series of repeated-measures analysis of variance (RM ANOVA) (one test per measure, totaling six tests for oral, four tests for written CAF, five tests for LIQ, with Bonferroni corrections for all post hoc tests) were computed. The between-subjects variable was context while the within-subject variable was time. Before running the tests, assumptions to violations of RM ANOVA were checked, such as normality of distribution, equality of variances, and data sphericity. Only sphericity was found to be partially violated. Thus, the values with Greenhouse–Geiser correction were reported to provide robust statistics. Also, data regarding pretest oral accuracy (Err/ASU), pretest breakdown fluency (P/T), and pretest written lexical complexity (D) were found not to be normally distributed, for which log transformations were implemented to follow up with parametric tests (Field, 2013). To report the magnitude of change, we calculated and interpreted two types of effect size (ES) indices: partial eta squared (η^2 , $\leq .001$ as small and $\geq .014$ as large effect sizes for between-subjects) for RM ANOVAs and Cohen's *d* values for paired samples *t*-tests ($\leq .40$ as small and ≥ 1.00 as large effect sizes for within-subjects) following the field-specific benchmarks suggested by Plonsky and Oswald (2014).

Quantitative data from the Likert-scale items in the LIQ were examined for descriptive statistics to determine means and SDs for each item. Later, these item-based values were grouped into a category defining the type of activity in the TL, such as listening. All the statistical analyses were computed using the Statistical Package for Social Sciences (SPSS) version 27.

Qualitative analysis

We followed an inductive content analysis approach (Schreier, 2012) to build a model by discovering patterns, categories, and themes in the interview data. Following two rounds of open-coding, we determined the contextual factors initiating, challenging, or affording L2 development. Interview data were explored through a systematic three-level analysis: data reduction (selecting, simplifying, and focusing on essential information to help emerge patterns and themes), data display (merging higher-order categories under themes), and conclusion drawing (and meaning checking) (Schreier, 2012). Our analyses yielded four major themes including several different categories. To ensure intercoder reliability, all incongruent parts were reviewed and revised. All participant names used are pseudonyms.

RESULTS

RQ1 THE EFFECT OF CONTEXT ON L2 DEVELOPMENT

Oral CAF development over time

Data from the 1-minute oral task were analyzed using CAF measures to examine differences in development between the three groups over time. Descriptive statistics (Table 2) showed that pretest CAF scores from the two sojourn groups were similar except for repair fluency and accuracy. The AH group had lower scores on all measures compared to the sojourn groups except for repair fluency, accuracy, and lexical complexity. Inspecting the means and SDs from the three groups across time, there was a great deal of variation in CAF gains across the groups.

TABLE 2. Descriptive statistics for oral production data

| Oral Production Data | | SA (n = 8) | | ELFSA (n = 24) | | AH (n = 15) | |
|-----------------------------------|------|--------------------------------|-------|--------------------------------|-------|--------------------------------|-------|
| | | M [CI] | SD | M [CI] | SD | M [CI] | SD |
| Fluency Speech Rate (W/T)* | Pre | 1.45 [1.34, 1.56] | .175 | 1.52 [1.32, 1.72] | .520 | 1.30 [1.16, 1.40] | .249 |
| | Post | 1.70 [1.47, 1.93] | .349 | 1.87 [1.64, 2.09] | .551 | 1.46 [1.28, 1.67] | .551 |
| Breakdown Fluency (P/ASU)* | Pre | .315 [.257, .377] | .953 | .341 [.276, .415] | .180 | .412 [.377, .478] | .106 |
| | Post | .240 ⁺ [.170, .312] | .107 | .232 ⁺ [.197, .268] | .090 | .347 ⁺ [.294, .406] | .114 |
| Repair Fluency (D/ASU) | Pre | 3.20 [2.34, 3.99] | 1.29 | 4.55 [3.33, 5.71] | 3.09 | 2.59 [1.66, 4.18] | 2.52 |
| | Post | 3.91[1.88, 6.33] | 3.43 | 4.52 [3.42, 5.77] | 2.95 | 3.71[2.20, 5.34] | 3.08 |
| Accuracy ⁺ (ERR/ASU)** | Pre | .691[.542, .837] | .235 | .464 [.276, .698] | .522 | .305 [.211, .419] | .208 |
| | Post | .328 ⁺ [.205, .476] | .210 | .255 ⁺ [.163, .368] | .272 | .393 [.275, .532] | .259 |
| Syntactic Complexity (CL/ASU) | Pre | 2.31 [1.78, 2.76] | .775 | 2.21 [1.93, 2.50] | .682 | 2.18 [2.00, 2.54] | .565 |
| | Post | 2.71 [2.13, 3.37] | .959 | 2.22 [1.91, 2.49] | .725 | 2.06 [1.75, 2.40] | .650 |
| Lexical Complexity (D) | Pre | 35.83[28.58, 45.29] | 10.62 | 36.28 [29.92, 42.10] | 14.96 | 38.32 [31.63, 40.27] | 11.57 |
| | Post | 36.88[31.01, 42.91] | 9.46 | 42.48 [37.66, 47.32] | 12.66 | 33.76 [29.89, 38.02] | 8.45 |

*Main effect significant at <.05 level.

**Main and interaction effect significant at <.05 level.

⁺A lower score indicates a better performance.

To investigate if these changes were statistically significant, CAF scores were entered into a series of RM ANOVA tests with time as the within-subjects factor (pretest-posttest) and context group as the between-subjects factor (AH, ELFSA, SA). Results yielded a significant main effect of time only for speech rate ($F(1,44) = 20.074, p = .000$, partial $\eta^2 = .313$), suggesting improvement overtime for the groups as a whole with a large effect size. Considering breakdown fluency, results also indicated a main effect for time only ($F(1,44) = 14.519, p = .000$, partial $\eta^2 = .248$), with a large effect size. Results for repair fluency were insignificant.

Regarding spoken accuracy, a significant time*group interaction ($F(2, 44) = 3.176, p = .042$, partial $\eta^2 = .134$) with a large effect size was detected. To examine how groups performed differently on this measure, a paired-samples t-test was performed for each group. Results indicated that only the two sojourn groups significantly developed over time on spoken accuracy (see Figure 1): (1) SA ($t(7) = 4.027, M$ difference [CI] = .363, [.150, .576], $p = .005$) with a large effect size (d [CI] = 1.63 [.45, 2.75]) and (2) ELFSA ($t(23) = 2.142, M$ difference [CI] = .210, [.007, .411], $p = .043$) with a medium effect size (d [CI] = .48 [.016, .97]). No significant effects were found for syntactic or lexical complexity (see Table 3 for a summary of oral results).

