THE SECOND TIME AROUND: 
THE EFFECT OF FI UPON RETURN FROM SA

Victoria Monje Sangüesa

Màster: Lingüística Teòrica i Aplicada
Edició: 2015-2016
Directors: Dra. Carmen Pérez Vidal  Angelica Carlet
Any de defensa: 2016

Col·lecció: Treballs de fi de màster

Programa oficial de postgrau
"Comunicació lingüística i mediació multilingüe"

Departament de Traducció i Ciències del Llenguatge
Acknowledgements

I would like to thank my two TFM supervisors, Dr. Carmen Pérez-Vidal and Angelica Carlet in the first place. I am grateful to Carmen for having accepted to guide my work and frame it within the SALA project. She bestowed me with the chance of investigating a topic of my entire interest. I especially thank her for having put in me contact with my co-supervisor, Angelica Carlet, to whom I owe most of what I have learnt and who became my main pillar of support during the process of writing my TFM. Not only did she always put her expertise at my service and advised me academically, but she also helped me to cope with pressure whenever I needed it.

I gratefully acknowledge Dr. Joan Carles Mora for having accepted to be the vocal of the tribunal at the defence of my TFM, whose expertise in the field and the goal of meeting his academic expectations have been key in keeping my motivation high during the writing process.

I am also thankful to my parents, Victoria and Valentín, my boyfriend Vipin Singh and my good friend Camila González, who have caringly supported me throughout the process and proved to be patient, caring and understanding as I coped with pressure. Lastly, I cannot forget about some of my classmates, with whom I have shared the joys and sorrows of being a master student.

Victoria Monje

Barcelona, June 2016
Abstract

The present study has aimed at assessing the effects of an 8-week FI period lacking explicit attention to phonology preceded by a 3-month SA term on the VOT production of English voiceless stops in initial stressed position by Catalan/Spanish EFL learners through two tasks. Speech samples were collected from 13 participants upon their return from SA (T1) and immediately after an 8-week period of FI (T2). Three native speakers served as a baseline group. Results failed to yield a significant effect of the FI period preceded by a SA term on informants’ VOTs. Proficiency level was examined too: it was found to play no significant role in the VOT production of participants. The effects of factors reported to influence VOT production such as speaking style, vowel height and place of articulation were also explored. Analyses revealed that they did significantly affect VOT production of voiceless stops for the non-native group. The native showed the same numerical tendency.

Key words

Voice onset time (VOT), L2 phonological acquisition/development, formal instruction (FI), study abroad (SA), L2 production, pronunciation.
TABLE OF CONTENTS

Acknowledgements i
Abstract ii
Table of contents iii
List of figures and tables v

1. Introduction 1

2. Literature Review 3
   2.1. Contexts on focus: FI and SA 3
       2.1.1. Teaching pronunciation in formal instruction 3
       2.1.2. SA on an Erasmus programme 6
   2.2. FI and SA: a research perspective 8
       2.2.1. FI and SA characterised 8
       2.2.2. An overview of research findings 12
   2.3. L2 phonological acquisition 17
       2.3.1. Speech Learning Model (SLM) 17
       2.3.2. The phenomenon investigated: VOT 18

3. Research questions and hypotheses 23
   3.1. Research questions 23
   3.2. Hypotheses 24
       3.2.1. Hypothesis for RQ1 24
       3.2.2. Hypothesis for sub-RQ1.1 25
       3.2.3. Hypothesis for sub-RQ1.2 25
       3.2.4. Hypothesis for sub-RQ1.3 25
       3.2.5. Hypothesis for sub-RQ1.4 25
4. Methodology 26
   4.1. Design 26
   4.2. Participants 27
       4.2.1. Non-native speakers 27
       4.2.2. Native speakers 30
   4.3. Data collection instruments 30
       4.3.1. Carrier sentence task 31
       4.3.2. Text to be read aloud 32
       4.3.3. Carlet-SALA questionnaire 32
       4.3.4. X-Lex / Y-lex vocabulary size test 32
   4.4. Data collection procedure 33
   4.5. Analysis 34

5. Results and discussion 35
   5.1. RQ1 35
       5.1.1. Sub-RQ1.1 37
       5.1.2. Sub-RQ1.2 41
       5.1.3. Sub-RQ1.3 42
       5.1.4. Sub-RQ1.4 44

6. Summary and conclusions 46

References 49
Appendices 61
List of figures and tables

**Figure 1** Phonetic category classification of English and Spanish consonants. Zampini, M. L., & Green, K. P. (2001) in Yayas (2007).

**Figure 2** Proficiency level according to vocabulary size.

**Figure 3** Self-reported language dominance.

**Figure 4** Participants’ language background.

**Figure 5** Participants’ times abroad.

**Figure 6** Mean VOT measurements (ms) at both testing times (T1, T2).

**Figure 7** Mean VOT measurements (ms) as a function of proficiency level.

**Figure 8** Mean VOT measurements (ms) as a function of task (text vs. words).

**Figure 9** Mean VOT measurements (ms) as a function of vowel height averaged across time (T1+T2).

**Figure 10** Mean VOT measurements (ms) as a function of place of articulation averaged across time (T1+T2).

**Table 1** Vocabulary size following common European Framework for Reference: X_Lex Score equivalences.

**Table 2** Mean VOT measurements (ms) at both testing times (T1, T2).

**Table 3** Mean VOT measurements (ms) as a function of proficiency level.

**Table 4** Mean VOT measurements (ms) as a function of task (text vs. words).

**Table 5** Mean VOT measurements (ms) as a function of vowel height averaged across time (T1+T2).

**Table 6** Mean VOT measurements (ms) as a function of place of articulation averaged across time (T1+T2).
1. Introduction

The widespread assumption that living in a country where the target language is spoken necessarily yields positive results for L2 development is closely intertwined with the current popularity that Study Abroad (SA) programmes enjoy. The ERASMUS exchange programme is surely the best-known example in the European context. As backed up by research on SA, the above mentioned assumption does not always hold true, at least not for all language skills indistinctly. Conversely, there exists the widespread idea that Formal Instruction (FI) in an EFL context does not guarantee progress in oral speech development.

In such respect, L2 phonological acquisition warrants particular focus, as it is often assumed that L2 oral speech development will necessarily improve as a function of SA, whereas it will not as a consequence of FI. Moreover, there is a research gap in the domain of L2 phonological acquisition, as investigation has been rather scarce and results conflictive. In turn, phonological development seems to be one of the most challenging aspects of L2 acquisition for learners, although such a lack of success is highly likely to be ascribable to a lack of a consistent pedagogical methodology in the teaching sphere, as backed up by research on oral speech development in FI learning contexts.

Therefore, there is little evidence to support the aforementioned assumption that SA necessarily entails linguistic gains for the learners, namely when it comes to L2 phonological development. The Study Abroad and Language Acquisition (SALA) project is precisely devoted to measuring the impact of a SA on sojourners’ linguistic and non-linguistic progress, including phonological development. This dissertation has been conducted within the framework of this research project (see Pérez-Vidal, 2014 for
a fully-fledged presentation of the project and its scientific results), which has hence bestowed an excellent chance to the author of this TFM to canalise her personal and academic interest in the acquisition of EFL phonology. In addition, it is attempted to reduce the present research gap in the domain of L2 phonological acquisition. The studies which build up the SALA project have mainly looked at the effects of SA periods following FI ones. The present study aims at providing a symmetrical view thereof, which is where the main contribution of this TFM to the SALA project lies.

Specifically, this study focuses on the effects of a FI period preceded by a SA term on L2 production by a group of undergraduate Catalan/Spanish bilinguals. It is intended as a follow-up study based on Mora’s (2008) SALA study, with a similar population, but with a reversed view. Mora (2008) examines the effects of a SA period following a FI period and the subsequent retention effects measured 15 months upon return from abroad. The opposite view is taken here: the effect of an 8-week FI period following a SA period is analysed. As for the specific abilities on focus, Mora (2008) studies L2 production, with voice onset time (VOT) of English voiceless plosive consonants used to measure it. In addition, he also deals with phonemic contrasts to test for perception accuracy. In this investigation, production as tested through VOT durations of English voiceless plosive consonants is examined. As L2 VOT production is the sole phenomenon investigated, an attempt has been made to look at it in as much depth as possible.

This TFM is structured into 6 sections. The introduction opens (section 1), followed by a literature review (section 2). The research questions and hypotheses are then stated in section 3. Section 4 addresses the methodology. Following that, the analysis of results is presented (5). Lastly, the conclusions close the study (6).
2. Literature Review

This section has been divided into the different but intertwined components of the issue and the phenomenon examined in this study: the effects of a period of FI preceded by a SA period on the duration of VOT of voiceless plosive consonants in the case of a group of advanced Catalan/Spanish bilingual EFL learners. The present section is organized in three parts. The first one provides a contextualisation of the learning contexts on focus in the present study (2.1). In the second one (2.2), a characterisation of both learning contexts and an overview of the main research findings in the field are presented. Lastly, the third one (2.3) deals with the phonological side of the work.

2.1. Contexts on focus: FI and SA

The relevant learning contexts for the present investigation are contextualised in this section, which consists of two parts. Firstly, the FI context is analysed (2.1.1), as its potential effects on VOT production are to be tested in the present study. Given that this TFM verses on phonology, pronunciation teaching is the main aspect considered. Since SA is the type of learning context preceding the one on focus, a contextualisation of it is likewise presented (2.2.2).

