Universitat de Lleida

Document downloaded from:

<http://hdl.handle.net/10459.1/59507>

The final publication is available at:

<https://doi.org/10.1111/ijal.12151>

Copyright

(c) Wiley, 2016

1. Introduction

The effects of the study abroad context on second or foreign language (L2) development have been examined over a few decades (Freed 1995). Nevertheless, this learning context has recently become the central focus of many studies in SLA (Collentine 2009; Author 2011). This interest in the effects of SA on L2 development does not come as a surprise given that SA experiences have become very popular, probably because of the need to function in a multilingual world imposed by globalization. What is surprising is the lack of literature on the effects of short SA experiences on adolescents given the large number of teenage students engaging in (summer) SA programs. The scarcity of studies that focus on L2 pronunciation as a result of SA is also remarkable given that one of the main reasons why learners go abroad is to improve their L2 oral skills in general and pronunciation in particular (Allen 2010). Moreover, pronunciation is an essential aspect of L2 oral production, as inaccurate pronunciation can impede communication (Arteaga 2000), even when the speaker has a good command of grammar and vocabulary (Lord 2005).

The aim of the present study is to contribute to the SA literature and the L2 pronunciation literature by comparing the effects of learning context, SA vs. intensive course at home (AH), on the L2 pronunciation accuracy of a group of teenage learners of English. Pronunciation accuracy was assessed acoustically through voice onset time (VOT) measures in the production of voiceless stops, and perceptually, through L1-English raters' assessment of overall degree of foreign accent (FA).

2. Literature review

2.1 The effects of SA on L2 development

Several decades of research on SA as a learning context indicate that the benefits for L2 development are not consistent across all populations, SA program schemes and linguistic domains. For example, whereas both adults and children have been found to improve in L2 fluency in SA abroad programs varying greatly in length (Freed 1995; Author1 2013; Martinsen 2010), for L2 phonological development a much more complex picture emerges with some studies finding improvement in pronunciation accuracy for children after stays as short as two months (Author1 2014) and others finding no clear improvement after relatively long stays (Díaz-Campos 2004; Mora 2008; see also Lara, Mora and Pérez-Vidal 2015). In general, SA has been found to have a positive impact in L2 development for vocabulary growth (Dewey 2008; Foster 2009; Milton and Meara 1995), the acquisition of pragmatic and sociolinguistic skills (Regan Howard and Lemée 2009) and listening skills (Beatty, Valls-Ferrer and Pérez-Vidal 2014; Cubillos, Chieffo and Fan 2008; Dyson 1988), whereas its effects on writing, grammar and reading are unclear, with various studies reporting either improvement (Pérez-Vidal and Juan-Garau 2009; Sasaki 2009; Authors 2012; Kinginger 2008) or lack of it (Collentine 2004; DeKeyser 1991; Dewey 2004; Freed, So and Lazar 2003).

Researchers interested in the SA context have also assessed the L2 gains of learners going abroad against those of an AH comparison group. In such studies (Freed 1995; Trenchs-Parera 2009) the SA learners are normally found to obtain greater gains than their AH peers, reportedly because of the differential amount of input the SA and AH groups receive. Although there is variation in the extent to which SA learners engage in L2 activities and social networks (Dewey et al. 2012; Kinginger 2008), the SA context theoretically offers many more opportunities for meaningful language use and greater exposure to authentic L2 input than the typical AH context, which is limited

to the 3-4 hours of class a week. Given such large differences in what SA and AH contexts typically offer in terms of language exposure and practice, it is hardly surprising that many studies report an advantage for the SA over the AH learning context. However, intensive/immersion AH programs, where learners generally obtain at least 14 hours a week of L2 exposure, may provide a more adequate learning context against which to measure the linguistic benefits of SA. Previous research comparing SA and intensive/immersion AH learner groups appear to disconfirm the superiority of SA over AH in triggering L2 gains (Freed, Segalowitz and Dewey 2004; Authors 2011, 2014). However, this research is still very scarce, and to the best of our knowledge has not yet investigated L2 pronunciation, which is the objective of the present study.

2.2 The effects of SA on L2 phonological acquisition

Previous research investigating the effect of SA on L2 phonology is relatively limited and has so far produced inconclusive results. The findings of previous research point in several directions, some have found improvement in favor of the SA group (Author1 2014; Højen 2003; Stevens 2011), others have found differences in favor of the AH group (Author2 2008; Author2 2014), and still others have found no (clear) differences between SA and AH participants (Avello and Lara 2014; Avello et al. 2014; Díaz-Campos 2004; George 2014) for a variety of target phonological features such as vowels (Author2 2008; Simões 1996; Stevens 2011), consonants (Díaz-Campos 2004; George 2014; Stevens 2001), or perceived foreign accent (Author1 2014; Avello et al. 2012).

~~This~~ These unclear results are probably due to the inherent difficulty of assessing gains over short periods of time in L2 phonological development. At the segmental level, improving the production accuracy of difficult L2 vowels and consonants is dependent upon learners developing accurate representations for L2

sounds (Flege 1995), and such changes may take longer and may need larger amounts of authentic input than that typically provided by short-term SA periods (ranging from a few weeks to several months). Besides, measuring segmental improvement in perception and production requires selecting specific target L2 contrasts or dimensions which, at best, can represent only one choice over many other perceptual and production dimensions across which phonological acquisition might be measured. Thus, for example, a study may choose to focus on L2 vowel production and fail to detect gains in accuracy and for the same participants find improvement in laryngeal timing (VOT) (e.g. Avello and Lara 2014). That is, despite the need to focus on phonetic dimensions specific to a given L1-L2 combination, the choice of dimension tends to be motivated by predictions regarding cross-language differences presenting problems for acquisition rather than predictions based on what the L2 learning context may offer in terms of phonological development. A common finding of SA research investigating phonetic accuracy in L2 speech perception and production is that certain aspects improve and others do not (Díaz-Campos 2004; Højen 2003), or they improve only for some learners; for example, those with a lower pre-SA level of proficiency (Author 2014), and not for others (Simões 1995). Concerning the methodology to investigate L2 phonological development in SA, most studies have used foreign accent (FA) ratings as a means of assessing overall improvement in pronunciation accuracy (Avello, Mora and Pérez-Vidal 2012; Author 2014). In general, research examining specific aspects of phonetic accuracy is less frequent and has focused primarily on segmental production (Avello and Lara 2014; George 2014; Stevens 2001, 2011) rather than on segmental perception (Author2 2014).