Written CAF development over time

Data from the written task were analyzed for CAF measures, and those values were utilized as dependent variables in the statistical tests. Table 4 summarizes the mean and SDs for all measures across the two data times. Inspecting the mean changes over time for

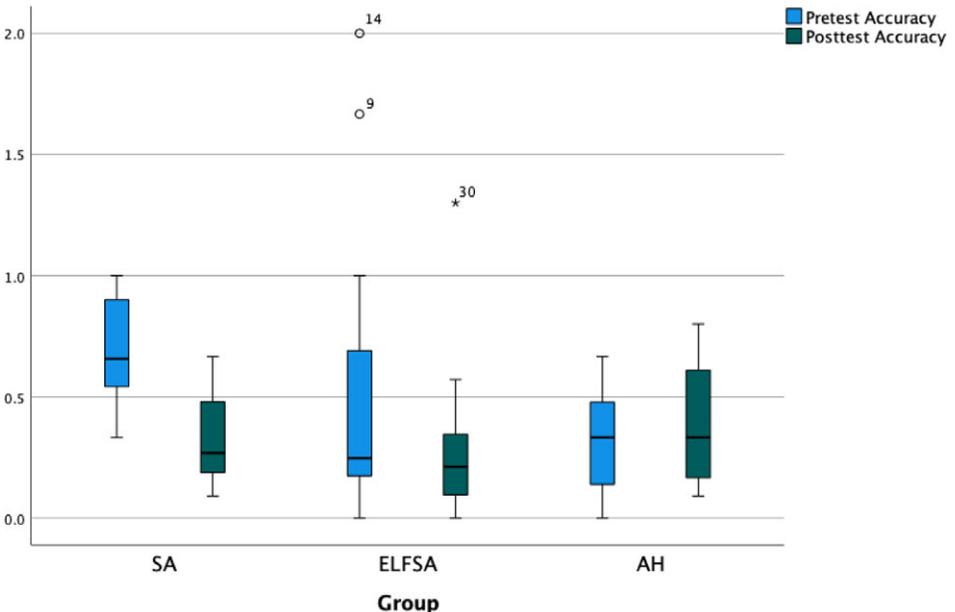


FIGURE 1. Boxplots for pre- and postoral accuracy (ERR/ASU).
 Note: Lower scores indicate a better performance.

TABLE 3. Results summary for oral development

| Spoken Fluency | Lexical Complexity (D) | Accuracy (ERR/ASU) | Syntactic Complexity (CL/ASU) |
|--|------------------------|--------------------------------|-------------------------------|
| Speech Rate (W/T) Time only SA = ELFSA = AH | None | Interaction SA = ELFSA > AH | None |
| Breakdown Fluency (P/T) Time only SA = ELFSA = AH | | | |
| Repair Fluency (D/T) None | | | |

The = means no significant differences found; > means outperformed.

written fluency, SA and AH demonstrated gains from pretest to posttest as opposed to ELFSA. Considering accuracy, AH and ELFSA had fewer errors in the posttest. For syntactic complexity, only AH had gains on the posttest. For lexical diversity, only the sojourn groups had higher means in the posttest.

To explore if these changes were statistically significant, a series of RM ANOVAS were computed. Results of written fluency indicated a main effect for time ($F(2,44) = 2.873, p = .002$, partial $\eta^2 = .196$) with a large effect size, suggesting significant gains on this measure regardless of context (see Figure 2). No significant differences were found on measures of written accuracy, and syntactic and lexical complexity (see Table 5 for a summary of written results).

RQ2: AMOUNT AND TYPES OF L2 CONTACT

To examine potential differences between the types and amount of English contact across the three contexts, descriptive statistics from the LIQs were determined from the 13 6-point Likert-scale items (6: everyday, 5: four or five times a week, 4: two or three times a week, 3: once a week, 2: once in every 2 weeks, 1: never) and one “self-reported English use hours,” provided in Table 6 and depicted in Figure 3. Data from each item were grouped into a category defining the type of contact activity as (1) listening, (2) reading, (3) writing, (4) speaking, and (5) internet use. To exemplify, questions about reading in English for leisure and academic purposes (e.g., *How frequently did you read something for leisure?* and *How frequently did you read something for academic purposes?*) were merged and an average mean and SD were calculated per type of activity for each administration time (T1-T2-T3-T4).

Comparing across the four skill-areas, the main group difference is speaking. As expected, the AH group reported speaking in English very infrequently (see Table 6) with very few native (NS) and nonnative speaker (NNS) contacts (see Table 7). In the ELFSA group, the amount of contact with a NS was as low as the AH, but they reported the highest English interaction with NNSs. All groups reported similar internet use in English. Although the AH group had the lowest score for overall activities in English, they reported the highest amount of self-reported English hours due to their 18 hours of EMI per week.

RM ANOVAs were conducted to investigate differences in the amount and type of contact across time. Results indicated a significant main effect for time ($F(2,44) = 4.853$,

TABLE 4. Descriptive statistics for written production data

| Written Production Data | | SA (n = 8) | | ELFSA (n = 24) | | AH (n = 15) | |
|--------------------------------|------|--------------------------------|-------|--------------------------------|-------|--------------------------------|-------|
| | | M [CI] | SD | M [CI] | SD | M [CI] | SD |
| Fluency (W/T)* | Pre | 14.33 [11.83, 16.70] | 3.79 | 17.20 [15.11, 19.39] | 5.61 | 12.68 [12.07, 13.95] | 2.04 |
| | Post | 18.05 [13.23, 22.33] | 7.03 | 17.65 [15.52, 20.01] | 5.54 | 16.92 [14.04, 21.10] | 6.89 |
| Accuracy ⁺ (ERR/TU) | Pre | .544 [.283, .890] | .473 | .445 [.310, .586] | .354 | .359 [.260, .515] | .301 |
| | Post | .464 ⁺ [.292, .664] | .284 | .402 ⁺ [.293, .509] | .275 | .243 ⁺ [.167, .462] | .301 |
| Syntactic Complexity (CL/TU) | Pre | 2.01 [1.72, 2.28] | .440 | 2.28 [2.05, 2.53] | .620 | 2.32 [2.00, 2.46] | .587 |
| | Post | 2.15 [1.85, 2.43] | .456 | 2.39 [2.18, 2.62] | .548 | 2.11 [1.89, 2.32] | .429 |
| Lexical Complexity (D) | Pre | 52.14 [44.69, 58.61] | 11.37 | 68.26 [61.78, 74.75] | 16.79 | 63.25 [57.27, 70.24] | 12.11 |
| | Post | 60.38 [53.21, 67.23] | 11.02 | 74.61 [67.67, 81.44] | 17.86 | 62.99 [57.79, 68.47] | 11.43 |

*Main effect significant at <.05 level.