2.1.1. Teaching pronunciation in formal instruction

In this section, the reasons for the absence of explicit attention to pronunciation in the EFL classroom are explored. EFL learners often find themselves learning their target language in FI contexts which tend to be the sole or main providers of opportunities to practice the target language. This affects of proficiency level, but even more so oral
production, as learners may find opportunities to be exposed to the target language via listening to music, watching films, i.e. listening, or reading books, internet literature, i.e. reading, while speaking, just as much as writing, is often only practiced in class. This section focuses on the extant literature dealing with pronunciation instruction in FI contexts. Sadly but realistically, what such literature is stressing, is the lack of systematic attention paid to the oral dimension of language. The opposite happens with grammar, a domain which has traditionally received much attention in the FI context. In Howard’s (2005: 501) words: “strong emphasis is typically placed on the learners’ grammatical skills during classroom instruction, such that the classroom may facilitate greater feedback on this skills”.

Indeed, pronunciation is often neglected in the ESL classroom despite its importance and interconnection with the four linguistic skills (Darcy, Ewert & Lidster, 2012). It is not the case that explicit instruction on pronunciation is always given to learners in the EFL classroom. For instance, Lord (2010) compared the production of Spanish plosives by a group of English speakers who had received explicit phonetic instruction after a SA period in Mexico with that of another group who did not receive FI. With these results at sight, the author concludes that SA in combination with FI is preferable than either learning context in isolation if significant gains in phonological development are to be accrued. Despite this author’s positive results, similar studies are still scarce.

Moreover, phonology is considered one of the most difficult skills to acquire and develop for Spanish learners of English (Calvo Benzies, 2014). L1 interference, incoherent relation between spelling and pronunciation and other non-linguistic factors such as motivation, age and amount of exposure have been pointed at as factors to which such difficulties are ascribable (Calvo Benzies, 2014; Darcy, Ewert & Lidster, 2012; Wijk, 1996).
Darcy, Ewert & Lidster (2012) conducted a survey of teachers’ practices and beliefs about pronunciation in the Intensive English Program at Indiana University. According to their results, both learner and instruction related factors block pronunciation enhancement. Motivation and fossilization are good examples of learner related factors, whereas lack of real or meaningful interactions and ineffective instruction are instances of instruction related factors.

These authors provide a list of reasons which make the teaching of pronunciation complex. There is no well-established systematic way of deciding what to teach, and when and how to do it (Derwing & Foote, 2011). Practices and purposes are often contradictory. The ordering of essential pronunciation features is unclear (Prator, 1971; Jenner, 1989; Derwing, Munro & Wiebe 1998), teachers are frequently undertrained (Baker, 2014; Derwing, 2010; Foote, Holby & Derwing, 2011; Murphy, 2014) and lack of teaching materials results in low confidence among teachers.

Pronunciation is likewise often neglected from syllabuses, which leads teachers to the belief that spending time on it is unnecessary. There also exists the myth that pronunciation is not for beginners (Zielinski & Yates, 2014), as they are too busy learning new grammar and vocabulary. Darcy, Ewert & Lidster (2012) and Calvo Benzies (2014) emphasize the need that pronunciation should be taught systematically at different levels of proficiency, from beginners to the most advanced learners.

Lastly, Gordon & Darcy (2012) advocate for the usefulness of drawing explicit focus on form in pronunciation instruction, to which Saito & Lyster (2012) add the need of corrective feedback in order for such instruction to become effective. The lack of success in the acquisition of L2 pronunciation might simply be related to the amount of attention it receives in the L2 classroom.
Another line of research that proves the effectiveness of feedback is the one of high variability phonetic training, which provides learners with immediate feedback. Although this issue goes beyond the present work, it cannot be neglected that perceptual training is a good alternative to pronunciation instruction in order to explicitly draw the learners’ attention to L2 relevant cues. The positive results of phonetic training have been shown in several different studies (Aliaga-García & Mora, 2009; Bradlow, 2008; Cebrián & Carlet, 2014; Flege, 1995b; Huensch & Tremblay, 2015; Kartushina et al., 2015; Logan et al., 1991).

To our knowledge (see Pérez-Vidal, 2014), the type of FI our informants were exposed to lacked such attention. They have not received explicit instruction on phonology and pronunciation. It is important to distinguish both kinds, this has been the aim of the present subsection. A situation similar to the one experienced while they were abroad, however the massive exposure and input might counterbalance lack of explicit instruction, an issue further explored in the forthcoming sections.

2.1.2. Study Abroad on an ERASMUS programme

The ERASMUS programme, standing for European Community Action Scheme for the Mobility of University Students, is aimed at favouring academic mobility among university students and professors within the state members of the Economical European Space, Switzerland and Turkey. It is part of the larger Socrates programme, an educational initiative promoted by the European Commission which started in 1987, replaced by Socrates II from 1999. Lifelong Learning is the third phase of the Socrates II programme, to which the Erasmus sub-programme more specifically belongs, together with Comenius, Leonardo da Vinci and Grundtvig, with a budget of 7 million
Euros for the period of 2007-2013 (Pérez-Vidal, 2009). More than 3 million students have benefited from EU Erasmus grants since the exchange scheme’s launch in 1987.

According to The Erasmus impact study regional analysis (2016) undertaken by CHE Consult for the European Commission, the top five reasons to go on Erasmus are the opportunity to live abroad, to improve and widen career prospects in the future, the opportunity to develop soft skills i.e. adaptability, taking initiative, proactivity, the opportunity to learn/improve a foreign language and the opportunity to meet new people. These were considered important by at least 90% of Erasmus students, with only minor differences between the individual regions. In Southern Europe, 9% more students saw mobility as a way to improve and broaden their career prospects than in Western Europe. Similarly, 9% more students in Southern Europe than in Northern Europe considered the opportunity to improve their language skills a compelling reason to go abroad. Improved career prospects is an important reason to go abroad for a 93% of Southern European students.

Southern European students have the largest variety of reasons to go abroad. A large difference between Southern Europe and Northern Europe was especially evident for “opportunity to learn / improve a foreign language”. However, at the same time, non-mobile students in Southern Europe found more barriers not to study abroad. This in part indicates that only the most dedicated students become mobile, as a consequence of relatively high institutional barriers to participation (recognition issues, insufficient access to information and support, financial constraints, etc.). These students can therefore be considered the most curious and convinced of the positive effects of a period abroad (The Erasmus Impact Study: regional analysis, 2016).
2.2. FI and Study Abroad: a research perspective

The present section aims at providing as clear as possible a picture of a FI context following a SA context. In order to do so, it has been divided into two sub-sections. Firstly, FI is presented and contrasted with SA (section 2.2.1). Following that, an overview of the available literature on the effects of FI and SA on oral speech production is presented (2.2.2). Regarding this section, to our knowledge, no previous study has tackled the analysis of the effects of FI following SA, in particular with respect to oral speech development, In this respect, Mora’s (2008) examination of retention effects including the FI period must not be obliterated. However, potential improvement immediately following the FI period was not addressed as it is here. For this reason, the literature on the impact of each of these contexts is revised separately in the sub-sections which follow.

2.2.1. FI and SA characterized

Context of learning has been reported to be one of the most important variables in the study of SLA that might explain some differences in L2 development among learners (Collentine & Freed, 2004; Collentine, 2009; Freed, 1995; Mora, 2008). When SA and FI are contrasted, they are said to differ mainly in the quantity and quality of input they offer to the learner, the SA setting being the one that bestows them with more advantages in that respect. These increased opportunities for interaction and constant exposure ground the assumption that SA is more likely to lead learners to enhanced L2 knowledge than FI. The present section aims at explaining why this is not exactly the case when it comes to phonological development.
The fact that phonology does not usually improve while abroad might be better understood from two ideas: the fossilization phenomenon and the concept of length of residence (LoS). In these respects, Dalton & Seidlhofer (1994: 9) argue that many non-native adults display a high command of their L2, often because they have lived in the target language community for a long time. Despite that, many do still retain a distinct foreign accent revealing phonetic traits of their L1. More specifically, Flege & Frieda (1997) conducted a study with native Italian participants who had immigrated to Canada at early ages. They were found to “speak English with detectable foreign accents even though they began learning English as children and had spoken English for 34 years on average” (1997: 169).

Turning now to the characterisation of SA and FI, Pérez-Vidal (2014) divides them into macro-level features and micro-level features in order to thoroughly account for them. The macro-level features refer to sociolinguistic and psycholinguistic variables such as input, output, interaction and negotiation of meaning. These are held as key factors in SLA research at present for SLA to take place (McDonough & MacKey, 2013). Considering the nature of such factors, SA settings can be described as naturalistic learning contexts in which exposure to both the target language and culture is constant through immersion, which potentially makes available to them massive amounts of input, output and interaction opportunities. If the same factors are applied to FI contexts, one finds that input is poor, and so are the opportunities for authentic interaction. Inevitably then, output practice and the type of learning processes which are fostered by FI are different from those obtained in SA. The micro-level features include the individual variables on the basis of which learners may show significant variation in their abilities to avail themselves of such opportunities for practicing the target language.
The differences between both learning contexts can be better understood along two axes: the sociolinguistic axis and the language acquisition axis. It is the first one which explains that both language and culture are available to the learner in the SA context in contrasting quality and quantity with FI, as the learner can take part in a wide range of situations, speech events adopting diverse social roles (Kasper & Rose, 2002). Along the language acquisition axis, the mechanisms and processes at play in FI and SA differ.