However, whether the measure of phonological development is one of phonetic accuracy in segmental production or one based on overall FA ratings, results are mixed,

with some studies reporting improvement during SA (Author1 2014; Stevens 2001, 2011) and others not (Díaz-Campos 2004; Author2 2008). The range of populations used (children vs. adults), the variety of perceptual and productive dimensions examined (e.g. segmental discrimination, VOT in stop consonant production, overall FA ratings, among others), the elicitation instruments used (reading aloud, sentence repetition), the length of the SA period (from 3 weeks to 12 months) and the characteristics of the AH comparison group (regular FL class, intensive course), are all factors likely to explain the mixed results obtained so far.

Considering the previously mentioned gaps in the SA literature, the aim of the present study is to examine the effects of a very short sojourn (18 days) on L2 pronunciation and do so including an immersion AH comparison group. No studies to date have compared phonological development in these two contexts, despite the fact that short-term immersion (both in the SA and in the AH context) is quite frequent, especially for younger learners. The research questions that guide this study are the following:

1. Do learners of English improve their L2 pronunciation (as assessed by VOT and FA measures) differentially as a function of learning context: a short study abroad (SA) experience in an L2-speaking environment vs. an intensive course at home (AH)?
2. To what extent are learners' changes in VOT related to raters' perception of FA?

3. Method

The present study assessed learners' changes in pronunciation through VOT and FA measures in sentences elicited through a delayed sentence imitation task. VOT, or the period of voicelessness between the release of an unvoiced stop consonant (/p, t, k/)

and the onset of voicing of the following vowel, is characteristically very short in Spanish (0-15 msec) and relatively long in English (30-90 msec; e.g. Yavas and Wildermuth 2006), so learners' VOT was expected to increase over time as their pronunciation gets closer to English. Acoustic measurements of VOT were performed manually from time-amplitude waveforms and spectrograms in Praat. FA measures were obtained by averaging the FA ratings of a panel of 27 native English raters who judged the degree of FA in the learners' sentences on a 7-point scale, which was expected to diminish over time.

3.1 Participants

3.1.1 L2 Learners

The L2-English learners in the present study were 36 Catalan/Spanish bilinguals differing in the type of learning experience they engaged in. Fourteen learners (n= 4 males, n= 10 females) engaged in a 3-week SA experience in the south of England, whereas 22 learners (n= 14 males, n= 8 females¹) enrolled in a one-month intensive English course in their home country (AH). The mean age in the SA context was 15.14 (standard deviation 1.35), with students' ages ranging from 12 to 17. In the AH context the students' ages ranged from 14 to 17, and the mean age was 14.91 (standard deviation 0.97). The proficiency level of the students in the two contexts ranged from intermediate to upper-intermediate, according to the English courses they were enrolled in. Also, most of the learners in the two contexts had been learning English at school since they were 6 years old or younger (66.5% in the SA context and 76,2% in the AH context). Additionally, the majority of the learners in the SA context (70%) had had

¹ It must be pointed out that, despite differences in the gender profile of the two groups, statistical analyses showed that gender was not an influential variable in the results. Similarly, other characteristics associated with differences between males and females (e.g., motivation) were comparable between groups.

previous SA experiences; and similarly, a high percentage of the learners in the AH program (57%) had been previously enrolled in an intensive summer course.

Once in England, SA learners were administered a placement test and were distributed into classes according to their level. They attended classes for 3 hours a day 5 days a week, where they were mixed with other international students. All of their teachers were native speakers of British English and so were the monitors who were in charge of the afternoon activities that the participants attended for 2 hours a day. These activities were varied and included sports or arts and crafts. Although there were no planned activities on the weekends, participants were encouraged to interact with other international students. Regarding accommodation type, six students stayed with a family, three in a student dorm with other international students, and five experienced both types of accommodation (they switched either from dorm to family or the other way round).

As for AH learners, the intensive EFL course took place in the participants' home country (Spain). This course was offered in a private institution and all the teachers were native speakers of British English. The course lasted for four weeks and learners received a total of 90 hours of instruction in English (4.5 hours a day). The students were also placed in groups depending on their initial proficiency level. The classes followed a textbook and, apart from the activities included there, the teachers also used communicative activities that promoted interaction. The AH learners lived with their own families and therefore exposure to the L2 mostly came from the intensive course (in addition to the out-of-class hours devoted to homework and to any additional private exposure to the L2 through internet or any social media).

3.1.2 Raters

A group of 27 L1-English speakers (n= 9 males, n= 18 females) aged 19 to 23 was recruited at a university in the UK to act as raters. They auditorily judged a set of sentences produced by the L2 learners for degree of FA on a 7-point FA scale and filled out a questionnaire about their experience with languages and foreign-accented English. None of the raters reported any hearing problem.

