Perceived foreign accent and comprehensibility in the oral production of adolescent learners of English. Study Abroad vs. At Home learning contexts

Carmen del Río San Román

TESI DOCTORAL UPF / 2013

DIRECTORES DE LA TESI

Dra. Carmen Pérez-Vidal (Departament de Traducció i Ciències del Llenguatge, Universitat Pompeu Fabra)

Dra. Maria Juan-Garau (Department de Filologia Espanyola, Moderna i Clàssica, Universitat de les Illes Balears)

DEPARTAMENT DE TRADUCCIÓ I CIÈNCIES DE LLENGUATGE

Universitat Pompeu Fabra Barcelona

A mi madre y a mi padre

A Fernando

Acknowledgements

First and foremost I would like to thank my two supervisors, Carmen Pérez-Vidal and Maria Juan-Garau, for their trust in me and for their guidance and understanding throughout this process. Carmen, you are a wonderful motivator. I would not have written this dissertation without your encouragement and advice. Maria, thank you for your generous dedication and the time spent reading drafts of this dissertation. I really appreciate all the support you have given me from beginning to end.

I would also like to sincerely thank the heads of the school in Valencia, where the participants' data collection was carried out. Thanks are due to all student participants and teachers who made the project possible. Special thanks go to my friend Esperanza, head of the Foreign Languages department at this school, who helped me coordinating data collection for three academic years, and sorting out further questions related to this research. Thank you for your dedication and time, and also for your support and encouragement. Furthermore, I would also like to thank the teachers who volunteered their time and participated in the second collection data, making this study possible. Thank you for your good disposition, effort, dedication and comments.

Next I would like to thank many people who gave me good and generous advice during the design of second collection data. Special thanks go to Prof. Pavel Trofimovich from Concordia University for his generous sharing of materials and thoughts for designing the listeners' data collection for this research. I am also particularly indebted to Dr. Juan Maria Garrido (UPF), Dr. Lucrecia Rallo (UIB), and Jorge Rico (CSIC) for their generous assistance with the analyses of speech samples, which made the listeners' perception task design of this research possible. Special thanks go to Andrés Chandía (UPF), who voluntarily helped me with the online expriment design and all technical issues. I would also like to thank Dr. Joan Carles Mora (UB) for the time he dedicated to helping me with all statistical analyses (online and offline). With this regard, special thanks go to my colleague Pilar Avello. Pilar, thank you so much for your time and great efforts to explain things clearly and simply. I would not have been able to complete all this work without your dedication and help. I would also like to thank Dr. Natalia Fullana (UB) for all her advice during the data analyses. Thank you so much to other colleagues and researchers in the field who answered to my questions and shared their comments, previous experiences and suggestions with me, some of them without even knowing me personally. Thank you, Prof. Tracy Derwing (University of Alberta), Dr. Àngels Llanes (UDL), Dr.

Francisco Gallardo del Puerto (UNICAN), Prof. Nivja de Jong (Utrecht University) and Anastazija Kirkova-Naskova.

I have had the privilege to work and share my time and experience with so hard-working and generous colleagues at the UPF for the last two years. Margalida, thank you for your support and advice at the beginning of this period. Pilar, Rebecca, thanks for so many hours of dedication and work; for your comments about pilot studies conducted before the final one was ready. Thanks for your effort and patience. Thanks for being there. It has been a pleasure to share many hours with Elisa Barquin in the office and Isabel Tejada next door... Thanks for your encouragement and enthusiasm even in the long distance.

To my parents and sister, thank you for a lifetime of support, trust, encouragement, and unconditional love. Fernando, thank you for your patience, good humour, support, and love, especially during the final stages of this PhD. Finally, to Carmen and our second baby... great survivors of this experience.

Carmen del Río San Román
Barcelona, November 2013

Abstract

The aim of the present research study is to examine the contrasting effects of two learning contexts, study abroad (SA) vs. at home (AH), on the language development of Spanish adolescent learners of English. It focuses on the dimensions of foreign accent (FA) and comprehensibility in the target language. First, we explore learners' linguistic progress in these speech dimensions by assessing the impact of a 3-month SA programme on the extemporaneous speech of a group of 25 learners, and compare their results with those obtained by a control group of 31 learners receiving classroom instruction at home. For that purpose, speech samples were collected for each group longitudinally at two different points in time, before (Pre-test) and after (Post-test) the SA and the formal instruction (FI) context, respectively. A group of non-native listeners (n=12) were asked firstly to assign ratings to the samples by means of Likert scales, and then they were asked to report on the aspects that had affected their ratings most for each of the two speech dimensions analysed. Secondly, on the basis of this data, we explore the relationship between FA and comprehensibility by analysing the correlations between FA and comprehensibility scores at Pre-test and Post-test for the SA and AH participants and the information reported from the group of listeners. Results indicate that SA participants obtained significantly greater gains in FA than the AH group. The findings also suggest that the SA context was more beneficial than the AH context in terms of comprehensibility development, since the percentage of learners improving their comprehensibility scores during SA was significantly larger than the percentage of learners improving their scores in the AH context, and SA learners obtained larger comprehensibility gains than AH learners, although such improvement was not significant. In addition, results indicate that SA participants with initial lower levels of native-like accent and/or comprehensibility obtained significantly greater gains than their peers at home in both dimensions. Further analyses show significant large positive correlations between the two speech dimensions at the two testing times in the case of both groups of participants, suggesting that the more native-like accent, the greater the comprehensibility. Regarding the aspects that listeners heeded when rating speech samples, pronunciation is the main aspect taken into account when assessing L2 learners' FA and comprehensibility. Listeners in our study did not leave aside reference to accent or native-likeness in their comprehensibility ratings.

Resumen

El objetivo de este estudio es examinar el efecto de dos contextos de aprendizaje distintos, estancia en el extranjero (ES) y de instrucción formal en el aula, en el desarrollo lingüístico de adolescentes españoles aprendices del inglés. El estudio se centra en las dimensiones de acento extranjero y comprensibilidad del inglés como segunda lengua. En primer lugar, analizamos el progreso que los aprendices hacen en el grado de acento extranjero y de comprensibilidad examinando el impacto que tiene una estancia de tres meses en el extranjero sobre la producción oral de 25 aprendices, y comparando estos resultados con los obtenidos por parte de un grupo de control formado por 31 aprendices que reciben instrucción formal en el aula. Para ello, se recogieron muestras de habla de cada uno de los grupos en dos tiempos, antes y después de la estancia en el extranjero y del período de instrucción formal, respectivamente. Un grupo de 12 jueces no-nativos del inglés evaluaron el grado de acento extranjero y de comprensibilidad de las muestras de habla utilizando escalas de Likert, e indicaron los aspectos que más habían influido en las valoraciones realizadas para cada dimensión. En segundo lugar, a partir de estos datos, examinamos la relación entre acento extranjero y comprensibilidad. Por una parte analizamos las correlaciones entre las puntuaciones en ambas dimensiones de los dos grupos de participantes en los dos tiempos de recogida. Por otro lado, examinamos la información proporcionada por el grupo de jueces en cuanto a sus evaluaciones. Los resultados indican que el grupo ES obtuvo ganancias significativas mayores en cuanto al grado de acento extranjero, y que dicha estancia proporcionó también mayores beneficios en cuanto a comprensibilidad, dado que el porcentaje de aprendices que mejoraron sus puntuaciones en esta dimensión fue significativamente mayor que el porcentaje de alumnos que mejoraron en el contexto de instrucción formal, y dado que los resultados también señalaron mayores ganancias para el grupo ES, si bien esta mejora no resultó ser significativa. Asimismo, los resultados revelan que los participantes del grupo ES con un nivel inicial menor de acento nativo y de comprensibilidad obtuvieron ganancias significativas mayores en ambas dimensiones que los participantes en el contexto de instrucción formal. Asimismo, se hallaron correlaciones altas significativas entre el grado de acento y de comprensibilidad en los dos tiempos y para ambos grupos de participantes, sugiriendo que a mayor acento nativo, mayor comprensibilidad. En cuanto a los aspectos que más influyeron en las evaluaciones del acento y de la comprensibilidad de las muestras, los jueces señalaron que la pronunciación fue el factor más influyente en ambos aspectos. En sus valoraciones, los jueces relacionaron el grado de acento extranjero con la comprensibilidad de la producción oral en inglés de los aprendices.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AH:	At-home learning context
CLIL:	Content and Language Integrated Learning
CEFR:	Common European Framework of Reference
EFL:	English as a Foreign Language
L1:	First/native language(s)
FA:	Foreign Accent
FI:	Formal Instruction
NS:	Native Speaker
NNS:	Non-Native Speaker
SLA:	Second Language Acquisition
L2:	Second/non-native language
RP:	Received Pronunciation
SA:	Study Abroad
UIB:	University of the Balearic Islands
UPF:	Universitat Pompeu Fabra

Introduction

This research study examines language development in the oral production of English as Foreign Language (EFL) adolescent learners, and focuses on foreign accent (FA) and comprehensibility. Although many EFL learners view the development of their speaking ability as a priority, oral skills are not usually given sufficient attention in formal instruction settings. Given this situation, many EFL learners take part in study abroad (SA) programmes in order to develop or increase proficiency in a foreign language (especially oral skills) and learn about the culture. Moreover, learners may benefit from other positive side-effects deriving from this SA period. SA experiences may increase their motivation to learn the foreign language, foster socialization with other people, challenge their cultural perceptions and beliefs, and broaden their minds.

Since study abroad experiences are frequently encouraged for foreign language learners, and many families make an effort in sending their children abroad, the question of whether SA participants really benefit from this experience is one of practical interest. State-of-the-art accounts during the past two decades suggest that the learning context plays a crucial role in the process and outcome of second language (L2) learning (Collentine & Freed, 2004; Collentine, 2009; Freed, 1995a; Kinginger, 2009; Llanes, 2011; Pérez-Vidal, forthcoming), and the impact of SA periods on several areas of L2 development has been the focus of research in several studies.

To our knowledge, the impact of SA context on the development of foreign accent and comprehensibility has not been examined simultaneously in previous research and even less so with adolescent learners. The present study aims, in the first place, to fill this gap by examining longitudinal changes in FA and comprehensibility in the oral production of two groups of participants (SA and AH). The participants in the SA group consisted of 25 adolescent learners of English who were native Spanish speakers and took part in a 3-month SA programme. The AH group included 31 adolescent learners of English who were native Spanish speakers and followed classroom instruction at their home institution during the same period of time. We also compared FA and comprehensibility development and gains between the two groups of participants in order to determine whether one learning context was more beneficial than the other. We also take into account onset level when measuring progress in each of the contexts compared for both of the speech dimensions analysed. Within group variability may be as important as between groups differences.

A second contribution of the present study has to do with the participants' profile. So far most studies have used undergraduate students as participants in their research, and thus empirical evidence examining the effects of SA in the case of adolescent learners is scarce (Marriott, 1995; Fisher & Evans, 2000; Lapkin, Hart & Swain, 1995; Llanes & Muñoz, 2009). However, every year many adolescents take part in SA programmes with the objective of improving their language proficiency. Why have adolescent learners been left out of much of SA research? A possible reason may be the difficulty in obtaining data. Our study makes a contribution to this under-researched area by examining the effects of SA context involving adolescent EFL learners.

A third aim of this research is to explore in further depth the two dimensions of oral production that make up the focus of our study: foreign accent and comprehensibility. Previous research has supported the assumption that a non-native accent does not necessarily preclude successful communication (Derwing & Munro, 2009; Munro & Derwing, 1999). However, as noted by Trofimovich & Isaacs (2012), many educators, researchers, policy makers, and people in general equate non-native speakers' accents with their ability to communicate effectively. The present study aims to confirm previous results showing that a non-native accent does not necessarily impede successful communication, and support the claim that while related, accent and comprehensibility are two partially independent dimensions. We are interested in examining the relationship between the L2 learners' degree of foreign accent and comprehensibility, and identifying the aspects of L2 speech affecting perceived foreign accent and comprehensibility ratings reported by a group of listeners.

A fourth issue in our discussion deals with the listeners who took part in our study. In previous studies English native speakers have been asked to assess learners' FA and comprehensibility holistically (Gallardo del Puerto, 2005; Magen, 1998; Munro & Derwing, 1995; Trofimovich & Isaacs, 2012). English non-native instructors are frequently responsible for teaching EFL in the AH context in Spain and, consequently, in charge of assessing EFL learners' oral productions. For this reason, in this study the listeners were 12 native speakers of Spanish, who were teaching EFL at the secondary education level in Spain.

In line with Derwing and Munro's (2005) claim for future research, we believe more attention should also be given to the mutual intelligibility of non-native speakers of a language, especially at a time when English is increasingly used as an international language (Jenkins, 2000; Kachru, 1992). The debate of native versus non-native pronunciation, and the

(un)importance of achieving a native-like accent will be present throughout this dissertation. Moreover, a contrastive analysis between our non-native listeners' contributions and the ratings and comments provided by native-listeners in previous studies will be conducted, making it possible to establish a comparison between the two groups of listeners (native and non-native speakers of English) when rating EFL learners' foreign accent and comprehensibility.

In sum, we would like to point out that one of the reasons and motivations for this dissertation has been to contribute to the link between research and language teaching practice. In line with other researchers (Derwing & Munro, 2005), we find it necessary to pursue a mutually enriching and collaborative relationship between these fields. Thus, we would like to respond to the claims made by many recognized linguists and researchers who have asked for this greater collaboration between researchers and practitioners to encourage more classroom-relevant research. In this sense, we would also like to highlight the ecological validity of this research choice in that it reflects what really goes on in classrooms at least in the Spanish context and possibly in quite a number of other countries as well.

The dissertation that follows is divided into two parts and eight different chapters. The first three chapters constitute Part 1, and provide the theoretical background for the empirical study presented in Part 2, which includes the five remaining chapters.

Chapter 1 highlights the practical importance of SA research and reviews the growing body of research on language acquisition in this context, in general, and oral production, in particular. More specifically, we focus on studies that compare SA and AH learning contexts, and offer a review of the handful of studies involving adolescent participants. Furthermore, we explore the factors affecting oral production development in a SA context, paying special attention to participants' initial language level. Chapter 2 presents the definition and the factors affecting the dimensions of foreign accent and comprehensibility. Moreover, it offers a review of previous research in these speech dimensions, with a special focus on the design and method of these studies. Chapter 3 reviews some of the main issues in the field of pronunciation teaching today, paying special attention to the objectives of pronunciation instruction, problems in pronunciation teaching, and the assessment of pronunciation.

The remaining chapters constitute Part 2 of this dissertation and present the design and findings of our empirical study. Chapter 4 provides a brief introduction to the study and presents the objectives and research questions. Chapter 5 describes the design of this research study and gives

detailed information on the participants, the learning contexts, the data collection processes, and the analyses undertaken to address the research questions. Chapter 6 presents the results of the analyses conducted in order to respond to the three research questions. Results are organized around the three research questions formulated in Chapter 4. These results are then discussed in detail, and in relation to previous research, in Chapter 7. Finally, Chapter 8 presents the conclusions from this study and the implications of the results, and suggests topics for future research based on our findings.

Part I

Chapter 1

The Effects of Study Abroad on Adolescent Learners' Oral Production

One of the key elements affecting second language acquisition is context of learning, an idea which has been particularly emphasised from the field of study abroad research. A large number of studies during the past decades have suggested that the learning context plays a crucial role in the process and outcome of L2 learning (Coleman, 1998; Collentine & Freed, 2004; Collentine, 2009; Freed, 1995a; Kinginger, 2009; Llanes, 2011; Pérez-Vidal, forthcoming). As Collentine (2009) observes, learning context is one of the most important variables affecting the nature of second language acquisition and the extent to which learners acquire an L2. The quality and quantity of L2 input, and the type of practice of the L2 account for some of the differences between the two types of contexts considered in our research: study abroad (SA) and at-home (AH) learning contexts.

The typical AH learning context in Spain involving adolescent learners can be described as follows. In Spain English Language is a compulsory subject at school from the age of six. Learners are generally exposed to the L2 three hours a week, but the quantity and quality of the input and the interaction in the L2 is regulated by the teacher. The L2 classroom takes place in the home institution of the learners and input is usually limited to classroom hours. Students only interact with their teacher (usually not a native English speaker) and their peers.

SA experiences, on the other hand, provide the L2 learners with extensive and varied exposure to input, interaction, practice, and feedback in natural, real-life situations (Pérez-Vidal & Juan-Garau, 2011; Pérez-Vidal, forthcoming). Participants in SA programmes have been usually exposed to classroom teaching in their home country before studying abroad. The most common objective for study abroad participants is to develop or increase proficiency in a foreign language, but there are other positive side-effects deriving from this SA experience (e.g. fostering learners' motivation to learn the foreign language, socializing with other people, challenging their cultural perceptions and beliefs, and broadening their minds).

The question of whether SA experiences really benefit language proficiency is one of practical importance given the fact that L2 learners are usually encouraged to take part in SA programmes. The impact of SA periods on both linguistic and non-linguistic areas of L2 development has been the focus of research in numerous studies over the last two decades. Moreover, some of these studies have examined whether a SA context provides greater gains than regular classroom learning in terms of linguistic skills.

Most of the SA research into learning context has examined its effects on L2 development in adult learners (i.e. undergraduate students). Official numbers reported by the European Commission show that a great number of university students take part in these programmes: in the 2011-12 academic year, 252,827 students went to another European country to study or train, which represented a year-on-year increase of 9%[1].

No official figures are published reporting the number of children or adolescents taking part in SA programmes, but it is a fact that many parents and institutions believe in the importance of SA experiences for the development of proficiency in the foreign language, and every summer many adolescents are enrolled in SA programmes. However, very few institutions offer the students at secondary education levels the possibility to take part in a SA programme during the academic year (as it is the case of the participants in this research).

In the next section (section 1.1), we provide an overall picture of the growing body of research on language acquisition in the SA context,

[1] Every year, the European commission compiles statistics from the national agencies that run the Erasmus Programme in the participating countries and publishes an annual statistical overview online: http://ec.europa.eu/education/erasmus/doc/stat/erasmus1112_en.pdf

focusing on the studies that have examined the effects of SA experiences on learners' oral production, and looking at the results obtained when comparing SA and AH learning contexts. A review of the handful of studies involving adolescent participants will also be presented in section 1.2. Finally, section 1.3 will explore the factors which affect oral production development in a SA context.

1.1 The effects of SA on oral production development

Both non-linguistic and linguistic aspects related to L2 language learning processes and outcomes have been the focus of recent SA research. Several studies have analysed non-linguistic aspects, such as motivation, affective and cultural factors, and have documented the beneficial impact of SA experiences (Allen, 2010; Allen & Herron, 2003; Ismail, Morgan & Hayes, 2006; Trenchs-Parera & Juan-Garau, forthcoming). With regard to linguistic skills, studies examining listening comprehension have concluded that SA experiences are beneficial for participants' listening skills development (Allen and Herron, 2003; Cubillos, Chieffo & Fan, 2008; Dyson, 1988; Llanes & Muñoz, 2009; Beattie, Valls-Ferrer & Pérez-Vidal, forthcoming). However, gains after a SA period in comparison to other learning contexts are not so clear in the case of reading comprehension skills. While some studies have reported no significant differences in reading comprehension after SA experiences (Dewey, 2004; Davidson, 2010), others have (Lapkin, Hart & Swain, 1995). Likewise, mixed results have also been reported in the case of written production. Positive outcomes as a result of an SA experience have been documented in several studies (Barquin, 2012; Pérez-Vidal & Juan-Garau, 2009; Sasaki, 2004, 2007, 2009, 2011), whereas previous research by Freed, So & Lazar (2003) did not support this claim. A number of studies have reported lexical gains during SA experiences (Collentine, 2004; Dewey, 2008; Foster, 2009; Ife, Vives & Meara, 2000; Llanes, 2010; Llanes & Muñoz, 2009; Milton & Meara, 1995; Serrano, Tragant & Llanes, 2012). Mixed results have been reported in the area of morpho-syntax. While some studies have documented no significant gains after a SA period (Collentine, 2004; DeKeyser, 1991), others have provided evidence of the opposite (Howard, 2001, 2005). More recently, Juan-Garau, Salazar-Noguera and Prieto-Arranz (forthcoming) have examined the effects of FI and SA periods on the lexico-grammatical competence of 57 advanced learners of English, measured through cloze and sentence rephrasing tests. The data was collected at three data collection times: T1 (at the beginning of FI), T2 (after FI and prior to SA), and T3 (after a three-month stay in an English-speaking university). Results showed visible lexico-grammatical gains, with learners gradually approaching native-speaker performance levels after both FI and SA. The

authors emphasized the positive and complementary impact of FI and SA learning contexts on learners' grammatical development in English (Lord, 2010; Pérez-Vidal, 2011[2]). Finally, positive results have been obtained in the area of pragmatics and sociolinguistics (Kinginger & Farrell, 2004; Lapkin, Hart & Swain, 1995; Marriott, 1995; Regan, 1995, 2005), with some exceptions (Siegal, 1995).

Let us explore in greater depth the effects of SA on the linguistic area focus of our research: oral production. This has been the most investigated aspect in SA research, and it is the area which benefits the most from a SA experience (Davidson, 2010; Freed, 1995b; Lafford, 1995; Lennon, 1990; Llanes & Muñoz, 2009, 2013; Segalowitz & Freed, 2004; Valls-Ferrer, 2011). Different dimensions of oral production have been examined in previous studies (e.g. fluency, accuracy, complexity, narrative skills, pronunciation, foreign accent and comprehensibility). In the next pages a summary of this research is provided, setting the justification and contribution of our study, as well as the context for our discussion later.

Oral fluency, accuracy and complexity

One of the earliest studies investigating oral fluency changes during a SA period was Lennon's (1990) research. After examining oral production data collected from four German EFL learners before and after a 6-month period abroad, he concluded that there was significant improvement in terms of both perceived and utterance fluency. Since Lennon's study did not offer comparative data with an AH control group, it is difficult to determine whether the learners might have made a similar improvement with 6-months exposure to English at home. However, subsequent comparative studies examining oral fluency among undergraduates learning an L2 in different contexts have suggested that Lennon's findings were valid, and that SA contexts promote greater gains than AH contexts in terms of oral fluency (Freed 1995c; Freed, Segalowitz & Dewey, 2004; Juan-Garau & Pérez-Vidal, 2007; Llanes, 2010; Segalowitz & Freed, 2004; Trenchs-Parera, 2009; Valls-Ferrer, 2011).

One of the first comprehensive collections of research on language acquisition and context of learning is Freed's (1995a) volume, which offers two comparative studies supporting Lennon's results (Freed, 1995c; Lafford, 1995). Freed (1995c) analyzed Oral Proficiency Interview (OPI) data from 15 American students learning French as an L2 before and after

[2] In relation with the SALA project[2], Pérez-Vidal (2011) suggested the *Combination of contexts hypothesis*, which states that it is not one learning context or another in isolation that is most beneficial (i.e.: SA, FI, CLIL), but rather the combination of the three.

16 weeks studying abroad, and compared these results with those obtained by a group of 15 American students learning French as an L2 at home. She examined speech samples in terms of both utterance fluency and perceived fluency, and concluded that in the post-test data the SA participants obtained higher gains in terms of utterance fluency, and were perceived as more fluent by native listeners.