Skill acquisition theory might better account for such differences, as DeKeyser (2007: 214) argues. It distinguishes three stages: declarative knowledge, proceduralization and automatization. SA is an ideal setting to conduct to the third stage. He also spells out that repeated careful and deliberate production of instances of rules is required for practice with declarative knowledge, which can take place in the classroom. However, in order for declarative knowledge to become procedural knowledge, large amounts of practice are not required. In contrast, feedback is. The classroom is still the optimal environment for this to occur. However, the SA enhances the process from (early) proceduralization to automatization, as great many instances of practices are required for this, and the sociolinguistic context while abroad makes it possible. That functional knowledge of the grammar required for automatization to take place is assumed to be mastered at intermediate levels.

As stated by Valls-Ferrer (2011: 3), different linguistic outcomes can be expected for each learning context simply because the linguistic experiences they provide to learners differ. SA spurs the enhancement of certain skills which are normally difficulty to teach in FI (fluency, for instance). The latter, in turn, tends to focus on aspects such as metalinguistic awareness and grammar.
However, there remains a very important aspect to be investigated, as already mentioned above, and Collentine & Freed (2004) emphasize: the quantity and quality of negotiating encounters students avail themselves of while abroad, whereas this is known for FI. Here is when the micro-level features or the learner individual variables come into play. Individual differences on L2 phonological attainment have recently started to attract attention of research. The results indicate that success in L2 production is highly subject to inter-speaker variability due to an interplay of several factors (Mora, 2014). It is learners who seek to engage in interactional encounters who should be in a better position to learn. Attitude and motivation are the most related attributes to that ability (Dörnyei & Skehan, 2003). Similarly, having a high motivation to learn the L2 makes learners more likely to seize opportunities to interact as much as possible with L2 natives and therefore gain access to very rich pragmatic input (Hassall, 2003).

Even though the differences between the learning contexts of FI and SA are generally agreed upon and it seems clear that the quality and quantity of the input received while abroad is consistently superior than in FI, the benefits brought about by the input received in a SA context might not be that clear, after all. Firstly, the quality of the input would be irrelevant did the learners not take advantage from it. In Bleyhl’s (2009: 137) words: “learning is what the learner makes of intake”, understanding “intake” as the part of the input actually processed by the learner (Archibald, 2005). Hence, comprehensible input is a necessary condition for improved L2 knowledge, but not a sufficient condition (Larsen-Freeman & Long, 1991; Long, 1996). Importantly, in-country residence does not guarantee quality input or interaction, as the learner should engage in the L2 environment (Moyer, 2009), which, applied to the domain of L2 phonology acquisition, is in turn related with Ming’s (2010) view that context of learning per se does not grant enhanced L2 production.
The fact that input in a SA context may seem not suffice to reach success in L2 phonological acquisition might be linked to the fact that the processing demands learners have to face leave them little room to focus on form. As opposed to FI, SA is a meaning-oriented context. Other limitations to the quality and quality of the SA input are the (frequent) use of L1, fossilization and lack of feedback.

The learner’s onset level before the SA period must also be taken into account. There is a fair acceptable degree of agreement in that the learners’ initial L2 level will influence the gains accrued (if any) during their experience abroad (Collentine, 2009). This author refers to that notion as the threshold level. It has been found that learners at a lower level make greater progress during SA than their higher level counterparts (Carroll, 1967; Brecht, Davidson and Gingsberg, 1995; Lapkin, Hart and Swain, 1995). Conversely, higher level learners will gain relatively less.

2.2.2. An overview of research findings

This study focuses on the students’ ability to make use of the progress accrued as a result of their sojourn abroad, once back in the FI classroom. As already mentioned, no prior study has looked into such an issue with a focus on L2 speech development. Mora (2008) looked at retention effects which included the FI period, but did not look at progress right after the FI period as we do.

The first empirical studies on effects of study/stay abroad sprung in the US in the early 90s (Freed 1900; DeKeyser, 1991; Lafford 1995). The work produced in Europe relative to the field is only recent. Some studies compare-immersion abroad and at home immersion (DeKeyser, 1991; Freed, Segalowitz & Dewey 2004), where contrasting results have led to too diverse a field with contradicting evidence not always easy to
interpret. Apart from low statistical power and ceiling effects, Sanz (in Pérez-Vidal, 2014) points at design as a potential source of this contradictory evidence, as learners analysed in studies comparing SA vs. AH contexts are different, since they might vary namely in motivation. In order to control for this, the SALA project has been conceived with a within-subject longitudinal design. The population it works with is likewise more representative of the sociolinguistic reality in Europe.

A brief summary of the main findings in the 9 studies which build up the SALA project is presented hereafter. Mixed results have been obtained for listening, with Allen & Herron (2003) and Kinginger (2008) having found significant improvement after a six-week and a whole term sojourn in France, respectively, while, in contrast, Tanaka and Ellis (2003) reported only minor gains. Conversely, substantial linguistic benefits are reported to take place as a function of SA as far as vocabulary is concerned (Collentine 2004; DeKeyser 1991; Milton & Meara, 1995). Regarding grammar, greater gains are obtained as a function of FI at home than at SA contexts, where gains are limited (Collentine 2004; Isabelli 2001; Isabelli & Nishida 2005).

As a broader conclusion from the findings condensed above, Pérez-Vidal (2014: 29-30) states that skill development is not linear and for all skills alike in a SA context, just as it is not in a FI context. Oral skills show substantial progress after an SA period. However, this is not so in the case of phonological development. In fact, the literature versing on L2 phonological development and SA is scarce and has failed to show a clear superiority of SA over FI, again yielding conflicting results (Stevens 2001; Díaz-Campos 2004; Mora 2008; Avello 2013), with Højen (2003) as an exception, who looks at foreign accent.
Some of the previously mentioned studies versing on L2 phonological development are insightfully described down below. For instance, Díaz-Campos (2004) analysed a sample of 46 American participants with Spanish as their L2. The semester-long study consisted of 46 learners of Spanish in two learning contexts: a formal university AH classroom group at the University of Colorado at Boulder, and a SA group (n= 26) in Alicante, Spain. In that piece of research it is indicated that a SA context may not have as much influence on certain phonological skills of learners as does the number of years that a learner studies the L2. The effects of the SA and AH experiences are compared, without expecting significant gains from either group on the phonological features analysed (a variety of Spanish consonants), as the treatment period might not have been long enough for such benefits to accrue. Both groups improved during the treatment period. The SA group displayed a general tendency toward enhanced pronunciation. However, neither group improved on the production of intervocalic voiced fricatives, one of the most difficult consonantal phenomena of Spanish. It can hence be concluded from this study that nativelike phonological control is apparently available in both the AH and the SA context, at least for the variables considered in Díaz-Campos (2004).

If we turn to the remaining studies of the few existing which examine phonological development and improvement at a segmental level and which compare the impact of different learning contexts on gains in the acquisition of segmental L2 production, they seem to present contradictory outcomes. For instance, Sanz, Morales-Front, Nagle and Moorman (2013) find a positive effect of SA on L2 phonological acquisition after analysing the production of Spanish plosives of a group of American learners after a 6-week SA period in Spain. However, in Højen (2003), Simões (1996) and Avello (2010) it was not possible to assert such linguistic gains. Højen (2003) analysed utterances produced by two groups of Danish learners. He found that length of stay (LoS) had a
significant effect on foreign accent reduction, whereas the AH group displayed no improvement in such respect. Nevertheless, the groups presented no differences at the segmental level. Simões (1996) looked at the production of vowel quality by a small sample of L2 Spanish learners after a short-term programme in Latin-America. Similarly, Avello (2010) examined the production of English vowel quality and duration by a reduced group after a 3-month SA programme.

Conversely, two studies belonging to the SALA project which look at the combined effects of FI and SA on learner’s phonological acquisition present contrasting results. As indicated previously, the project has a longitudinal pre-test/post-test design: data were collected at four different times: upon student’s enrolment (T1), after two terms of formal instruction (about 80 hours) (T2), after a SA term (T3) in an English-speaking country (this included about 40 hours of FI), and 15 months later after a two-term period without instruction/exposure to English.

Pérez-Vidal, Juan-Garau & Mora (2011) do find significant improvement in the perception and production of English vowel and consonant minimal pairs after FI, but not after SA. Lord (2010), whose study has been previously referred to, obtains similar results. This confirms in turn the SALA project’s Combination and Complementarity of Contexts Hypothesis. This enables the SALA researchers to posit that the different types of practice available in each context yield different proficiency gains which, combined, result in enhanced linguistic, motivational and cultural benefits (Pérez-Vidal, 2014).

On the other hand, also within the SALA project, Mora (2008) looked at the effects of a SA period preceded by a FI period on L2 phonological development, more specifically at the perception of vowel contrasts and at the production of English voiceless plosives. In order to examine participants’ perception, a categorical AX discrimination test
containing 135 English word pairs expressing 9 phonemic contrasts through minimal pairs was administered. It contained 108 minimal pairs (15 per contrast) and there were 27 distractors. The target contrasting phonemes in the minimal pairs were presented in a variety of phonetic environments and in a randomized order. The phonemic contrasts were selected on the basis of their relative difficulty for Catalan-Spanish speakers (Cebrián, 2006; Mora, 2005), as predicted by Flege’s (1995a) Speech Learning Model (SLM) or Best’s (1995) Perceptual Assimilation Model (PAM), the former of which shall be tackled further ahead. Significant gains were found only after the FI period.

As for production, speech samples were obtained from a reading aloud task for subsequent acoustic analysis. VOT was the segmental measure under scrutiny. Only voiceless stops in word-initial position and followed by a stressed vowel were included in the VOT analysis. A total of 960 measurements were obtained as a result of 15 word-initial voiceless stops produced by 16 valid subjects at 4 data collection times. He found slight non-significant positive effects of SA on VOT duration in voiceless stops by Catalan/Spanish speakers after a period of FI. The latter, in turn, was reported to have had very little influence on the VOT production of those informants.