⁺A lower score indicates a better performance.

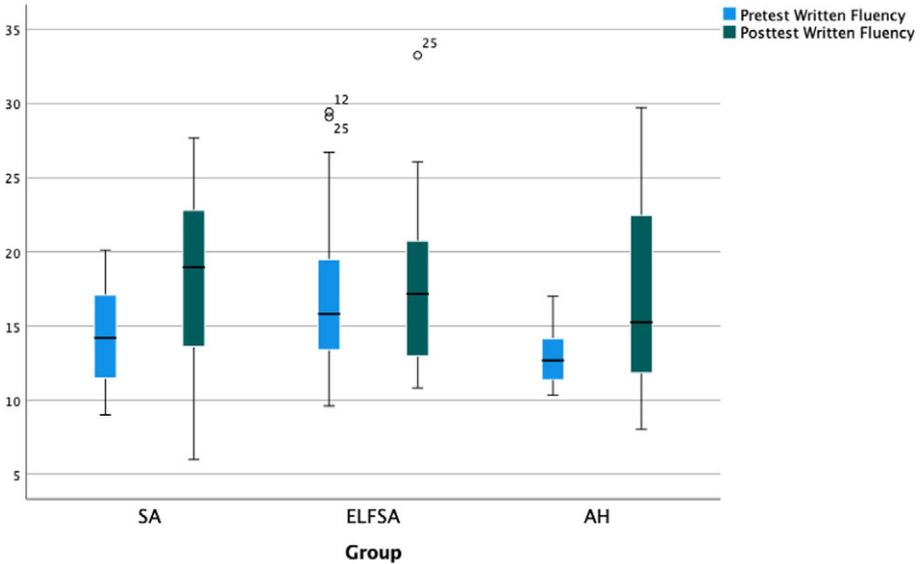


FIGURE 2. Boxplots for pre- and postwritten fluency (W/T).

TABLE 5. Results summary for written development

| Fluency (W/T) | Lexical Complexity (D) | Accuracy (ERR/TU) | Syntactic Complexity (CL/TU) |
|------------------------------|------------------------|-------------------|------------------------------|
| Time only SA = ELFSA = AH | None | None | None |

The = means no significant differences found.

$p = .003$, partial $\eta^2 = .099$) with a large effect size on the total amount of English use (average of all activities) across T1–T4, suggesting that the groups increased their English use similarly over time.

The five types of L2 contact were also compared to examine if the groups had significantly different types of contact across T1–T4. Results of the RM ANOVAs indicated a significant time*group interaction on English use for writing (academic, leisure, e-mailing) ($F(2,44) = 3.453, p = .012$, partial $\eta^2 = .115$) and speaking ($F(2,44) = 7.727, p = .000$, partial $\eta^2 = .131$) with large effect sizes. To examine these differences further, a series of pairwise and multiple comparisons with Bonferroni correction were performed. Results indicated that the AH participants significantly increased their writing in English from T1 to T4 ($t(14) = -3.912, M$ difference [CI] = $-.283, [-.437, -.128], p = .002$) with a medium effect size (d [CI] = $.94 [0.34, 1.50]$).

For the speaking results (Table 8), no differences were found between the sojourn groups, but they had significantly higher spoken interaction over time compared to AH with a large effect size (d [CI] = $.3.008 [2.079, 3.936]$). It should be noted that only the significant pairwise and multiple comparisons are reported. Table 9 summarizes the LIQ results for all types of activities.

TABLE 6. Descriptive statistics for the Likert-scale items in the LIQ for English (L2) use