Raters were asked about the amount of foreign-accented English they are exposed to in their every-day life, their familiarity with the Spanish/Catalan accent(s), their teaching experience with Catalan/Spanish native speakers, and difficulty to understand Spanish/Catalan-accented English. As a group, raters had an average exposure to foreign accented English in their every-day life (6 were always exposed to accented English, 8 were often exposed to accented-English, 7 reported being sometimes exposed to accented-English, and 6 were rarely exposed to accented English), and had very little experience teaching English to Catalan/Spanish NSs (18 had never taught Catalan/Spanish speakers, whereas 9 had done so). Moreover, raters reported being overall relatively unfamiliar with the Spanish accent ($M= 3.15$, $SD=1.89$) on a 7-point Likert familiarity scale (1= very familiar; 7= very unfamiliar), and found Spanish-accented English relatively easy to understand ($M= 2.7$, $SD=1.06$) on a 7-point Likert comprehensibility scale (1= very easy to understand; 7= very difficult to understand).

3.2 Materials, instruments and procedure

3.2.1 Delayed sentence imitation task

3.2.1.1 Procedure

In order to assess changes in the learners' pronunciation, we elicited speech samples through a delayed sentence imitation task that learners performed twice, at the

beginning (more specifically, the second day), and the penultimate day of their respective program. The time elapsed between the pre- and post-test was 18 days for both learner groups.

3.2.1.2 Instrument

In the delayed sentence imitation task, learners were auditorily presented with 20 different sentences (once only) for imitation spoken by a native speaker of English (see Appendix A). The sentences were based on a picture story (Heaton 1966). The target sentences varied in length, ranging from five to eight syllables and included a variety of phonetic and phonological features that were especially challenging for Catalan/Spanish native speakers, such as word-initial and word-final consonant clusters disallowed by Spanish phonotactics (e.g. /str-/ in *strong*), contrasting vowel phonemes that tend to be assimilated to single Spanish native vowel categories (*beat* /bi:t/ vs. *bit* /bit/ for Spanish /i/), English aspirated unvoiced stops, or English phonemes that do not exist in Spanish/Catalan (e.g., [h], or [v]) or that are differently distributed (e.g. [r]). Before performing the task, the learners were allowed to become familiar with the story that the sentences were based on for 20 seconds and then the pictures were taken away. In the instructions (which were in the students' L1) learners were asked to repeat the sentences as accurately as possible. One second after each sentence there was a beep sound, and the students had 6 seconds to repeat the sentence. It is considered that the time that had elapsed between the target sentence and the learner's repetition (2 seconds) together with the intervening beep sound would erase any previous sensory memory trace. In order to repeat the sentence the participants had to "remember" the original sentence and then repeat it, which means that the repetition they produced came from their long-term memory representations for the words and sequences of sounds they repeated. In other words, it was hoped that this delayed elicitation procedure would prevent learners

from repeating the sentences directly from sensory memory, so that the learners' pronunciation would more closely reflect their underlying L2 phonological representations. Before the target sentences were presented, the learners listened and repeated three trial sentences so that they became familiar with the procedure. This task had a total duration of 5 minutes (including the trial phase and the instructions) and was administered individually. Participants' productions were digitally recorded on Marantz PDM620 solid-state recorders using their built-in microphones at 44.1 KHz (16 bit) for subsequent acoustic analyses.

3.2.2 VOT measures

VOT was measured from the onset of the release burst of the stop consonant (visible as an amplitude peak in the waveform) to the first positive peak of periodic energy indicating the onset of vocal cord vibration for the following vowel. We were expecting L2 exposure and instruction to result in learners' production of English unvoiced stops with increased VOT durations. A mean VOT measure per learner was obtained by averaging, separately for each one of the two testing times, the VOT of initial stops in words selected from 5 of the 20 sentences each one of the participants had recorded (*Pete, Paul, put, time, tall, kicked*). The sentences were the following: *The tall boys are playing football; My friend Pete kicked the ball; Pete and Paul are very strong; Rick put his arm in the hole; We waited for a long time*. These sentences were selected because they contained unvoiced stops in addition to other features that were challenging for the population under study (see previous section).

3.2.3 Foreign accent rating task

Given the FA task duration constraints imposed by the number of learners, only 3 out of the 5 sentences selected for the VOT measures were included in the rating task

(*My friend Pete kicked the ball, Pete and Paul are very strong, We waited for a long time*). The most noticeable features that were likely to trigger an FA in these sentences were the lack of contrast between the tense-lax vowel pairs /i:/-/ɪ/ (*Pete-kicked*) and /ɔ:/-/ɒ/ (*Paul-long*), as these are likely to assimilate to the single Spanish vowel categories /i/ and /o/, the pronunciation of the simple past allomorphs /t/ and /ɪd/ in *kicked* and *waited* (especially because of the word-final cluster in the former), the VOT of the unvoiced stops (*Pete, Paul, kicked, and time*) and the vowel reduction processes in the weak form of grammatical words such as those of *and* and *for*, as Spanish does not reduce vowels in the unstressed position.

The FA rating task was designed in *Praat* (Boersma and Weenink 2013) and contained 3 sentences from each learner recorded at the pre- and post-test (3 sentences x 2 testing times x 37 learners = 222 sentences). Before presentation to the raters, the sentences were normalized for peak and mean amplitude and were also low-pass filtered (50Hz) to eliminate any low-frequency noise. Three tokens of each one of the three sentences varying in strength of accent were first presented to listeners in a short practice session to familiarize them with the procedure and the range of accents they would hear. Raters were instructed to assess the sentences for degree of FA by first clicking on the sentence they had heard (which was always visible on the screen) and then by clicking on one of the numbers on a 7-point scale appearing on the screen (1= No Foreign Accent / 7= Very Strong Foreign Accent). After the practice phase, the listeners rated the sentences in 3 separate blocks, one per sentence, each containing 72 fully randomized trials (36 learners x 2 testing times). Listeners were instructed to listen specifically for FA features and to ignore repetition accuracy errors or dysfluencies.