Mora (2008) also tackled retention effects after 15 months, as mentioned above. Within that period, 6 months of FI were included, followed by 6 months of no FI instruction (see Mora, 2008: 244). However, this author analyses the entire period as one, with no data collection following the FI period. Hence, no consideration seems to be necessary in this study to his retention results.
2.3. L2 phonological acquisition

In this section Flege’s (1995a) Speech Learning Model (SLM) is briefly described (2.3.1), as it provides a sound explanation as for why EFL learners tend to produce intermediate VOT values between their L1 and their L2. On the basis that perception precedes production, it explains the L2 phonological acquisition phenomenon from the notion of L1-L2 transfer. The present study hence follows the SLM. The phenomenon under investigation is tackled thereafter (2.3.2).

2.3.1. Speech Learning Model

The predictions made by SLM about L2 vowel and consonant production are mostly of a perceptual basis, although other causes for inaccurate production are not ruled out. In the SLM it is assumed that “the phonetic systems used in the production and perception of vowels and consonants remain adaptive over the life span” (Flege, 1995a: 233), casting thus into doubt the validity of the well-known Critical Period Hypothesis (CPH) for SLA. Although the SLM defends that earlier is better when it comes to L2 phonological acquisition, the reason to question the CPH lies in the notion of input: it is not age of onset learning by itself that accounts for success in L2 phonological acquisition, but rather the fact that children are more likely to receive better quality input than adult L2 learners. SLM attributes L2 phonological errors mostly to incorrect perception. Nevertheless, other causes are not discarded.

Flege (1995a) claims that the L1 and L2 categories coexist in a common phonological space, inevitably influencing each other, leading hence to a bidirectional interlanguage interaction. In this sense, the further apart an L2 sound is perceived to be from an L1 sound in that phonological space, the more likely it is to be discerned. In contrast, if the
L1 and L2 sounds are close to each other, category assimilation is said to take place. However, if there are cues which differ from one language to the other, learners might be sensitive to them. This is explained through the notion of categorical perception, because “even if listeners perceive two speech sounds as belonging to the same category, they subconsciously perceive a difference, as stimuli that fit better into a given category are easier and faster to process” (Bach, 2012: 25). This would ultimately lead to merged categories or intermediate values between the L1 and the L2. This is the case of voice onset time (VOT), which shall be discussed in the next section.

2.3.2. The phenomenon investigated: VOT

Abdiraman & Koffi (2012: 1) explain that consonants [p, b, t, d, k, g] are classified as stops because they are produced by momentarily stopping the pulmonic air before releasing them. The airstream is blocked when two articulators come together. They further state that some characteristics of such consonants are studied by phoneticians by measuring the time lag between closure and the release of closure. The acoustic term for that lapse of time is voice onset time (VOT).

Abramson & Lisker (1964) provided the most generally accepted definition for voice onset time (VOT): “the interval between the release of the stop and the onset of glottal vibration, that is, voicing” (1964: 389). Another interesting definition is that offered by Yavas (2011), who describes VOT as “the duration of the period of time between the release of a plosive and the beginning of vocal fold vibration”.

Although there are “many acoustic cues signaling the one phonetic dimension of voicing” (Whiteside & Marshall, 1998: 1), VOT is considered to be the single most adequate physical basis for separating the homorganic stop categories of English and a
number of other languages (Abramson & Lisker, 1964) and thus “the most important feature in stop consonant classification” (Bach, 2012: 12).

VOT is hence “a key temporal feature in spoken language” (Whiteside & Marshall, 1998: 1). In English, Spanish, and many other languages, the plosive minimal pairs /p-b/, /t-d/ and /k-g/ are distinguished through the primary phonetic dimension of voicing.

“Stops with VOT values longer than some boundary value are classified as voiceless, while stops with VOTs shorter than the boundary value are classified as voiced” (Benki, 2005: 240). In fact, plosives are among the most common sounds of human languages (Yavas, 2007; Yava & Wildermuth, 2006). However, they are not always phonetically realized in the same way. There actually exist three main patterns of VOT in the world languages:

a) Negative VOT, pre-voicing or voice lead, which takes place when the voicing starts before the release of the stop, in an interval of time oscillating between -125 and -75ms.

b) Zero VOT or short voicing lag occurs when the release of the stop is simultaneous to the vocal fold vibration, in a time lapse ranging from 0ms to +35ms.

c) Positive VOT or aspiration happens when the voicing starts after the release of the stop, that is, there is a delay of the onset of the preceding vibration, which lasts between 35ms and 100ms.

Types a) and b) are characteristic for voiced and voiceless plosives, respectively, in Romance languages, whereas type c) is found in Germanic languages, which display
longer VOT values. For the purpose of this paper, only the language-specific differences in VOT between Spanish and English will be discussed.

VOT values for voiceless stops in Spanish are between zero and 10ms (Abramson & Lisker, 1964), corresponding to pattern b) above, whereas voiced stops display negative VOT values, as in pattern a). Contrastingly, English voiceless stops are produced with a long lag of 30ms or longer (pattern c), whereas in English voiced stops, voicing begins simultaneously with the release (pattern b). These differences are represented in the figure below, adapted by Yavas (2007) from Zampini & Green (2001).

![Figure 1. Phonetic category classification of English and Spanish consonants.](image)

Yavas (2007) explains from this figure that there is an overlap between English voiced stops and Spanish voiceless stops at the phonemic level. In contrast, English voiceless stops find no equivalent in the Spanish system. This explains the difficulty Spanish learners face when acquiring the long lag of English voiceless stops. They normally produce English voiceless stops without their characteristic aspiration.
English voiced stops in utterance-initial position are simultaneously released with the onset of voicing, resulting in zero VOT values (Caisse, 1982; Docherty, 1992). The same consonants are produced with voice lead in Spanish, resulting in negative VOT values of less than -40ms (Benki, 2005).

Although there is no absolute value for each plosive, some authors (Kent & Read, 1992; Toribio, Bullock, Botero & Davis, 2005) indicate that the standard VOT patterns in English are 55ms for /p/, 70 ms for /t/ and 80ms for /k/. Importantly, these values apply only to stressed syllables in word-initial position (Reis & Nobre-Oliveira, 2007), as stress is a factor of variation in VOT production. Normally, stops in unstressed syllables display lower VOT values than their stressed counterparts (Antoniou et al., 2014; Abramson & Lisker, 1967). Other factors influencing VOT production are tackled further ahead in this section.

In contrast, this value is of 30ms for the VOT of word initial /p t k/ in Romance languages (Schwarzhaupt, 2012; Yavas, 2007). ESL learners with a Romance language as their L1 tend to produce English word initial voiceless stops with a duration longer than 30ms, but shorter than typical native values. These intermediate values for VOT arising from cross-language influence have been reported in many studies (Alves & Zimmer, 2015; Flege, 1987; Flege, Frieda, Walley, & Randazza, 1998; Mora, 2008; Reis & Nobre-Oliveira, 2007; Schwartzhaupt, 2014; Wrembel, 2011; 2013; Yavas, 2007).

In this respect, Alves & Zimmer (2015) suggest that their Brazilian Portuguese participants might perceive the characteristic aspiration of English voiceless stops as a detail to make them reduce their degree of accent, but not as a necessary cue to establish phonological distinctions. Hence, it could be the case that VOT does not take priority for learners due to its allophonic character. This explanation is closely intertwined with
the previously discussed notion of primacy of meaning over form in Avello (2013). From this perspective, the authors ground the need of FI for learners to prioritise VOT, regardless of their proficiency level, since, otherwise, they argue, they will continue to turn to acoustic cues relevant in the distinction of voicing in their L1.

Stress has been mentioned as a factor influencing VOT production. However, “perhaps the most influential factor in VOT variation is the speaking rate – the faster it is, the smaller the measurements tend to be” (Reis & Nobre-Oliveira, 2007: 399). This is also reported by Mora (2008) and Bach (2012), who on this basis conclude that speaking style has an effect on pronunciation, an idea which comes originally from Labov (1972). This can therefore pose challenges to VOT studies, given the difficulty to account for the variety in the speaking rate of informants.

There are two other important variables which have been found to affect VOT realization. One the one hand, there is the well-known effect of place of articulation on VOT (Cho & Ladefoged, 1999). “VOT increases as the place of articulation progresses farther back in the oral cavity (bilabial to alveolar to velar)” (Yava & Wildermuth, 2006: 260), which explains that velar stop consonants have the longest average VOT (Cho & Ladefoged, 1999). On the other hand, the height of the vowel preceding the target stop might affect its VOT value, with longer VOT values found in the context of high vowels than low vowels (Flege, Frieda, Walley, & Randazza, 1998; Yava & Wildermuth, 2006).

To sum up, the present investigation is a follow up of Mora’s (2008) study with a reversed view, as the effects of an 8-week FI period preceded by a SA period on L2 phonological development constitute the main focus of interest here. VOT is hence taken as the specific measure for testing segmental production, similarly to Mora (2008). However, VOT constitutes the only phenomenon explored in the present study,
unlike in Mora (2008), who also looks at the perception of vowel contrasts, as said. For this reason, VOT is more thoroughly analysed than in Mora (2008). Specifically, the factors found to influence VOT production mentioned above are taken into account in the present study (i.e., speaking style, vowel height and place of articulation). To recall, he found that SA had non-significant positive effects on the VOT duration of voiceless stops by Catalan/Spanish speakers only after the FI period.