| | Time 1 | | | Time 2 | | | Time 3 | | | Time 4 | | | | | | | | | | | | | | |
|---|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|---------------------|------------|
| | ELFSA | | AH | ELFSA | | AH | ELFSA | | AH | ELFSA | | AH | | | | | | | | | | | | |
| | SA (n = 8) | (n = 24) | (n = 15) | SA (n = 8) | (n = 24) | (n = 15) | SA (n = 8) | (n = 24) | (n = 15) | SA (n = 8) | (n = 24) | (n = 15) | | | | | | | | | | | | |
| | M | SD | | | | | | | | | | | | |
| | M [CI] | | M [CI] | | M [CI] | | M [CI] | | M [CI] | | M [CI] | | | | | | | | | | | | | |
| <i>Listening (academic & leisure)</i> | 4.63 | 1.22 | 4.65 | 1.10 | 4.53 | 1.23 | 4.44 | 1.05 | 4.69 | .73 | 4.80 | .92 | 4.50 | 1.20 | 4.65 | .83 | 4.97 | 1.22 | 4.44 | .82 | 4.64 | .67 | 5.23 | .82 |
| | [3.81, 5.37] | | [4.21, 5.08] | | [3.87, 5.10] | | [3.69, 5.06] | | [4.40, 4.98] | | [4.33, 5.20] | | [3.56, 5.19] | | [4.31, 4.98] | | [4.33, 5.50] | | [3.87, 4.94] | | [4.35, 4.90] | | [4.77, 5.60] | |
| <i>Reading (academic & leisure)</i> | 4.00 | 1.07 | 4.62 | 1.24 | 4.67 | 1.08 | 3.88 | .88 | 4.23 | .79 | 4.73 | .86 | 4.13 | .74 | 4.25 | .92 | 4.70 | 1.66 | 4.13 | .64 | 4.33 | .97 | 4.93 | 1.02 |
| | [3.31, 4.74] | | [3.56, 4.56] | | [4.16, 5.20] | | [3.31, 4.43] | | [3.94, 4.52] | | [4.30, 5.13] | | [3.62, 4.62] | | [3.92, 4.60] | | [4.17, 5.23] | | [3.68, 4.56] | | [3.95, 4.72] | | [4.43, 5.40] | |
| <i>*Writing (academic, leisure, e-mailing)</i> | 3.63 | .93 | 4.09 | 1.19 | 2.87 | .92 | 4.04 | .86 | 4.08 | .94 | 3.29 | 1.13 | 4.08 | .71 | 3.92 | .98 | 3.64 | 1.20 | 4.17 | .99 | 4.11 | 1.04 | 3.73 | .93 |
| | [3.04, 4.20] | | [3.60, 4.55] | | [2.42, 3.11] | | [3.46, 4.58] | | [3.71, 4.43] | | [2.73, 3.82] | | [3.62, 4.50] | | [3.53, 4.28] | | [3.04, 4.20] | | [3.50, 4.75] | | [3.68, 4.50] | | [3.27, 4.18] | |
| <i>*Speaking (academic/personal, service encounters)</i> | 4.38 | 1.30 | 5.31 | .87 | 2.13 | 1.04 | 5.31 | .88 | 5.00 | 1.12 | 2.43 | 1.29 | 5.19 | .84 | 4.98 | 1.11 | 2.10 | 1.15 | 5.25 | .93 | 4.90 | .94 | 2.33 | 1.18 |
| | [3.50, 5.19] | | [4.94, 5.60] | | [1.66, 2.67] | | [4.69, 5.81] | | [4.54, 5.42] | | [1.87, 3.13] | | [4.56, 5.69] | | [4.52, 5.35] | | [1.63, 2.73] | | [4.32, 5.75] | | [4.52, 5.23] | | [1.83, 2.93] | |
| <i>Internet use (surfing & social media)</i> | 4.38 | 1.48 | 4.88 | 1.12 | 4.78 | 1.15 | 4.19 | 1.30 | 5.17 | 1.03 | 5.03 | 1.11 | 4.50 | 1.00 | 5.13 | 1.01 | 5.23 | 1.10 | 4.56 | 1.11 | 5.10 | 1.01 | 5.23 | 1.12 |
| | [3.62, 4.83] | | [4.34, 5.19] | | [3.60, 4.67] | | [3.67, 4.92] | | [4.53, 5.24] | | [3.75, 4.73] | | [3.87, 5.00] | | [4.44, 5.17] | | [4.04, 5.09] | | [3.83, 5.25] | | [4.50, 5.27] | | [4.09, 5.02] | |
| *Total amount of English use (average of all activities) M, SD, and M [CI] | 4.15 | .29 | 4.55 | .17 | 3.71 | .21 | 4.34 | .25 | 4.58 | .14 | 3.99 | .18 | 4.44 | .27 | 4.52 | .15 | 4.08 | .20 | 4.48 | .25 | 4.57 | .14 | 4.24 | .18 |
| | [3.67, 4.69] | | [4.18, 4.89] | | [3.35, 4.10] | | [3.92, 4.78] | | [4.32, 4.85] | | [3.62, 4.37] | | [4.03, 4.84] | | [4.24, 4.79] | | [3.68, 4.50] | | [4.02, 4.87] | | [4.28, 4.84] | | [3.85, 4.60] | |
| <i>Daily hours of English use</i> | 3.63 | 2.00 | 5.29 | 3.20 | 4.53 | 3.68 | 4.75 | 3.11 | 4.90 | 3.28 | 5.60 | 4.03 | 4.88 | 3.40 | 4.77 | 3.19 | 5.80 | 3.86 | 4.50 | 3.42 | 5.13 | 2.75 | 6.00 | 3.98 |
| | [3.66, 4.73] | | [4.18, 4.86] | | [3.33, 4.09] | | [3.94, 4.79] | | [4.32, 4.84] | | [3.59, 4.36] | | [4.02, 4.85] | | [4.22, 4.79] | | [3.67, 4.50] | | [3.99, 4.90] | | [4.28, 4.83] | | [3.84, 4.58] | |

*Significant at <.05 level.

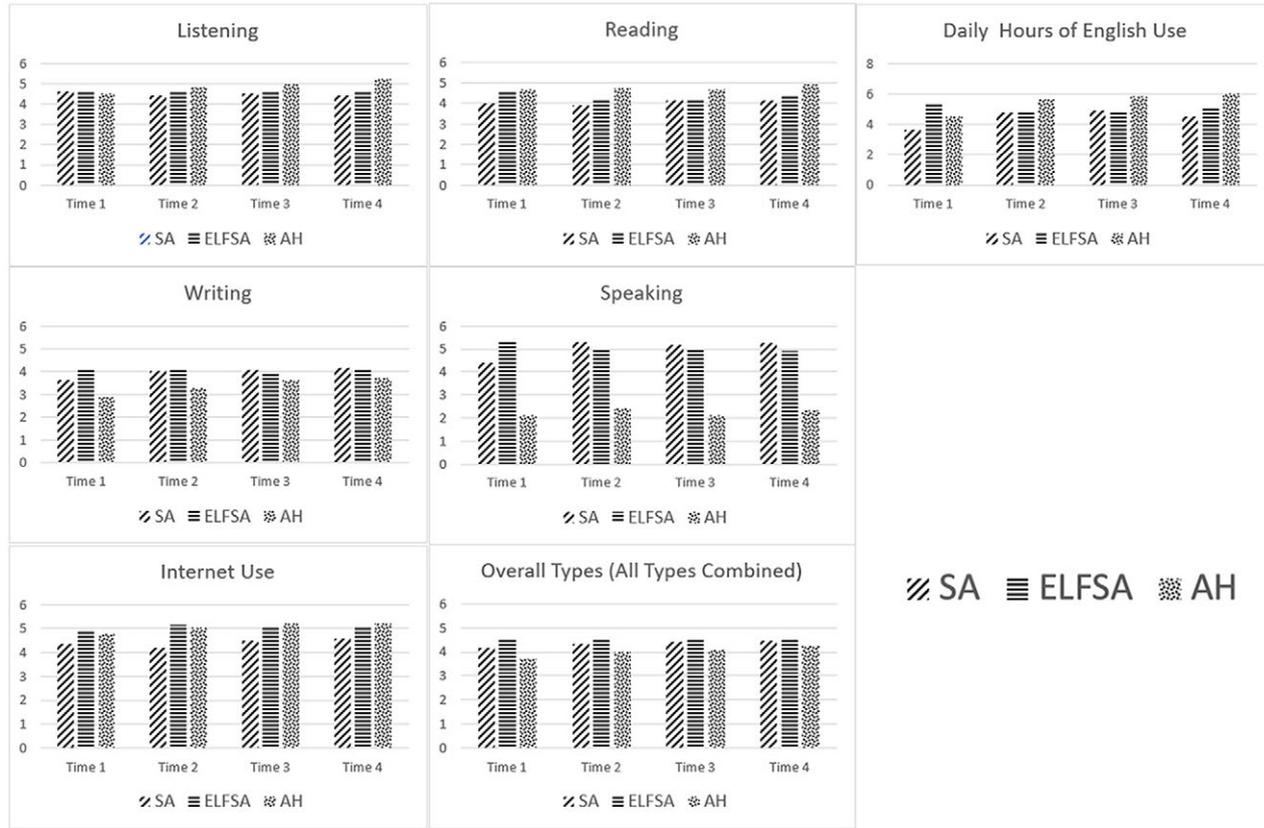


FIGURE 3. LIQ types from time 1 to time 4.