Lafford (1995) analyzed communicative strategies used by SA and AH learners. Her participants included two SA groups and one AH. All of the participants were American undergraduates from Arizona State University. The first SA group consisted of 13 American students who spent a semester in Mexico; the second SA group was made up of 16 students who spent a semester in Spain, and 16 students served as the control group (AH). In the post-test SA participants produced more words than the AH participants and used a broader range of communicative strategies, which she interpreted as an indication of greater improvement in oral fluency. In a later study, Lafford (2004) reported that the SA group used significantly fewer communication strategies than the AH group after the stay, which indicated that SA participants were able to maintain a conversation without recurring to these strategies as frequently.

Although their research did not offer data from a control group AH, Towell, Hawkins and Bazergui (1996) also reported significant improvement in fluency in the oral production of British undergraduates learning French after one year abroad. Likewise, Yager (1998) reported that 22 out of the 30 American undergraduates in his study improved significantly after seven weeks abroad in Mexico.

A special monograph on SA research[3] was edited by Collentine and Freed in 2004. In this volume evidence of the benefits of SA experiences in oral fluency was also reported by Segalowitz and Freed (2004) and Freed, Segalowitz and Dewey (2004). Segalowitz and Freed (2004) examined gains in oral fluency after one semester comparing the oral production by a group of 40 native English-speaking undergraduates learning Spanish in one of two different contexts (SA: n=22; AH: n=18). Their results showed greater gains in some aspects of fluency in the case of SA participants. In the second study, Freed, Segalowitz and Dewey (2004) analyzed various aspects of oral fluency by 28 American undergraduate students learning French in three different contexts: SA, AH, and an intensive summer immersion programme (IM). Their research also documented greater gains in several aspects of fluency for participants abroad in France than for

[3] Published in *Studies in Second Language Acquisition*

their AH counterparts, but this improvement was even greater in the case of participants in a domestic immersion programme.

In the same year Segalowitz et al. (2004) published their study comparing differences in the linguistic gains made by two groups of native English-speaking students from the United States learning Spanish in one of two different contexts of learning: SA (n=26) and AH (n=20). Data collection was carried out at the beginning and at the end of the semester (13 weeks). With regard to oral proficiency and fluency, the SA group showed gains whereas the AH did not, but they did not show greater gains in the pronunciation of some target words. Both SA and AH participants demonstrated gains in the pronunciation of some items but not in others after one semester, showing no advantage for the SA group.

More recently, Valls-Ferrer (2011) examined the development of L2 fluency and rhythm in the extemporaneous speech of 30 advanced L2 learners during a SA period. She documented that a 3-month SA period is highly beneficial for advanced EFL learners' development of L2 fluency and rhythm.

A more reduced number of studies have focused on the impact of SA on oral accuracy exclusively, yielding contradictory results. Collentine (2004), for instance, compared grammar in the oral productions of two groups of American undergraduates learning Spanish in two different contexts (AH vs. SA), and documented greater gains in the case of AH participants. In contrast, Isabelli and Nishida (2005), for example, examined the development of the Spanish subjunctive in subordinate clauses, measuring the frequency and accuracy of production in an oral interview task based on the OPI. The authors analyzed longitudinal development in a group of 29 SA participants who spent a year abroad in Spain. Data was collected at 3 different times: month 0 (prior to SA); month 4 of SA; and month 9 of SA. The authors also collected comparative, cross-sectional data from two groups of AH learners: 16 students at the end of their 5th semester of university-level Spanish, and 16 students at the end of their 6th semester of university level Spanish. Isabelli and Nishida (2005) reported no differences in the use of the subjunctive when comparing the AH learners in the 5th or 6th semesters of formal study. They found that SA participants showed significant improvement in the use of the subjunctive from month 0 to month 4, and then showed continued, though more moderate, improvement from month 4 to month 9. More recently, Juan-Garau (forthcoming) examined the effects on oral accuracy of a three-month SA period on 43 advanced-level English learners. Her study analysed the gains over their soujourn abroad and compared these results with the ones obtained from a preceding and

subsequent formal instruction environment. Her research also examined whether learners with a lower or higher pre-departure level made most progress in accuracy abroad. Results documented an advantage for the SA context, where learners became more target-like, with long-lasting effects. Students with lower pre-departure levels benefited the most from SA.

Several studies have examined fluency, accuracy and complexity measures in the same research. Allen and Herron (2003) examined the linguistic and affective outcomes of SA and reported significant gains in different dimensions. Two judges rated both listening and speaking tasks according to four criteria (amount of communication, comprehensibility, fluency, and quality of communication). The participants of this study were 25 undergraduate students learning French and enrolling in a summer SA programme in France. It is difficult to have a clear idea of what these authors meant by 'amount of communication' and 'quality of communication' since no definition or explanation is provided in their article. We have included their research in this section, however, since they concluded "grammatical correctness of speech" was "the least improved quality" (p. 382).

A handful of recent studies analyzing language development in Catalan-Spanish speaking EFL learners participating in ERASMUS exchanges have examined the impact of SA experiences on the acquisition of fluency, accuracy and complexity, presenting different results (Juan-Garau & Pérez-Vidal, 2007; Trenchs-Parera, 2009; Pérez-Vidal & Juan-Garau, 2011; Mora & Valls-Ferrer, 2012; Pérez-Vidal, et al. 2012). All of these studies offered comparisons between groups including the same subjects from the SALA project at different data collection times. Speech samples were collected at different points in time before and after a 6-month formal instruction (FI) period, and a 3-month SA period in an English-speaking country. The oral data was obtained from a guided interview conducted in pairs, and an oral role-play also in dyads. Juan-Garau and Pérez-Vidal (2007) examined overall oral proficiency, assessed as fluency, accuracy, formulaic speech and complexity in a role-play task. Their results confirmed the hypothesis that after a SA period participants improve their linguistic oral skills in terms of fluency (words per clause or sentence), accuracy (grammatical and lexical errors) and complexity (clauses per sentence and dependent clauses per clause). They reported substantial gains in oral fluency for the period spent abroad compared to no gains during the formal instruction period at home. More recently, Pérez-Vidal and Juan-Garau (2011) confirmed these results presenting further evidence of significant improvement in participants' oral fluency, accuracy and syntactic complexity after SA. They concluded that participants showed greater gains after the SA period than after the AH

period in terms of oral accuracy and formulaic language. Oral and written accuracy showed a steady non-significant increase throughout the AH period preceding the SA period, but grammatical and lexical complexity only improved after the SA period. Trenchs-Parera (2009) analyzed data from 19 Catalan/Spanish speaking EFL undergraduates and showed that after a short-term sojourn abroad, but not after an AH period, participants showed a significant improvement in overall oral performance. A recent study by Mora and Valls-Ferrer (2012) analyzed oral production and compared results obtained in SA and AH learning contexts. Results documented robust fluency gains after the SA, moderate improvement in accuracy and lack of gains in complexity, whereas no gains were reported after the formal instruction period AH. Pérez-Vidal, et al. (2012) explored the differential impact of a SA experience and a formal instruction context on learners' development of oral and written production. The participants in their research were a group of 29 Spanish undergraduate students (aged 17-25) taking part in a SA programme in an English-speaking country. As for the L2 learners' oral production, they examined measures of fluency, complexity and accuracy, and reported improvement across the three data collection times, particularly after the SA period. Their results showed a significant improvement in accuracy in oral productions after SA but not in grammatical complexity.

Serrano, Llanes and Tragant (2011) examined L2 written and oral production in terms of fluency, syntactic and lexical complexity, and accuracy at different data collection times in two different groups of learners. The SA group consisted of 25 undergraduates from Spain learning English in the UK. The AH group consisted of 106 participants receiving intensive or semi-intensive instruction in their home universities in Spain. The SA group obtained better results than the AH group in the post-test in fluency and lexical complexity.

In Serrano, Tragant and Llanes' (2012) research the participants were 14 Spanish-speaking undergraduate students (aged 20-24) who enrolled in a UK university for one year as part of the Erasmus European Exchange Programme. Oral and written data was collected at three points in time: at the beginning of their stay abroad around the last week of September (T1), before they returned to their home country for the Christmas Holidays (T2) and in the month of May (T3). The samples were analyzed in terms of fluency, syntactic complexity, lexical richness, and accuracy. Results showed that a few months abroad may result in gains in oral performance, but not in written production, which is slower and seems to take place after one semester abroad. They concluded that learners benefited first from the SA context in terms of fluency and lexical richness (T2), and this progress extended to accuracy some time later (T3). Their analysis

revealed that a few months abroad might be sufficient for some gains in oral performance to occur.

Narrative skills

Another aspect related to students' improvement in their oral production is narrative abilities. Some studies exploring fluency and proficiency as well as grammatical features have reported improvement in narrative ability (Collentine, 2004; Isabelli-García 2003). Collentine (2004) examined the development of grammar and lexical abilities comparing two groups of American undergraduates: SA vs. AH. The SA group consisted of 26 learners of Spanish enrolled in a SA programme in Spain, while the AH group consisted of 20 participants. The author documented that the AH context led to more grammatical and lexical gains, whereas the SA group achieved better narrative abilities and produced more semantically dense language. For her part, Isabelli-García (2003) examined the oral performance of three undergraduate American students who enrolled in a 5-month SA programme in Argentina. The objective of her research was to measure the development of two aspects of oral communication skills (OCS): fluency and performance in the oral functions of narration, description and supporting an opinion. Results showed that "the OCS that were developed in the semester abroad were mostly those corresponding to: 1) uncomplicated, basic communicative tasks, such as a simple narration and description, and 2) increase in utterance length" (Isabelli-García, 2003: 171).

Oral accuracy: pronunciation

Pronunciation development during a SA programme has been the focus of research of a modest number of studies. Although SA has been considered an optimal L2 learning context for the development of oral skills, research has not provided consistent evidence of significant benefits of this learning context in the area of pronunciation accuracy.

Simões (1996) analyzed as a measure of 'fluency' "the number of accurately pronounced syllable nuclei in sequences of words found acceptable" in the Spanish discourse of a group of five adult learners of Spanish taking part in a 5-week SA programme in Costa Rica. He reported significant changes in two out of five participants' productions. Lord (2010) observed the effects of the SA context in the production of Spanish plosives by a group of 8 learners of Spanish taking part in an 8-week SA programme in Mexico. Half of the participants received a Spanish phonetics and pronunciation course prior to their SA (instruction group) whereas the other group did not (control group). Lord concluded that SA

promoted gains in L2 phonological acquisition, especially in the case of the instructed group. Therefore, she pointed out that it was the combination of the two contexts that was most beneficial for phonological development (Pérez-Vidal, 2011; Juan-Garau, Salazar-Noguera and Prieto-Arranz, forthcoming). However, the number of participants in these studies do not allow for generalizations. Moreover, neither Simões (1996) nor Lord (2000) contrasted the gains obtained by a SA group with those experienced by a control group receiving formal instruction at their home institution. Thus, it remains unknown whether or not the SA context was more beneficial than other learning settings.

Studies involving a larger number of participants and examining pronunciation development learning an L2 in different contexts appeared some years later (Díaz-Campos, 2004; Højen, 2003; Mora, 2008). These studies brought different results regarding the impact of SA on pronunciation development. Højen (2003) examined the oral productions of 14 young adult native speakers of Danish who spent 3-11 months in Southern England. Participants reported that they had had no (or limited) amount of exposure to conversational English before the SA experience. The author reported that production at the segmental level did not improve significantly. Díaz-Campos (2004) compared the performance of 26 American undergraduates taking part in a 10-week SA programme in Spain and a group of 20 American undergraduates taking Spanish classes in the regular programme at university. His results showed that the SA participants did not outperform the AH group. Actually, the results from this research showed striking differences between the two groups. However, in a follow-up study, Díaz-Campos (2006) found significant differences in pronunciation, with SA learners performing better than AH participants. Mora (2008) analyzed the perception and production of sound units by a group of Spanish/Catalan bilingual learners after a formal instruction period at their home university, and after a 3-month SA programme. Learners showed gains in their ability to auditorily discriminate between English contrastive sound units after the two different contexts of learning, although it was only after the AH instruction period that these gains were significant. As in Díaz-Campos (2004), significant gains in perception were not greater after the SA term than after the AH period. As for the learners' ability to produce English voiceless oral stops, participants showed a slight (non-significant) improvement after the SA period, while no effect of formal instruction was found. Avelló, Mora and Pérez-Vidal (2012) examined the impact of a 3-month SA period on the pronunciation of 23 Catalan/Spanish undergraduates learning English as an L2 by means of phonetic measures at the segmental level, and stress. A significant decrease of pronunciation errors was found in the learners' speech production after SA. In contrast,

no strong effect of SA was documented by Avello and Lara (forthcoming), who analyzed two groups of undergraduate Spanish/Catalan learners of English participating in a 3-month and 6-month SA programme, respectively, and found no significant improvement in segmental production accuracy.

Foreign accent and comprehensibility

Studies interested in examining the effects of SA learning context on the development of foreign accent and comprehensibility are quite recent and still scant (Allen and Herron, 2003; Avello, Mora and Pérez-Vidal, 2012; Højen, 2003; Martinsen, 2010). Contrary to research on other dimensions of L2 learners' oral productions, none of these studies offers a comparison between the SA group and a control group AH.

Allen and Herron (2003) analysed linguistic and affective outcomes after a summer SA programme in the oral productions of 25 American university students learning French. The authors reported gains in comprehensibility after a SA period. This was "the most improved quality" of the four dimensions ("amount of communication", comprehensibility, fluency, and "quality of communication") rated by two judges. As previously mentioned, Højen (2003) examined pronunciation accuracy in the oral productions of 14 young adult native speakers of Danish who spent 3-11 months in Southern England. Moreover, he also analysed FA development in the participants' productions, which were rated by 10 native English listeners. Højen (2003) reported better perceived FA ratings after the SA period. For his part, Martinsen (2010) asked the raters in his research to evaluate pronunciation and comprehensibility of L2 learners' speech, along with other language aspects (i.e. fluency, grammar, and vocabulary). The participants in this study were 45 American university students who participated in a summer abroad programme in Argentina. The raters gave a score for each of these dimensions according to a rubric created by the researcher, and a final punctuation for the speaking task was obtained. This final punctuation served to compare gains in overall speaking proficiency in the speech of 45 American undergraduate students after a 6-week SA programme in Argentina. His research, however, does not analyze the pronunciation aspect and the comprehensibility dimension in particular, but offers a general picture of oral language development based on the overall ratings. Overall results show that participants in the SA programme had modest yet highly significant improvements in their oral language skills over the course of their SA. As previously reported, Avello, Mora and Pérez-Vidal (2012) examined the impact of a 3-month SA programme on a set of phonetic measures in the oral productions of 23 Catalan/Spanish

undergraduates learning English. In the same study, the authors also assessed perceived FA development in the participants' productions. Thirty-seven proficient non-native listeners rated the L2 learners' speech samples. The authors found a slight, non-significant improvement in perceived foreign accent after SA.

The impact of a SA context on the development of the speech dimensions of foreign accent and comprehensibility has not been examined simultaneously in any previous study. The present research aims to fill this gap by exploring the development of these two aspects after a SA period, and comparing the results with those obtained by a group of learners staying AH.

1.2 The effects of SA on adolescent learners

Studies presented so far have examined the effects of SA on adult learners' oral production. SA research examining the effects of the learning context on other age groups of participants is scarce. In the next pages, we present a review of research on language acquisition involving children and adolescents in a SA abroad context.

To our knowledge, the only published studies exploring the effect of SA context on children are the ones conducted by Llanes (2012), and Llanes and Muñoz (2013). Llanes (2012) examined short- and long-term effects of a SA period on children aged 11. The participants were 16 Catalan/Spanish learners of English. Nine participants enrolled in a 2-month SA programme in Ireland, and seven remained AH. Oral and written data was collected from the two groups of participants at three different points in time: (1) one week before the SA group departure to the target language country (pre-test); (2) the week after their return from the host country (post-test); (3) 12 months after the post-test had been administered (delayed post-test). SA participants showed greater gains than AH participants. These gains seem to benefit oral skills rather than written skills in the short term. As for the long-term effects, almost none of the variables analyzed turned out to be statistically significant from the post-test to the delayed post-test. However, SA participants obtained higher scores in the delayed post-test in most of the variables. These findings suggest that the L2 gains obtained in an SA experience have a long-term effect confirming Regan's findings (2005), and contrary to results reported by Pérez-Vidal and Juan-Garau (2009) and Howard (2009).

Llanes and Muñoz (2013)[4] conducted a study examining the oral and written production of two groups of children and two groups of adults learning English as an L2 in two different contexts: SA and AH. Measures of fluency, lexical richness, complexity and accuracy were analyzed from the oral and written samples at two different times by children and adult participants in both contexts of learning. The adult learners were aged 19-33, and the group of children were 10-11 years and were studying 5th / 6th year of Primary Education in Spain. Results showed that both children and adults performed better after the SA than after the AH period, particularly in the case of oral skills, and found that children benefited more from the SA context than adults as measured in comparative gains. However, adults outscored children in absolute gains.

Research examining the effect of SA context on the language development of adolescent learners is also scant and studies usually report results after short periods abroad (e.g. school exchanges or short summer stays). Moreover, all studies involving adolescent learners report results from a SA group of participants, with no AH control group.

To our knowledge, the first published studies with adolescents participating in SA programmes are those of Marriott (1995) and Lapkin, Hart and Swain (1995). Marriott (1995) studied the acquisition of politeness forms in Japanese by a group of eight Australian adolescents and young adults spending one academic year in Japan. The learners lived with Japanese host families and attended regular school with native Japanese students. Therefore, they had a large number of opportunities for exposure to the L2, with a lot of contact with native speakers. Marriott reported considerable gains in polite formulaic expressions after the SA experience. Lapkin, Hart and Swain (1995) considered 119 English-speaking adolescents from three Canadian provinces who participated in a bilingual exchange programme. Results showed considerable gains in dictations and speaking after three months in Quebec. Like the students in Marriott's research, the participants in this programme lived with their host families and attended their secondary schools, so they had contact with the host community and many opportunities to interact with native speakers.

Fisher and Evans (2000) examined the effects of a school exchange experience on learners' language proficiency and attitudes. The SA group consisted of 68 British adolescent learners of French (aged 13-14) spending from 6 to 11 days abroad. Participants completed listening, reading, speaking and writing tests before and after the exchange. Results

[4] The data used in this piece of research came from previous work by Llanes (2010).

showed considerable gains in language proficiency after the SA period, particularly in listening and writing skills. Lesser gains were made in speaking, and gains in reading were not discernible. Production data also revealed important gains in grammatical aspects.

Llanes and Muñoz (2009) also reported large gains in fluency for students who spent 3-4 weeks abroad. They analyzed the oral production of 24 Catalan/Spanish adolescent and young adult (from 13 to 22) learners of English through oral fluency and accuracy measures. Considerable gains on most measures were reported. Moreover, the authors also claimed that proficiency level affected the intensity of learners' progress, as participants with initial lower proficiency level showed comparatively greater gains in vocabulary and oral accuracy and fluency.

As we can see foreign accent and comprehensibility are two dimensions which have not been examined together in any previous study involving adolescent learners in a SA context. The present study aims to fill this gap in SA research, analyzing the effects of SA on these dimensions and comparing the results with those obtained by a control group AH.

1.3 Factors affecting oral production development in a SA learning context

As we have seen not all empirical learning-context research has suggested significant gains for students going abroad over students staying at their home institutions receiving formal instruction (Collentine, 2004; DeKeyser, 1991; Díaz-Campos, 2004; Freed, Segalowitz & Dewey, 2004; Mora, 2008; Segalowitz et al., 2004; Serrano, Llanes and Tragant, 2011). Despite the popular belief that the most efficient way to learn an L2 is spending time abroad, SA contexts have not been found uniformly more beneficial for language development than the AH context.

Factors other than the learning context may affect the learning process: the length of stay abroad (Dwyer, 2004; Ife, Vives & Meara, 2000, Schmidt-Rinehart & Knight, 2004), individual differences (Brecht, Davidson & Ginsberg, 1995; Kinginger, 2008; Segalowitz & Freed, 2004; Siegal, 1995; Simões,1996), language contact and practice while abroad (Brecht & Robinson, 1993; Freed, Segalowitz & Dewey, 2004; DeKeyser, 2007; Isabelli, 2001; Juan-Garau & Pérez-Vidal, 2007; Segalowitz & Freed, 2004), the degree of proficiency prior to the SA or initial proficiency level (Brecht, Davidson and Ginsberg, 1995; Freed, 1995b; Ife, Vives & Meara, 2000; Lapkin, Hart & Swain, 1995; Llanes & Muñoz, 2009; Milton & Meara, 1995; Segalowitz & Freed, 2004), and the design of the SA programme (2012; forthcoming).

Studies such as Milton and Meara (1995) and Martinsen (2010) offer an insight into different factors which may affect language development while abroad to some extent. For her part, Pérez-Vidal and Juan-Garau (2009) and Pérez-Vidal (2011) suggested a characterisation of SA based on three parameters: the 'macro-level features', the 'micro-level features' and the 'programme features'. In the next pages we will describe this characterisation in order to better understand the variables at play in oral production development in a SA learning context.

1.3.1 The 'macro-level features' of SA

The 'macro-level features' of SA consist of the external features that define the SA learning context and differentiate it from other contexts, such as formal instruction or Content and Language Integrated Learning (CLIL). Second language acquisition (SLA) has identified input, output, feedback, interaction and practice as some specific context-dependent or external factors which influence L2 development and characterize SA learning context.

Participants in SA programmes are exposed to higher-quality input than their AH peers. Moreover, SA participants have the opportunity to practice the L2 more than their AH peers, as the former get more opportunities for output and meaningful interaction, which takes place in a wider variety of contexts, with a larger number of interlocutors (including native-speakers).

Cognitive approaches to L2 development within SLA research, such as skill acquisition theory, have emphasized the importance of practice for L2 acquisition development. Skill acquisition theory distinguishes three main stages in skill development: declarative, procedural and automatic (DeKeyser, 2007a). Practice has a crucial role in transforming declarative/explicit knowledge into procedural /implicit knowledge, which can be later transformed into automatized knowledge.

In the declarative stage, the L2 learner is presented with explicit information on the structure, rules and patterns of the L2 ('knowledge that'). In the next stage, initial practice allows the learner to act on this declarative knowledge and use it so that it becomes a behavioural routine, or 'procedural knowledge' ('knowledge how'). Once declarative knowledge has turned into procedural knowledge and after a large amount of varied practice, procedural knowledge may become automatized. As noted by DeKeyser (2007c), automatization is a complex concept, which involves a decrease in "the time required to execute a task (reaction time),

the percentage of errors (error rate), and the amount of attention required (and hence interference with/from other tasks)" (DeKeyser, 2007c: 98-99).

SA and AH learning contexts represent different stages of this declarative-proceduralized-automatized knowledge path towards L2 acquisition. A formal instruction setting AH typically involves large amounts of declarative knowledge, limited real language contact and few opportunities for meaningful interaction in the L2. Productive oral skills often have a secondary place in AH contexts, where productive written skills and receptive skills are the focus. Many learners and teachers confirm the fact that oral skills are less practiced in the EFL classrooms, where most of the time is devoted to other aspects of the language (grammar, vocabulary, writing and reading). In contrast, the SA context offers the kind of input and practice opportunities which may enhance the proceduralization –and ultimately automatization– of declarative knowledge previously acquired in the AH context.