The consistency of listeners' ratings was assessed through intra-class correlation coefficients. Cronbach's alpha was high (.919), indicating a high level of agreement

among raters. Consequently, a single FA score from each learner was obtained by first averaging, separately for each one of the two testing times, the 27 ratings per sentence, and then by averaging the mean ratings of the three sentences.

3.2.4. Questionnaires and observations

In order to obtain information regarding the students' experiences in the two contexts (mostly about language use in and out of class), as well as some background characteristics, we designed questionnaires for the students to fill in at the end of their respective programs. Moreover, we also spent one day observing the learners in the two contexts: in the class and breaks in the AH context; and in class, breaks, the dining room and during an afternoon activity in the SA context. We also informally interviewed some SA participants while we were together in the dining room. The information obtained from these instruments will be used in the discussion to interpret our findings.

4. Results

Overall the results show that both learner groups improved their pronunciation after their instruction periods, both for VOT and degree of FA (see Table 21)². At pre-test, the SA group appeared to produce unvoiced stops with shorter VOT and was judged to have a slightly weaker FA than the AH group, the difference being significant in the former case, (VOT: $t(34) = -2.14$, $p = .039$ for VOT) but not in the latter (FA: $t(34) = 0.291$, $p = .773$) (see Figure 1). The size of the gains computed for the SA and AH learner groups by subtracting pre-test from post-test scores shows a substantially larger improvement in both VOT and FA for the SA group than for the AH group, that

² Before conducting any inferential statistical analyses, normality of distribution was checked for all the variables with no violations noted.

is, VOT increases indicating a greater degree of aspiration, whereas FA scores decrease, indicating a reduction in degree of perceived FA. This suggests that despite the relatively large inter-learner variation, the SA context appeared to have a stronger impact on learners' pronunciation.

TABLE 21

FIGURE 1

In order to assess the differential contribution of the learning contexts to the improvements in pronunciation observed (RQ1), we submitted learners' VOT durations and FA ratings to an Analysis of Covariance (ANCOVA) because of initial differences between the two groups in the pre-test, with *Learner Group* (SA vs. AH) as fixed factor, the post-test VOT scores as dependent variable, and the pre-test VOT scores as covariate. The results of the ANCOVA suggest that the learners in the SA context significantly improved their VOT to a larger extent than those in the AH context did ($F(1,33)= 7.08, p= .012, \eta^2= .177$). We explored the results further by comparing pre-test VOT to post-test VOT separately for each learner group through paired-samples *t*-tests. These analyses showed that the SA group had improved significantly from pre-test to post-test ($t(13)= -4.473, p= .001, r= .546$), producing longer, more target-like VOT durations (the effect of the difference being large), while the AH group had not ($t(21)= -.741, p= .467, r= .04$). Therefore, we can conclude that for the VOT measure, the SA learning context made a significantly stronger contribution to the improvement of VOT than the AH learning context did.

In order to examine differences in the degree of perceived FA, we performed another ANCOVA with *Learner Group* (SA vs. AH) as the fixed factor, the post-test FA scores as dependent variable, and the pre-test FA scores as covariate. In this case,

there was no significant difference between learner groups: ($F(1,33)= 1.38, p= .248, \eta^2=.040$). Next, we compared gains from pre- to post-test in degree of FA for the two groups separately. The results of the paired-samples t -tests suggest that while the difference between pre-test and post-test FA scores approached significance for the SA group ($t(13)= 1.945, p= .074, r= .159$), it was well beyond significance for the AH group ($t(21)= 1.06, p= .326, r= .081$). These results suggest that, although there were no significant differences between the two contexts in FA scores at post-test, the learners in the SA context reduced their degree of FA to a larger extent than those in the AH context.

Next, given that SA learners produced significantly more target-like VOTs at post-test and both SA and AH learner groups had reduced their FA, we explored the extent to which VOT and FA scores might be related to one another (RQ2). This was done in order to assess the extent to which judges might have been sensitive to differences in VOT, a characteristic cross-language difference in the production of unvoiced oral stops in English and Spanish, when judging degree of FA. In order to do so, we ran *Pearson-r* correlations between the learners' mean VOT durations and their perceived FA ratings both at pre-test and post-test. These analyses showed that the raters' degree of perceived FA in the speech samples was unrelated to how accurately learners produced unvoiced oral stops (all $rs < .15$ and non-significant). This indicates that raters did not base their accentedness judgements on the learners' production of oral stops alone, but were also likely to have perceived inter-subject differences in pronunciation accuracy based on other phonetic dimensions known to be difficult for Spanish learners of English, such as vowel quality and duration, vowel reduction processes and rhythmic structure, or linking phenomena across word boundaries. Since such phonetic dimensions may be independent of VOT, raters' judgements of

accentedness are likely to be unrelated to inter-learner differences in VOT. It is beyond the scope of the present study to acoustically measure sources of accentedness other than VOT, a task we would like to undertake in future research. However, the AH and SA periods may have affected pronunciation accuracy in VOT, as well as other dimensions of pronunciation accuracy, so that learners may have improved in pronunciation accuracy in general, including VOT. We explored this possibility by correlating gains in VOT with gains in FA. VOT gains were significantly but only weakly correlated with the amount of FA reduction ($r = .35$, $p = .034$), suggesting that, despite considerable inter-subject variation, there was a trend among learners who had obtained a larger increase in VOT durations to be also more likely to reduce their degree of FA to a larger extent (Figure 2). That is, VOT durations were not directly related to degree of FA, but learners achieving greater FA reduction also achieved greater gains in VOT. Interestingly, individual data analyses show that the proportion of learners who improved their VOT and obtained benefits in overall FA reduction was much larger in the SA context (92.8% and 71.4%, respectively) than in the AH context (63.6% and 54.5%, respectively), further confirming the group analyses reported above and suggesting that the SA context was superior to the AH context in improving L2 pronunciation.