As for proficiency level, by the time the execution of the present TFM started, the possibility of testing participants prior to their sojourn abroad was no longer available. For that reason, participants’ proficiency level was tested both after the SA and the FI periods. This apparent limitation to the study is not held as such, however, because, in contrast with the aforementioned factors, proficiency level does not necessarily influence VOT production (Alves & Zimmer, 2015). In turn, this is coherent with Collentine’s (2009) notion of threshold level, previously referred to.

3. Research questions and hypotheses

Against this background, and in order to tackle the study above describe and to shed light on the issue of L2 phonological acquisition while contributing to the SALA project, the research questions presented in the following subsection have been identified, and the subsequent hypotheses have been established.

3.1. Research questions

Specifically, this study attempts to answer the following main research question:
1) Does an 8-week FI period immediately following a 3-month long SA period have a significant effect on the VOT production of voiceless plosive consonants by Spanish/Catalan EFL participants?

In addition, proficiency level (measured at T2) is addressed, as well as factors influencing VOT production such as task effect (i.e., speaking style), vowel height and place of articulation. In order to shed light on these issues, the research sub-questions below have been identified:

1.1. Do results vary when proficiency level as measured through vocabulary size at T2 is taken into account?

1.2. Are there differences in the VOT values as a function of task type (i.e. speaking style) as measured through words produced in two different tasks (text reading-aloud task vs. carrier sentence task)?

1.3. Are there differences in the VOT values as a function of the height of the vowel following the target voiceless plosive consonant?

1.4. Does place of articulation have an effect on VOT duration?

3.2. Hypotheses

This section separately presents the hypotheses held for each of the research questions presented above.

3.2.1. Hypothesis for RQ1

It is hypothesised that the 8-week FI period preceded by a SA term will have a positive though not necessarily significant impact on the duration of the VOT production of
voiceless stops by the advanced Spanish/Catalan learners tested. More specifically, informant are expected to produce intermediate VOT values, higher than those characteristic of their L1, but never reaching typical native values, consistently with the literature.

3.2.2. Hypothesis for sub-RQ1.1
No significant differences in VOT values are expected as a function of proficiency level as tested both after the SA and the FI periods, consistently with Alves & Zimmer (2015). Any differences found triggered by this variable are expected to result in larger improvement by low level learners than by more advanced ones, following Collentine’s (2009) notion of threshold level.

3.2.3. Hypothesis for sub-RQ1.2
The VOT values obtained from the text to be read aloud are expected to be shorter than those gathered from the carrier phrase task, as VOT values tend to decrease in natural speech in contrast with words uttered in isolation (Bach, 2012; Labov, 1972; Mora, 2008).

3.2.4. Hypothesis for sub-RQ1.3
Longer VOT values are expected in the context of high vowels as opposed to low ones, consistently with Flege, Frieda, Walley & Randazza (1998).

3.2.5. Hypothesis for sub-RQ1.4
It is also expected that higher VOT values will be obtained for /k/ than for the other stops as a function of place of articulation. In turn, the shortest VOT values are expected to be obtained for /p/, according to Yava & Wildermuth (2006).
4. Methodology

In this section the methodological issues of the present study are presented. Firstly, the design is provided (4.1), followed by a description of the participants (4.2). Data collection instruments are subsequently described (4.3), followed by the data collection procedure (4.4). Lastly, the analysis of results is presented (4.5).

4.1. Design

This investigation has been conducted within the framework of the longitudinal SALA project. It has adopted a pre-test/post-test design.

It aims at revealing whether a FI period preceded by a SA period affects EFL VOT production of plosive consonants in the same way as it does when it is the FI period that precedes the SA one, as Mora (2008) found. Accordingly, we have hypothesized that this time, a positive though not necessarily significant difference will obtain after FI in the VOT plosive consonant values produced by our participants.

Time 1 (T1) in our data collection corresponds to the beginning of the period of FI, and also the end of the SA period. So it is placed immediately after participants’ arrival from their SA. Then, after 8 weeks of FI, Time 2 (T2) data collection, takes place. The data collected will allow us to test whether the VOTs produced at pretest and posttest significantly differ as a function of the FI context, preceded by a SA term. The impact of FI (with no explicit attention to pronunciation) is the independent variable and VOT duration of voiceless plosive English consonants the dependent variable. Additional independent variables explored are proficiency level, and speaking style, vowel height and place of articulation.
4.2. Participants

Two groups of participants were administered the same tasks for the data collection of the present investigation. The VOT durations of English voiceless plosives produced by a group of undergraduate students from a university in Barcelona (4.2.1) were measured. They were contrasted against those of a group of native speakers, who served as a baseline for the study (4.2.2).

4.2.1. Non-native speakers

A group of 17 undergraduate students from a university in Barcelona was recruited for the present study. Of those, there was a total of 13 valid subjects ($n=13$, 11 females and 2 males)\(^1\), as four of them had not spent their SA period in an English-speaking country and were hence discarded.

Participants belong to the institution’s intact groups, which have been organized on the basis of an entrance online test pitched at a B2 - C1 level. However, in order to check on the real homogeneity of the group, a vocabulary test was administered. It was found out that our participants actually differ in their proficiency level, as by means of an $X/Y_{\text{lex}}$ vocabulary size test. Scores are divided into different ranges which correspond to the proficiency levels set by the Common European Framework of Reference for Languages (CEFR). See Table 1 for correspondences.

---

\(^1\) The fact that the vast majority of participants are females is due to the high number of females taking the degree informants were recruited from. As a consequence, it was not possible to have a balanced amount of males and females. This also prevented possible gender effects.
According to the vocabulary size test results, shown in Figure 2, 46% of the participants have a C level, four of which fall within the range of a C2 level, whereas the score of the remaining two corresponds to the C1 level. The 39% of learners who attained a B level include two learners who fell within the range of a B2 level, with the remaining three having obtained a B1 level. A learner scoring at an A1 level and another one at an A2 level build up the 15% scoring within the A level range.

As for language dominance, all participants are Spanish/Catalan bilinguals. However, not all of them report to feel equally dominant in both languages (see Figure 3). More than a half feels Catalan dominant, according to the answers they provided in the questionnaire they were administered.
As for language background, all participants report knowing a third language (either French or German) apart from their mother tongues (Spanish/Catalan) and EFL. Only participant 2 reports to be the only foreign language she knows (see Figure 4).

Remarkably, participants 1 and 9 report French to be the native language of one of their parents and using it at home. This is likewise the case for participant 6 with Vietnamese. However, this informant does not report to consider himself fluent in this language, unlike the former two.

As for their destinations, they have all been in the UK except for participant 8, who went to the USA. For more than the half, it was their first time in an English-speaking country (see Figure 5).
4.2.2. Native speakers

Five native speakers were originally recruited to take part of the present study. Due to schedule constraints, two of them failed to attend the recording session and thus, the remaining three served the present investigation as a baseline group (n=3, 2 females, 1 male, mean age = 24.8). They all share a similar linguistic background, as they are linguistic majors and have a high command of two foreign languages, namely Spanish and French. Two participants are speakers of American English and the third participant is a speaker of Hiberno-English. For the purpose of this investigation, it has been considered that VOT values shall not differ as a function of language variety among native speakers, as they produce native-like values at any rate.

4.3. Data collection instruments

Two tasks were designed for the data collection: a carrier-sentence read-aloud task and a short-text read-aloud task. These were aimed at obtaining the VOT durations of the consonant stops for subsequent analysis. A questionnaire was likewise designed with
the purpose of gathering relevant information about the participant’s language background as well as about their SA (data having been collated in the previous section). Lastly, an X_lex / Y_lex vocabulary size test was administered to participants with the goal of determining their proficiency level according to the CEFR (reported on likewise in the previous section).

4.3.1. Carrier sentence task

This task was conceived to test the participants’ VOT production of word-initial stops in stressed syllables. Thirty-one monosyllabic words starting with a voiceless stop (p, t, k) were selected for the test. There were 29 distractors, resulting in a total of 60 words.

Vowel height was taken into account since it is an influencing factor in VOT production, as discussed in the literature review. Of the 11 words starting with /p/, the stop was followed by a high vowel in 5 of the items (peach, pill, pear, pin, pig) and by a low vowel in 6 of them (pub, pan, park, pup, part, pun). Of the 10 words starting with /t/, the stop was followed by a high vowel in 5 instances (tear, tip, two, ten, tent) and by a low vowel in 5 of them (tan, tuck, touch, tart, toss). As for the 10 words starting with /k/, the stop was followed by a high vowel in 5 cases (key, could, kill, kilt, kit) and by a low one in the remaining 5 (cod, card, cot, cap, cut). The 60 items were randomized and displayed in a PowerPoint presentation by means of a carrier-sentence task (see Appendix I).
4.3.2. Text to be read aloud

The goal of the text was to take into account the effect of speaking style, as mentioned in previous sections. In order to do so, 12 of the items starting with a voiceless stop present in the carrier task were selected in order to be included in a text to naturally be read aloud. The stop was followed by a high vowel in 6 of the instances (key, ten, could, pill, two, peach) and by a low vowel in 6 of them (tan, park, pub, cod, tart, card). The text was printed and physically handed to the participants for them to read it aloud (see Appendix II).

4.3.3. Carlet-SALA questionnaire

The language background questionnaire Carlet-SALA is a combination of the questionnaire used in Carlet’s PhD dissertation (ongoing) and the SALA questionnaire on SA conditions (see Appendix III). Its goal is to obtain relevant information both about the language background of our informants and about their experience while abroad.