DeKeyser (2007b) posited that learning an L2 abroad provides more opportunities for practice in real-life situations, leading to the automatization of L2 skills. Formal instruction contexts, on the other hand, foster the development of declarative knowledge to a larger extent. DeKeyser (2007b) claimed that participants need to have gone through minimal proceduralization in the classroom before going abroad and that "the transition in skill acquisition that *should* coincide with going abroad is automatization" (p.217). To benefit fully from the SA context of learning, learners should have acquired "functional knowledge of the grammar that is assumed to be known at an intermediate level" (DeKeyser, 2007b:217) so that the process of proceduralization can be completed and make progress towards automatization.

Participants in SA programmes usually have some declarative proceduralized knowledge of the L2, and after numerous hours of practice, they automatize certain aspects of the L2. Automatization may account for the better performance of SA participants compared to AH learners. L2 learners usually have fewer opportunities to practice the L2 when they are AH, especially as far as oral practice is concerned. This may explain why AH participants in many studies above mentioned are outperformed by SA participants in oral production measures. SA participants usually take less time to carry out an oral narrative task, for instance, as reported by Llanes and Muñoz (2013). Practice may also explain why a SA context does not enhance learners' writing skills to the same extent (Serrano, Tragant & Llanes, 2012), since SA learners may have very limited practice in L2 writing while abroad. L2 learning

contexts vary in terms of the quantity and quality of the L2 input they provide, and the opportunities they offer for learner's output and interaction with other speakers (native and non-native).

In line with DeKeyser (2007b), Llanes and Muñoz (2013) pointed out that the improvement in oral skills may be partly explained by the role of practice in L2 learning. They also referred to the Interaction Hypothesis (Long, 1981) and the Output Hypothesis (Swain, 1985) in order to explain why AH participants in their study obtained lower oral gains than SA participants. Language learners need to be active learners when receiving language input: only listening to new language structures will not lead to successful language learning, which requires input and output.

However, practice in a SA context may be more reduced than expected. Learners' ability to actively engage in meaningful interactions with native speakers will affect their success in taking advantage of opportunities for optimal practice while abroad. This idea relates to SA micro-level features, which will be discussed in the following sub-section.

1.3.2 The 'micro-level features' of SA

The 'micro-level features' refer to the individual learner characteristics that interact with the external features and determine learner's success in their L2 development (e.g. age, personality, aptitude, motivation, and ability for taking advantage of L2 contact opportunities). Learners' involvement in the L2 context can facilitate opportunities for interaction, and thus gains in overall language development. However, the degree of learner involvement and, therefore, the degree of contact with the L2 and interaction with other L2 speakers while abroad may be affected by different factors. Some learners interact with L2 speakers more than others and the success in this endeavour is related to learner characteristics (e.g. openness, anxiety, and motivation). Recent research on SA has focused on individual differences, since they may affect the quantity and quality of interaction with other native speakers throughout the SA period. An extended revision of the literature on motivation, anxiety, and learning strategies is offered by Dufon and Churchill (2006).

On the other hand, the type of learning context may have an effect on some of these variables, such as motivation and anxiety. Fisher and Evans (2000:15), for instance, reported that in the L2 country, learners "felt they were able to use a French accent more confidently and reported that speaking in a French accent became almost involuntary". This contrasts with situations in formal instruction AH, where learners may not feel comfortable speaking the L2 with a native-like accent. Once immersed in

the L2 environment, learners feel less embarrassed about making mistakes and less anxious about speaking (Coleman, 1996; Fisher and Evans, 2000), so more opportunities for L2 practice arise.

Several studies have explored learners' interaction with their host families and outside of the home (Campbell, 1996; Kinginger, 2004; Kinginger & Farrell, 2004; Kinginger and Whitworth, 2005; Knight and Schmidt-Rinehart, 2002; Levin, 2001; Li, 2000; Schumann, 1997; as cited in Dufon and Churchill, 2006). Adaptation to the L2 environment is considered a key element in this process leading to interaction. Students' living arrangements and their influence on L2 development have been analyzed in some studies (Rivers, 1998; Schmidt-Rinehart & Knight, 2004) and results have shown that more interaction leads to more practice and greater insights into the culture. Other studies have claimed the host culture itself may affect the learners' approach to interaction (Isabelli-García, 2003; Kinginger and Whitworth, 2005; Twombly, 1995; Wilkinson, 1998; among others). Previous studies have shown that students' attitude toward the people of the target country affects language learning (Coleman, 1996; Fischer and Evans, 2000; Gardner and Lambert, 1972; Schumann, 1976).

Learners' readiness and ability to take advantage of language contact and practice opportunities has been identified as an important factor for SA success. This may explain why a given context may be more advantageous for some learners' than others. In this sense, it is not the contexts *per se* but the quantity and quality of contact within the context that promotes language development and gains (Freed, Segalowitz & Dewey, 2004).

Apart from contextual and individual factors, Pérez-Vidal (2011) also notes that the design of the SA programme may play a role in the gains learners obtain during the SA period. In the following section, some of the most important features of SA programmes are presented.

1.3.3 The 'architecture' of the SA programme

Pérez-Vidal (2011) proposes a set of features that should be taken into consideration in the design of a SA programme, given that they may influence on SA outcomes. The author identifies eight main programme variables: (1) length of stay abroad, (2) L2 language setting, (3) 'other' target language contact (leisure, extra-curricular...), (4) pre-departure language level, (5) pre-departure preparation, (6) point in the curriculum, (7) academic out-of-class assignments and (8) re-entry conditions at the home institution.

As can be seen, some of the variables are related to academic programme aspects (e.g.: learners' pre-departure preparation, the point in the curriculum at which SA takes place, academic out-of-class assignments, and re-entry conditions at the home institution). The remaining factors are linked to L2 input conditions (e.g. length of SA, L2 language setting, and L2 contact) and initial language level, which refers to learners' onset level or L2 proficiency before the SA period. Let us focus next on the latter for the purpose of our study.

1.3.3.1 Initial language level

Learners' involvement in interaction may also be conditioned by their degree of proficiency in L2 at the beginning of the SA period. Students with higher-proficiency of the language may find it easier to interact with other L2 speakers and native speakers in the L2 context, whereas learners with a lower degree of competence may have difficulties in understanding native-like speech and in being understood.

Some programmes require learners to have achieved a specific level of proficiency in the L2 before the SA period. This is associated with the 'threshold hypothesis', which formulates that there may be a threshold level in L2 development at which learners benefit more from SA programmes. As previously pointed out, DeKeyser (2007) suggested that learners should have a specific L2 proficiency level in order to fully benefit from a SA period. In other words, learners should have a proficiency level that is high enough to enable them to participate in meaningful interactions so as to benefit from large amounts of practice that is required to automatise previously acquired procedural knowledge. Likewise, Collentine (2009) pointed out that SA participants must have reached a certain 'threshold level' of proficiency prior to going abroad in order to take full advantage of the input and opportunities for learning. On the other hand, it seems that once participants have 'crossed' the threshold, the higher students will gain relatively less, in accordance with the normal learning curve.

Previous research has documented that learners with higher pre-departure scores obtained more gains in oral performance after the SA period (Brecht et al, 1995; Segalowitz & Freed, 2004). Brecht et al. (1995) reported that grammar and reading achievement scores before the SA were significant predictors of gains in speaking proficiency, listening and reading after the SA. They reported positive correlations between grammar and reading scores before the SA and SA gains in speaking, listening and reading. Likewise, Segalowitz and Freed (2004) examined a

wide range of cognitive and linguistic factors which would affect gains during SA. They found that a certain level of ability in word recognition and lexical access processing abilities was necessary for oral proficiency and fluency to develop abroad significantly.

However, results from a considerable body of research have concluded that learners with lower proficiency benefit more from SA periods. Although learners' with a lower initial level may have a harder time at the beginning of their SA period, they may obtain more gains from such an experience than students with a higher level or in AH context.

Carroll's study (1967) was the first to suggest that a short SA experience was beneficial for the improvement of the L2, and that such an SA learning experience favoured weaker learners. This claim has been supported by a number of studies exploring oral production development in a SA context (Freed, 1995c; Juan-Garau, forthcoming; Lapkin et al., 1995; Llanes & Muñoz, 2009; Marriott, 1995; Regan, 1995; Valls-Ferrer, 2011). Milton and Meara (1995) reported that those students obtaining greater linguistic gains during their SA were those who had started at relatively low levels of fluency and suggested sending students abroad when they are at these comparatively low levels of fluency. Likewise, Lapkin et al. (1995) reported negative correlations between SA gains in all four linguistic skills and pre-departure scores, indicating that those learners with lower initial level obtained greater gains. Marriott (1995) confirmed the hypothesis that learners with initial lower proficiency made the greater gains, and Freed (1995c) also reported that less advanced proficiency level students improved slightly more in fluency at the end of the SA programme. In the same line, Regan (1995) concluded that contact with natives did not make a great difference to the acquisition of certain structural elements in the case of advanced learners. More recently, Llanes and Muñoz (2009) documented that SA participants with lower initial proficiency level showed comparatively greater gains in using L2 words (vocabulary) and in producing more accurate and fluent speech. Valls-Ferrer (2011) also reported larger gains in fluency during the period abroad for those learners who had started with a relatively lower fluency level. Juan-Garau (forthcoming) documented that the L2 learners who had performed less accurately before the SA period were generally the ones to exhibit more gains upon return.

In line with Llanes' (2011) claim, it is important to shed some light on this controversial issue because of the pedagogical and economical implications for SA programmes:

> "If we knew what the ideal pre-departure L2 level is, we could prepare our learners of the L2 so that they were equipped with the necessary

> knowledge and skills before going abroad. In this way, they would benefit more from their SA and, therefore, this would imply a more efficient use of money and time as many participants have reported when they have come back and feel their L2 skills have not improved". (Llanes, 2011:108).

As we have seen throughout this chapter, different language-learning contexts can lead to variability in gains (or no gains) in particular aspects of oral production. Contexts differ in terms of the learning opportunities they offer, but it is also worth remarking also the importance of what the L2 learners bring to the learning context. Individual differences and learners' prior linguistic and cognitive knowledge will somehow affect the potential benefit they may obtain from the language development opportunities provided in each context.

1.4 Contribution of the present study

Throughout the last two decades many studies have analysed the impact of a SA period on L2 learners' oral production. As previously seen in our review of the literature in section 1.1 of this chapter, most of them have examined oral fluency development, and a reduced number of studies have examined gains in other aspects associated with the speaking ability (i.e. phonological accuracy, lexico-grammatical aspects, communicative competence, narrative abilities). Research on the impact of SA learning context on the speech dimensions of foreign accent and comprehensibility is scant. The first objective of this research is to contribute to this area by examining the effects of a SA experience on the L2 learners' foreign accent and comprehensibility. Moreover, we will compare the results obtained by the participants in the SA with those obtained by a control group remaining at their home schools in Spain (AH).

Section 1.2 has offered an overall picture of SA research involving adolescent learners. Again, to our knowledge, no studies have analysed foreign accent and comprehensibility development in adolescents'L2 oral productions during a SA experience. Therefore, the second objective of this dissertation is to offer a first study involving adolescent learners who have taken part in a SA programme, and compare their results with those obtained by adolescent learners staying AH.

The factors involved in L2 oral production development in a SA context have been discussed in section 1.3 of this chapter. Our aim is to contribute to the research on initial learner proficiency and explore the impact of SA on the speech dimensions of foreign accent and comprehensibility in the case of teenagers with different initial levels. Is foreign accent and

comprehensibility development different for participants with different initial levels?

In short, the present study aims to shed new light on the SLA field with a longitudinal study examining the development of the speech dimensions of foreign accent and comprehensibility in adolescent learners in two different learning contexts.

Chapter 2

Foreign Accent and Comprehensibility

As we have seen in the previous chapter, research on the oral production of L2 learners has examined several dimensions of non-native utterances, such as fluency, accuracy, complexity, narrative skills, pronunciation, foreign accent and comprehensibility. In this study we will focus on the speech dimensions of foreign accent and comprehensibility.

In the first two sections of this chapter we will present the definitions of foreign accent (FA, also referred to as *accentedness* in the literature) and comprehensibility adopted in this study. We will then explore the factors influencing these dimensions in L2 learners' oral productions. In section 2.3 we will review previous research in these domains, and the contribution of our research will be presented in the last section of this chapter.

2.1 The construct of foreign accent

Lippi-Green (1997:42) defines accents as "loose bundles of prosodic and segmental features distributed over geographic and/or social space", and highlights the fact that every speaker has an accent (either of his/her L1, or other languages he/she may speak). Thus, Lippi-Green considers the distinction between two kinds of accents: L1 accent and L2 accent. L1 accent refers to the native variety of English spoken: "every native speaker of US English has an L1 accent, no matter how unmarked the

person's language may seem to be" (Lippi-Green, 1997:43); and L2 accent is considered "the breakthrough of native language phonology into the target language" (Lippi-Green, 1997:43).

A well-known and recognized framework of reference on foreign accent research is the series of studies conducted by Munro and his colleagues (Derwing & Munro, 1997, 2009; Munro and Derwing, 1995a, 1999). In these studies (and others) accentedness is defined as the listeners' perception of how closely the pronunciation of an L2 utterance resembles that of a native speaker of English. Munro (1998) provided the following definition of foreign-accented speech:

> "nonpathological speech produced by second language (L2) learners that differs in partially systematic ways from the speech characteristic of native speakers of a given dialect" (Munro, 1998:139).

The most clear indication that someone is an L2 user is the production of foreign accented speech. Whereas an L2 learner can avoid certain phenomena in syntax or vocabulary when speaking, avoidance is virtually impossible in phonology. Speech research suggests that people tend to be highly sensitive to even tiny divergences from the pronunciation patterns of their L1. Flege (1984) proved that phonetically untrained listeners were able to detect FA in short samples of speech, including phrases, words, individual segments, and even segment portions of speech as short as 30 milliseconds in length. For their part, Munro, Derwing and Burgess (2003) found that listeners could identify accentedness even in a single word presented backward.

Foreign accent is a complex phenomenon of the language that affects speakers and listeners in both perception and production. Within the EFL/L2 setting, a foreign accent may have several consequences for the speaker. Discriminatory treatment and attitudes toward accented speakers have been documented in a wide range of studies over the past years (see reviews in Lippi-Green, 1997; Munro, 2003). More recently, Lima (2011) explored American undergraduates'perceptions of International Teaching Assistants (ITAs) language competence, and revealed that students tended to react negatively to ITAs even when the ITAs were highly proficient in English. Most complaints referred to poor English language proficiency and/or communicative competence, but other factors played a role in this perception (e.g. age, gender, nationality, teaching style, and accent, among others) (Lindemann, 2002; Rubin, 1992).

There are different reasons for reacting negatively to an accent. Several studies have shown that some native speaker listeners may activate prejudices against a particular group of people when they hear speech

patterns associated with that group. This is what Munro (2003) refers to as *accent stereotyping.* The discrimination of L2 speakers due to their accent seemed to foster the rise of accent reduction programmes promising to help L2 learners reduce or eliminate their foreign accent. However, is foreign accented speech really something problematic in itself? Some studies have claimed that accentedness should be subject to treatment, intervention, or even eradication in much the same way as a language pathology:

> "The goal of instruction in pronunciation is that the student (or patient) should learn to speak the language as naturally as possible, free of any indication that the speaker is not a clinically normal native" (Griffen, 1991:182).

Another negative consequence of foreign-accented speech is that it may lead to a loss of intelligibility. Listeners may experience difficulty in understanding speech that differs from the patterns of oral production to which they are accustomed. Moreover, accented speech may demand special effort for comprehension, so listeners may not develop a favourable attitude towards listening to their interlocutor. As some studies have stressed, listeners may take longer to process L2-accented speech than native speech (Munro & Derwing, 1995b). Other studies have analysed to what extent impatience, irritation, inexperience with accented speech or prejudice on the part of the listener account for negative reactions towards accented speech (Derwing, Rossiter & Munro, 2002; Kang & Rubin, 2009; Lindemann, 2002; Rubin, 1992; Wrembel, 2010).

Although accented speech can have consequences for communication, including reduced intelligibility, research exploring the relationship between foreign accent and comprehensibility has reported that heavily accented speech can often be perfectly intelligible (Derwing & Munro, 1997; Munro & Derwing, 1995a, 1999).

Another interesting question is formulated by Lippi-Green (1997) when she analyzes the objective of some L2 instructors, teachers, and even L2 learners who aim at losing their foreign accent when they speak an L2:

> "Putting aside the question of personal freedoms protected under the law, putting aside the issues of social identity, is it true that it is not within the power of the individual to change their language?" (1997:45).

We all have accents, native and non-native, and it is possible to adjust an accent to some degree. We may reduce the degree of foreign accent, but we cannot remove it. It is true that some learners are better at acquiring an L2 pronunciation, and we may have heard of cases of people, who as

adults acquire a second variety of their L1, or another language with excellent native accent. However, this seems to be restricted to a very small number of highly motivated individuals (Moyer, 2004) and to learners with special aptitude (Ioup, Boustagi, El Tigi, & Moselle, 1994). More frequently we find L2 speakers who never lose their foreign accent, but are able to speak and be understood, as we find native speakers of a language with marked regional or social accents. Given this situation, the next question we will address in the following section of this chapter is why some L2 learners acquire a more native-like accent than others.

2.1.1 Factors affecting foreign accent

When addressing the question of which factors affect the degree of foreign accent, we should take into account aspects related to the speaker and the listener. A great number of studies have been concerned with the identification of variables affecting degree of accentedness and the study of the relationship between these variables and perceived accentedness (see Hayes-Harb & Watzinger-Tharp, 2012; Piske, MacKay and Flege, 2001, for an overview).

Overall degree of foreign accent is likely to vary as a function of different factors directly related to the speaker (the L2 learner), such as: age of learning (and age of arrival in the L2 country), length of residence, quantity and quality of L2 contact, formal instruction, motivation, gender, L1/L2 language use, and cognitive abilities. In the next pages we provide an overview of the main findings related to individual speaker factors affecting degree of foreign accent.

Age of L2 learning

Age of L2 learning has been considered an important variable affecting the degree of a native-like accent acquisition (Flege, Munro and MacKay, 1995; Fullana, 2005; Major, 2001). The limited ability to acquire a native-like pronunciation at some time in our lives is explained by the critical period hypothesis, which establishes that virtually everyone who acquires an L2 after a certain critical period will have a foreign accent.

One of the first allusions to the optimal age to learn an L2 can be found in Penfield's work during the 1950s (as cited in Dechert, 1995:79), claiming that the approximate ideal age ranged from nursery school to before age 10 or 14. A decade later, Lenneberg (1967) supported the existence of an optimal period for L1 acquisition which extended from about age 2 to age 12. The Critical Period Hypothesis was first proposed for L1 acquisition but it is also claimed to exist for L2 acquisition (Long 1990; Neufeld,

1979, 1980, 1988; Patkowski, 1990, 1994; Scovel, 1969). Although Lennneberg concentrated on L1 language acquisition, he also documented the fact that L2 development usually ended after the ages of 12 or 13. He suggested a relationship between foreign accent and starting age of L2 learning, whereby a child learning to speak the L2 at the age of 3 or 4 was able to speak it with no foreign accent. Children starting to learn the L2 at puberty would fail to speak it accent free. Thus, according to Lenneberg (1967) learning to speak an L2 between the ages of 4 and 12 enabled the learners to speak the L2 accent free.

It is now widely accepted that the acquisition of an L2 after early childhood results in speech that differs from that of the native speakers (Long, 1990; Oyama, 1976; Scovel, 1988). As previously mentioned, most adults who learn an L2 will speak it with an accent. The main argument supporting this view is that knowledge of the sound system of the L1 influences the perception and production of the phonetic patterns of the L2 (Flege, 1995; Werker & Polka, 1993). Research in favour of a critical period indicates that when acquiring L2 phonology, the younger the learner, the more native-like the pronunciation (Flege, 1991; Flege, Munro & Mackay, 1995; Munro, Flege & Mackay, 1996; Piske, MacKay & Flege, 2001). However, the age when the critical period ends is still debated. Researchers show considerable disagreement on the cut-off age. Long (1990) claimed it is 6 or 7, and Flege (1991) concluded the age of L2 learning at which foreign accents become perceptible is long before puberty (between the ages of 5 to 8). In contrast, Patkowski (1994) suggested it is slightly later (around 15 years), and Scovel (1988) indicated it lasts through puberty.

Counterevidence to the widespread belief that children typically pronounce foreign languages without an accent has also been documented (Seliger, Krashen & Ladefoged, 1975; Flege, 1988; Flege, Frieda, & Nozawa, 1997). In addition, older learners have been judged native-like when producing speech samples in the foreign language in some studies (Bohn & Flege, 1992; Bongaerts, 1999; Bongaerts, Planken & Schils, 1995; Bongaerts, Van Summeren, Planken & Schils, 1997).

A number of studies have examined the effect of age of onset in formal instruction on the degree of foreign accent in Spanish EFL learners (Fullana, 2006; Gallardo del Puerto, García Lecumberri & Cenoz, 2005; García Lecumberri & Gallardo del Puerto, 2003). García Lecumberri and Gallardo del Puerto (2003) explored foreign accent and intelligibility in a group of 60 Basque schoolchildren who differed in starting age of exposure to English (4, 8 and 11 years). On the basis of ratings assigned by a native English listener, they concluded that older students were

perceived as having more native-like accent and being more intelligible. They posited that early starting age was not a variable which promoted L2 sounds acquisition in a formal instruction learning context, since they found that participants who had started to learn English later obtained higher scores in FA and intelligibility. In other words, 11-year-old beginners were perceived as producing less accented and more intelligible than 8-year-old and 4-year-old starters. In a later study, Gallardo del Puerto, García-Lecumberri & Cenoz (2005) confirmed previous findings (García Lecumberri and Gallardo del Puerto, 2003) on the perception of English phonemes in formal contexts. Their study reported an inverse relationship between foreign accent and age of L2 learning. In line with Singleton & Ryan (2004), older learners in their research outperformed younger learners. Fullana (2006) examined FA in the oral production of 281 Spanish learners of English in a formal setting, who differed in terms of onset age of L2 learning and in amount of formal exposure to the L2. Listeners in her research were 7 female native speakers of General Canadian English, who assessed participants' production of English segments in an imitation task. She found that age of onset of L2 learning and exposure "were not conclusive determinants for perceiving and producing English sounds in a native-like manner in a formal language learning context" (p. 59).

Length of residence

Apart from age of learning, previous research has also focused on the relationship between L2 experience or amount of L2 exposure and degree of foreign accent. Most studies on FA have analysed long-term immersion contexts, so the variable of L2 exposure has been typically indexed as length of residence (LOR). Studies have offered conflicting results. Whereas a number of studies have reported an influence of LOR on degree of L2 foreign accent (Asher & García, 1969; Flege & Fletcher, 1992; Flege, Munro & MacKay, 1995; Purcell & Suter, 1980), others have not found an effect of LOR (Flege, 1988; Moyer, 1999; Oyama, 1976; Thompson, 1991).

Piske, MacKay & Flege (2001) suggested that LOR effects on perceived FA may depend on learners' initial proficiency level. Thus, LOR might have an initial and rapid effect on the pronunciation of low-level learners, but in the case of more advanced learners, LOR exposure would be unlikely to result in a significant FA reduction (Flege, 1988).