Note: For all results except *Daily Hours of English Use*, the items are coded as 6: everyday, 5: four or five times a week, 4: two or three times a week, 3: once a week, 2: once in every 2 weeks, 1: never.

TABLE 7. Descriptive statistics for interlocutor types

| | SA (n = 8) | | ELFSA (n = 24) | | AH (n = 15) | |
|------------------------|--------------|-----|----------------|-----|--------------|-----|
| | M | SD | M | SD | M | SD |
| Interaction in English | M [CI] | | M [CI] | | M [CI] | |
| NS Interlocutor | 4.71 | .68 | 1.86 | .90 | 1.42 | .46 |
| | [4.28, 5.19] | | [1.51, 2.19] | | [1.24, 1.69] | |
| NNS Interlocutor | 3.95 | .52 | 4.64 | .71 | 2.49 | .42 |
| | [3.57, 4.28] | | [4.36, 4.90] | | [2.31, 2.71] | |

Note: This Likert-scale item appeared only in the last LIQ (T4) and were coded as 6: everyday, 5: four or five times a week, 4: two or three times a week, 3: once a week, 2: once in every 2 weeks, 1: never.

TABLE 8. Results of the multiple comparison on speaking from T1 to T4

| Groups | | Mean Difference | 95% CI | |
|----------|----|-----------------|--------|-------|
| Compared | | | Lower | Upper |
| ELFSA | SA | .016 | -.952 | .921 |
| | AH | 2.79* | 2.02 | 3.58 |
| SA | AH | 2.78* | 1.75 | 3.08 |

*Significant at .000 level.

TABLE 9. Results summary for LIQs

| Listening | Reading | Writing | Speaking | Internet use | All activities combined |
|-----------|---------|--------------------------------|--------------------------------|--------------|------------------------------|
| None | None | Interaction AH > SA = ELFSA | Interaction SA = ELFSA > AH | None | Time only SA = ELFSA = AH |

The = means no significant differences found, > means outperformed.

RQ3: INDIVIDUAL LEARNERS' EXPERIENCES ACROSS CONTEXTS

To investigate participants' lived experiences related to language contact, contextual features, and L2 development, three participants from each group (representing either low, medium, or high amount of contact as reported on the LIQ), were chosen for semistructured interviews. Four major themes were identified from the coded responses, with some context-specific themes. Table 10 displays the interviewer characteristics and Table 11 the distribution of themes across the three groups.

Initiator

One theme consistently discussed by participants in all three contexts was *the willingness to leave one's comfort zone* as an initiator to L2 interaction and development. For the sojourner groups, it was an informal prerequisite for participating in an international mobility program. Even in the AH context, participants acknowledged the importance of looking for opportunities to interact with speakers of the TL even if they were uncomfortable doing so. To exemplify, Hale, who reported the least amount of interaction in the LIQ from the AH

TABLE 10. Interviewee profiles

| Name*/ Group | Sex | Major | Pre-EIT (out of 120) | Amount of L2 Interaction (1st LIQ Mean out of 6.00) |
|--------------|-----|-------------------------|----------------------|---|
| Burcu/SA | F | Economics | 47 | 3.36 |
| Suzan/SA | F | ELT | 106 | 4.22 |
| Bilge/SA | F | ELT | 86 | 5.55 |
| Fuat/ELFSA | M | Economics | 111 | 2.18 |
| Can/ELFSA | M | Economics | 98 | 4.60 |
| Ada/ELFSA | F | Business Administration | 79 | 5.82 |
| Hale/AH | F | ACL | 114 | 2.73 |
| Seda/AH | F | ACL | 53 | 4.35 |
| Ebru/AH | F | ACL | 67 | 5.18 |

*All names are pseudonyms. ELT = English language teaching, ACL = American culture and literature.

TABLE 11. Themes

| Themes | Categories combined | SA | ELFSA | AH |
|--------------------------|---|----|-------|----|
| Initiator | Willingness to leave one’s comfort zone | ✓ | ✓ | ✓ |
| Challenges/Diminishers | Local English variety | ✓ | | |
| | Type of interlocutor (NS) | ✓ | | |
| Facilitators/Affordances | Local language/L3 | | ✓ | |
| | Amount of contact | | | |
| | Type of Interlocutor (NNS) | ✓ | ✓ | |
| | ELF | | ✓ | |
| | Coursework | | | ✓ |

group, stated that she was not ready to leave her *comfort zone* to study abroad or even to continue interacting with L2 speakers in online environments. Burcu, although staying with a host family in England, reported the least amount of interaction in the SA group. She claimed to have developed less than expected, because “it took three months to leave my *comfort zone*, adapt and start to learn.” Fuat, who chose to study in Germany and reported the least amount of English interaction in the ELFSA group, also described his context as one “forcing [him] to leave [his] comfort zone ... as [he has] no choice but to speak.” Conversely, reporting the highest amount of interaction in their respective groups, Ada (ELFSA) and Bilge (SA) described how they looked for opportunities to use the TL even if it meant leaving their *comfort zones*: “this [ERASMUS experience] is nothing like a holiday because it is not a short-term experience ... we are now far from our *comfort zones* and we have to do a lot of things in English” (Ada, ELFSA- Denmark).

Challenges/diminishers

Participants highlighted a variety of context-specific challenges/diminishers that might slow down L2 development. To exemplify, the SA participants stressed several difficulties understanding the local variety of English due to NS’s “thick British accent,” speech rate, and spoken features like weak forms or connected speech. These accent or variety-related issues of intelligibility were major categories for this theme. As Suzan

(SA) reported it became “easier to understand what they [NSs] say ... I know what *cheers* means when used instead of *thanks*.”

Additionally, NS interlocutors in SA were repeatedly mentioned as a source of anxiety, diminishing L2 interaction, and development. All SA participants complained about NS criticism toward their accents or use of English. Burcu stressed this theme and added that she is “afraid of making mistakes” when interacting with NSs, who mostly ask for clarification with “weird faces.” She reported being frequently asked where she came from as her “Turkish accent was weird for them” (Burcu, SA). Similarly, Suzan (SA) referred to her context as one where “no mistakes are tolerated.” Suzan (SA) thought that this would have never been the case if her instructor had been a NNS.

In ELFSA, all participants referred to challengers/diminishers related to communication in the local language, particularly during service encounters with locals. Ada described how being able to use basic greetings in Danish helped her connect with locals who were then more willing to switch to English: “If one cannot say hi in Danish, they will not approach you and later codeswitch into English once your Danish is insufficient to keep talking.” For Ada and others in the ELFSA group such experiences led them to start learning the local language at a rudimentary level. In contrast, Fuat, already a competent German speaker, actively sought more interaction opportunities in his L3.