FIGURE 2

5. Discussion and conclusion

The results of the present study have shown that the SA context (but not the AH context) contributed significantly to a more target-like production of oral stops and resulted in a more substantial reduction of FA in the learners' L2 pronunciation. In addition, learners who had improved the most in VOT at post-test also showed larger

improvements in FA, suggesting that changes in pronunciation resulting in overall FA reduction are likely to originate in more specific changes in production accuracy at the segmental (such as VOT) and supra-segmental levels.

The differential effect of SA and AH contexts on VOT may be explained on the basis of the greater sensitivity of a specific phonetic feature such as VOT to the type of massive exposure and production practice in the L2 characteristic of SA contexts. The present findings provide further evidence of the positive effects of SA on the segmental phonology of English, supporting previous research showing that VOT can be very sensitive to context of use and consequentially, gains are likely to be obtained during SA (e.g. Author2 2008). One may also consider the possibility that the native English teachers in the AH context might have developed a slightly more Spanish-like VOT through extensive exposure and use of Spanish than the SA teachers. However, the amount of VOT duration shortening attributable to L2 exposure and use in their L1-English is unlikely to be large enough to explain the differences in VOT gains between the SA and AH learner groups observed in the present study (see Amengual 2012).

Foreign accent, however, being a more holistic measure of L2 pronunciation accuracy based on judges' ratings, may be more resistant to showing large gains through short-term immersion. There are at least two possible reasons for this.

First, L2 phonological development in general, and during SA in particular, is characterized by large inter-learner variation (Author1 et al. 2012; Kinginger 2008) with some learners showing significant progress, while others not showing any improvement at all (Avello et al. 2012). In the present study, raters failed to detect any FA reduction in 28.6% of the SA learners despite their improvement in VOT. Secondly, overall FA in L2 speech arises as the consequence of a myriad of phonetic and phonological

deviations from the canonical segmental and supra-segmental properties of native speech, and it is therefore expected that learners should proceed through phonological development gradually by first acquiring certain L2 features and then others. Consequently, changes in L2 pronunciation accuracy affecting properties of specific L2 sounds may not translate into a change in overall degree of perceived FA. However, the larger improvement in FA reduction observed in the SA learner group is in line with the findings for VOT. This suggests, given the short length of the SA period examined in the present study, that the SA context is potentially a stronger contributor to L2 pronunciation development than the AH context.

One may speculate that the SA context provides many more opportunities for interaction with native speakers and greater exposure to authentic L2 input than the AH context (despite the individual variation found by previous research; i.e. Kinginger 2008; Dewey et al. 2012). The AH learners had 4.5 hours of class, including three breaks (10, 20, and 10 minutes respectively), and out-of-class practice was mostly limited to doing homework (85% of students reported so in the questionnaire). In our observations, we also confirmed that not all students used the L2 in class, and that almost no one used English during the breaks. On the other hand, the SA learners had at least 5 hours of monitored L2 exposure: 3 hours of class and 2 hours of afternoon activities. Additionally, they also used English at home (whether it was in a family or in a residence hall together with other international students), and they were exposed to the target language each time they were in public places such as the street, shops or public transport. Thus, a higher exposure to the L2 might have lead to higher opportunities to practice it.

Nevertheless, despite the wider range of opportunities to practice the L2 abroad, not all SA learners improved their pronunciation. In fact, when analysing the students'

self-reports, it became evident that not all learners used the target language to the same degree. Although English use was standard in the English class and also for most learners in the afternoon activities (which we could also confirm through our observations), there was high variability of self-reported use during weekends or free time during the week (5 students used it little or very little, 4 used it sometimes, and 4 often/almost always³). Additionally, those participants staying with a family at some point reported to use more English in their free time than those in a residence hall. In the informal interviews we performed during lunch with some of the SA students, they confirmed that those staying with a family used English more than those staying in dorms. In fact, some students said they switched accommodation type because of this. Considering overall language use during the stay, 5 participants claimed to have used their L1 more than English, 6 claimed to have used both to the same degree, and only 2 claimed to have used English more frequently than their native language. Finally, when asked about the nationality of the people they interacted with during their stay, out of the 12 students who responded, 7 used English more often with international people and 5 with English native speakers. It could also be observed that those who used English more often with native English speakers obtained greater VOT gains and FA gains than those who spoke more often with other international people (mean VOT gains = 15.74 vs. 9.04; mean FA gains = 0.55 vs. 0.08). Although these differences are not statistically significant (probably due to the small sample size in the two groups), they certainly show a tendency, especially in the case of VOT.

As can be inferred from the questionnaire data, students present great variability of profiles regarding L2 use abroad. Since we only have 14 participants in the SA

³ It must be noted that not all the students answered all the questions; that is why there are some cases where the *n* is lower than 14.

context, we could not perform inferential statistics with these variables as independent predictors of pronunciation gains. However, the information from the questionnaire, interviews and our observations can be used to illustrate that learners had different experiences while abroad, and, moreover, that a high proportion of them interacted mostly with international English speakers and not with native English speakers, which could arguably be helpful to develop other L2 areas, but possibly not pronunciation.