Quantity and quality of L2 contact

Flege and Liu (2001) reminded us that LOR in itself may not be an accurate index of L2 exposure, and that it is actually the amount of contact with native speakers that plays an important role. In line with this claim, Højen (2003) found a correlation between LOR and FA gains in L2 English during a SA period. This correlation was stronger between FA gains and an overall measure of total input which combined LOR and use of English while abroad (as reported by learners). He proved that it is not additional time spent in an L2 context that fosters native-like pronunciation, but the amount and quality of L2 input. As Piske (2007:306) noted, "progress in learning and L2 is dependent on both the quantity and the quality of the L2 input L2 learners receive".

Formal instruction

Some studies have suggested that many L2 production difficulties stem from perception. In this sense, appropriate perceptual training can lead to improvement in production (Bardlow, et al. 1997). Other studies have supported the claim that global (prosodic) or segmental content in classroom instruction fosters improvement in perceived accentedness and comprehensibility (Derwing, Munro & Wiebe, 1998).

However, research examining the influence of formal instruction on degree of foreign accent has not identified any instructional variables as significant predictors of learners' degree of foreign accent (Thompson, 1991; Flege, Munro & MacKay, 1995). Piske, MacKay & Flege (2001) remarked that the reason why instructional variables seem to have had little effect on degree of L2 foreign accent may be the little attention L2 pronunciation has received in most foreign language classrooms. Instructional variables may have a larger effect on degree of L2 foreign accent if the learners receive special training in L2 pronunciation:

> "L2 pronunciation receives little attention in most foreign language classrooms. This might explain why instructional variables seem to have had so little effect in the studies just cited" (Piske, MacKay and Flege, 2001: 200).

Given the fact that traditional teaching methods in FI seem to be limited in terms of learners' language development in general, other learning contexts, such as CLIL, appeared as an alternative to promote gains in the L2. A few studies have examined the impact of CLIL instruction on English pronunciation in Spanish learners of English (Gallardo del Puerto et al., 2009; Rallo & Juan-Garau, 2011). Gallardo del Puerto et al. compared two groups of Spanish-Basque learners, one receiving only FI

and another receiving both FI and CLIL instruction. Participants' oral productions were rated by native English listeners in terms of FA, intelligibility and FA irritation. CLIL participants were found to be significantly more intelligible and less irritating than non-CLIL learners, but no significant differences were found in terms of FA. Rallo and Juan-Garau (2011) examined the effects of CLIL context on intelligibility and foreign accent after one year of CLIL instruction, and compared the results obtained with those from a group of learners receiving formal instruction, and from a group of native-speakers. The speech samples produced by CLIL participants (n=64) were judged to be more intelligible and less accented than the samples by their peers in formal instruction (n=42), although both groups differed significantly from the native English speakers (n=15). However, no significant short-term improvement in intelligibility or FA was reported after one year of CLIL instruction, suggesting that one year of CLIL instruction might be insufficient to improve learners' oral production in L2.

Motivation

Some studies have documented the influence of the motivational variable on foreign accent development (Suter, 1976; Purcell & Suter, 1980). These studies reported 'strength of concern for L2 pronunciation accuracy' as a significant predictor of degree of L2 foreign accent. In contrast, other studies have concluded that factors such as 'professional motivation', 'integrative motivation' and 'concern for L2 pronunciation' have not automatically led to accent-free L2 speech (Flege, Munro & MackKay, 1995; Bongaerts, Summeren, Planken & Schils, 1997; Moyer, 1999).

Concern for L2 pronunciation can be developed during learners' interactions in the L2, especially if they experience communication breakdowns due to lack of pronunciation accuracy. In this sense, SA contexts may provide the L2 learners with an excellent setting for pronunciation awareness, even though gains in L2 pronunciation in SA contexts are also limited (as seen in Chapter 1 of this dissertation). With regard to formal instruction contexts or AH settings, Mora (under review) suggests that "the very limited exposure to authentic unaccented input as well as the scarce opportunities for meaningful interaction in the L2 typically prevent learners from developing the L2 phonetic system sufficiently to speak the L2 with good pronunciation" (see also Fullana, 2006; Gallardo del Puerto et al., 2009; Mora & Fullana, 2007; Piske, 2007).

Gender

Mixed results have been reported from studies examining the effect of gender on FA development. While some studies have documented a significant influence of gender on degree of FA (Asher and García, 1969; Thompson, 1991; Flege, Munro & MacKay, 1995), others have not identified this variable as a significant factor influencing degree of L2 foreign accent (Suter, 1976; Purcell & Suter, 1980; Flege & Fletcher, 1992).

L1/L2 language use

Several studies have concluded that language use patterns are also an important factor in L2 FA, especially regarding learners' use of their L1 while immersed in the L2 learning context (Flege, Munro & MacKay, 1995; Flege, Frieda & Nozawa, 1997; Piske et al., 2001). In these studies, L2 learners had to estimate L1 and L2 use in terms of the amount of time they spend using their L1 and L2 in different contexts, the amount of contact with L2 native speakers, or L1 and L2 proficiency.

Flege et al. (1995) concluded that language use patterns were an important predictor of foreign accent ratings for Italian learners of L2 English. In a follow-up study (Flege et al., 1997), the role of L1 use was further explored by creating two groups of early Italian/English bilinguals who were AOL-matched, but who differed in percentage of L1 use. The authors indicated an L1 use effect as the learners with higher L1 use were perceived to have a significantly stronger FA than the learners with lower L1 use. Piske et al. (2001) found that amount of continued L1 use (and age of L2 learning) affected degree of FA.

Cognitive abilities

Mora (under review) claims that L2 speech research has focused primarily on the investigation of age and experience-related effects on L2 speech perception and production to the detriment of other variables such as differences in cognitive ability. Do some people have a special aptitude for producing unfamiliar speech sounds? Some studies have investigated the extent to which mimicry ability (i.e. the ability to mimic unfamiliar speech sounds) affects degree of L2 foreign accent production and have concluded that this ability is a significant factor (Suter, 1976; Purcell & Suter, 1980; Thompson, 1991; Flege, Yeni-Komshian, & Liu, 1999). Other abilities, such as musical ability, have not been yet identified as affecting degree of L2 foreign accent (Piske, MacKay & Flege, 2001).

So far we have examined factors related to individual speakers which may affect the degree of FA in their L2 speech. Another factor influencing accentedness is the particular utterance features in the L2 utterances produced by the learners.

Some studies have examined the relation between the number of segmental errors in the L2 utterance and overall prosodic accuracy, and the perception of FA. Research has established that foreign accent ratings correlate with the frequency of segmental and prosodic divergences from typical native speaker patterns (Anderson-Hsieh, Johnson and Koehler, 1992; Brennan and Brennan, 1981; Magen, 1998; Munro and Derwing, 1998, Munro and Derwing, 2001). Magen (1998) examined the speech of native speakers of Spanish talking in English. Syllable structure, final /s/ deletion, consonant manner, and lexical and phrasal stress were aspects influencing accentedness ratings by native English listeners. In line with these results, Brennan and Brennan (1981) reported a significant correlation between accentedness and the frequency of several types of pronunciation errors identified in the speech samples; while Anderson-Hsieh, Johnson and Koehler (1992) found that measures of deviance from native English speech in segment production, prosody, and syllable structure affected pronunciation ratings. Speaking rate has also been reported to affect accentedness judgments in Munro and Derwing (1998; 2001) research. More recently, Trofimovich and Baker (2006) found that suprasegmentals contributed to foreign accent. More recently, Trofimovich and Isaacs (2012) also posited that accent is linked to aspects of pronunciation.

Speech sample properties other than acoustic aspects can also affect accentedness and intelligibility. Kennedy and Trofimovich (2008) reported that semantic context (i.e. listeners' familiarity with the topic of speech) affected listeners' ratings for accentedness, comprehensibility and intelligibility. These results supported Gass and Varonis' (1984) findings, which concluded that semantic context affected the intelligibility of L2 speech.

A third group of factors have been identified as affecting FA perception. Listener properties, such as language background, familiarity with the learner's L1, and cognitive abilities, may also affect the degree of perceived FA.

A number of studies have examined the relationship between listeners' L1 background or familiarity with the type of L2 speech they are listening to and the perceived degree of FA (Derwing and Munro, 2013; Flege, 1988; Gallardo del Puerto, Gómez Lacabex & García Lecumberri, 2007;

Kennedy & Trofimovich, 2008; Mackay, Flege and Imai, 2006; Munro, Derwing and Morton, 2006). Flege (1988) reported strong similarities in accent ratings of Taiwanese, Mandarin, and English native listeners who rated Mandarin-accented English speech. Mackay, Flege and Imai (2006) also documented a close relationship between native English speakers' and Arabic speakers' ratings of Italian accented English speech. Likewise, Munro, Derwing and Morton (2006) found that non-native listeners from diverse L1 backgrounds (Cantonese, Japanese, Spanish, and Polish) evaluated non-native speech samples from L2 speakers of those same languages similarly. Gallardo del Puerto, Gómez Lacabex and García Lecumberri (2007) compared FA, intelligibility and FA irritation ratings by two groups of listeners: native speakers of English who had little linguistic/phonetic training and non-native listeners, who were native speakers of Spanish, proficient in English and experienced teachers of English and phonetically/linguistically trained. With regard to FA ratings, results revealed that native listeners and non-native listeners rated FA similarly. Likewise, Kennedy and Trofimovich (2008), analysed comprehensibility and accentedness ratings of English utterances spoken by 6 English and 6 Mandarin speakers. The listeners in their research were 24 native speakers of English: twelve were half experienced with L2 speech, and the remaining twelve were not. Kennedy and Trofimovich (2008) concluded that both groups of listeners did not differ in their FA ratings of L2 speech samples. More recently, Derwing and Munro (2013) corroborated these findings. In their study, 10 non-native listeners from a diverse range of L1 backgrounds who were high proficiency L2 speakers of English and 34 native speakers of Canadian English responded to FA of the L2 speech samples in similar ways.

On the other hand, Isaacs and Trofimovich (2011) examined whether individual differences in phonological memory, attention control, and musical ability affected listeners' judgments of L2 speech. They concluded that only listeners' musical ability had a significant effect on listeners' accentedness ratings of non-native speech.

2.2 The construct of comprehensibility

Many L2 researchers and teachers would agree that intelligibility is the main aim in oral communication and L2 pronunciation instruction. Actually, the main objective of L2 learners in most of the cases is being able to communicate, and to be understood, rather than accent reduction (Derwing & Munro, 1997; Jenkins, 2000; Munro, 2008; Pennington & Richards, 1986). In most previous research a distinction between the constructs of intelligibility and comprehensibility has been made (Derwing & Munro, 1997, 2009; Munro & Derwing, 1995a, 1999). In

these studies intelligibility is defined as the extent to which a given utterance is understood by a listener, whereas comprehensibility refers to listeners' perception of how easily they understand an utterance.

In the present study, we have chosen the term 'comprehensibility' to refer to the construct which some studies have identified as 'intelligibility'. This choice is in line with Isaacs and Trofimovich (2012), who considered Levis' (2006) distinction between broad and narrow definitions of intelligibility to explain this choice. In its narrow sense, intelligibility refers to listeners' actual understanding of L2 speech (Munro & Derwing, 1999). It is often measured by examining listeners' accuracy of orthographic transcriptions of L2 speech, although other methods have also been used (e.g., comprehension questions, true-false statements, reaction times). In its broad sense, intelligibility is defined as listeners' ability to understand the speech and "is not usually distinguished from closely related terms such as comprehensibility" (Levis, 2006, p. 252) and has been typically measured through listeners' ratings of how easily they understand speech (Munro & Derwing, 1999).

A further reason for focusing on comprehensibility stems from the fact that several oral proficiency scales make use of the term 'intelligibility' (e.g. TOEFL, IELTS) when they are actually measuring listeners' comprehensibility, since they are asked to make subjective judgements through scalar ratings. As Isaacs and Trofimovich (2012) pointed out it is actually comprehensibility (not intelligibility) that is being assessed.

Therefore, in line with Trofimovich and Isaacs (2012), the construct of comprehensibility in this research falls under Levis' broad sense of intelligibility and reflects a common approach to assessing intelligibility in oral proficiency scales. The terms 'comprehensibility' and 'intelligibility' will be used interchangeably in this dissertation under this definition.

As previously mentioned, producing comprehensible speech is more than a matter of pronunciation. While it is true that some errors in pronunciation may affect speech comprehensibility, foreign accented speech does not impede comprehensibility. In the next section we will explore the aspects affecting the comprehensibility of speech, and previous research dealing with this issue.

2.2.1 Factors affecting comprehensibility of L2 oral speech

Research exploring the variables affecting the comprehensibility of L2 speech has focused on the utterance properties, and listener factors. Some studies have examined the impact of non-target-like suprasegmental patterns on speech comprehensibility and have reported the negative effects of suprasegmental errors on this dimension (Derwing and Munro, 1997, 2005; Hahn, 2004; Zielinski, 2008). Research has also indicated that segmental errors may affect comprehensibility (Munro and Derwing, 1995a; Munro and Derwing, 2006).

As previously discussed in section 2.1.2, speech sample properties other than the acoustic ones can also affect accentedness and comprehensibility. Kennedy and Trofimovich (2008) reported that semantic context affected listeners' ratings for accentedness and intelligibility. These results were in line with previous research conducted by Gass and Varonis (1984) who claimed that semantic context affected the intelligibility of L2 speech.

More recently, Isaacs and Trofimovich (2012) studied in greater depth the construct of comprehensibility, and explored the aspects of speech that affected L2 comprehensibility at different ability levels. Based on the analysis of 19 quantitative speech measures, listeners' judgments and introspective reports, five speech measures were identified that distinguished between L2 learners at different comprehensibility levels: "lexical richness and fluency measures differentiated between low-level learners; grammatical and discourse-level measures differentiated between high-level learners; and word stress errors discriminated between learners of all levels" (Isaacs and Trofimovich, 2012:476). Thus, not only pronunciation features of a foreign-accented speech, but also other language aspects affect speech comprehensibility (e.g vocabulary, grammar, discourse measures). These results are in line with Varonis and Gass (1982) whose previous research concluded that grammar and pronunciation interact to influence overall intelligibility. With regard to discourse features, Tyler (1992) also reported that a number of discourse factors influenced the overall comprehension of speech. In her research, a Mandarin speaker and a native English speaker prepared lectures which were later delivered by a native English speaker to university students. Discourse factors were identified as playing a significant role in overall comprehensibility.

On the other hand, the degree of comprehensibility does not reside exclusively in the L2 speaker or in the utterance itself. Listener characteristics also affect L2 speech comprehensibility. Thus, familiarity

with a topic, familiarity with non-native speech in general (and non-native accent in particular), and familiarity with a particular speaker have also been reported to influence listeners' comprehension of accented speech (Bent & Bradlow, 2003; Derwing & Munro, 1997; Gallardo del Puerto, Gómez Lacabex & García Lecumberri, 2007; Gass & Varonis, 1984; Hayes-Harb, Smith, Bent & Bradlow, 2008; Kennedy & Trofimovich, 2008; Wingstedt & Schulman, 1984). Gass and Varonis (1984) analyzed the effects of familiarity on native speakers' comprehension of L2 accented speech. Transcription tasks showed that familiarity with the topic, non-native speech, a particular accent, and a particular speaker all had an effect on intelligibility. In the same way, Wingstedt and Schulman (1984) (as cited in Derwing & Munro, 1997) also reported that familiarity with a particular accent facilitated comprehension, a claim which was also confirmed by Derwing and Munro's (1997) study. Likewise, Bent and Bradlow (2003) found that the relative intelligibility of each speaker depended on the language background of the listener. Hayes-Harb et al. (2008) concluded that having the same L1 background as the speaker can offer comprehension benefits particularly in the case of low-proficiency listeners. Gallardo del Puerto, Gómez Lacabex and García Lecumberri (2007) found that native-speakers and non-native listeners rated L2 learners' speech differently in terms of intelligibility. Non-native listeners judged participants' accents more intelligible than native judges. The authors argued that factors such as familiarity with Spanish/Basque-accented English speech and Spanish/Basque knowledge accounted for these findings as L2 learners' oral productions sometimes contained Spansih and/or Basque code-switching.

Kennedy and Trofimovich's study (2008) yielded mixed results regarding the influence of listener experience on the different speech dimensions. Whereas listeners with more experience in their exposure to L2 speech understood more speech samples from native and non-native speakers of English than listeners with less experience, they did not rate them differently in comprehensibility and accentedness.

In the same way, Munro, Derwing and Morton (2006) found that listeners with more exposure to particular L2 accents were not better at understanding speakers with those accents than those listeners who were less familiar with them. In more recent contributions, Munro, Derwing and Holtby (2012) and Munro (2013) suggested that there is no indication that sharing an L1 background equals large benefits for comprehension. Listeners from different L1 backgrounds tend to comprehend L2 speech similarly. Derwing and Munro (2013) also stressed the correlation between ratings from native and non-native listeners from a diverse range of L1 backgrounds.

Jun and Li (2010) examined the factors reported by three NS listeners and three NNS listeners when rating 7 ESL learners' speech for comprehensibility and accentedness. NNS raters reported segmental features more often that NS raters. They were also more aware of other pronunciation aspects such as intonation and linking. NS raters referred to comprehensibility more often than NNS while rating for accentedness. The NS listeners reported aspects such as 'attention' and 'extra effort' needed for comprehension and 'being frustrated' because of the incomprehensibility of an L2 speech sample. The NNS raters noticed specific pronunciation features more often than NS raters, while NS raters were more global in their assessments, paying special attention to speech comprehension.

Munro, Derwing and Holtby (2012) also remark that listeners' age may affect comprehension of L2 speech. In their research they found a significantly better performance by teens than younger children in comprehension of accented speech, suggesting "a beneficial effect of cognitive maturity on the comprehension of accented speech" (p. 238).

Mora (under review) suggests that cognitive skills such as phonological short-term memory, acoustic memory, and attention control may play an important role at the various stages involved in the processing of L2 speech, such as listeners' ability to perceive differences in degree of accentedness in L2 speech (Isaacs and Trofimovich, 2011). Mora (under review) reports that the only factor identified as contributing significantly to listeners' assessment of non-native speech is musical ability. Other cognitive variables such as phonological memory or attention control have not been found to play a role in listeners' subjective evaluations of speech (Isaacs & Trofimovic, 2010, 2011).

2.3 Research in foreign accent and comprehensibility: a focus on design and methods

The objective of this section is not to provide an extensive review of most published research on accent and comprehensibility, which has recently been done by Munro & Derwing (2011). Indeed, a considerable number of these studies have also been referred to in the first pages of this chapter. The main focus of the following pages is to explore what previous research has accomplished and those decisions which have been made regarding three main aspects: participants (L2 speakers, and listeners); tasks, and rating scales. Thus the design and methods of previous studies on accentedness and comprehensibility will be the focus in this section. These are questions which were thoughtfully considered for our study, and the decisions made with this regard have somehow conditioned the

type of results we have obtained, and will be referred to in the discussion in Chapter 7.

Most research in the field of FA and comprehensibility has been conducted with immigrant populations in learning contexts of long-term immersion in the L2 community (Derwing & Munro, 1997; Derwing, Munro & Wiebe, 1997, 1998; Isaacs & Trofimovich, 2012; Jun & Li, 2010; Kennedy & Trofimovich, 2008; Munro & Derwing, 1995, 1999; Trofimovich & Isaacs, 2012). A reduced number of studies have examined accentedness and comprehensibility in a formal instruction setting in Spain (Fullana, 2006; Gallardo del Puerto, Gómez Lacabex & García Lecumberri, 2007; Gallardo del Puerto, García Lecumberri & Cenoz, 2005; García Lecumberri & Gallardo del Puerto, 2003; Rallo & Juan-Garau, 2011), and a few more in periods of shorter immersion, such as those characteristic of SA learning contexts (Allen and Herron, 2003; Højen, 2003; Martinsen, 2010; Avello, Mora and Pérez-Vidal, 2012). Moreover, a group of native speakers are usually included to provide baseline data.

Most of these studies have a cross-sectional design in which oral data is collected at a single point in time. Common data elicitation techniques include controlled tasks (e.g. lists of words, sentences or paragraphs read aloud) or extemporaneous speech (e.g. picture story telling or personal narrative tasks). The use of controlled tasks has the advantage of controlling for grammatical or lexical errors which may affect listeners' judgements, and facilitates analyses in terms of comparisons between L2 speakers' productions (or of the same speaker at different points in time). However, as noted by Munro (2008:202) "the use of these controlled tasks may result in unnatural or better-than-normal speech". Samples of extemporaneous speech may thus also be collected so as to obtain more natural sounding oral data in a less controlled setting:

> "Unlike many other studies, ours used extemporaneous utterances rather than excerpts from reading passages or sentence stimuli. As a result, we examined accent and intelligibility under circumstances that better reflect naturally occurring speech" Munro & Derwing (1995a:92).

In L2 speech research, accentedness has typically been quantified by eliciting listeners' scalar judgements of the degree of foreign accent of a given speech sample (Munro & Derwing, 1999). This construct has often been measured on 9-point Likert scales (Derwing, Munro & Wiebe, 1997; 1998), although smaller scales have been used (e.g. 7-point in Avello, Mora & Pérez-Vidal, 2012; and Anderson-Hsieh, Johnson & Koehler, 1992; or 5-point in Hahn, 2004, and Rallo & Juan-Garau, 2011). The use of 9-point scales has also been generalized to measure comprehensibility.

Generally, only scalar endpoints are defined in these Likert scales (e.g. *extremely difficult to understand – extremely easy to understand*, for comprehensibility). As noted by Isaacs and Thomson (2013), despite the absence of descriptors in mid-scale range, ratings derived using 9-point scales have consistently reported high interrater reliability in empirical studies, even for inexperienced listeners (Derwing, Thomson & Munro, 2006).

Isaacs and Thomson (2013) explored the effects of rating scale length on listeners' judgments of L2 speech, and found no differences in mean scores obtained using 5-point versus 9-point scales. In addition, reports from listeners assigned to one of these rating scales revealed that while 5-point scales were too constraining for some raters, listeners in the 9-point condition were unable to meaningfully differentiate between so many scalar points. In line with their conclusions, the use of one rating scale or another is sample dependent, and what should be a priority is a more concrete definition of the constructs we aim at operationalising in the rating scales. In line with Isaacs and Thomson's (2013) suggestion, a focus on listeners' perceptions is required to understand holistic constructs in L2 pronunciation research that are defined in terms of listeners' perceptions. In particular, we should "investigate the nature of these perceptions, how they relate to the L2 speakers' productions, and linguistic and non-linguistic influences on listeners' judgements" (p. 157).

A frequent reflection among researchers using rating scales for measuring different speech dimensions is that even though listeners assign the same rating score to a speech sample, their rationale for doing so may be different. In other words, quantitatively equivalent ratings do not mean that raters have interpreted the construct in the same way, or taken the same aspects into account in the decision-making task. Listeners may assign the same scores for different reasons (Douglas, 1994, Turner & Upshur, 2002).

Almost all L2 pronunciation studies involving raters report some measure of interrater reliability, but only a few have explored listeners' own interpretation of the construct that they are rating (Brown, Iwashita & McNamara, 2005), or used qualitative methods to elucidate listeners' processes as they relate to L2 speech dimensions (Harding, 2008; Rossiter, 2009; Zielinski, 2008).