Facilitators/affordances

This theme emerged partially differently in the sojourn contexts. Only the SA interviewees described their context as one with “millions of” or “tons and tons of opportunities to use English” in daily life (Burcu, SA), which helped them overcome initial problems with local English varieties. Bilge (SA) attributed her improvement to the fact that “everybody speaks English.”

Another facilitator/affordance of both sojourn contexts was the availability of NNS interlocutors. Burcu (SA) indicated that she prefers “interacting with other international students because it is easier to comprehend their speech” and they are less “critical” of her pronunciation. ELFSA interviewees echoed the ease of talk with other “fellow sojourners” and “ERASMUS people” who “tolerate mistakes” (Fuat, Germany), do not look with “weird faces” (Ada, Denmark), and interact without “native speaker-based parameters and expectations” (Can, Portugal). NNSs created a space where lower proficiency learners felt motivated to communicate without constantly questioning whether their “speech is grammatically correct” (Ada, Denmark). Clarification requests by more expert NNSs were described as helping develop their English in contrast to what some described as demoralizing constant negative feedback from NSs. Hence, NNS interaction was a factor changing sojourners’ perceptions toward their mistakes, and likely facilitating their L2 development.

As a variety, ELF use in NNS-NNS interaction was the last facilitator/affordance that emerged in the sojourn contexts. ELFSA interviewees highlighted “less-complex sentence structure with more common vocabulary” (Can, Portugal). They also referred to the multilingual nature of their contexts, that everyone is indeed able to speak at least one other language than English, and that English is the “building bridge” between speakers of different L1s (Ada, Denmark). ELF also helped them build rapport with fellow sojourners, interact more using English, and end up with self-perception of development.

The basic condition for successful communication in “the international contexts [referring to ELF] is ... to use more chunks [formulaic language] regardless of native-speaker based parameters and expectations” (Can, Portugal). Many ELFSA participants also criticized the teaching curriculum and native-speakerist ultimate attainment expectations in Turkey and emphasized the friendly atmosphere they experienced in ELFSA where using communication strategies and alternative ways to convey meaning were accepted even when mistakes were made. Overall, all three ELFSA interviewees’ experiences appeared to help them adapt an ELF viewpoint to L2 use and ultimate attainment in the TL.

Considering the last facilitator/affordance, AH interviewees attributed their English development primarily to the intense amount of coursework in their EMI program, a theme not discussed by sojourners. They described spending long hours on campus using English to complete coursework including preparing and delivering presentations, reading class materials, and writing reports and essays. This academic orientation was the most remarkable difference between the AH and sojourner groups emerging from the interview data. Ebru (AH) referred to having occasional interactions with incoming ERASMUS students, but this did not come up as a major source of L2 development in the AH.

DISCUSSION

The purpose of this study was to explore how learning English abroad in ELFSA compares to learning English abroad in England and at home in Turkey. Due to the increased popularity of ELFSA, particularly in Europe, it is crucial to investigate the potential learning opportunities students have in this context. The current study is the first to compare two different sojourn contexts with intensive AH instruction. This is an important topic because many students are opting to learn English in non-Anglophone countries through exchange programs offering EMI. Furthermore, with Britain’s recent decision to leave the Erasmus program, ELFSA may become even more popular.

The findings of RQ1 indicated that time spent in an intensive AH program was equally beneficial to time spent abroad in terms of development in speech rate and breakdown fluency, supporting the results of Serrano et al. (2011). With increased exposure through EMI, AH participants in the current study still developed certain dimensions of their oral L2 performance without additional organized TL activities outside of their regular intensive EMI coursework. No differences were found between ELFSA and SA on oral fluency, supporting recent research (e.g., Llanes, 2019), demonstrating ELFSA as an equally favorable context for improving L2 oral performance.

The finding indicating an interaction effect for oral accuracy in favor of the sojourn groups lends support to studies such as DeKeyser (2007), Juan-Garau (2014), and McManus et al. (2021) showing evidence of accuracy gains as a result of time spent abroad. This finding may be linked to the authentic nature of spoken interactions in SA and ELFSA where sojourners have ample opportunities to proceduralize and later automatize declarative knowledge in the TL, as suggested by Skill Acquisition Theory (SAT) (DeKeyser, 2007). Additionally, the SA and ELFSA groups likely received more informal speaking practice, which could have benefitted them on the specific oral production task used in this study. In comparison, the AH group was practicing more formal oral tasks in their classes, such as giving presentations, and reported lower amounts of speaking on the LIQ. Similar to Juan-Garau’s (2014) findings, the AH participants did

not produce more targetlike speech despite having received plenty of explicit instruction and having the highest overall number of instructional hours.

None of the groups showed evidence of syntactic or lexical complexity development after a semester abroad or at home. This finding might be related to the amount of time spent in sojourn, in that the longer the SA experience the larger the gains in syntactic complexity would be (McManus et al., 2021). The SA participants, as revealed in the interviews, might have also preferred simpler language less prone to errors as a strategy to avoid communication breakdowns, which they mentioned often triggered clarification requests from more competent interlocutors. Additionally, these sojourners might have also sacrificed lexical diversity and sophistication for increased accuracy while interacting with English NSs. The nature of NNS-NNS interaction with less competent speakers might account for simpler sentence structures including high-frequency vocabulary, typical of ELF for successful communication (Jenkins et al., 2011).

Considering written gains, the only significant finding was a main effect for written fluency, a result also reported for the SA group in Mitchell et al. (2017). This finding may be due to the type of tasks and the amount of written practice experienced by all participants in their EMI programs, at home and abroad. Limited improvement in other areas of writing is a common finding in the research on SA and may be explained by the length of the study period (Juan-Garau, 2014) or that sojourners often prioritize spoken language development over written.

RQ2 focused on the amount and types of L2 contact. The LIQs indicated similar use and exposure to the TL across the two sojourn contexts. Although the total amount of spoken interaction was similar between SA and ELFSA, the ELFSA group reported the highest amount of English contact with NNSs, whereas the SA group reported the highest amount of English contact with NSs. Interestingly, the SA group also reported high amounts of interaction with NNSs, primarily other ERASMUS and international students while in England. This informal interaction, especially within NNS-NNS communication using ELF, likely contributed to the sojourn groups' significant improvement in oral accuracy. The intensive AH group reported the highest amount of overall L2 use in hours, yet their spoken contact was significantly less than the sojourn groups. Comparing the mean use of English across T1–T4, all groups significantly increased their TL contact over time. Taking the type of contact into consideration, AH had significantly higher written contact over time with the L2. Overall, despite these findings showing increased English contact, gains were somewhat limited. All groups improved in fluency (both spoken and written), but only the sojourn groups showed improvement in spoken accuracy.