It should also be pointed out that experience-related factors such as amount and quality of L2 input and use are not the only factors potentially contributing to L2 phonological development. Individual differences in aptitude-related factors, especially those enhancing cross-language perception skills and L2 phonological processing, may interact with learning context features and explain both the large individual variation in L2 pronunciation attainment typically associated with SA contexts and the differential pronunciation gains obtained by learners in SA and AH contexts. One such set of factors, poorly understood at present and largely under-researched to date in the domain of L2 phonology, include memory- and attention-related cognitive skills, such as phonological short-term memory, that have already proved influential in the development of grammatical skills (French and O'Brien 2008) and oral fluency (O'Brien et al. 2006, 2007). Few studies have examined the role of cognitive skills in L2 phonological development, and these have mainly investigated phonological memory (Aliaga-García, Mora and Cerviño-Povedano 2011; MacKay, Meador and Flege 2001; Cerviño-Povedano and Mora 2011), attention (Darcy, Mora and Daidone 2014; Guion and Pederson 2007) and inhibitory control (Darcy et al. 2016; Lev-Ari and Peperkamp 2013) with learner populations acquiring the L2 either through long-term immersion or in the foreign language classroom. Some studies exist that have examined these cognitive skills in an SA context. For example, Lord (2006) examined the role of

phonological memory in oral mimicry, García-Amaya (2012) on attention in L2 oral fluency, and Linck, Hoshino and Kroll (2008) on inhibition in language comprehension and production tasks. However, no studies to date have directly investigated the differential role of cognitive skills in SA and AH learning contexts for L2 phonological development. Future studies should address the role of individual differences in cognitive ability in L2 pronunciation development as they interact with the different input and learning conditions of SA and AH contexts. Such lines of research should be able to shed some light on the large inter-learner variation in L2 phonological attainment found in both contexts and the advantage of an SA context over an AH context for L2 pronunciation development.

To conclude, the present study revealed that a short SA experience has a positive impact on the L2 pronunciation development of a group of adolescents, especially for oral stops. It was also found that those participants showing greater improvement in VOT reduced their FA to a greater extent. However, these results must be interpreted with caution, as there was large inter-subject variability and this made it hard for us to test the contribution of experience-related factors (L2 use during free time, nationality of frequent interlocutors, type of accommodation, etc.) to the development of L2 pronunciation. Future studies should try to reduce the variability by focusing on one or two factors or should consider gathering larger sample sizes for inferential statistics to be performed considering multiple variables. Second, data extracted through the questionnaires are self-reported, which allows for a subjective appraisal of reality. Despite its limitations, the present study makes an important contribution to the field, as it is the first one to provide evidence for L2 pronunciation development in a very short SA (3 weeks) in the case of adolescents. This finding is in line with previous research reporting benefits of an overseas experience on L2 pronunciation in adults (Avello et al.

2012; Stevens 2001; 2011) and children (Author1 2014). However, comparison across studies is difficult due to differences in participants' age, proficiency level, elicitation methods and the measures used. For example, Avello et al. 2012 found a significant decrease in the number of pronunciation errors learners produced after SA, but only a slight non-significant reduction of FA based on a 7-point FA scale. Similarly, in Avello and Lara (2014) the FA reduced non-significantly on a 7-point scale (-0.23). In both these studies the participants were more advanced than the ones in the present study, they were older (university students) and spent a longer time abroad (3-months). In a similar study with an analogous population Mora (2008) also found a slight non-significant improvement between pre-SA VOT (41ms) and post-SA VOT (43ms). However, in these studies the average VOT ranged between 40-50ms and the task used to elicit the oral stop productions was a reading aloud task. In the present study learners appeared to be less target-like in their VOT production (20-30ms), improved considerably more on average (12ms), and the speech samples were elicited through a delayed imitation task. The less-target like VOT of the participants in the present study might have proved more sensitive to experiential factors during SA, as participants would have greater room for improvement. The delayed elicitation technique used might also have contributed to capturing improvement in VOT better than the reading aloud task used in other studies. In sum, the amount of pronunciation improvement and VOT gains obtained by different participants across studies is likely to depend on both the participants' L2 level and the measures and methods used to elicit their speech.

The present study has important pedagogical implications. The greater benefits in L2 pronunciation obtained in the present study by the SA group suggest that a formal instruction context (despite it being intensive) may not provide enough input quantity and output practice for general observable benefits to occur in pronunciation (Piske

2007). However, some remedial research-based pedagogical interventions are available that might help learners overcome the limitations of the AH context in enhancing L2 pronunciation (e.g. Thomson 2012). Several decades of lab research on phonetic training in L2 speech perception and production (e.g. Iverson and Evans 2009; Pisoni and Lively 1995) suggest that focused training on areas of special perceptual and productive difficulty may lead to robust long-lasting improvements in pronunciation in a relatively short period of time. Further research is needed in order to confirm whether training on certain L2 pronunciation areas AH might compensate for the lack of exposure and practice opportunities that are more typical of an SA context.

The results of the present study appear to add to the accumulated evidence that SA is superior to AH learning contexts in effecting gains in oral production skills, especially in the domain of L2 pronunciation. This would suggest, following well-established findings, that language acquisition, and pronunciation in particular, benefits from large amounts high quality input and practice (Flege, 2009). The SA tends to offer greater opportunities for such input and practice conditions to occur. However, whereas the AH context offers less opportunities for practice and massive exposure to authentic input, it appears to be more homogeneous and controlled in what it can offer than the SA context, where typically researchers have less control over the L2 usage variables that might be influencing oral performance over time. Thus, the large inter-learner variation reported in SA research as regards learners' amount of engagement in practice speaking activities providing L2 use in meaningful communicative situations (with native speakers and other L2 speakers) might partly explain the usually large variation found in the gains obtained through SA.