4.3.4. X_Lex / Y_lex vocabulary size test

As already mentioned, in order to refine our appraisal of the participants’ EFL proficiency, an X_Lex and Y_lex vocabulary size test was administered at data collection T2. It was decided to test the participants’ language proficiency through this lexical decision task following Souza (2015), who also studies L2 production. Vocabulary size should increase as language proficiency does, according to Milton (2010), which makes such a measure appropriate for our purposes.
Fairly enough, making claims upon the informants’ proficiency level by looking only at their lexical knowledge would result in oversimplification, since other areas such as grammatical and pragmatical knowledge, for instance, would be neglected. However, for the purpose of this investigation, vocabulary size was judged to be an adequate measure, provided that the main focus of the study is pronunciation, it was considered unnecessary to make participants undergo a time-consuming language proficiency test.

X_Lex (Meara, 2005) and Y_Lex (Meara & Miralpeix, 2006) serve to measure vocabulary range: they estimate how many words the test taker knows in the test language. Both tests measure knowledge of the most frequent words in the target language. The X_Lex test contains a total of 5,000 words, whereas the amount for the Y_Lex test is of 10,000, therefore reflecting a more advanced vocabulary knowledge.

They are administered via a computer in a ‘yes’-‘no’ format. A word is displayed on a computer screen and participants indicate whether they know its meaning or not by clicking on the corresponding answer (a happy or a sad face). Non-words are included in the test in order to increase reliability. If test takers claim knowing one of these non-existing words, they are penalized in the final score. The tests are self-paced and take around 4-8 minutes to be completed.

4.4. Data collection procedure

Participants were recorded in high quality sound-proof booths, using an Audacity recorder. They first signed the consent form, and then were instructed by the researcher to proceed by completing the tasks, which were administered in randomized order. Order 1 corresponded to the carrier sentence task being performed first, whereas with
order 2, participants started by reading the text. They were required to read as naturally as possible.

Additionally, at T1 the leaners completed the language background questionnaire and at T2 they performed the vocabulary size test. Participants were financially rewarded with 5 Euro for their contribution to the study at T2, in order to avoid mortality.

4.5. Analysis

This TFM explores the potential effects of FI following SA on the VOT production of voiceless stops in stressed word-initial contexts by EFL Spanish/Catalan learners. In order to do so, the analysis of both tasks was conducted as follows. VOT for the initial voiceless stops produced by the participants of both the 31 target words in the carrier task and the 12 target words in the text to be read aloud was analyzed with Praat, following Mora’s (2008: 247-248) measure: “the distance in milliseconds from sounds waveforms and wide-band spectrograms between the onset of the release burst and the first vertical striations corresponding to the onset of voicing in the vowel”.

Thirty-one VOT instances in the carrier sentence task produced by the 13 participants at the two data collection times resulted in a total of 806 VOT measurements for that task (403 at each time).

As for the analysis of the data extracted from the text, 12 VOT instances produced by the 13 participants when reading the text aloud at the two testing times resulted in a total of 312 VOT measurements for that task (156 at each time). A total of 1118 VOT measurements were hence obtained.
5. Results and discussion

The results of the analyses described above are presented in this section, around each research question and sub-question.

5.1. RQ1

Results for the main RQ1 are as follows. A reminder of the question is previously offered below.

<table>
<thead>
<tr>
<th>Does an 8-week FI period immediately following a 3-month long SA period have a significant effect on the VOT production of voiceless plosive consonants by advanced Spanish/Catalan EFL participants?</th>
</tr>
</thead>
</table>

In order to answer RQ1, Wilcoxon-tests were conducted. Given the sample size, nonparametric tests were performed.

As shown in Table 2 and Figure 6 below, participants displayed slightly longer VOT values at T2 than at T1. However, the 8-week FI period immediately following a 3-month long SA period was found to have no statistical significant effect on the VOT production of voiceless plosive consonants by advanced Spanish/Catalan EFL participants ($z=0.384, N=13, p=.3505$, one-tailed).

Similar results were also obtained by looking at the two tasks separately. As can be seen in Table 2 and Figure 6, participants displayed slightly longer VOT values at T2 than at T1 for both tasks. A further Wilcoxon-test was run in order to reveal whether this difference reached statistical significance. Again, the 8-week FI period immediately following a 3-month long SA term was found to have no statistical significant effect on the VOT production of voiceless plosive consonants by advanced Spanish/Catalan EFL participants.
participants in neither the text ($z = 0.314, N=13, p = .3765, \text{one-tailed}$) nor the carrier
sentence task ($z = 0.454, N=13, p = .325, \text{one-tailed}$).

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>NES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both tasks</td>
<td>51.05 (20.39)</td>
<td>51.54 (18.13)</td>
<td>65.47 (24.94)</td>
</tr>
<tr>
<td>Words</td>
<td>54.93 (23.55)</td>
<td>55.07 (20.13)</td>
<td>67.96 (29.87)</td>
</tr>
<tr>
<td>Text</td>
<td>40.99 (14.55)</td>
<td>42.40 (13.98)</td>
<td>59.06 (13.54)</td>
</tr>
</tbody>
</table>

Table 2. Mean VOT measurements (ms) at both testing times (T1, T2).

These results may be interpreted as follows. The lack of explicit focus on L2 phonology in the FI our participants received might-account for the fact that the slight lengthening of VOT values displayed at T2 failed to reach statistical significance. This view is supported by those studies stressing the need of explicit attention to L2 phonology in FI for the improvement of L2 production accuracy (Calvo Benzies, 2014; Darcy, Ewert & Lidster, 2012; Gordon & Darcy, 2012).
Hence, the answer to our research question is that an 8-week FI period preceded by a 3-month long SA term has no statistical significant effect on the VOT production of voiceless plosive consonants. Our hypothesis has not been confirmed as results are not significant, although this possibility was also contemplated. However, we have obtained a numerical tendency towards the native-like model in the VOTs of plosive consonants in initial position. Importantly, the native group always produced longer VOTs than the non-native participants. It must be noticed, however, that no statistical tests were run when comparing both groups due to their low numbers. For this reason, the native informants served the present investigation solely as a baseline group. This confirmed the SLM’s prediction and finding that EFL learners produce intermediate VOT values between their L1 and their L2 (Alves & Zimmer, 2015; Flege, 1995a; Flege, 1987; Flege, Frieda, Walley, & Randazza, 1998; Mora, 2008; Reis & Nobre-Oliveira, 2007; Schwartzhaupt, 2014; Wrembel, 2011; 2013; Yavas, 2007). As explained by the SLM, learners perceive the L2 sounds in relation to their pre-existing L1 categories. Therefore, this model accounts for the intermediate VOT values produced by our participants, whose interlanguage is in the process of moving to the target language values, notwithstanding that attaining them completely is not predicted by the SLM.

5.1.1. Sub-RQ1.1

Results for sub-RQ1.1 are presented in this section. Previous to that a reminder of the question is provided.

Do results vary when proficiency level as measured through vocabulary size at T2 is taken into account?
In order to determine whether VOT productions differed as a function of proficiency level (on a lexical basis only) a Mann-Whitney test was conducted. Participants ascribed an A and B proficiency level according to the CEFR (see section 4.2.1) were considered as the lower level group, whereas participants having scored within the C range built up the high level group. Data gathered at both times and displayed on Table 3 were averaged.

As can be observed in Table 3, the lower level group shows a larger improvement than the advanced group. In fact, the latter experiences a slight numerical decrease on VOT. Nevertheless, the VOTs produced by the high level group were closer to those produced by the natives, thus showing a numerical tendency that favours experience. However, these must tentatively be considered as tendencies, since no statistical significant difference was found in the VOT productions between the two groups (z= 1.286, N=13, p= .0995, one-tailed).

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NES</strong></td>
<td>65.47 (24.94)</td>
<td></td>
</tr>
<tr>
<td><strong>High Level Group</strong></td>
<td>60.11 (21.26)</td>
<td>55.83 (16.56)</td>
</tr>
<tr>
<td><strong>Low Level Group</strong></td>
<td>43.27 (20.12)</td>
<td>47.87 (18.96)</td>
</tr>
</tbody>
</table>

Table 3. Mean VOT measurements (ms) as a function of proficiency level.

In the absence of a non-parametric version for a 2x2 ANOVA, this parametric test was performed to further explore whether participants’ VOT differed as a function of proficiency level based on their vocabulary size. Before doing so, it was checked that the data were normally distributed by means of tests of normality, namely the Kolmogorov-Smirnov Test and the Shapiro-Wilk Test. Having met such assumption, time (T1 and T2) was introduced as the within-subjects factor and level (high. vs. low) as the between-subjects factor. The test revealed a non-significant effect of time
a non-significant effect of group \((F(1,11)= 1.534, p= .241)\) and a non-significant group by time interaction \((F(1,11)= 2.043, p= .181)\). The non-significant effect of time is explained by the fact that neither of the two groups showed a considerable improvement from T1 to T2, as previously reported for RQ.1. The non-significant effect of group indicates that proficiency level played no decisive role for gains in L2 phonological development to be accrued. The non-significant group by time interaction may be accounted by the fact that differences obtained by each group as a result of time were not dissimilar.

Interestingly, despite the non-significant results, it can be observed that the high level group obtained numerically higher and more native-like VOT values than the low level group at the outset of the study, that is, after the SA period. This result might point towards a tendency of language experience having a potential impact on L2 category learning, as predicted by the SLM. In these lines, the lower level group L1 categories would be influencing their ongoing L2 corresponding categories to a greater extent than the more robust L2 categories created by the high group would be influenced by their L1 categories.