RQ3 examined selected participants' experiences across the three contexts through questionnaire and interview data. These qualitative results underscored a major theme across contexts, leaving one's *comfort zone* as an initiator to creating opportunities to develop in the TL, a theme also discussed in Jackson (2008). The interviewees in all three contexts explained the lack or presence of practice opportunities to one's daring to leave her *comfort zone*. This *metacontextual* theme makes sense in today's global world as a language learner even in at-home contexts may still have multiple opportunities to get exposure to and practice English online or face-to-face. For example, Turkey is attracting a growing number of incoming ERASMUS exchange students and faculty, creating opportunities for AH students to interact with other English speakers. Yet, it depends on individual agency to benefit from such opportunities. For those who struggle taking risks or leaving their *comfort*

zones, the sojourn experience may not result in more gains than an AH experience for someone who sought out and seized every opportunity to interact using the TL.

Focusing on challenges or affordances of the three contexts, participants in SA highlighted issues of intelligibility due to the local variety of English. ELFSA interviewees described their context as one with a more relaxed and tolerated atmosphere helping them gain an ELF perspective and ownership of English. They became more tolerant of mistakes, lowered their nativelike ultimate attainment expectations, and enjoyed negotiating meaning with their interlocutors, factors prevalently echoed in the ELF literature (e.g., Jenkins et al., 2011). ELF might have particularly encouraged lower-level speakers to look for ways to use English to develop their skills. This result supports previous findings regarding the low-anxiety nature of ELF contexts compared to Anglophone contexts (Borghetti & Beaven, 2017; Martin-Rubió & Cots, 2018). Numerous input and interaction opportunities were available in both sojourn contexts, yet ELFSA brought a different perspective with its multilingual nature, adding flexibility to learners' self-perceptions about their development. ELFSA participants reported more opportunities to use English as compared to SA interviewees. Their comments suggest that ELFSA not only eased participation in ELF interactions but also promoted self-confidence and ownership (Kohn, 2018) of English more than the NS context. Interviewees also discussed the use of multiple languages in ELFSA and how international students use ELF as a mediator to negotiate meaning when they failed to communicate using the local language. Consequently, a sojourn is a personal experience and some participants, due to individual differences (IDs) such as motivation or personality, might develop more compared to others (Marijuan & Sanz, 2018). Finally, the coursework-oriented gains in intensive AH are no less valuable if one wants to develop academically in the L2, given the focused instruction and practice opportunities available in domestic EMI curricula.

LIMITATIONS AND FUTURE RESEARCH

This study contributed to our understanding of different contexts of learning, including two sojourn contexts (SA and ELFSA) and one intensive AH context, and their impact on L2 CAF development. Yet, there are several limitations of this study that should be acknowledged. The type of performance tasks selected were limited to free production prompts, lending themselves more easily to the analysis of fluency. Future research should include a variety of different tasks to provide a fuller picture of proficiency development. Also, a single measure (Err/ASU for oral and Err/TU for written production) was utilized to document general accuracy development. A detailed analysis of grammatical, lexical, and pragmatic errors would give further insights into context of learning and accuracy development. Furthermore, our study does not explore any IDs, such as personality or motivation. IDs might play a large role in participants deciding to study abroad or making the most of their sojourn experience.

This study also utilized multiple RM ANOVAs to investigate if changes in CAF and type and amount of TL contact were statistically significant. Though typical of CAF analyses in SLA, running multiple tests might inflate Type-1 error rates. Thus, additional studies are needed to corroborate these findings. Acknowledging such risks, we followed the norms of the field and investigated CAF measures utilizing

robust statistics with bootstrapped CIs and effect sizes. Considering the qualitative results, this study primarily draws on semistructured interview data to triangulate LIQ findings and give a voice to a selected subgroup of participants to explore the dynamics of their contexts. Additional qualitative data (e.g., participant observation) could provide a more complete picture of how learners interact with their environment when learning English.

CONCLUSION

The current study investigated the effects of three learning contexts on oral and written L2 development over time. In addition to traditional SA in an Anglophone country (England in this study) and intensive AH instruction, this study focused on an additional context, ELFSA, that is just beginning to gain attention. Due to the increased popularity of international exchange opportunities where sojourners can take EMI classes in non-Anglophone countries like Denmark, it is important to examine this context and the learning opportunities it affords for those who wish to improve their English abroad. Therefore, this study also investigated contextual characteristics of the three learning contexts qualitatively through participant interviews.

Results of this study suggest that ELFSA is as beneficial as SA and intensive AH for helping users of English make improvements in oral and written fluency after a semester abroad. Both sojourn groups improved significantly over time on oral accuracy, but otherwise, no differences were found between the groups. In light of the qualitative findings, ELFSA may provide some advantages over SA for certain students who wish to go abroad, such as providing a low-anxiety atmosphere to improve their English. Clearly, ELFSA is emerging as an appealing sojourn context that merits future research, particularly as we experience a multilingual turn in SLA (Mitchell, 2021).

DATA AVAILABILITY STATEMENT

The experiment in this article earned an Open Materials badge for transparent practices. The materials are available at <https://www.iris-database.org/iris/app/home/detail?id=york:938936>

NOTES

¹At the time of the study, Britain was still an EU member. From 2021 and onward, no British institutions will be involved in the ERASMUS+ program. Therefore, no traditional SA to Britain will be possible for learners of English under the ERASMUS scheme, which accentuates the significance of ELFSA even more. Ireland remains an option for those wishing to go to an Anglophone country.

²This does not rule out the possibility of encountering NNS interaction or ELF contact in an Anglophone country, or NS interaction in an ELF country.

³Official ERASMUS+ agreement document enabling mobility of students between two institutions.

⁴Please note that the type of interlocutor question appeared only in the last administration of the LIQ as a retrospective item.

⁵The EIT results demonstrate equally significant development for all participants regardless of context.