Finally, our results suggest that both learning contexts under study (SA and intensive courses AH) contribute to a betterment of L2 pronunciation, but that the SA

context was slightly more beneficial than the AH one. Previous studies comparing SA with AH programs with an equivalent period of exposure to the L2 have found no differences between these learning contexts (Serrano et al., 2011; Serrano et al., 2014), but none of these studies focused on L2 pronunciation. It could be the case that the aspects of pronunciation examined in this study are more sensitive to gains after a stay abroad than the L2 aspects examined in the previously mentioned studies (global measures of fluency, complexity and accuracy). However, in view of the contradictory results reported in the literature on SA, further studies are needed in order to determine what learning context is more beneficial for L2 learning in general and L2 pronunciation in particular.

References

Author1

Author2

Author3

Aliaga-García, C., J. C. Mora, and E. Cerviño-Povedano (2011) Phonological short-term memory and L2 speech learning in adulthood. *Poznań Studies in Contemporary Linguistics* 47: 1-14.

Allen, H. W. (2010) Language-learning motivation during short-term study abroad: An activity theory perspective. *Foreign Language Annals* 43: 27–49.

Amengual, M. (2012) Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition* 15: 517-530.

- Arteaga, D. L. (2000) Articulatory phonetics in the first-year classroom. *Modern Language Journal*, 84: 339–354.
- Avello, P., J. C. Mora, and C. Pérez-Vidal (2012) Perception on FA by non-native listeners in a study abroad context. *Research in Language* 10: 63–78.
- ___ and R. Lara (2014) Phonological development in L2 speech production during study abroad programmes differing in length of stay. In C. Pérez-Vidal, (ed.) *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins. 137-165.
- Barron, A. (2006) Requesting in Irish English and English English: A study of intra-cultural regional pragmatic variation. *LAUD paper 684*, Essen: Universität Essen.
- Beatty, J., M. Valls-Ferrer, and C. Pérez-Vidal (2014) Listening performance and onset level in formal instruction and study abroad. In C. Pérez-Vidal (ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts* (pp. 195-216). Amsterdam: John Benjamins.
- Boersma, P., and D. Weenink (2013) *Praat: doing phonetics by computer* [Computer program]. Version 5.3.52, retrieved from <http://www.praat.org/>.
- Cerviño-Povedano, E., and J. C. Mora (2011) Investigating Catalan learners of English over-reliance on duration: vowel cue weighting and phonological short-term memory. In K. Dziubalska-Kołaczyk, M. Wrembel, and M. Kul (eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010* (pp. 56-64). Volume I. Frankfurt am Main: Peter Lang.
- Collentine, J. (2004) The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition* 26: 227–248.

- ____ (2009) Study abroad research: Findings, implications and future Directions. In C. J. Doughty and M. H. Long (eds.), *Handbook of language teaching* (pp. 218-233.). Malden, MA: Blackwell.
- Collins, L., R. H. Halter, P. Lightbown, and N. Spada (1999) Time and the distribution of time in L2 instruction. *TESOL Quarterly* 33: 655–680.
- ____ and J. White (2011). An intensive look at intensity and language learning. *TESOL Quarterly* 45: 106-133.
- Cubillos, J. H., L. Chieffo, and C. Fan (2008) The impact of short-term study abroad programs on L2 listening comprehension skills. *Foreign Language Annals* 41: 157–185.
- Darcy, I., J. C. Mora, and D. Daidone (2014) Attention control and inhibition influence phonological development in a second language. In *Proceedings of the International Symposium on the Acquisition of Second Language Speech. Concordia Working Papers in Applied Linguistics*, 5, 115-129.
- ____ (2016). [The role of inhibitory control in second language phonological processing. *Language Learning*, 66 \(3\).](#)
- DeKeyser, R. M. (1991) Foreign language development during a semester abroad. In B. F. Freed (ed.), *Foreign language acquisition research and the classroom* (pp. 104-119.). Lexington, MA: D. C. Heath.
- Dewey, D. P. (2004) A comparison of reading development by learners of Japanese in intensive and domestic immersion and study abroad contexts. *Studies in Second Language Acquisition* 26: 303–327.
- ____ (2008) Japanese vocabulary acquisition by learners in three contexts. *Frontiers: The Interdisciplinary Journal of Study Abroad* 15: 127–148.

- ____ J. Bown, and J. Eggett, (2012) Japanese Language Proficiency, Social Networking, and Language Use during Study Abroad: Learners' Perspectives. *The Canadian Modern Language Review* 68.2: 111-137.
- Díaz-Campos, M. (2004) Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition* 26: 249-273.
- Dyson, P. (1988) *The year abroad*. Report for the Central Bureau for Educational Visits and Exchanges, Oxford University Language Teaching Centre.
- Flege, J. E. (1995) Second language speech learning: Theory, findings and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229-273). Timonium, MD: York Press.
- ____ (2009) Give input a chance! In T. Piske and M. Young-Scholten (eds) *Input Matters in SLA* (pp. 175-190). Bristol: Multilingual Matters.
- Freed, B. F. (1995) What makes us think that students who study abroad become Fluent? In B. F. Freed (ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam: John Benjamins.
- Foster, P. (2009) Lexical diversity and native-like selection: The bonus of studying Abroad. In B. Richards et al., (Eds.), *Vocabulary studies in first and second language acquisition* (pp. 91-106). Hampshire, UK: Palgrave Macmillan.
- Freed, B. F., S. So, and N. A. Lazar (2003) Language learning abroad: How do gains in written fluency compare with gains in oral fluency in French as a second language? *ADFL Bulletin* 34: 34-40.
- Freed, B. F., N. Segalowitz, and D. P. Dewey (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.