Apart from the rating scale length, individual listeners' characteristics may affect the rating process. As pointed out in previous sections, several studies have explored the influence of listeners' features, such as

linguistic background, experience with L2 accented speech, or familiarity with the type of L2 speech they are asked to rate.

Listeners' FA and comprehensibility ratings have been traditionally collected from groups of native speakers (Derwing & Munro, 1997; Isaacs & Thomson, 2013; Munro & Derwing, 1995a, 1999; Trofimovich & Isaacs, 2012). Some studies have analysed the perception of L2 speech by non-native listeners (Hayes-Harb & Watzinger-Tharp, 2012) and a number of studies have included both native and non-native listener groups for L2 speech assessment purposes (Avello, 2013; Derwing & Munro, 2013; Rossiter, 2009; Valls-Ferrer, 2011). The objective of exploring the perception of L2 learners speech by different groups of listeners is related to concerns regarding to what extent listeners' L2 speech perception may vary as a function of speech sample properties, or as a function of listener differences (Munro, 2008). This is an interesting issue which relates to pronunciation assessment, and the question of whether L2 proficient non-native speakers are as valid as native-speakers to assess L2 learners' oral performance. The evidence available from the studies analysing the perception of L2 accented speech by groups of native and non-native listeners with different L1 backgrounds seems to indicate that although factors such as prejudices and familiarity can influence L2 speech rating of accentedness and comprehensibility, stimulus properties are likely to be more relevant than listeners' factors (Derwing & Munro, 2013). In line with these results, research has indicated that native and proficient non-native listeners respond to aspects of L2 speech samples in similar ways (Derwing & Munro, 2013; Flege, 1988; Gallardo del Puerto, Gómez Lacabex & García Lecumberri, 2007; Kennedy & Trofimovich, 2008; MacKay, Flege & Imai, 2006; Munro, 2013; Munro, Derwing & Holtby, 2012; Munro, Derwing & Morton, 2006). On the other hand, some recent studies have indicated that non-native listeners may be more severe judges than the native listeners (Avello, 2013; Derwing & Munro, 2013; Valls-Ferrer, 2011).

2.4 Contribution of the present study

The present study contributes to research in the field of the L2 speech dimensions of accentedness and comprehensibility from different and new perspectives. To begin with, it adds to the reduced number of studies which have examined these dimensions in the speech production of adolescent learners of English in a formal instruction setting in Spain (Gallardo del Puerto, Gómez Lacabex & García Lecumberri, 2007; García Lecumberri & Gallardo del Puerto, 2003; Rallo & Juan-Garau, 2011). Moreover, to our knowledge, it is the first longitudinal study comparing

the development of these speech dimensions in two groups of teenager L2 learners (SA vs. AH).

It has been claimed that work on perceived accentedness and comprehensibility with non-native listeners is still insufficient (Derwing and Munro, 2011; Isaacs and Trofimovich, 2012). The present study provides data regarding perceived foreign accent and comprehensibility from a group of non-native listeners. We seek to confirm the correlation between accentedness and comprehensibility posited in previous research, and examine the aspects influencing perceived accentedness and comprehensibility according to listeners' reports, given the fact that ratings can be assigned for different reasons. Thus, comparisons with previous studies including native listeners will be made in order to support (or not) previous findings.

Chapter 3

Pronunciation in a Formal Instruction Context

One of the pillars of this dissertation is the multidimensional nature of L2 speech. As we have seen in chapters 1 and 2, a wide range of studies have explored different aspects of oral production (i.e. fluency, accuracy, complexity, narrative skills, pronunciation, foreign accent and comprehensibility). However, this multidimensional view is not consolidated to the same extent in the field of L2 teaching and assessment.

For example, as Munro and Derwing (1995a) note, we may find pronunciation assessment scales ranging "from *not accented, perfectly comprehensible* at one endpoint to *accented and difficult to understand* at the other" (1995a:92). Thus, the claim made by some researchers arguing that foreign accent and comprehensibility are two partially independent dimensions may not be reflected in some assessment rubrics and actual assessment practice, who very often conflate the different aspects.

In the following pages we consider some relevant issues related to current pronunciation teaching practices, within the framework of this dissertation: foreign accent and comprehensibility as two related but independent dimensions of L2 speech. Special attention is given to the following questions: the objectives of pronunciation instruction; problems in pronunciation teaching; and assessment of pronunciation.

3.1 The objectives of teaching pronunciation

Two main principles have traditionally led the discussion about the objective of pronunciation instruction: the *nativeness principle* vs. the *intelligibility principle* (Levis, 2005). The *nativeness principle* aims at native-like pronunciation for L2 speakers, whereas the *intelligibility principle* considers intelligibility as the primary objective.

Many researchers and practitioners describe the objective of pronunciation teaching in terms of "intelligibility" and claim that pronunciation teaching should aim at language intelligibility and not at native-like pronunciation:

> "Very few teachers today would claim that a pronunciation that is indistinguishable from that of a native speaker is necessary or even desirable for their learners. Instead, it is generally accepted that intelligibility is the most sensible goal".
> (Kenworthy, 1987:13)

> "Probably the most obvious, justifiable and pressing goal in the area of phonology is intelligibility. For beginning students, this is the most immediate need, as no communication can take place without a certain level of mutual intelligibility among speakers". (Pennington, 1996:220)

Most L2 pronunciation research does not consider accent reduction to be the goal for communicative teaching (Derwing, 2008; Thomson, 2013). The interest in teaching pronunciation is not centred on the nuances of particular speech sounds, but getting the L2 learners up to a level of competence which allows them to deal with everyday communication situations or, to put it in Gimson's words, learning the L2 language "for the practical purposes of everyday communication" (1994:270):

> "The goal of teaching pronunciation to such learners is not to make them sound like native speakers of English. With the exception of a few highly gifted and motivated individuals, such a goal is unrealistic. A more modest and realistic goal is to enable learners to surpass the threshold level so that their pronunciation will not detract form their ability to communicate" (Celce-Murcia, Brinton & Goodwin, 1996:8).

Although most L2 learners do not consider having a native-like pronunciation as a priority, there might be L2 learners who aim at achieving perfect native-like pronunciation for different reasons (e.g. professional reasons, building up a determined 'self-image', integrative motivation, etc.). In contrast, we may find L2 learners preferring to retain something of their L1 accent when speaking in English (Porter & Garvin, 1989). Dalton and Seidlhofer (1994:7) note that "pronunciation is so much a matter of self-image that students may prefer to keep their accent

deliberately, in order to retain their self-respect or to gain the approval of their peers". Actually this is what we usually find as teachers in the context of foreign language classroom when our students tend to avoid sounding 'English', since this can result in their peers making fun of them or joking about their 'native like accent' (Fisher and Evans, 2000).

As previously mentioned in Chapter 2 of this dissertation, several studies have concluded that comprehensibility is far more important than acentedness in terms of communicative success (Derwing & Munro, 1997; Munro, 2008; Munro and Derwing, 1995a; Munro and Derwing, 1999):

> "Although strength of foreign accent is indeed correlated with comprehensibility and intelligibility, a strong foreign accent does not necessarily cause L2 speech to be low in comprehensibility or intelligibility" (Munro and Derwing, 1995a:92).

If comprehensibility is the main objective of pronunciation instruction, the degree of foreign accent in L2 learners' oral productions should be of minor concern, and accent reduction should not be a priority. Rather, those aspects of L2 speech that appear to interfere with listeners' comprehension of the learners' production should be the focus. The question is: which aspects do seem to affect comprehensibility?

A research priority is to identify the aspects of L2 speech that hinder comprehensibility from those that, while noticeable or irritating, do not impede understanding the message (Munro, 2008; Isaacs & Trofimovich, 2012). Little empirical research has examined the particular aspects of foreign-accented speech which affect comprehensibility (Munro and Derwing, 1995a, 1999; Isaacs and Trofimovich, 2012; Trofimovich and Isaacs, 2012; Zielinski, 2008). Moreover, opinions of a particular L2 speaker's pronunciation problems may vary from listener to listener since familiarity with accented speech and individual differences in the ability to comprehend L2 speech may influence foreign accent and comprehensibility perception (Gass & Varonis, 1984; Munro & Derwing, 1999).

A consideration of learners' pronunciation errors and how these may hamper successful communication is a useful basis on which to assess why it is important to deal with pronunciation in the classroom. The inaccurate production of a phoneme can lead to misunderstanding, and an L2 learner who consistently mispronounces a range of phonemes can be extremely difficult to understand for a speaker from another language community. This can be very frustrating for the learner who may have a good command of grammar and lexis but has difficulty in understanding and being understood by a native speaker.

In line with Derwing and Munro (2009), we believe it is appropriate to work on those aspects of accent which may affect comprehensibility. Some studies have indicated that pronunciation training can help L2 speakers produce more intelligible speech. Derwing, Munro, & Wiebe (1998) examined perceived accentedness, comprehensibility and fluency in the oral productions of L2 learners of English. The learners were assigned to one of these conditions: (1) no specific pronunciation instruction group; (2) global instruction group, who received instruction with a focus on features such as speaking rate, intonation, rhythm, projection, word stress, and sentence stress; (3) segmental instruction group, who received instruction to improve their production of individual sounds. Their research concluded that even though the two groups receiving instruction in pronunciation showed significant improvement in accentedness and comprehensibility on the sentences, only the group receiving global instruction showed improvement in comprehensibility and fluency in the narratives.

In line with these results, Munro and Derwing (1999) reported that "prosodic errors appear to be a more potent force in the loss of intelligibility than phonetic errors" (1999:285). These findings are in controversy with the actual situation in the EFL classroom, where much pronunciation practice and error correction focuses on the segmental level.

More recently, Trofimovich and Isaacs' (2012) and Isaacs and Trofimovich's (2012) explored the linguistic aspects which affect foreign accent and comprehensibility. In the former, Trofimovich and Isaacs (2012) examined the linguistic aspects of L2 speech related to accent and comprehensibility. They concluded that both dimensions were related to many speech measures, but that "four categories uniquely distinguished accent from comprehensibility, with all categories specific to the dimension of phonology (i.e. vowels and consonants, syllables, sounding nativelike, and rhythm)" (p. 9, whereas comprehensibility was mainly linked to grammatical accuracy and lexical richness. Although it is true that speaking involves pronunciation, it is worth highlighting that L2 speech comprehensibility is linked to vocabulary and grammar. In the latter, Isaacs and Trofimovich (2012) studied in greater depth the construct of comprehensibility, and explored the aspects of speech that affected L2 comprehensibility at different ability levels. As already referred to in section 2.2.1, based on the anaylysis of 19 quantitative speech measure, and listeners' judgments and introspective reports, the authors identified five speech measures that distinguished between L2 learners at different comprehensibility levels: "lexical richness and fluency measures differentiated between low-level learners; grammatical

and discourse-level measures differentiated between high-level learners; and word stress errors discriminated between learners of all levels" (Isaacs and Trofimovich, 2012:476). Thus, it is interesting to highlight that not only pronunciation features of a foreign-accented speech, but also other language aspects affect speech comprehensibility (e.g vocabulary, grammar, discourse measures).

On the other hand, it is worth noting that the above mentioned studies included English native speakers as listeners of L2 learners' oral production. As noted by Isaacs and Trofimovich, further research with groups of non-native listeners is needed in order to confirm these results.

3.2 Teaching English pronunciation in Spain

Kelly (1969) featured pronunciation as the "Cinderella" area of foreign language teaching. He noted that Western philologists and linguists had studied grammar and vocabulary much longer than pronunciation, which began to be studied systematically shortly before the beginning of the twentieth century. In line with this view, Derwing and Munro (2005) reported the lack of attention to pronunciation teaching, and Pavón (2000) and Llurda (2002) confirmed that EFL pronunciation has traditionally been given little attention in the particular case of Spain.

Several reasons for this lack of focus on pronunciation practice have been given (Celce-Murcia, Brinton & Goodwin, 1996; Derwing and Munro, 2005; Kelly, 2000; Llurda, 2002). Kelly (2000) argued that very often pronunciation tends to be neglected; and if not, it tends to be reactive to a particular problem that has arisen in the classroom rather than being strategically planned. The fact that pronunciation tends to suffer from neglect may not be due to teachers' lack of interest in the subject but rather to a feeling of doubt as to how to teach it. Lack of knowledge of phonetics and lack of formal preparation to teach pronunciation are two of the most cited problems.

A recent survey on English pronunciation teaching practices conducted in different European countries (Henderson et al., 2012) reported that teachers in Spain give greater importance to pronunciation in relation to other language skills. However, insufficient time and resources are spent on this language aspect. Teachers in this survey argued two reasons for this: (1) the difficulty for both teachers and students; (2) the fact that teachers are bound by curricular demands and by the need to obtain results. The urgent need for specific pronunciation training for teachers in Spain has been called for frequently (Donovan, 2001; Levey 1999, 2001; Pavon, 2001; Pavon and Rosado, 2003) and as Derwing and Munro

(2005:389) noted "expecting teachers to rely solely on intuition is unrealistic and unfair".

A related concern for non-native teachers of English may be their own pronunciation of English, and the question we would like to address is whether or not it is necessary to be a native speaker to teach pronunciation. While being a native speaker of the language helps, this is not enough. Proper knowledge of language instruction, confidence and good oral skills (together with an intelligible accent -whether native or not) may endow the (non-)native teacher with sufficient abilities to teach pronunciation. Moreover, **s**ome native speakers (of any language) may not communicate effectively.

Another concern is that teachers may rely too heavily on published pronunciation materials without regard to their own students' problems. As Derwing and Munro (2005) suggest, most materials may have been designed without a basis in pronunciation research findings. While many researchers and teachers consider intelligibility to be the goal, pronunciation curricula, materials and classroom practice still often reflect the nativeness principle (Derwing, 2008; Derwing and Munro, 2009).

Another relevant issue within the framework of this dissertation deals with the choice of a model of English for EFL learning. So far we have been referring to foreign accent and comprehensibility, but we have not considered which English variety should serve as the referent to measure the degree of accentedness or comprehensibility of L2 learners' speech. The choice of an English pronunciation model for EFL teaching/ learning is complex and may respond to different criteria (Mompeán, 2008).

Geographical and cultural proximity between EFL learners and the territories where English is an L1 has been one of the reasons for choosing one variety over another. In the case of Spain this would explain why British English has been the preferred model.

The availability of published material featuring a given variety and the availability of teachers speaking that target model also affects the English variety which L2 learners are exposed to. This fact, however, is sometimes paradoxical. In the case of Spain, for instance, where Received Pronunciation (RP) is the proposed English pronunciation model, we may find EFL teachers who do not speak RP.

Another pedagogical reason that can be considered is the difficulty in teaching a specific model. Although RP is the preferred model in Spain,

this is not the easiest accent for an L2 learner to acquire (Jenkins, 2000; Mompeán, 2008).

Mompeán (2008) notes that social and psychological criteria have also been associated with the choice of a pronunciation model. Some studies have reported that EFL learners show more positive attitudes towards accents like RP or GA than towards non-native accents of English (Chiba, Matsuura & Yamamoto, 1995; Dalton-Puffer, Kaltenboek & Smit, 1997; Smith and Dalton, 2000). In a previous study by Mompeán (2004), he inquired about the English accent Spanish university students studying English language/linguistics wanted to learn. He concluded that 71% wished to learn a British accent, whereas 17% wanted to acquire an American accent or an Irish accent (12%). Mompeán argued that the preference was generally associated with the belief that a British accent "excelled in purity, correctness, and beauty over other accents". These findings contrast with Henderson et al's research (2012) showing that RP is preferred by teachers although they recognize that GA might be more popular amongst L2 learners.

On the other hand, it is inevitable to refer to English as an International Language (EIL) and English as a Lingua Franca (ELF) when dealing with varieties of English. With these terms we will refer to English as a language of wider communication, spoken by a large number of native speakers, but also by native speakers of other languages (McKay, 2002). Actually Jenkins (2000:1) reports the fact that "for the first time in the history of the English language, second language speakers outnumber those for whom it is the mother tongue, and interaction in English increasingly involves no first language speakers whatsoever".

If we think in terms of communicative criteria for the election of an English pronunciation model, the idea on the mutual intelligibility between non-native speakers has to be considered. As Derwing and Munro (2005:392) observe "more attention should also be focused on the mutual intelligibility of NNSs of English. This is especially important at a time when English is increasingly used as a lingua franca around the world (Jenkins, 2000; Kachru, 1992)".

Within this worldwide context of use, claims in favour of teaching pronunciation to "develop the learner's pronunciation sufficiently to permit effective communication with native speakers" (Roach, 1991:6) are out of place. Today many interactions in English involve communication wholly between non-native speakers, and it seems inevitable that more traditional EFL pronunciation goals would be put in question (i.e. the nativeness principle). However, still today the majority

of EFL teacher training and textbooks in Spain persist with phonology syllabi that assume a "native-speaker" interlocutor.

Jenkins (2000) observed the relative failure of English language teaching pedagogy to adjust its methodologies to the changing patterns of use, in which the aim of learning is more often to be able to use English as a lingua franca in communication with other non-native speakers, than as a foreign language in communication with native speakers. Given this context, we question again to what extent it is necessary to teach L1 pronunciation norms to learners who are likely to communicate with speakers of English others than the RP variety speakers.

3.3 Pronunciation assessment

Most often speaking evaluation in EFL classrooms (e.g. at the secondary education level) is usually done impressionistically, with teachers assigning a single score on a 10-point scale. Very frequently teachers do not use assessment rubrics of any type or instructions. Rather, teachers rely on their experience assessing learners' oral productions, and assess them within the framework of the rest of oral deliveries in the group.

In the context of official assessment scales, we find that pronunciation is considered in different ways. We may find that accentedness is not included as an assessment criterion. In the benchmark-level descriptors of the Common European Framework of Reference (CEFR), there is no reference to pronunciation:

Proficient User	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
Independent User	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a

		wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

Table 3.1: Level descriptors of the CEFR

In other speaking scales of high-stakes tests used for academic purposes (e.g. TOEFL, IELTS) pronunciation description are often too vague to have a clear and coherent idea of the construct. As noted by Isaacs and Trofimovich (2012) band 4 in the IELTS Speaking Band Descriptors, for instance, reads: "Uses a limited range of pronunciation features; attempts to control features but lapses are frequent; mispronunciations are frequent and cause some difficulty for the listener".

In the TOEFL Internet-based test (iBT) Integrated Speaking Rubrics, "intelligibility" is associated with "pronunciation", "intonation" and "pacing":

Score	Delivery
4	Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.
3	Speech is generally clear, with some fluidity of expression, though minor difficulties with pronunciation, intonation, or pacing are noticeable and may require listener effort at times (though overall intelligibility is not significantly affected).
2	Speech is basically intelligible, though listener effort is needed because

	of unclear articulation, awkward intonation, or choppy rhythm/pace; meaning may be obscured in places.
1	Consistent pronunciation, stress, and intonation difficulties cause considerable listener effort; delivery is choppy, fragmented, or telegraphic; frequent pauses and hesitations.
0	Speaker makes no attempt to respond OR response is unrelated to the topic.

Table 3.2: TOEFL Internet-based test (iBT) Integrated Speaking Rubrics: Delivery

In the speaking performance assessment scale for Level B2 by University of Cambridge (i.e. First Certificate of English) "intelligible pronunciation" is linked with "intonation", "stress" and "individual sounds articulation":

B2	Pronunciation
5	• Is intelligible. • Intonation is appropriate. • Sentence and word stress is accurately placed. • Individual sounds are articulated clearly.
4	*Performance shares features of Bands 3 and 5.*
3	• Is intelligible. • Intonation is generally appropriate. • Sentence and word stress is generally accurately placed. • Individual sounds are generally articulated clearly.
2	*Performance shares features of Bands 1 and 3.*
1	• Is mostly intelligible, and has some control of phonological features at both utterance and word levels.
0	*Performance below Band 1.*

Table 3.3: Level B2 by University of Cambridge

Even the use of the term "pronunciation" is not used in a consistent way across these scales:

> "Whether the term refers solely to segmental features (i.e., errors that involve individual sounds) or also encompasses other aspects of speech, including suprasegmental features (e.g., word stress, rhythm, intonation), needs to be clearly spelled out to facilitate the interpretation of the scale descriptors for both raters and test users (Isaacs and Trofimovich, 2012:478).

Another shortcoming of L2 speaking scales is that they may conflate accentedness and comprehensibility (Harding, 2013; Munro & Derwing, 1995a; 1999) even though previous research has shown them to be partially independent dimensions (Derwing & Munro, 2009). For instance,

the band descriptors of the CEFR Scale of Phonological Control Scale[5] reads:

C2	No descriptor available.
C1	Can vary intonation and place sentence stress correctly in order to express finer shades of meaning.
B2	Has a clear, natural, pronunciation and intonation.
B1	Pronunciation is clearly intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur.
A2	Pronunciation is generally clear enough to be understood despite a noticeable foreign accent, but conversational partners will need to ask for repetition from time to time.
A1	Pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by native speakers_used to dealing with speakers of his/her language group.

Table 3.4: CEFR Scale of Phonological Control Scale

The language of some current pronunciation assessment scales (e.g. CEFR Phonological Control Scale) demonstrates a shift away from the nativenesss principle, since it includes statements in criteria focusing on intelligibility ("understood"). However, accentedness and comprehensibility are grouped together in some band descriptors. For example, A2 band descriptor reads "[…] clear enough to be understood despite a noticeable foreign accent… ", and B1 considers "… clearly intelligible even if a foreign accent is sometimes evident…". As already mentioned, some studies support that pronunciation features which influence accentedness do not necessary interfere with intelligibility and comprehensibility (Derwing & Munro, 1997), or as Jenkins (2000) observes, native accents might not be the most intelligible. Language testing practices can benefit from this type of research, and consider accentedness and comprehensibility as different aspects of L2 speech, which should be assessed separately.

A further limitation is the relativistic wording in three band descriptors. The scale bands in the CEFR Scale of Phonological Control Scale, for instance, include terms such as "correctly" (in C1 band descriptor) or "clear" (in descriptor B2). What do these words exactly mean? How do listeners/raters from different language backgrounds interpret "correctly" or "clear"? Do they interpret these terms in the same way?

Still less clearly defined are the numeric Likert scales with no band descriptors, used in assessment practice or research. In research, for

[5] http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key_reference/Overview_CEFRscales_EN.pdf

instance, the 9-point numerical comprehensibility scale is commonly used ranging from *extremely difficult to understand to extremely easy to understand* at scalar points, with no further definition provided to listeners (Derwing, Munro, & Thomson, 2008). Interrater reliability is high using this rating procedure but, as cited in Isaacs and Trofimovich (2012:478) "reliability is a necessary but insufficient condition for validity" (Cohen, Manion, & Morrison, 2000:105).

As already stressed, even when listeners assign the same rating to a speech sample, their rationale for doing so may be different (Isaacs and Thomson, 2013). Moreover, Douglas (1994) suggested that similar ratings may represent qualitatively different learner performance. Despite the interest and importance of this matter, the number of qualitative studies examining the bases of raters' judgements is limited (Isaacs & Thomson, 2013; Harding, 2008; Rossiter, 2009; Zielinksi, 2008).

A clearer description of the bands in rating scales, in general, and of the construct of comprehensibility, in particular, seems to be necessary within the framework of pronunciation assessment discussed so far. Such operationalization would benefit raters' assessment task (Isaacs and Trofimovich, 2012).