Looking more closely to the performance obtained for both groups over time (see Figure 7), the results point towards a numerical non-significant tendency for improvement in the lower level group. In fact, the tendency points towards the other direction for the high level group. These results, even though drawn from tendencies only, may suggest that the high level group had reached their ceiling VOT values during the SA period, whereas the lower level group still had room for improvement. A potential reason for this is that the L2 categories created by the high level group for the target segments are more robust than those of the low level group. Therefore, the FI period following the SA might have been more effective in enhancement of L2 VOT
production for the low level group than for their more advanced counterpart. This numerical non-significant tendency found in our data is in line with Collentine’s (2009) notion of threshold level.

![Figure 7. Mean VOT measurements (ms) as a function of proficiency level.](image)

Hence, the answer to our research question is that proficiency level has shown not to play a significant role in the VOT production of English plosive consonants in initial stressed position by Catalan/Spanish EFL learners, at least as far as the effects of an FI period preceded by a SA are concerned. In this case, the absence of statistical significance confirms our hypothesis, based on Alves & Zimmer’s (2015) conclusion that learners’ proficiency level does not guarantee per se the complete attunement of the acoustic cue on focus, as the VOTs produced by either group failed to reach native-like values. Importantly, a numerical tendency pointing at ceiling effects in the high level group was obtained.
5.1.2. Sub-RQ1.2.

This section presents the results for sub-RQ1.2. Previously, a reminder of it is given.

Are there differences in the VOT values as a function of task type (i.e. speaking style) as measured through words produced in two different tasks (text reading-aloud task vs. carrier sentence task)?

In order to explore whether the VOT values obtained significantly differ as a function of task type, a Wilcoxon-test was performed. As shown in Table 4 and Figure 8, the VOTs produced by means of the carrier sentence task (words), were higher than those obtained through the text for both the non-native and the native groups.

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T1+T2</th>
<th>NES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words</td>
<td>54.93</td>
<td>55.07</td>
<td>55.00</td>
<td>67.96</td>
</tr>
<tr>
<td>Text</td>
<td>40.99</td>
<td>42.40</td>
<td>41.69</td>
<td>59.06</td>
</tr>
</tbody>
</table>

Table 4. Mean VOT measurements (ms) as a function of task (text vs. words).

Figure 8. Mean VOT measurements (ms) as a function of task (text vs. words).
It was found that the VOT durations produced when reading the text were significantly shorter than those obtained at the carrier sentence task at both times ($z= 2.830, N= 13, p= .0025$, one-tailed) as well as at T1 ($z= 2.621, N=13, p= .0045$, one-tailed) and at T2 ($z= 2.900, N= 13, p= .002$, one-tailed).

Hence, the answer to our research question is that speaking style significantly affects the VOT production of Catalan/Spanish learners of English. Despite the lack of statistical comparison between the native and non-native groups, the numerical values obtained for the natives suggest that speaking style has affected both our groups of participants similarly. Our hypothesis is thus confirmed. Our data hence support Labov’s (1972) original idea that speaking style does have an effect on pronunciation and more specifically on VOT production, as also found by Mora (2008) and Bach (2012), confirming hence that in natural speech, VOT values tend to decrease, whereas they tend to get closer to standard values when produced in (quasi-)isolation.

5.1.3. Sub-RQ1.3.

Results for sub-RQ1.3 are as follows. A reminder of the question is offered previously.

Are there differences in the VOT values as a function of the height of the vowel following the target voiceless plosive consonant?

With the purpose of determining whether there are differences in the VOT values as a function of the height of the vowel, another Wilcoxon-test was conducted. As observed in Table 5 and Figure 9, VOTs produced preceding a high vowel were longer than those followed by a low vowel both for the native and the non-native speakers. The test
revealed that this difference did reach statistical significance for the non-native group ($z = 3.180, N=13, p= .0005$, one-tailed).

<table>
<thead>
<tr>
<th></th>
<th>Non-natives</th>
<th>NES</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Vowel</td>
<td>56.52 (18.04)</td>
<td>71.51 (25.75)</td>
</tr>
<tr>
<td>Low Vowel</td>
<td>46.25 (18.96)</td>
<td>59.72 (24.18)</td>
</tr>
</tbody>
</table>

Table 5. Mean VOT measurements (ms) as a function of vowel height averaged across time (T1+T2).

Figure 9. Mean VOT measurements (ms) as a function of vowel height averaged across time (T1+T2).

Hence, the answer to our research question is that vowel height does have a significant effect on VOT production of Catalan/Spanish learners of English. As for the natives, they did numerically respond to this same pattern, although this was not statistically confirmed for this group. The hypothesis held regarding this sub-RQ is confirmed by the data and therefore go in line with Flege, Frieda, Walley & Randazza (1998)’s and Yava & Wildermuth (2006) findings that vowel height does influence VOT production.
5.1.4. Sub-RQ1.4.

The results for sub-RQ1.4 are presented in this section. A reminder of it is given before.

**Does place of articulation have an effect on VOT duration?**

In order to determine whether VOT values differed as a function of place of articulation, three Wilcoxon-tests were run on the data obtained by the non-native speakers. As shown by Table 6 and Figure 10, /k/ displayed the longest VOT values, with /t/ in the second place and /p/ having triggered the shortest durations for the non-native group. However, natives produced slightly longer VOTs for /t/ than for /k/. In turn, VOT values for /p/ were the shortest for this group.

<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-natives</td>
<td>61.62 (20.17)</td>
<td>57.50 (22.09)</td>
<td>35.78 (15.26)</td>
</tr>
<tr>
<td>NES</td>
<td>70.44 (24.68)</td>
<td>72.79 (22.81)</td>
<td>54.01 (27.28)</td>
</tr>
</tbody>
</table>

Table 6. Mean VOT measurements (ms) as a function of place of articulation averaged across time (T1+T2)

![Figure 10. Mean VOT measurements (ms) as a function of place of articulation averaged across time (T1+T2)](image-url)
The test revealed that VOT values obtained for /k/ were significantly different from those obtained for /p/ ($z = 3.180$, $N=13$, $p = .0005$, one-tailed). However, VOT values obtained for /k/ did not significantly differ from those obtained for /t/ ($z= 1.223$, $N=13$, $p=.1105$, one-tailed). In turn, the VOT values obtained for /t/ did significantly differ from those obtained for /p/ ($z= 3.180$, $N=13$, $p=.0005$, one-tailed).

The fact that only VOT values for /k/ as opposed to those from /t/ did not to reach statistical significance might be explained through Alves & Zimmers’ (2015) observation that aspiration is a cue learners pay attention to. Specifically, it is more salient in some places of articulation than in others. This lack of significance might in turn be linked to the very slight numerical difference between /k/ and /t/ produced by the native group. This seems to indicate that our participants are in the process of creating L2 categories for the target segments, as their performance shows certain similarities with that of the baseline group. However, their VOT durations never reach those produced by the natives, as predicted by the SLM. Accordingly, the possibility for our informants to establish phonetic categories for English voiceless stops seems to be discarded.

The hypothesis for this sub-RQ based on Yava & Wildermuth (2006) that place of articulation makes VOT increase as it moves back in the oral tract is confirmed.
6. Summary and Conclusions

In this study we have aimed at examining and measuring the effects of an 8-week FI period, with no specific training in the learners’ L2 phonological abilities, preceded by a 3-month long SA term, on the VOT production by Catalan/Spanish EFL learners.

Our analyses of results have revealed that the 8-week FI period preceded by a 3-month long SA term undergone by our participants had no statistical significant effect on their VOT production of English voiceless plosive consonants, although a tendency towards improvement was obtained. We have explained such finding as due to the absence of explicit instruction on L2 phonology which characterised their FI period.

As for proficiency level, our data show that it does not significantly affect VOT production of English plosive consonants in initial stressed position. However, the data do show that it is the lower level group of participants who displayed a slight tendency towards improvement, possibly due to having more room for improvement. Their more advanced counterpart, on the other hand, had most likely reached their ceiling VOT performance during their SA period as their values at the outset of the study were pointing towards the native values. Regarding speaking style, vowel height and place of articulation, results reveal that these three independent factors significantly affected the VOT productions of our informants. For what concerns speaking style, words in (quasi-)isolation (carrier sentence task) displayed significantly higher values than those produced in natural speech (text to be read aloud). We might interpret these results by stating that the more attention is paid when uttering words, the more likely the sounds produced are to be accurate. As for vowel height, VOT values produced before a high vowel were significantly longer than those produced preceding a low vowel. Finally, concerning place of articulation, VOTs for /k/ displayed the highest values, with /t/ in
the second place and values for /p/ being the shortest. It must be stressed that only the /k/ vs. /t/ comparison failed to reach statistical significance. The results concerning place of articulation can be understood through the SLM. Our participants seem to be in the process of creating L2 categories for voiceless stops, as they performed similarly to the native group, namely regarding the lack of difference between /k/ and /t/. This in turn suggests that aspiration is more salient in some places of articulation than in others, Alves & Zimmer (2015).

From these results we can point at the tentative conclusion that explicit instruction on L2 phonology in the FI context would seem to be a factor which may contribute to the improvement of L2 phonological acquisition in EFL learners. In such circumstances, attention to acoustic cues should be more feasible than when students are in highly demanding communicative contexts such as those in SA, the context which our participants have experienced prior to the data collection conducted in this study. This is an issue which invites further research. A further study should measure the effects of a post-stay-abroad FI period including either explicit instruction on L2 phonology or L2 phonetic training and focusing either on the VOT of English stops, the focus of the present study, or on other relevant acoustic cues.