French, L. M., and I. O'Brien (2008) Phonological memory and children's second language grammar learning. *Applied Psycholinguistics* 29: 463-487.

García-Amaya, L. (2012) Second Language Fluency and Cognition: The Study of Spanish Second Language Development in an Overseas Immersion Program and an At-Home Foreign Language Classroom. Unpublished PhD Thesis. Department of Spanish and Portuguese, Indiana University.

Guion, S.G., and E. Pederson (2007) Investigating the role of attention in phonetic learning. In O.-S. Bohn and M. Munro (eds.), *Language Experience in Second Language Speech Learning* (pp. 57-77). Amsterdam: John Benjamins.

Heaton, J. B. (1966) *Composition through pictures*. Essex, UK: Longman.

Højen, A. (2003) Second language speech perception and production in adult learners before and after short-term immersion. Unpublished PhD Thesis. University of Aarhus, Aarhus, Denmark.

Kinging, C. (2008) Language learning in study abroad: Case studies of Americans in France. *Modern Language Journal Monograph Series*, volume 1.

Lara, A. R., J. C. Mora, and C. Pérez-Vidal (2015) How long is long enough? L2 English development through study abroad programmes varying in duration. *Innovation in Language Learning and Teaching* 9.1: 46-57.

Lev-Ari, S., and S. Peperkamp (2013) Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics* 41: 320-331.

- Linck, J. A., N. Hoshino, and J. F. Kroll (2008) Cross-language lexical processes and inhibitory control. *The mental lexicon* 3: 349.
- Lord, G. (2005) (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course, *Hispania*: 557-567.
- ____ (2006) Defining the indefinable: Study abroad and phonological memory abilities. In Carol A. Klee and L. Timothy (eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages*, Somerville, MA: Cascadilla Proceedings Project.
- MacKay, I. R. A., D. Meador, and J. E. Flege (2001) The identification of English consonants by native speakers of Italian. *Phonetica* 58: 103-125.
- Martinsen, R. (2010) Short-Term Study Abroad: Predicting Changes in Oral Skills. *Foreign Language Annals* 43.3: 504-530.
- Milton, I. and P. Meara (1995) How periods abroad affect vocabulary growth in a foreign language. *ITI Review of Applied Linguistics* 107: 17-34.
- Mora, J. C. (2008). Learning context effects on the acquisition of a second language phonology. In C. Pérez-Vidal, M. Juan-Garau, and A. Bel (eds), *A Portrait of the Young in the New Multilingual Spain* (pp. 241-263). Clevedon: Multilingual Matters.
- O'Brien, I. , N. Segalowitz, J. Collentine, and B. F. Freed (2006) Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics* 27: 377-402.

- ____ N. Segalowitz, B. F. Freed, and J. Collentine (2007) Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition* 29: 557-581.
- Pérez-Vidal, C. and M. Juan-Garau (2009) The effect of study abroad on written Performance. *Eurosla Yearbook* 9: 269–295.
- Piske, T. (2007) Implications of James E. Flege's research for the foreign language classroom. In O-S Bohn and M.J. Munro (eds.), *Language Experience in Second Language Learning* (pp. 301-314). Amsterdam: John Benjamins.
- Pisoni, D. B., and S. E. Lively (1995) Variability and invariance in speech perception: A look at some old problems in perceptual learning. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*. Timonium, MD: York Press.
- Regan, V., M. Howard, and I. Lemée (2009) *The acquisition of sociolinguistic competence in a study abroad context*. Clevedon, UK: Multilingual Matters.
- Sasaki, M. (2009) Changes in English as a foreign language students' writing over 3.5 years: A sociocognitive account. In R. M. Manchón (ed.), *Writing in foreign language contexts: Learning, teaching, and research*. Clevedon, UK: Multilingual Matters.
- 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, 2, 133-143.
- Serrano, R., Tragant, E., & Llanes, À. (2014). Summer English courses abroad vs. "at home". *English Language Teaching Journal*, 68, 4, 397-409.

- Simoes, A. (1996) Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. *Hispania* 79: 87–95.
- Stevens, J. (2001). Study abroad learners' acquisition of the Spanish voiceless stops. *MIFLC Review* 10: 137–151.
- Stevens, J. (2011) Vowel duration in second language Spanish vowels: Study abroad versus at-home learners. *Arizona Working Papers in SLA & Teaching* 18: 77–104.
- Thomson, R. I. (2012). Improving L2 listeners' perception of English vowels: A computer-mediated approach. *Language Learning* 62: 1231-1258.
- Trenchs-Parera, M. (2009). Effects of Formal Instruction and a Stay Abroad on the Acquisition of Native-Like Oral Fluency. *Canadian Modern Language Review* 65.3: 365-393.
- Yavas M., and R. Wildermuth (2006) The effects of place of articulation and vowel height in the acquisition of English aspirated stops by Spanish speakers. *International Review of Applied Linguistics in Language Teaching* 44: 251–263.

Appendix A Elicited Imitation Task (in bold are the sentences that were used for the present study)

1. **The tall boys are playing football.**
2. **My friend Pete kicked the ball.**
3. John couldn't shoot the ball.
4. **Pete and Paul are very strong.**
5. The ball fell into a hole.
6. The hole was big and deep.
7. **Rick put his arm in the hole.**
8. **but he couldn't get the ball.**
9. There was a snake near the tree.
10. This snake was staring at us.
11. We need to think what to do!
12. **Soon John started to run.**
13. **We waited for a long time.**
14. Good! John is back!
15. The bowl is full of water.
16. What is this water for?
17. The bowl was big and heavy.
18. We put the water into the deep hole.
19. Yes, we could see our ball.
20. We started playing football again.