3.4 Contribution of the present study

It is widely agreed that the objective of L2 pronunciation instruction should be to help learners be intelligible to their interlocutors. However, L2 teachers have received little guidance on how to do this (Derwing & Munro, 2009; Isaacs and Trofimovich, 2012). Further research is needed to gain better understanding of what is meant by intelligibility. To answer this question this research explores the features of L2 speech which affect comprehensibility by analysing listeners' reports. Previous research has observed that not only specific pronunciation features but also other aspects (i.e. vocabulary, grammar, and discourse) are associated with this dimension of L2 speech. Levis (2005) pointed out that the principle that L2 learners should aim to be understandable to their interlocutors is fundamentally incompatible with the idea that L2 learners should aim to acquire a native-like accent, eradicating all traces of their L1. Rating scales need to reflect this reality. Thus, the present research explores comprehensibility so that a description of this construct in rating scales is provided with greater precision. At the same time, reports from non-native speakers of English rating L2 speech samples for FA and comprehensibility will help us to elucidate whether this type of listeners leave aside reference to accent or nativelikeness in their comprehensibility

ratings, as it has been the case in previous studies involving native listeners of English.

Language learners often show considerable interest in pronunciation practice, since they feel it will help them to communicate better. Paradoxically, even though both teachers and learners regard this aspect as important, pronunciation is often neglected in EFL formal instruction context. In addition, the multidimensional perspective of L2 learners' speech adopted in this research may be under-considered (even ignored) in some teaching practices.

These and other issues discussed throughout this chapter have been vaguely explored in the context of EFL teaching in Spain. However, we believe it is necessary to gain a better understanding of the actual EFL classroom context, so that we are able to make suggestions for future research. To address this question, we designed a questionnaire for the teachers participating in our research as listeners, which was also completed by the teachers who taught the learner participants but did not participate in the rating experiment as judges (see Chapter 5). This questionnaire aims to shed light on some of the research questions we will analyze (e.g. study the contrast between teachers' beliefs and opinions before and after participating in the experiment).

The questionnaire provided us with detailed and relevant information about the listeners who rated accentedness and comprehensibility in L2 learners' speech productions. At the same time it gathered information to build an idea of the context in the EFL classroom in Spain. The questions in the survey dealt with different aspects: (1) personal information, (2) linguistic profile, (3) training, (4) views about EFL teaching, (5) teachers' individual teaching practice, and (6) assessment of pronunciation. In line with the whole approach of the present research, this questionnaire has been designed to link research, language learning and teaching practice to the greatest possible extent.

PART II

Chapter 4

Objectives and Research Questions

4.1 Introduction to the study

In Chapter 1 of this dissertation we have reviewed the body of research on language development in SA contexts. As we have seen, certain aspects are relatively under-researched in comparison to others. Several studies have analysed the degree of foreign accent in non-native English speakers' oral production (Derwing & Munro, 1997; Gallardo del Puerto et al, 2005, 2007; Magen, 1998; Munro & Derwing, 1995, 1999; Rallo Fabra & Juan-Garau, 2011; Trofimovich & Isaacs, 2012), but only a few have analysed this dimension involving learners of English in a SA context (Allen and Herron, 2003; Højen, 2003; Avello, Mora and Pérez-Vidal, 2012). On the other hand, a number of studies have examined the comprehensibility of non-native English speakers' oral production (Derwing & Munro, 1997; Derwing, Munro & Wiebe, 1998; Isaacs & Trofimovich, 2012; Jun & Li, 2010; Kennedy & Trofimovich, 2008; Munro & Derwing, 1995, 1999; Rallo Fabra & Juan-Garau, 2011; Trofimovich & Isaacs, 2012), but to our knowledge no study has explored the impact of SA on this dimension.

As we have seen, research on the perceived constructs of foreign accent and comprehensibility involving groups of SA participants is scant, and participants in existing studies were undergraduate students. Research with adolescent learners taking part in SA programmes is also very scarce (Marriott, 1995; Fisher and Evans, 2000; Lapkin, Hart & Swain, 1995; Llanes and Muñoz, 2009). Thus, to our knowledge, no study to date has examined the perception of foreign accent and comprehensibility with

adolescent learners of English. The present study aims to shed light on the development of these speech dimensions in the extemporaneous speech of adolescent EFL learners during a SA period and compares these results with those obtained by a control group in an AH learning context.

A second objective of this research study is to examine the relationship between L2 learners' degree of foreign accent and comprehensibility. We aim to identify and draw comparisons between the factors underlying the accentedness and comprehensibility ratings reported by the non-native listeners participating in our research (following some of the analyses conducted by Trofimovich and Isaacs, 2012). As pointed out by Isaacs (2010), knowledge of the factors influencing comprehensibility in L2 speech can help teachers to set instructional objectives, integrate pronunciation with the teaching of other skills, and take these questions into account in their assessment practice in the EFL classroom.

The following sections outline the specific objectives and research questions formulated for this study.

4.2 Objectives

The main objective of this study is to explore the constructs of foreign accent and comprehensibility in two different learning contexts (SA vs. AH). This study examines a sample of oral narratives from a group of adolescent EFL learners in secondary education. The speech samples were collected longitudinally before (pre-test) and after (post-test) the SA period. Participants' oral productions were evaluated in terms of perceived foreign accent and comprehensibility. We also included speech samples collected from native speakers (NS) as baseline data to assess listeners' ratings.

More specifically, we aim:

1. To examine longitudinal changes in FA and comprehensibility in the oral production of a group of adolescent learners of English after a 3-month SA period.
2. To examine longitudinal changes in FA and comprehensibility in the oral production of a group of adolescent learners of English after a 3-month formal instruction period in an AH context.
3. To determine whether one learning context is more beneficial than the other in terms of FA and comprehensibility gains.
4. To examine the effect of initial foreign accent level on FA development in SA or AH learning contexts.

5. To examine the effect of initial comprehensibility level on comprehensibility development in SA or AH learning contexts.
6. To explore the relationship between the constructs of foreign accent and comprehensibility.
7. To identify the aspects influencing non-native listeners' accentedness and comprehensibility ratings.

4.3 Research questions

The objectives above mentioned led us to formulate three principle research questions, which guided the analysis and discussion presented in chapters 6 and 7. The first research question focuses on FA and the second one deals with the analyses of comprehensibility. The third research question explores the relationship established between the two speech dimensions. Each research question is reformulated into one or more sub-questions:

RQ1: Do participants improve in terms of foreign accent after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is foreign accent development different for participants with different initial foreign accent level?

RQ1a: Do participants improve their FA ratings after a 3-month SA period, or a 3-month FI period in an AH context?

RQ1b: Is one learning context (SA vs. AH) more beneficial than the other? In other words, do SA participants obtain greater gains in foreign accent development than the AH group?

RQ1c: Is foreign accent development different for participants with different initial foreign accent level in each group (SA and AH)?

RQ2: Do participants improve in terms of comprehensibility after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is comprehensibility improvement different for participants with different initial comprehensibility level?

RQ2a: Do participants improve their comprehensibility ratings after a 3-month SA period, or a 3-month FI period in an AH context?

RQ2b: Is one learning context (SA vs. AH) more beneficial than the other? In other words, do SA participants obtain greater gains in comprehensibility development than the AH group?

RQ2c: Is comprehensibility development different for participants with different initial comprehensibility level in each group (SA and AH)?

RQ3: To what extent are foreign accent and comprehensibility related speech dimensions?

RQ3a: To what extent do degree of FA and comprehensibility correlate?

RQ3b: To what extent do non-native listeners report the same linguistic aspects as having influenced their ratings when evaluating FA and comprehensibility?

Chapter 5

Method

This chapter presents the methodology of the current study and is divided into different sections. Section 5.1 presents the design of the study. The methodological approach taken in this research involves production and perception tasks. Section 5.2 presents a description of the participants and the data collection instruments and procedures for the production task. The production task elicited speech samples from different groups of learners (SA and AH), and the speech samples obtained from this task served as the stimuli for the perception task. The objective of the perception task was to examine perceived foreign accent and perceived comprehensibility by a group of listeners. Section 5.3 provides a description of the participants and the data collection instruments and procedures for the perception task, together with an explanation of the different measures and data analyses carried out afterwards.

5.1 Design

The present study has a longitudinal, pre-test/post-test design. Data was collected from participants at two different times over 7 months. The data collection procedure comprised 3 academic years (2008-09; 2009-10; 2010-11) since data was collected from two consecutive cohorts of students at the same home institution. The first data collection (T1 or Pre-test) took place before finishing the academic year previous to the SA period, that is, in June 2009 or June 2010. SA and AH participants were

tested again after their return from a 3-month SA or the corresponding period AH (T2 or Post-test), that is, in January 2010 or January 2011.

Table 5.1 illustrates the data collection times for SA and AH groups, indicating the time and number of participants recruited for the study:

	Group	Data collection	T1 in 2nd year	T1 in 4th year
SA	**Cohort 1** (n=10)	T1: June 2009 *SA period: Sept 09-Dec 09* T2: January 2010	8	2
	Cohort 2 (n=15)	T1: June 2010 *SA period: Sept 10-Dec 10* T2: January 2011	11	4
			SA (n=25)	
AH	**Cohort 1** (n=2)	T1: June 2009 *AH period: Sept 09-Dec 09* T2: June 2010	2	0
	Cohort 2 (n=29)	T1: 2008-09 *AH period: Sept 10-Dec 10* T2: January 2011	20	9
			AH (n=31)	

Table 5.1: Time and number of SA and AH participants

SA and AH participants completed a battery of exam-like tests at the two collection times (T1 and T2). These tests covered different linguistic skills and aspects: listening, reading, speaking, writing, grammar and use of English. Before starting the battery of oral and written tests, participants completed a set of three questionnaires which gathered information about (1) their linguistic profile, (2) the teaching/learning practice in their English lessons at school, and (3) attitudes and motivation to learn English.

A group of native speakers (NS) of English was also recruited to provide baseline data. Data from NS was collected as part of the COLE research project[6], a coordinated research project between *Universitat Pompeu Fabra* (Barcelona, Spain) and *Universitat de les Illes Balears* (Spain). COLE project researchers in the Balearic Islands collected the data from NS.

[6] *Contexto, contacto y nivel de competencia en la adquisición del inglés como lengua extranjera. Gradación en el desarrollo de habilidades orales y escritas a niveles medios* HUM2007-66053-C02-02/FILO.

As mentioned above, the design of the study responds to a longitudinal type of research, described by Menard (2002) as "research in which a) data are collected at two or more distinct time periods; b) the subjects or cases analysed are the same or are comparable (i.e. drawn from the same population) from one period to the next; and c) the analysis involves some comparison of data between periods" (cited in Dörnyei, 2007: 79). In previous SA studies examining acquisition in SA and AH learning contexts (Juan-Garau & Pérez-Vidal, 2007; Trenchs-Parera, 2009; Pérez-Vidal & Juan-Garau, 2011; Valls-Ferrer, 2011; Mora & Valls-Ferrer, 2012; Barquin, 2012; Pérez-Vidal et al. 2012), a within subjects design was used, that is, the same group of participants were compared in the two different contexts (SA and AH). As Valls-Ferrer (2011:119-120) noted in her research, "[...] the threat to internal validity comes when comparing the same group of participants in the two different contexts where 'carry over effects' can occur from the FI period to the SA". To address this question the introduction of a control group receiving FI while the SA participants are in the L2 country seems to be an appropriate option.

However, as Milton and Meara (1995) suggested, a SA study should ideally include two groups of participants which "should have roughly the same degree of exposure to the L2 over a similar time span" (p.18) and this is almost impossible to occur. Moreover, as noted by Sanz (forthcoming) it may be the case that "[…] SA/AH comparisons end up comparing apples and oranges, because students who choose to go abroad are different from students who choose staying in their home institutions". In line with Sanz's claim, studies focusing on CLIL learning context have also emphasized the existence of a 'natural' selection when referring to differences between CLIL and non-CLIL control groups. As Bruton (2011:237) reported: "a consensus voiced by numerous CLIL teachers was that it was generally the children of higher socio-economic-status parents who chose the CLIL option in their schools". In addition, he referred to several studies corroborating this 'natural' selection and arguing that parents and children who choose CLIL are generally the more motivated, and that CLIL participants are very likely to take extra English classes outside school and show higher L2 proficiency than the non-CLIL participants (Alonso et al., 2008; Lagabaster & Sierra, 2009; Ruiz de Zarobe & Lasagabaster, 2010; San Isidro, 2010).

Given the vast number of variables affecting SA research, Rees & Kappler (2008:98) pointed out that "it is unlikely that any study will be able to secure genuine 'control' groups, but researchers should at least attempt to make the best comparisons possible". In line with this suggestion and other recommendations in previous SA studies asking for further research including control groups, we will examine language development in SA

participants and compare their results with the ones obtained by their peers AH, who followed standard formal instruction AH during the same period of time.

5.2 Production task

In this section, we will provide further information about the participants in our study and their learning context, and explain in detail the data collection procedure from SA and AH participants, as well as the preparation of speech samples from the three group of participants (SA, AH and NS). These speech samples were used in the subsequent perception experiment with listeners, which will be described in detail in section 5.3.

5.2.1 Participants and learning contexts

The participants in this study were Spanish adolescent learners of English in two different contexts: SA (n=25) and AH (n=31). Moreover, data from a group of adolescent English native speakers (NS) (n=15) was also used in the perception task. The total number of participants in the three groups (SA, AH, NS) was 71.

5.2.1.1 SA group

The participants in the SA group were 25 adolescent learners of English who were native Spanish speakers (13 females, 12 males). The SA participants were from Valencia, Spain, and studied at the same semi-private school in this city. All of them were between 12 and 15 years old at Pre-test ($M_{age\ T1}$ = 12.96 years), and between 13 and 15 years old at Post-test ($M_{age\ T2}$ = 13.52 years). All participants had started learning English at school in their third year of primary education (i.e. at the age of 7-8) on a 60-minute weekly basis, and received up to 3 hours per week of subsequent EFL instruction at school. In addition, all participants except two[7] had attended at least three years of extra-curricular EFL lessons outside the school (68% reported more than five years). All participants reported normal hearing, and none had any detectable speech disorder.

The SA participants embarked upon their SA in the first term of their academic year (from September to December). The length of stay abroad was similar for all participants (3 months). For all of them the SA period was optional, and most of them chose a British or Irish school as their

[7] Participants' code: ANBA (who had attended extra-curricular English lessons for 6 months), and MARI (who had attended lessons for 1 year).

home institution while abroad[8]. The learners differed in terms of their current academic year at their school: 19 participants were at the beginning of their second year in secondary education, and 6 were beginning their fourth (and final) year of compulsory secondary education.

In a background questionnaire about their linguistic profile which participants completed at T1 (see Appendix 1), SA participants indicated whether they had participated in a similar SA programme before. As illustrated in Figure 5.1 below, most participants (80%) had already spent shorter periods of time in an English speaking country (half of them taking part in a two or three week summer course abroad); four students[9] (16%) reported having had a previous 3-month SA experience in an English speaking country, and only one participant[10] (4%) had not studied English abroad before:

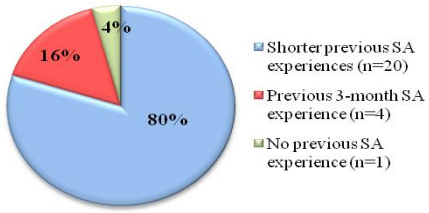


Figure 5.1: SA participants and pevious experience in an English speaking country

SA participants were asked to complete an online questionnaire upon return[11] (see Appendix 2). In this questionnaire SA participants[12] reported information on different aspects: (1) personal information; (2) general information about their stay abroad; (3) information about their study abroad period; (4) language use and L2 contact while abroad; (5) their impressions of L2 development after SA; (6) information about SA preparation and follow-up.

The answers to this questionnaire provided us with information about different aspects. For instance, answers to questions 15 and 16 gave us details of the type of formal instruction that participants received in their SA schools. During the 3-month SA period, the participants were enrolled

[8] 18 participants spent their term abroad in England, 3 in Ireland and 1 participant went to the USA. There is data missing from 3 participants.
[9] Participants' code: ALOB, CAFO, MASA, SARI.
[10] Participant's code: ALRU.
[11] Title of the questionnaire: *Cuestionario a alumnos de ESO que han realizado una estancia en el extranjero durante el curso académico*.
[12] Due to methodological shortcomings, only 13 out of 25 SA participants completed the online questionnaire after the SA period.

in the common school subjects, which were taught in English (i.e. Maths, Biology, Chemistry, Art, etc.). Seventy percent of the participants followed their lessons together with the English native teenagers at their school, and 30% reported that apart from having lessons with English native speakers, they also attended lessons which were exclusively for exchange students like them. These lessons were also taught in English. With regard to L2 use outside the classroom, 85% of the participants reported that they also took part in extracurricular activities after the school, which were carried out in English (question 17). As far as language use is concerned, question 21 in the background questionnaire asked participants to calculate the approximate amount of time of English use, and L1 or other languages use on a daily basis. Figure 5.2 below illustrates the proportion of English use during their SA period reported by SA participants:

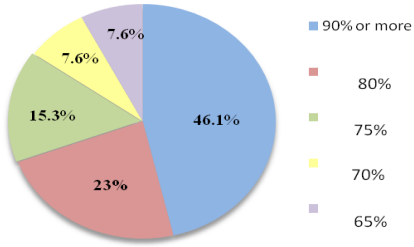


Figure 5.2: Percentage of SA participants and percentage of L2 use reported

Teachers in the AH institution (n=4) were asked to complete an online questionnaire (see Appendix 4 and next section for a description of this questionnaire). In the last section of this questionnaire, teachers were asked about the benefits of SA programmes for L2 learners' language development, in general, and for the students at their home school, in particular (see Question 42 and Question 43 in the questionnaire). With regard to the aspects they considered that students in general improve most during a SA sojourn (Q42), all of them agreed that learners gain in their speaking skills. Seventy-five percent also reported that learners develop their listening skills as well, and 50% pointed out that pronunciation and cultural aspects are language areas that learners tend to develop while abroad. As for the students in their school who spend one academic term abroad, all the teachers reported that in their opinion, learners show a lot of progress in listening skills and 75% indicated that students also exhibit a lot of improvement in speaking skills. On the contrary, writing and grammar were considered as two language aspects were students show little progress after the SA period. Figure 5.3 illustrates the results for Q42:

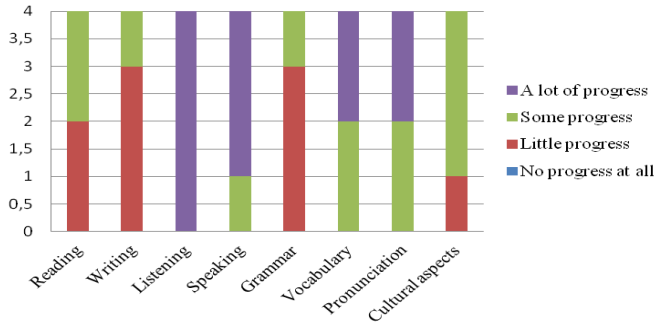


Figure 5.3: Answers provided by teachers at the AH institution (n=4) regarding Q42: *Which aspects do you think students improve most during a SA programme?*

5.2.1.2 AH group

The participants in the AH group were 31 adolescent learners of English who were native Spanish speakers (27 females, 4 males). All participants were from Valencia, Spain, and studied at the same semi-private school as the SA group. The AH participants were between 12 and 15 years old at the beginning of data collection ($M_{age\ T1}$ = 13.06 years) and between 13 and 15 years old at the second data collection time ($M_{age\ T2}$ = 13.61 years). All participants had started learning English at school in their third year of primary education (i.e. at the age of 7-8) on a 60-minute weekly basis, and received up to 3 hours per week of subsequent EFL instruction at school. In addition, all participants except four[13] had attended at least three years of extra-curricular EFL lessons outside the school (64.5% reported more than five years). All participants reported normal hearing, and none had any detectable speech disorder.

In a background questionnaire about their linguistic profile completed at T1 (see Appendix 1), AH participants indicated whether they had participated in a similar SA programme before. As illustrated in Figure 5.4 below, about half of the participants (54.8%) had not studied English abroad before; nine participants (29%) had already spent shorter periods of time in an English speaking country (only 2 of them reported it was for learning purposes, while the rest of learners reported it was on holiday), and 5 participants[14] (16.1%) reported having had a previous 3-month SA experience:

[13] Participants' code: MABA, (who had attended extracurricular English lessons for two years), MASA (who had not attended any extra-curricular English lessons before), and PAIZ and SACU (who had attended lessons for 1 year).
[14] Participants' code: ALPE, EDRO, EVCO, JAME, PAIZ.

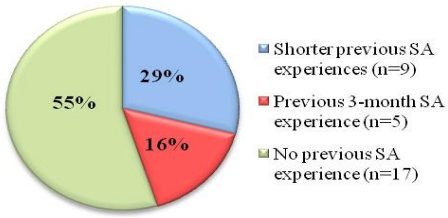


Figure 5.4: AH participants and previous experience in an English speaking country

During the 3-month SA period, the AH participants followed formal instruction at their school and received 3 hours of English lessons per week. AH participants were distributed in different class groups, so they had different teachers. However, the instructors teaching at the same level (i.e. 2nd year) delivered the same contents and assessed students on the same aspects.

In order to have a more detailed description of the AH learning context, we analyzed both learners' and teachers' answers to a questionnaire they completed during the research period. SA and AH participants completed a questionnaire at pre-test in which they were asked to provide information regarding the teaching/learning practice in their English lessons at school (see Appendix 3). A general picture of the language aspects[15] worked during the lessons and L2 use in the English classroom was built by analysing the participants' answers to questions 5, 6, 7, 8 and 9 in the questionnaire. This description will be provided in the following pages of this section. Information regarding the contents of the English lessons AH and the L2 use in the classroom was also collected from the teachers responsible of the English lessons at mainstream secondary education in the AH institution (n=4). Before the perception task, the teachers answered an online questionnaire (see Appendix 4) covering different aspects: (1) personal information; (2) linguistic profile; (3) teacher training; (4) views, attitudes and beliefs about EFL teaching; (5) their own daily classroom practice; (6) evaluation of pronunciation; and (7) SA programmes.

As shown in Figure 5.5 below, learners reported that from their point of view, writing and grammar were the most important aspects for their teachers. According to 71.4% of learners, writing was the most important aspect for their teacher, followed by grammar (58.1%) and vocabulary (52.7%). Oral skills and pronunciation were on the other extreme of this assessment. Pronunciation was regarded as 'not important' or 'not very

[15] A recent version of these questions included 'cultural aspects' as an item to be assessed. The teachers' questionnaire included this new version.

important' by 30% of the participants, and listening and speaking were considered as 'not important' or 'not very important' by 26.7% and 25% of the students, respectively.