Moreover, we can also state that proficiency level does not prove to determine the extent to which L2 phonological development accrues. In addition, apparent ceiling effects were found in the advanced group, whereas their lower level counterpart seemed to have more room for improvement. These must tentatively be taken as tendencies. Despite the lack of statistical confirmation for these tendencies, they leave a door open to further research, namely with a larger population so that more robust claims can be made.
We can also tentatively conclude, concerning the three other factors examined (speaking style, vowel height and place of articulation), that they do significantly affect VOT production of voiceless plosives by Catalan/Spanish learners. The native group displayed a similar numerical pattern, which indicates that those three factors affected both groups in a similar way. However, natives always produce higher VOT values for English voiceless stops, confirming the fact that EFL learners tend to produce intermediate VOT values between their L1 and their L2 for the same consonants.

The present study has aimed at making a contribution to the SALA project by providing a study never undertaken before. Moreover, we have also sought to contribute to reduce the research gap present in the field of L2 phonology acquisition.

This study has several limitations. As already mentioned, when work on this TFM started, it was no longer possible to test participants prior to their experience abroad, for which VOT measurements before SA could not be obtained. One other issue is the measurement of proficiency level. Testing should have been conducted at both data collection times, and not only at T2. Time constraints prevented this from happening. In addition, numbers of the population analysed were low, both for the non-native and the native groups. This has prevented us from drawing general conclusions and stresses the fact that even conclusions related to our population must be taken with caution. Finally, the natives who served a baseline group are not monolingual, so their VOT values might have been influenced.

We all like to have a nice visiting card.

Pronunciation is that of a second language.
References


and “at home”. In *Second language acquisition in a study abroad context*, ed. B. Freed, 97-121. Amsterdam: John Benjamins.


Milton, J. (2010). The development of vocabulary breadth across the CEFR levels: A common basis for the elaboration of language syllabuses, curriculum guidelines,
examinations, and textbooks across Europe. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Eurosla Monographs Series 1: Communicative Proficiency and Linguistic Development* (pp. 211-232).


Murphy, J. (2014). Teacher training programs provide adequate preparation in how to teach pronunciation. In L. Grant (Ed). *Pronunciation myths: Applying second language*
research to classroom teaching. (pp. 188-224) Ann Arbor: University of Michigan Press.


Appendices

Appendix I. Carrier sentence task.

I say _______

I say _______ now

I say _______ twice

1. peach 11. pill 21. pear
2. key 12. ten 22. cot
3. tear 13. park 23. pin
4. pub 14. tart 24. two
5. could 15. card 25. kit
6. tan 16. kill 26. toss
7. pan 17. pup 27. cap
8. tuck 18. kilt 28. pun
9. touch 19. part 29. tent
10. cod 20. tip 30. pig

31. cut
Appendix II. Text to be read aloud.

Read aloud task

Read the text twice. First, silently on your own, and then aloud for the researcher to record it.

That Sunday morning, as he slammed the door shut behind him, John realised he had left his keys in the house. Sam, his only housemate, wouldn’t be back till ten that night. ‘Could I be more stupid? And I have forgotten my pills too’, he thought to himself.

He then decided to work on his tan and go to the park, where he met two friends. They all went to the local pub and had a few beers, cod fritters, a tart and some peach for desert. When he was finally feeling a bit better about it all, he realised he had no money or credit card on him.
Appendix III. Carlet-SALA questionnaire

Welcome back to UPF! This questionnaire is designed to help us describe the conditions of your stay abroad, as part of our study on its effects on your English. Please be as open and complete as possible in providing your responses. Your feedback is very valuable! If you would like to discuss anything that the questionnaire does not address, please feel free to add those comments or set an appointment. We would appreciate if you could take some time to answer the following questions. Please do not spend too much time pondering each answer. Rather, state your sincere reaction after reading each statement. It is very important for us to obtain your true opinions. Thank you very much for your contribution to this project.

A. LANGUAGE BACKGROUND

1. Native language____________________________ Place of birth_____________________
2. Mother’s native language:____________________ Place of birth ____________________
3. Father’s native language:____________________ Place of birth ____________________
4. Are you Spanish/Catalan bilingual (learnt both languages before the age of 5)? □ Yes □ No
   Do you regard yourself as:
   A) Dominant in Spanish, because______________________
   B) Dominant in Catalan, because______________________
   C) I’m not Sp/Cat bilingual

5. What language do you use at home?
   A) Catalan & Spanish
   B) Catalan
   C) Spanish
   D) Other (specify which :) _____________________________

6. What is your second language (=learnt after the age of 5)?
   A) English
   B) Spanish
   C) Catalan
   D) Other (specify which :) _____________________________

7. What is your third language (=learnt after you second language)?
   A) English
   B) Spanish
   C) Catalan
D) Other (specify which :) ________________________
E) I don’t have a third language

8. I consider myself fluent in the following languages:
   A) Catalan
   B) Spanish
   C) English
   D) Other (specify which :) ________________________________

B. STAY ABROAD CONDITIONS

1. Your trips to an English speaking country in the past:
   This was my first trip to an English-speaking country ___.
   This trip is: the second ___ / the third ___ / the fourth ___ / or more___ / or more ___ / for me.

2. Place of study for the Stay Abroad:

   Fill in the name of your host University, Department.
   (a) University of _______________________________________
   (b) Department, School of______________________________

3. How many other students from the UPF were in your same destination? _____

4. Did you spend time together with them while abroad?

   ___ Most of the time       ___ A little bit          ___ Little

5. How many classes did you have per week?

6. Of those, how many classes were taught in English?

7. Accommodation. Put a cross if applicable.
   ___ Home-stay in family:
     ___ with children
     ___ no children

   ___ Hall of residence:
     ___ Single room
     ___ Double room with native speaker
     ___ Double room with non-native speaker from Spain
     ___ Double room with non-native speaker from another country
     ___ Apartment
65

___ Individual
___ Shared with native speakers only
___ Shared with natives and non-natives speakers
___ Shared with native Catalan/Spanish speakers
___ Shared with speakers of other languages
___ Other, please specify

8. Did you work while you were abroad?
   ___ Full time   ___ Part time   ___ Less than part time

9. What kind of job did you have?

10. How did you find it?

11. Did you use English in your job?
    ___ All the time   ___ Most of the time   ___ Little

12. Did you travel around in the area while you were abroad? ___ Yes   ___ No

13. How often did you make contact with your family and friends in Spain? Put a cross if applicable.
    ___ More than once a day
    ___ A few times a week (every other day)
    ___ Once a week to a few times a month
    ___ About once a month
    ___ None

14. Relationships:
    How often did you carry out any of the following activities in your leisure time during your study abroad?
    1 - Very often  2 - Often  3 – Sometimes  4 – Seldom  5 - Never
    ___ Studying, doing class work of any sort of work in English on your own
    ___ Studying with someone else.
    ___ Being with English-speaking people in general.
    ___ Being with English-speaking friends.
    ___ Being with friends of a variety of nationalities who spoke English.
15. How often did you?

___ Read a newspaper
___ Watch television
___ Listening to the radio
___ Read fiction books
___ Watch films
___ Go to the theatre
___ Write emails or letters
___ Other English learning activities

16. In what ways did you try to improve your English?

1. __________________________________________________

2. ___________________________________________________________________

3. ___________________________________________________________________

17. Self-observation

To what extent are the following statements applicable to your stay abroad experience?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>Not so</td>
<td>Neither</td>
<td>Fairly</td>
<td>Most applicable</td>
</tr>
</tbody>
</table>

___ I felt quite nervous every day.
___ I was exhausted.
___ I often became ill.
___ I felt very lonely.
___ I often felt frustrated.
___ I was anxious about my future.

18. Degree of difficulty or stress

How difficult, or stressful did you find the following ten features in your study abroad experiences?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very easy</td>
<td>Fairly easy</td>
<td>Unable to say</td>
<td>Fairly difficult or stressful</td>
<td>Very difficult or stressful</td>
</tr>
</tbody>
</table>

___ English language
___ Academic situation (in general).
___ Living conditions (i.e., clothes, food, housing etc.)
___ Human relationships (teachers, classmates and others.)
__ Leisure and hobby activities
__ Culture, custom, and habits.
__ Financial conditions
__ Physical health conditions
__ Mental health conditions?
__ Your job if you had any
__ Other stress factors. (Please describe anything you find very difficult or stressful beside those mentioned above) _________________________________________________

19. Perceived English language improvement
How much do you think your English has improved after your stay abroad?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>Not so</td>
<td>Neither</td>
<td>Fairly</td>
<td>Most applicable</td>
<td></td>
</tr>
</tbody>
</table>

| _I have improved my overall skills in English. |
| _I have improved my speaking skills in English. |
| _I have improved my listening skills in English. |
| _I have improved my reading skills in English. |
| _I have improved my writing skills in English. |
| _I have improved my vocabulary in English. |
| _I have improved my English grammar. |
| _I have improved the skills needed for a translator. |

20. The growth and change of yourself.
Are there any ways in which you find yourself different after you stay in an English speaking country?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td>Not so</td>
<td>Neither</td>
<td>Fairly</td>
<td>Most applicable</td>
<td></td>
</tr>
</tbody>
</table>

| _I have become more able to challenge difficulties. |
| _My academic skills have improved and I am more knowledgeable. |
| _I can make friends with people easily. |
| _My English has improved very much. |
| _My cross-cultural adjustment ability has increased. |
| _My personality has become more cheerful and outgoing. |
| _I have become more confident in my own ideas. |
| _Other changes. If you find yourself different in other aspects, please describe in details. |
21. Conclusion:
How satisfied are you with your study abroad experience in general?
Choose one number from 5 (greatest satisfaction) to 0 (no satisfaction).

5  4  3  2  1  0

* * *

Thank you again for your help