REFERENCES

- Borghetti, C., & Beaven, A. (2017). Lingua francas and learning mobility: Reflections on students' attitudes and beliefs towards language learning and use. *International Journal of Applied Linguistics*, 27, 221–241. <https://doi.org/10.1111/ijal.12123>
- Borràs, J., & Llanes, À. (2019). Re-examining the impact of study abroad on L2 development: A critical overview. *The Language Learning Journal*, 49, 527–540.
- Cairns, D. (2017). The Erasmus undergraduate exchange programme: A highly qualified success story? *Children's Geographies*, 15, 728–740.
- DeKeyser, R. M. (2007). Study abroad as foreign language practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208–226). Cambridge University Press.
- ERASMUS+ Programme Guide (2021). https://ec.europa.eu/programmes/erasmus-plus/resources/programme-guide_en
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191.
- Field, A. (2013). *Discovering statistics using SPSS* (3rd ed.). Sage.
- Firth, A. (1996). The discursive accomplishment of normality: On “lingua franca” English and conversation analysis. *Journal of Pragmatics*, 26, 237–259.
- Foster, P., Tonkyn, A., & G. Wigglesworth. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21, 354–375.
- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26, 275–301.
- González, C. R., Mesanza, R. B., & Mariel, P. (2011). The determinants of international student mobility flows: An empirical study on the Erasmus programme. *Higher Education*, 62, 413–430.
- Jackson, J. (2008). *Language, identity, and study abroad: Sociocultural perspectives*. Equinox.
- Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Language Teaching*, 44, 281–315.
- Juan-Garau, M. (2014). Oral accuracy growth after formal instruction and study abroad: Onset level, contact factors, and long-term effects. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 87–109). John Benjamins.
- Kalocsai, K. (2014). *Communities of practice and English as a lingua franca: A study of students in a central European context*. Walter de Gruyter.
- Kaypak, E., & Ortaçtepe, D. (2014). Language learner beliefs and study abroad: A study on English as a lingua franca (ELF). *System*, 42, 355–367.
- Kohn, K. (2018). MY English: A social constructivist perspective on ELF. *Journal of English as a Lingua Franca*, 7, 1–24.
- Köylü, Z. (2016). *The influence of context on L2 development: The case of Turkish undergraduates at home and abroad* [Unpublished doctoral dissertation]. University of South Florida.
- Köylü, Z. (2021). The ERASMUS sojourn: Does the destination country or pre-departure proficiency impact oral proficiency gains? *The Language Learning Journal*. Advance online publication. <https://doi.org/10.1080/09571736.2021.1930112>
- Llanes, À. (2019). Study abroad as a context for learning English as an international language: An exploratory study. In M. Howard (Ed.), *Study abroad, second language acquisition, and interculturality* (pp. 136–155). Multilingual Matters.
- Llanes, À., Arnó, E. & Mancho-Barés, G. (2016). Is a semester abroad in a non-English-speaking country beneficial for the improvement of English? The case of Erasmus students using English as a Lingua Franca. *Language Learning Journal*, 44, 292–303.
- Llanes, À. & Muñoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 63, 63–90.
- Llanes, À., & Serrano, R. (2017). The effectiveness of classroom instruction “at home” versus study abroad for learners of English as a foreign language attending primary school, secondary school and university. *Language Learning Journal*, 45, 434–446. <https://doi.org/10.1080/09571736.2014.923033>

- MacWhinney, B. (2000) *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.
- Marijuan, S., & Sanz, C. (2018). Expanding boundaries: Current and new directions in study abroad research and practice. *Foreign Language Annals*, 51, 185–204.
- Martin-Rubió, X., & Cots, J. M. (2018). Self-confidence amongst study abroad students in an “English as a lingua franca” university. *Language Awareness*, 27, 96–112.
- McGregor, J. (2021) An investigation of L2 learning peer interactions in short-term study abroad. In W. Diao and E. Trentman (Eds.) *Language learning in study abroad: The multilingual turn* (pp. 72–96). Multilingual Matters.
- 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*, XIX, 97–116.
- McManus, K., Mitchell, R. F., & Tracy-Ventura, N. (2021). A longitudinal study of advanced learners’ linguistic development before, during, and after study abroad. *Applied Linguistics*, 42, 136–163.
- Mitchell, R. (2021). Language and student mobility in Europe. In R. Mitchell & H. Tyne (Eds.) *Language, mobility and study abroad in the contemporary European context* (Pp. n.a.). Routledge.
- Mitchell, R., Tracy-Ventura, N. & McManus, K. (2017). *Anglophone students abroad: Identity, social relationships, and language learning*. Routledge.
- 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.
- Ortega, L., Iwashita, N., Rabie, S. and Norris, J.M. (1999) *A Multilanguage Comparison of Measures of Syntactic Complexity* [Funded Project]. University of Hawaii, National Foreign Language Resource Center.
- Pérez-Vidal, C. (Ed.). (2014). *Language acquisition in study abroad and formal instruction contexts*. John Benjamins.
- Pérez-Vidal, C. (2017). Study abroad and ISLA. In S. Loewen and M. Sato (Eds.) *The Routledge handbook of second language instructed second language acquisition* (pp. 339–360). Routledge.
- Pérez-Vidal, C., & Barquin, E. (2014). Comparing progress in academic writing after formal instruction and study abroad. In C. Pérez-Vidal (Ed) *Language acquisition in study abroad and formal instruction contexts* (pp. 217–234). John Benjamins.
- Pérez-Vidal, C., & Llanes, À. (2021). Linguistic effects of international student mobility in European perspective. In R. Mitchell & H. Tyne (Eds.), *Language, Mobility and Study Abroad in the Contemporary European Context* (pp. 22–33). Routledge.
- Plonsky, L., & F. L. Oswald (2014) How big is “big”? Interpreting effect sizes in L2 research, *Language Learning*, 64, 878–912.
- Sanz, C. & Morales-Front, A. (2018). *The Routledge Handbook of study abroad research and practice*. Routledge.
- Schreier, M. (2012). *Qualitative content analysis in practice*. Sage.
- Segalowitz, N. & Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition. *Studies in Second Language Acquisition*, 26, 173–199.
- 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.
- Serrano, R., Llanes, À., & Tragant, E. (2016). Examining L2 development in two short-term intensive programs for teenagers: Study abroad vs. “at home.” *System*, 57, 43–54.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510–532. <https://doi.org/10.1093/applin/amp047>
- Taguchi, N. (2015). *Developing interactional competence in a Japanese study abroad context*. Multilingual Matters.
- Tanaka, K. (2007). Japanese students’ contact with English outside the classroom during study abroad. *New Zealand Studies in Applied Linguistics*, 13, 36.
- Turkish National Agency [Türk Ulusal Ajansı]. (2015). Yayınlar ve istatistikler: (Publications and statistics: Research based on analysis of youth in action). https://www.ua.gov.tr/media/sxlipj0l/27_tu-rkiye-ulusal-raporu-2015-min.pdf