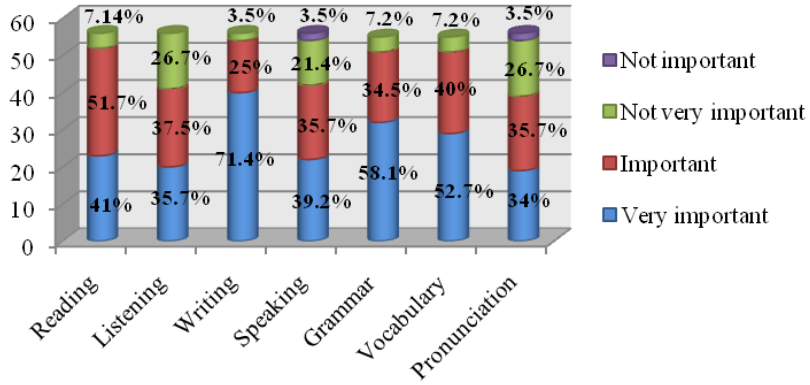


Figure 5.5: Answers provided by SA and AH participants (n=56) regarding Q5: *In your opinion, how important is each of the following language aspects for your teacher?*

The results obtained in question 5 were in line with the answers provided by the learners regarding question 6, where they were asked about the amount of time devoted throughout one term to each of the aforementioned language skills and aspects. As illustrated in Figure 5.6 below, about 70% of the learners pointed out that grammar was present in every/nearly every session. Fifty-seven percent of the participants also noted that writing skills and vocabulary were very frequent aspects in their English lessons. In line with the results obtained in the previous question, speaking and pronunciation were reported as the aspects which were given less attention in everyday lessons. Forty-nine percent of the learners considered 'speaking' was never or rarely worked on during the lessons, and about 52% of the participants indicated the same situation happening for pronunciation practice.

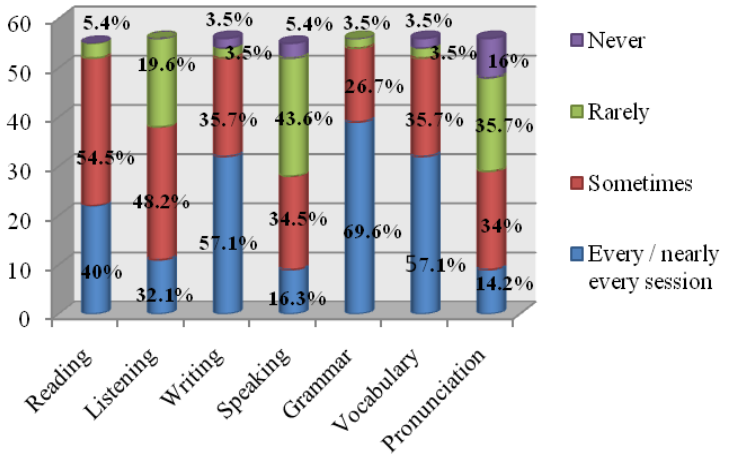


Figure 5.6: Answers provided by SA and AH participants (n=56) regarding Q6: *How much time is devoted to each of the following language aspects throughout one term?*

As mentioned in Chapter 3 of this dissertation, speaking is usually reported as the language skill which is given less time for practice in formal instruction in Spain. Answers from the learners' questionnaire have just confirmed this situation. In addition, teachers responsible for EFL lessons of the SA and AH participants corroborated this information. We asked the teachers in the AH context to consider the amount of practice of reading, listening, writing, speaking, grammar, vocabulary, pronunciation and cultural aspects in their lessons throughout one term (see question 23 in Appendix 4). Teachers' answers revealed that most of them (75%) included reading, listening and writing skills practice in every lesson, while the percentage lowered to 50% when referring to speaking skills. As for the language aspects (grammar, vocabulary, pronunciation and cultural aspects), cultural aspects were identified as the item which was given less attention in the formal instruction context, followed by pronunciation. On the contrary, grammar and vocabulary were regarded as language aspects which received most attention by most teachers. Table 5.2 below shows a summary of the teachers' answers to question 23:

	Reading	Listening	Writing	Speaking	Grammar	vocabulary	Pronunciation	Cultural aspects
Teacher 1	A	A	A	A	A	A	A	S
Teacher 2	F	F	F	F	F	F	S	R
Teacher 3	A	A	A	F	A	A	A	S
Teacher 4	A	A	A	A	A	A	A	F
(A=always; F=frequently; S=sometimes; R=rarely; N=never)								

Table 5.2: Teachers' answers to question 23 (n=4): *Consider your teaching throughout one term. How often do your students practice the following aspects?*

In their answers to questions 30-32, teachers indicated that they did not devote enough time to pronunciation in their lessons and pointed out that teaching pronunciation was not an easy task. The large number of students in the classrooms made pronunciation practice difficult for 50% of the teachers.

As far as language use in the classroom is concerned, both learners and teachers agreed on the fact that teachers usually used English during the lessons (see Figure 5.7 and figure 5.8):

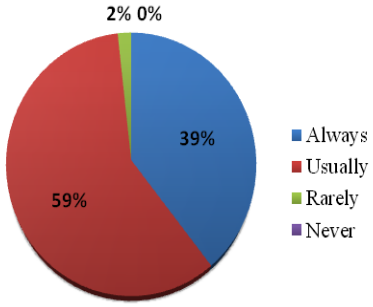


Figure 5.7: Answers provided by SA and AH participants (n=56) regarding Q7: *Does your teacher speak English during the lessons?*

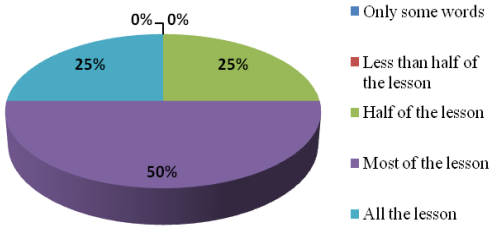


Figure 5.8: Answers provided by EFL teachers (n=4) regarding Q20: *How much English do you speak during the lessons?*

However, learners' use of English in the classroom was less frequent. 50% of the teachers reported that learners spoke in English most of the lesson, while the remaining 50% reported that students did so during half or less than half of the lesson. Learners provided more negative answers regarding the use of English among themselves. Question 9 in the questionnaire asked the learners whether they used English in student-student interactions during the lessons. 64.2% answered that they never spoke English with other classmates. It is worth remarking that the percentage of AH learners answering 'never' was higher (74%) than the percentage of SA learners (52%). In addition, 12% of the SA participants reported they 'always' or 'sometimes' spoke in English with their peers, while only 3% of the learners in the AH group indicated that they 'sometimes' spoke in English in this type of interaction. Figure 5.9 below illustrates these differences between the two groups of participants:

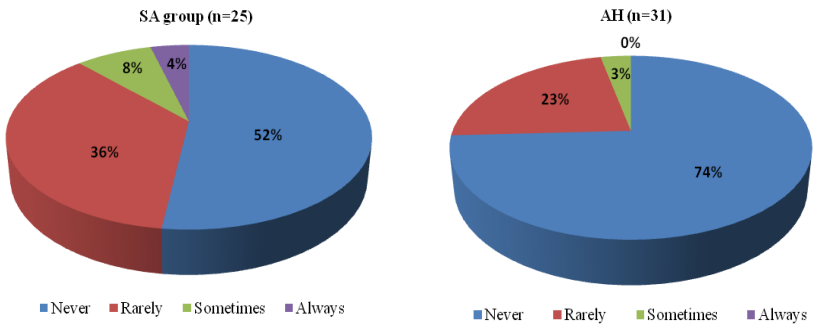


Figure 5.9: Answers provided by SA (n=25) and AH (n=31) participants regarding Q9: Do you speak English with your classmates during the lessons?

With regard to student-teacher interaction, 75% of the teachers reported that some of their students spoke in English to them, and 25% indicated that most of the students did so. The answers from learners revealed that 34% always spoke English to their teachers, 48.2% usually, 16% rarely, and 1.8% reported that they did not use English with their teachers. The results of this question were quite similar between the two groups of participants (SA and AH) as shown in Figure 5.10 below:

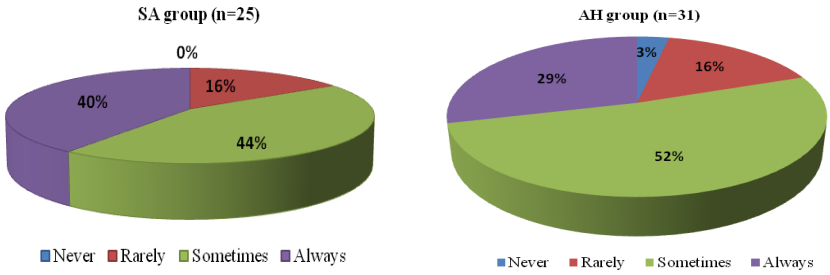


Figure 5.10: Answers provided by SA (n=25) and AH (n=31) participants regarding Q8: *Do you speak English with your teacher during the lessons?*

5.2.1.3 NS group

This group was formed by 15 English native speakers (10 females, 5 males) who were attending their second year of compulsory secondary education at a state school in Majorca (Spain). Two of these participants were born in England and had arrived in Majorca 5-6 years before to the time of testing. The rest of students in this group were early English-Spanish bilinguals. All speakers in this group were between 13 and 14 years old at data collection time.

5.2.2 Data collection: instruments and procedure

As for the speaking tests, participants were asked to carry out three different tasks: (1) reading aloud a short text; (2) telling an oral narrative from a picture story; (3) performing a role play in pairs. The individual oral narrative test was used for the purpose of the present study.

The speakers' extemporaneous speech was elicited using a six-frame pictures story about a bank robbery[16] (see Appendix 5). The speakers first studied the picture story for about one minute and then were recorded individually. High quality digital recordings were made at the learners' schools on different days.

The participants' oral productions were then digitized. Table 5.3 shows the duration of the full productions for each group at the different collection times:

[16] ©UIB-SALA COLE Oral Narrative.
©UPF-SALA Projecte de Recerca Coordinat.

	Range duration	Mean duration
SA (n=25)	T1: 0:39-1:47	1:10
	T2: 0:27-1:47	0:59
AH (n=31)	T1: 0:36-2:19	1:16
	T2: 0:35-2:13	1:06
NS (n=15)	T0: 0:42-1:35	0:59

Table 5.3: Range duration and mean duration
of full productions for all groups of participants

A shorter excerpt was extracted from the middle-end part of each narrative. Therefore, the content of the speech samples was kept relatively consistent across speakers. The first few seconds of the narrative were excised from the recordings by eliminating all dysfluencies (e.g. false starts) and by using natural pauses to demarcate the end of each excerpt. The preparation of speech samples for the perception task was conducted with Praat software (Boersma & Weenink, 2009). The excerpts from the two time periods (T1 and T2) for the SA and AH participants, and from T0 in the case of the NS group, were then normalized for peak intensity and randomized for presentation to the listeners. This procedure is consistent with previous studies using ratings of speech samples from the same task (Rossiter, 2009; Trofimovich & Isaacs, 2012; Derwing et Munro, 2013).

Table 5.4 below shows the range duration and mean duration of the final selected samples at the two data collection times for all groups of participants:

	Range duration	Mean duration
SA (n=25)	T1: 0:16-0:38	0:22
	T2: 0:13-0:30	0:21
AH (n=31)	T1: 0:15-0:35	0:22
	T2: 0:13-0:30	0:19
NS (n=15)	T0: 0:12-0:22	0:18

Table 5.4: Range duration and mean duration
of selected samples for all groups of participants

A total of 127 speech samples were obtained from the three groups of participants in the study. The SA group (n=25) recorded the oral narrative at two data collection times (25x2=50 speech samples), the AH group (n=31) recorded the story at two data collection times (31x2=62 speech samples), and the group of baseline native speakers (n=15) produced the speech samples once (15 speech samples).

5.3 Perception task

Measures of perceived degree of FA and comprehensibility were obtained from 12 listeners who performed a rating task. In the following sub-sections a description is provided of the listeners group, the instruments and the procedure for the rating task, and the analyses conducted with the data obtained from the perception task.

5.3.1 Participants: listeners

The speech samples were rated by 12 native speakers of Spanish/Catalan teaching EFL in mainstream secondary education in Spain (males = 1; females = 11). They ranged in age from 29 to 46 years (M_{age}=36.75). All listeners reported normal hearing.

Before carrying out the rating task, listeners completed the same online questionnaire for EFL teachers referred to in section 5.2.1.2 (see Appendix 4). As for their linguistic profile, seven listeners reported Spanish as their mother tongue, four considered Catalan and Spanish as their mother tongue and one listener reported that his mother tongue was Catalan. They also reported familiarity with British and American accents, and were highly familiar with the Spanish/Catalan-accented speech they had to assess, as they shared the learners' L1 background. When asked about the variety of English they usually spoke/wrote, 44.6% reported English as an International Language[17], 33.3% selected American English and 25% British 'RP'.

All the listeners had a BA either in English Studies, English Philology or Translation and Interpreting, and 3 of them also had a PhD in linguistics. Their EFL teaching experience ranged from 4 to 25 years ($M_{teaching\_experience}$=12.6). They rated their own knowledge of phonetics/phonology in English in a scale from 1 to 5, and the results indicated a mean of knowledge of 3.8. The same result was obtained when they were asked to rate their own pronunciation of English (M=3.8). Among the reasons given for not rating their pronunciation of English as 'excellent', listeners pointed out the following:

> T1: "When I speak English I don't pay so much attention to my pronunciation, so I may make some pronunciation mistakes. There might be words that are not of common use and I might

[17] 'English as an International Language' was defined in the questionnaire as "the English language as a global means of communication, not related to any specific geographical variety".

mispronounce them by mistake. There might be a sound which is hard for me to pronounce in English, for example glottalizations".

T2: "Teaching methods used when I was in primary and secondary school. No early exposure to the real English language. No early interaction in English".

T3: "I have never studied phonetics at university. It has always been in my "to do" list".

T4: "It lacks lustre".

T5: "I have a very bad ear for music, phonetics, and so on".

T6: "Lack of practice".

T7: "I consider my pronunciation quite acceptable although far from native-like pronunciation which I consider excellent, especially when it comes to the articulation and length of certain vowel sounds which do not exist in Spanish".

T8: "Problem with vowel phonemes, intonation and rhythm".

As for previous specific training in English pronunciation, 66.6% reported they had had some training before, while 33.3% indicated they had not. With regard to specific training in how to teach pronunciation, about 60% said they had not taken any course, while 40% had. When they were asked which specific training in EFL teaching they would like to receive, about 70% answered that they would like to have a course in phonetics and phonology (in rhythm and intonation, in particular), and also in how to teach pronunciation. Teachers were also asked about their feelings when teaching English pronunciation. About 75% of the listeners indicated that they felt motivated, 'comfortable', and 'confident', and 33.3% also reported lack of confidence and preparation to carry out this task (some of them reported being motivated but untrained).

We did not design any test or activity to assess listeners' general English language competence but we asked them to complete question 7 in the questionnaire (Appendix 4). This question asked the teachers to rate their level of reading, listening, writing, speaking, grammar, vocabulary, different aspects of pronunciation and cultural aspects in English assigning one of the following qualifications to themselves: *excellent, good, acceptable, poor*. Table 5.5 below shows the overall results for this question:

	Reading	Listening	Writing	Speaking	Grammar	Vocabulary	Pronunciation of individual sounds	Word stress	Rhythm	Intonation	Cultural aspects
Exc.	**16.6**	16.6	**16.6**	25	**33.3**	16.6	41.6	33.3	25	25	8.3
Good	**83.3**	66.6	**83.3**	58.3	**66.6**	75	50	50	33.3	50	66.6
Accept.	0	16.6	0	16.6	0	8.3	8.3	16.6	**25**	**25**	**25**
Poor	0	0	0	0	0	0	0	0	**16.6**	0	0

Table 5.5: Percentage of listeners for each item and punctuation.

As can be observed in Table 5.5, all the listeners rated themselves as having an excellent or good command of writing, reading and grammar. Only the word stress item received a poor punctuation by 16.6% of the listeners. Actually, 41.6% of the listeners regarded word stress as a poor or acceptable area, followed by intonation and cultural aspects, which were considered as acceptable by 25% of the listeners.

5.3.2 Data collection: instruments and procedure

Foreign accent and comprehensibility measures were obtained by means of a rating task. In previous studies, listening sessions were held in a quiet room trying to accommodate the listeners' schedules, and raters heard the speech samples individually and marked their accentedness and comprehensibility ratings in a response sheet (Derwing, Munro & Thompson, 2008; Derwing & Munro, 1997; Derwing & Munro, 2013; Munro & Derwing, 1999; Munro, Derwing & Morton, 2006). When the listeners were asked to write down the features of speech that they attended to when scoring, they were asked to do so into separate preformatted boxes for the different speech dimensions (Rossiter, 2009; Trofimovich & Isaacs, 2012). Other studies have presented the listeners with perception and rating tasks created and run with Praat software (Avello, Mora & Pérez-Vidal, 2012).

In the present study, the rating task was created and presented to the listeners using the software e-learning platform Moodle[18]. A course called *Foreign Accent and Comprehensibility* (see Appendix 6) was designed. In the main page of the course, there was a welcome message and an introduction to the online rating task providing information about the context of the experiment and the procedure. In the context section, listeners could download the six-frame pictures story about a bank robbery (see Appendix 5). They were instructed to view the cartoon story on which the oral narratives were based to minimize familiarity effects. This eliminated the possibility that the first few samples might be rated differently because the listeners did not initially know what to expect. In the procedure section a pdf document was posted providing full detailed instructions for the rating task (see Appendix 7).

Next the listeners heard the speech samples produced by the SA group (n=25) and the AH group (n=31) at Pre-test and Post-test, and by the group of baseline native speakers (n=15), who had been recorded once. Listeners heard the 127 stimuli in randomized order and assigned ratings using separate seven-point Likert-type scales for accentedness (1= heavy foreign accent, 7= native-like accent) and comprehensibility (1=extremely difficult to understand, 7=extremely easy to understand). As indicated in the instructions for the listeners, accentedness was defined as how different they thought the speaker sounded from a native speaker of English, if at all; and comprehensibility as how easy or difficult the sample was to understand. Listeners were instructed to use the whole scale over the course of the experiment. Figure 5.11 below shows an example of the rating task window on Moodle:

[18] A pilot experiment was conducted successfully with two different listeners.

Figure 5.11: Rating task window on Moodle (sample only for rating)

This methodology has been widely used in research on L2 speech production analysing constructs such as perceived FA, comprehensibility and fluency. A 9-point scale has been most commonly used in this type of studies, in which participants usually differed greatly in proficiency level. However, a 7-point scale was deemed more appropriate for the data in the present study, taking into account the smaller degree of variability in our speech samples (SA and AH participants with a similar age and proficiency level), as compared to other FA and comprehensibility studies.

The listeners were also asked to type in the aspects of speech that had influenced their comprehensibility ratings for 20 of the speech samples. They were instructed to write down the aspects of speech that they had found more striking and that they had taken into account when rating comprehensibility. They could use bullet points and report their impressions in English, Spanish and/or Catalan. Figure 5.12 shows an example of this specific part:

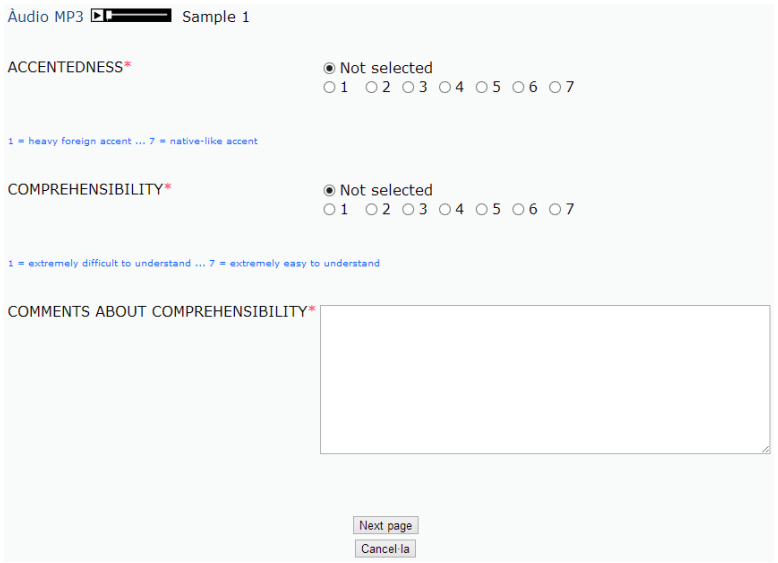


Figure 5.12: Rating task window on Moodle
(sample for rating and typing in comments about comprehensibility)

The whole rating experiment was a self-paced task. The listeners could play each speech sample as many times as needed and rate either accentedness or comprehensibility dimension first. After rating a sample, they had to click on the "Next page" button to listen to the following speech sample. Four samples were provided as rating practice at the beginning of the rating experiment so as to allow listeners to become familiar with the procedure.

The 127 speech samples were organised in 15 parts (with 8 or 9 speech samples each). Given that this was an online rating task, listeners could organize themselves at their convenience. The only restriction was that once they started whichever part of the experiment, they had to carry it out until the end. They could have a break or stop the experiment after finishing any part.

After completing the rating experiment, the listeners were asked to summarise their listening experience by answering a short online questionnaire (see Appendix 8)[19]. The main objective of this questionnaire was gaining insight into the aspects of speech that had affected listeners' ratings for accentedness and comprehensibility.

[19] The final version of this questionnaire was inspired on previous materials created and kindly shared by Ph.D. Pavel Trofimovich.

5.4 Data analyses

Longitudinal development in L2 learners' speech production was assessed through listeners' FA and comprehensibility ratings, which provided subjective measures of perceived degree of accentedness and comprehensibility. Listeners' ratings were extracted from the online rating experiment and transferred to an SPSS data editor. Statistical analyses were performed using the SPSS package (version 19).

Before conducting any statistical tests, we explored the data and checked rating consistency for FA and comprehensibility ratings by examining the inter-rater correlation coefficient. Normality was also ascertained by means of a Shapiro-Wilk test following Larson-Hall's (2010:84) recommendations for small sample sizes (under 50).

Longitudinal changes in FA and comprehensibility in the oral production of the same groups of participants (SA and AH) were examined. Mean FA and comprehensibility ratings for each learner at both Pre-test and Post-test were submitted to a mixed-design analysis of variance (ANOVA) with *group* as the between-subjects factor and *time* as the within-subjects factor. In order to assess whether listeners perceived any differences in FA and comprehensibility between T1 and T2, additional paired-samples t-tests were conducted. We also compared mean FA and comprehensibility ratings of the two groups at T1 and T2 by performing independent-samples t-tests.

In order to find out whether one learning context was more beneficial for FA/comprehensibility development than the other, we analysed FA and comprehensibility gains obtained by each group of participants. Independent-samples t-tests comparing the mean FA/comprehensibility gains of the two groups were conducted. Additional Chi-square analyses were ran in order to assess whether there was an association between learning context and FA/comprehensibility improvement.

Furthermore, we analyzed the effect of initial FA/comprehensibility level on FA/comprehensibility development over time. Correlations between FA/comprehensibility scores at T1 and FA/comprehensibility scores at T2 were conducted for the SA and AH groups. Additional correlations were also run between FA/comprehensibility at T1 and amount of gains to find out whether there was a relationship between initial FA/comprehensibility level and gains in that dimension.

So as to ascertain the differences in FA/comprehensibility development and gains among participants, we examined the differences in

FA/comprehensibility development between participants with higher FA initial level and participants with lower FA/comprehensibility level. To address this question, further mixed-design analyses of variance and additional t-tests were undertaken for each speech dimension and each group of participants. Given the reduced number of participants in the high initial level FA group AH, non-parametric techniques were used in the analyses involving this group of participants. Mann-Whitney U-tests were conducted as the non-parametric alternative to the t-test for independent samples, and Wilcoxon Signed Rank-tests were the alternative non-parametric option to the paired-samples t-test. We finally considered whether the two groups with the same initial FA or comprehensibility level (low or high) obtained the same gains in the two learning contexts (SA and AH). To answer this question, we compared the difference in amount of gains between the two low or high initial FA/comprehensibility groups by running independent-samples t-tests (or the alternative non-parametric option, Mann-Whitney U-tests).

In the last part of our research we examined the relationship between the participants' degree of FA and comprehensibility. Correlations between FA scores and comprehensibility scores were run in order to check for the existence of a relationship between the two dimensions and its strength. Moreover, the comments reported by the listeners in the online questionnaire they completed after the rating experiment were subjected to qualitative analyses in order to determine which aspects had influenced their FA ratings and which ones had affected their comprehensibility ratings. To gain insight into the comprehensibility dimension, further qualitative analyses were undertaken examining the data reported by the listeners in the rating experiment on Moodle when they were instructed to type in their comments for 20 of the speech samples.

In the following chapter, we present the results obtained with the measures and procedures just described.

Chapter 6

Results

This chapter presents the results of data analyses examining the speech dimensions of foreign accent and comprehensibility in the oral production of two groups of adolescent EFL learners (one spending a 3-month period abroad in an English speaking country, and the other receiving formal instruction in mainstream education at home). The chapter is divided into three main sections corresponding to the three research questions presented in Chapter 4.

Section 6.1 examines longitudinal changes in foreign accent in the oral production of two groups of participants (SA and AH). We compare FA development between the two groups, in order to determine whether one learning context is more beneficial than the other for improvement in FA. We also consider initial level of foreign accent, and analyse the patterns of longitudinal improvement observed at the group level for participants with higher and lower initial FA levels in both contexts (SA and AH). The first research question (RQ1) is addressed in this section:

RQ1: Do participants improve in terms of foreign accent after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is foreign accent development different for participants with different initial foreign accent level?

In section 6.2 longitudinal changes in comprehensibility in the oral production of the same groups of participants are examined. We also

compare comprehensibility development between the two groups (SA and AH) in order to find out whether one learning context is more beneficial than the other for improvement in comprehensibility. We analyse initial level of comprehensibility and examine the patterns of longitudinal improvement at the group level for participants with higher and lower initial comprehensibility levels in both contexts (SA and AH). The second research question (RQ2) is addressed in this section:

RQ2: Do participants improve in terms of comprehensibility after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is comprehensibility improvement different for participants with different initial comprehensibility level?

In the final section of this chapter (section 6.3) we explore the relationship between the participants' degree of foreign accent and comprehensibility. Moreover, we report and compare the aspects that listeners indicated as influencing their FA and comprehensibility ratings. The third research question (RQ3) is addressed in this section:

RQ3: To what extent are foreign accent and comprehensibility related speech dimensions?

All statistical analyses reported in this chapter were conducted using SPSS version 19, and the terminology and conventions used to describe and report statistical tests are in accordance with the recommendations of Pallant (2005). A criterion of $p<.05$ was adopted for statistical significance.

6.1 Changes in FA after SA and AH periods

In this section we explore data to address RQ1 and examine longitudinal changes in two groups of participants' FA after experiencing a SA and an AH context respectively:

RQ1: Do participants improve in terms of foreign accent after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is foreign accent development different for participants with different initial foreign accent level?

A total of 127 accent ratings[20] were obtained from the three groups of participants in the study: SA (n =25), AH (n=31), and NS (n=15). Inter-rater consistency was checked for these FA ratings by means of an intra-class correlation coefficient (ICC) analysis. Results yielded a high Cronbach's Alpha for the listeners' ratings at T1 (α=.983) and at T2 (α=.981), indicating a high degree of agreement among the twelve listeners at both rating times.

Normality was ascertained by means of a Shapiro-Wilk tests, which is the most powerful test for small sample sizes (n< 50), as noted by Larson-Hall (2010:84). These analyses indicated that the FA scores for the two of non-native participants at the two data collection times followed a normal distribution (statistic>.9; p>.05).

In line with previous research (Derwing & Munro, 2013), the mean of FA scores for native participants in our research (M=6.71, SD=0.41) indicated that listeners had recognized them during the rating task, and had assigned them high scores on the 7-point rating scale.

6.1.1 Longitudinal improvement in FA

After exploring the data and examining both inter-rater reliability and normality, we conducted the statistical analyses in order to answer the first research sub-question:

RQ1a: Do participants improve their FA ratings after a 3-month SA period, or a 3-month FI period in an AH context?

A first exploration of the data suggested that listeners' ratings pointed towards some improvement in foreign accent in the case of the SA group participants between Pre-test (*M*=3.17, *SD*=1.03) and Post-test (*M*=4.01, *SD*=1.10), while ratings assigned to the AH group were more similar between Pre-test (*M*=2.44, *SD*=0.70) and Post-test (*M*=2.68, *SD*=0.75). Table 6.1 presents the mean FA ratings of each group at the two data collection times:

[20] The listeners heard the speech samples produced by the SA group (n=25) at two data collection times (25x2=50 speech samples), the AH group (n=31) at two data collection times (31x2=62 speech samples), and a group of baseline native speakers (n=15), who produced the speech samples once (15 speech samples).

	Group	Mean	SD
Mean T1	SA (n =25)	3.17	1.03
	AH (n=31)	2.44	0.70
Mean T2	SA (n=25)	4.01	1.10
	AH (n=31)	2.68	0.75

Table 6.1: Mean FA ratings for SA group (Pre-test and Post-test) and AH group (Pre-test and Post-test)

One-sample t-tests were conducted in order to find out whether mean FA ratings at Pre-test and Post-test for the two groups of non-native participants (SA and AH) were significantly lower than the native speakers' mean FA ratings (M=6.71). Results showed that non-native participants and native participants differed significantly in terms of FA at Pre-test [$t(24)=-17.156$, $p<.001$, $d=4.51$ for the SA group; $t(30)=-33.604$, $p<.001$, $d=3.25$ for the AH group], and they remained so at Post-test [$t(24)=-12.236$, $p<.001$, $d=7.44$ for the SA group; $t(30)=-29.784$, $p<.001$, $d=6.66$ for the AH group].

Mean FA ratings for the two groups of learners at both Pre-test and Post-test were submitted to a mixed-design analysis of variance (ANOVA) with *group* (SA and AH) as the between-subjects factor and *time* (Pre-test and Post-test) as the within-subjects factor. Table 6.2 shows the results of the ANOVA evaluating changes in FA ratings over time:

Time	F(1)=35.43	p<.001	η2=.396
Group	F(1)=29.12	p<.001	η2=.282
Time x Group	F(1)=10.62	p=.002	η2=.164

Table 6.2: Results of mixed-design ANOVA with *group* (SA and AH) and *time* (Pre-test and Post-test) as factors

Results from this analysis yielded a significant main effect of *time* [$F(1)=35.43$; $p<.001$; $\eta 2=.39$], *group* [$F(1)=29.12$; $p<.001$; $\eta 2=.28$], and *time* x *group* interaction [$F(1)=10.62$; $p=.002$; $\eta 2=.16$].

Additional paired-samples t-tests showed that there was significant improvement in FA between T1 and T2 for the SA group [$t(24)=-5.32$, $p<.001$, $\eta^2=.54$], and also for the AH group [$t(30)=-2.38$, $p=.024$, $\eta^2=.15$]. The eta squared values indicated large effect sizes for both groups of learners, although the effect size was greater for the SA group. Figure 6.1 illustrates the mean FA ratings assigned to each group at the two testing times:

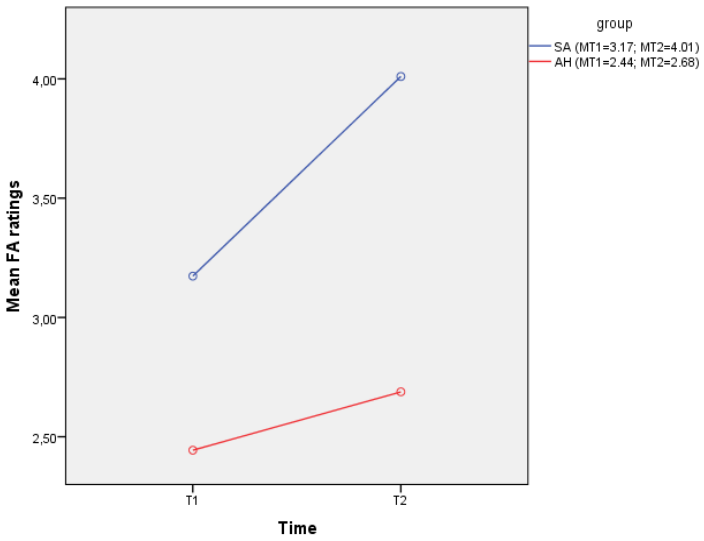


Figure 6.1: Mean FA ratings at T1 and T2 for SA and AH groups

Results from an independent-samples t-test comparing the FA ratings of the SA and AH groups at T2 showed a significant difference between FA ratings at T2 between the two groups [$t(40.76)= 5.11$, $p<.001$, $\eta 2= .32$]. The eta squared value indicated a large effect size. As can be observed in Figure 6.1 above, FA ratings between the two groups already differed at T1. An independent-samples t-test comparing the FA ratings of the two groups (SA and AH) at T1 indicated that the SA group and the AH group already differed significantly at T1 [$t(54)= 3.13$, $p=.003$, $\eta^2= .15$]. The eta squared value also showed a large effect size, although smaller than at T2.

6.1.2 Gains in FA: a comparison between SA and AH outcomes

Another aim of this research was to comparte FA development over the course of the SA and AH periods, as formulated in sub-question 1b:

RQ1b: Is one learning context (SA vs. AH) more beneficial than the other? In other words, do SA participants obtain greater gains in foreign accent development than the AH group?

To address this question, we analysed the FA gains obtained by each group. A new variable, 'FA gains', was created, which represented the amount of FA gains obtained by each learner. The value resulted from subtracting the mean ratings obtained at T1 from the mean ratings

obtained at T2 (T2 minus T1). Subsequently, the mean of FA gains for each group was calculated. Table 6.3 shows these results:

	FA gains	
Group	**Mean**	**SD**
SA (n = 25)	0.83	0.78
AH (n = 31)	0.24	0.57

Table 6.3: Mean FA gains for SA and AH groups

As illustrated in Figure 6.2 below, data showed that the SA group ($M_{FAgains}$=0.83) obtained greater gains than the AH group ($M_{FAgains}$=0.24). In order to check whether this difference was statistically significant, an independent-samples t-test comparing $M_{FAgains}$ of the two groups (SA and AH) was conducted. This analysis revealed that the SA group obtained significantly greater gains than the AH group [$t(54)= 3.25$, $p=.002$, $\eta^2=.16$]:

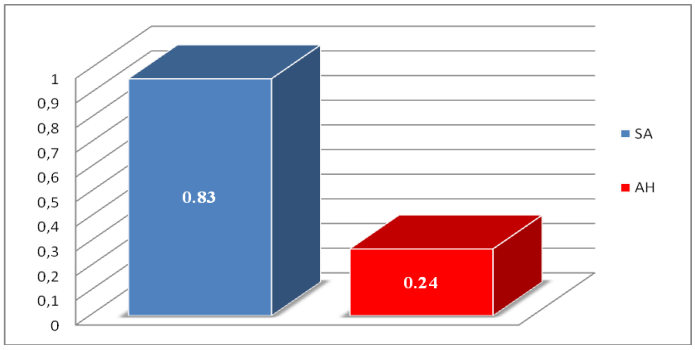


Figure 6.2: Mean FA gains for SA and AH groups

An additional chi-square analysis was conducted in order to assess whether there was an association between learning context and FA improvement. A new categorical variable was created, 'FA improvement', with two categories (yes/no). Learners were assigned to one of the two categories based on the presence of improvement in their FA score (FA gains>0) or lack thereof. Table 6.4 illustrates the percentage of learners in each group who improved their FA score between T1 and T2:

	FA Improvement	
	no	**yes**
SA (n=25)	4 16.0%	21 84.0%
AH (n=31)	12 38.7%	19 61.3%

Table 6.4: Percentage of participants improving FA after SA or AH period

As observed in table 6.4, most of the learners in the SA group improved their FA ratings (84.0% improvers versus only 16.0% non-improvers), whereas the difference between the percentages of improvers and non-improvers in the case of the AH learners was smaller (61.3% versus 38.7%). The chi-square test statistic did not reveal a significant difference between the proportion of SA learners who improved their FA ratings and the proportion of AH learners within a 95% confidence interval [$\chi^2(1)=3.497$, $p=.061$], although the reported p-value closely approached the .05 significance level. Actually, previous studies have referred to this type of p-values between .06 and .10 as 'marginally significant' (Højen, 2003, Fullana, 2005; Valls-Ferrer, 2011). Cramer's phi ($\varphi=-.250$, $p=.061$) suggested a low effect size. The odds ratio indicated that SA learners were 3.31 times more likely to improve their FA ratings than AH learners.

6.1.3 Effect of initial FA level

This section examines the effect of initial FA level on FA development over time. We observed whether the participants' FA level at T1 conditioned FA development during the 3-month SA and AH periods. We addressed the following sub-question:

RQ1c: Is foreign accent development different for participants with different initial foreign accent level in each group (SA and AH)?

Correlations between FA scores at T1 and FA scores at T2 were conducted for each group. The results of this analysis are reported in Table 6.5:

	SA (n=25)	**AH** (n=31)
Pearson	.730	.694
Sig.	<.001	<.001

Table 6.5: Pearson correlations between FA level at Pre-test and FA level at Post-test for SA and AH groups

Pearson correlation coefficients indicated large significant positive correlations between FA scores at T1 and T2 in both groups, suggesting quite a strong relationship between the ratings obtained before and after the SA and AH periods: the higher the FA scores at the pre-test, the higher the FA ratings at the post-test in both contexts. Likewise, those participants with lower FA scores at T1 obtained lower FA scores at T2.

Additional correlations were conducted between FA scores at T1 and amount of gains in order to test whether there was a relationship between initial FA level and gains in FA. In other words, did participants with

lower initial FA level obtain greater gains than those with higher initial FA level and vice versa? Table 6.6 shows the results of this analysis:

	SA (n=25)	**AH** (n=31)
Pearson	-.286	-.324
Sig.	.166	.076

Table 6.6: Pearson correlations between initial FA level and amount of gains for SA and AH groups

Pearson correlation coefficients yielded negative correlations between initial FA level and amount of gains in both groups (small in the case of the SA group, and moderate in the case of the AH group), suggesting that the higher the initial FA rating (i.e., the more native-like the accent), the lower the gains; and the lower the initial FA rating (i.e., the worse the accent), the higher the gains. Neither of these correlations reached significance though. Therefore, in order to gain a better understanding of possible initial level effects on FA development and gains among participants, we analyzed the differences in FA development between participants with higher initial FA level and participants with lower initial FA level.

We considered initial FA level as a function of participants' FA scores at the beginning of the study (T1). We followed the following grouping criterion in order to form two low initial FA level groups (one for the SA group and the other for the AH group), and two high initial FA level groups (one for the SA group and the other for the AH group)[21]:

1. Low initial FA level group (Low-iFA): participants obtaining a FA score <3.5 at Pre-test;

2. High initial FA group (High-iFA): participants obtaining a FA score ≥3.5 at Pre-test.

Table 6.7 shows the participants' distribution following this grouping criterion:

	SA	**AH**
Low-iFA group	n=17	n=28
High-iFA group	n=8	n=3

Table 6.7: Distribution of SA and AH participants in Low and High initial FA level groups

[21] Other criteria for dividing students into two level groups were considered (e.g., using a Median Split, or selecting the ten participants with the highest/lowest FA scores at T1). However, neither of these criteria allowed us to have SA and AH groups which did not differ statistically in terms of initial FA level.

The mean FA ratings at T1 and T2 for each group are indicated in Table 6.8 below:

		Mean	SD
Low- iFA-SA	T1	2.59	0.47
(n=17)	T2	3.55	0.76
High- iFA-SA	T1	4.39	0.78
(n=8)	T2	4.96	1.14
Low- iFA-AH	T1	2.27	0.49
(n=28)	T2	2.54	0.62
High- iFA-AH	T1	4.00	0.41
(n=3)	T2	4.00	0.65

Table 6.8: Mean FA ratings for Low and High initial FA level groups (SA and AH) at Pre-test and Post-test

The normality of the distribution of scores at T1 for all sub-groups was assessed. Shapiro-Wilk tests indicated that the FA scores for all initial FA level groups followed a normal distribution (statistic>.9, p>.05).

In the next sub-sections we present the results of the analyses examining the effect of initial FA level on FA development after the SA and AH periods. We compare the two low initial level groups (and the two high initial level ones), in order to explore the impact of learning context on FA development in participants with different initial FA levels. Parametric tests were conducted in the following analyses except for the tests using data from the High-iFA-AH sub-group. Non-parametric tests were run with this data given the small number of participants in this sub-group (n=3).

6.1.3.1 Initial FA level and SA outcomes

The first question that we examined was whether there were any differences in FA development between the two sub-groups in the SA group depending on their initial FA level. An independent-samples t-test comparing M_{FA} at T1 between the Low-iFA-SA and High-iFA-SA groups was conducted in order to confirm that the two groups differed significantly at T1. Differences between these two groups were highly significant [$t(23)$=-7.147, p=<.001, η^2= .68].

Subsequently, mean FA ratings for the two SA groups at both Pre-test and Post-test were submitted to a mixed-design analysis of variance (ANOVA) with *group* (High-iFA-SA and Low-iFA-SA) as the between-subjects factor, and *time* (Pre-test and Post-test) as the within-subjects factor. Table 6.9 shows the results of the ANOVA evaluating changes in FA ratings over time:

Time	F(1)=20.99	p<.001	η2=.477
Group	F(1)=33.34	p<.001	η2=.592
Time x Group	F(1)=1.34	p=.258	η2=.055

Table 6.9: Results of mixed-design ANOVA with *group* (Low-FA-SA and High-FA-SA) and *time* (Pre-test and Post-test) as factors

Results from this analysis yielded a significant main effect of *time* [$F(1)=20.99$, $p<.001$, $\eta^2=.477$], and *group* [$F(1)=33.34$, $p<.001$, $\eta^2=.592$], while there was not a significant *time* x *group* interaction [$F(1)=1.34$, $p=.258$, $\eta^2=.055$].

Additional paired-samples t-tests with testing time as the within-subjects factor showed that there was significant improvement between T1 and T2 for the Low-iFA-SA group [$t(16)=-5.573$, $p<.001$, $\eta^2= .65$], the eta-squared value indicating a large effect size. No significant improvement was found for participants in the High-iFA-SA group ($p=.122$). Figure 6.3 illustrates mean FA ratings assigned to each sub-group at the two testing times:

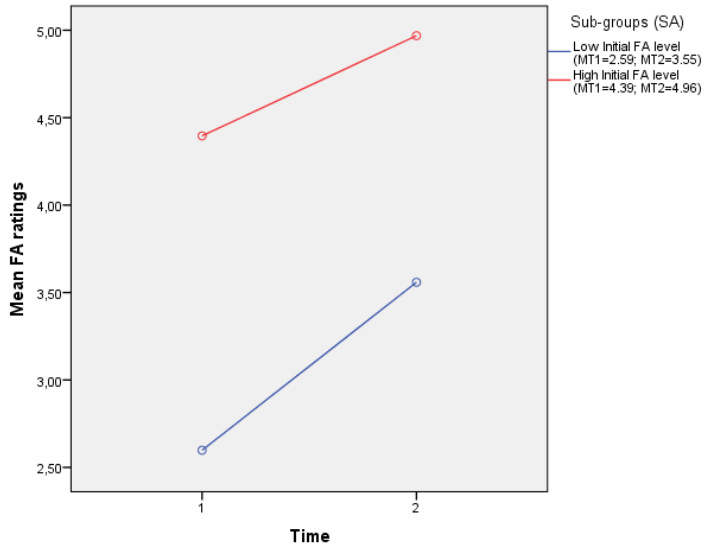


Figure 6.3: Mean FA ratings at T1 and T2 for High and Low initial FA level groups (SA)

As we have already reported at the beginning of this sub-section, results showed a significant difference in FA ratings between the low and high level SA sub-groups at T1. An independent-samples t-test also showed a significant difference in FA ratings between the two sub-groups at T2

[$t(23)$=-3.677, p=.001, η^2=.37]. The eta squared value also indicated a large effect size, although smaller than the large effect size found at T1.

As indicated in the ANOVA (see Table 6.9 above) there was no *time* x *group* interaction, which indicates that the effect of time was similar for the two groups, that is, that the change in FA scores over time for the two groups (Low-iFA-SA and High-iFA-SA) was similar. Figure 6.3 illustrates that the two groups experienced a parallel increase in their FA ratings.

The analysis of FA gains obtained by each sub-group confirmed this. Table 6.10 shows the mean of FA gains for each sub-group:

	FA gains	
Group	**Mean**	**SD**
Low-iFA-SA (n = 17)	.96	.710
High-iFA-SA (n = 8)	.57	.920

Table 6.10: Mean FA gains for Low-iFA-SA and High-iFA-SA groups

As illustrated in Figure 6.4 below, data showed that the Low-iFA-SA group ($M_{FAgains}$=.96) obtained greater gains than the High-iFA-SA group ($M_{FAgains}$=.57). However, an independent-samples t-test comparing $M_{FAgains}$ of the two groups revealed this difference was not significant (p=.258):

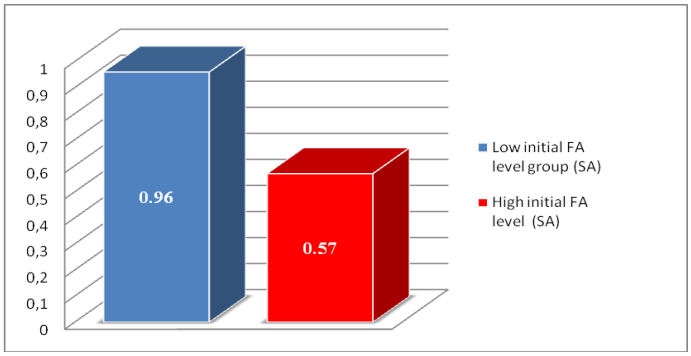


Figure 6.4: Mean FA gains for Low-iFA-SA and High-iFA-SA groups

An additional Chi-square analysis was conducted in order to assess whether there was an association between initial level and FA improvement. Table 6.11 illustrates the percentage of learners in each group who improved their FA score between T1 and T2:

	FA Improvement	
	no	**yes**
Low-iFA-SA (n=17)	3 17.6%	14 82.4%
High-iFA-SA (n=8)	1 12.5%	7 87.5%

Table 6.11: FA improvement of Low and High groups (SA)

The chi-square test statistic was not significant (p=.743), indicating that the proportion of High-iFA-SA learners who improved in FA was not significantly larger than the proportion of Low-iFA-SA learners.

6.1.3.2 Initial FA level and AH outcomes

A Mann-Whitney U-test confirmed that the FA rating of the Low-iFA-AH and High-iFA-AH sub-groups differed at T1 (Z=-2.81, p=.005, r=.505). Figure 6.5 illustrates the mean FA ratings assigned to each sub-group at the two testing times:

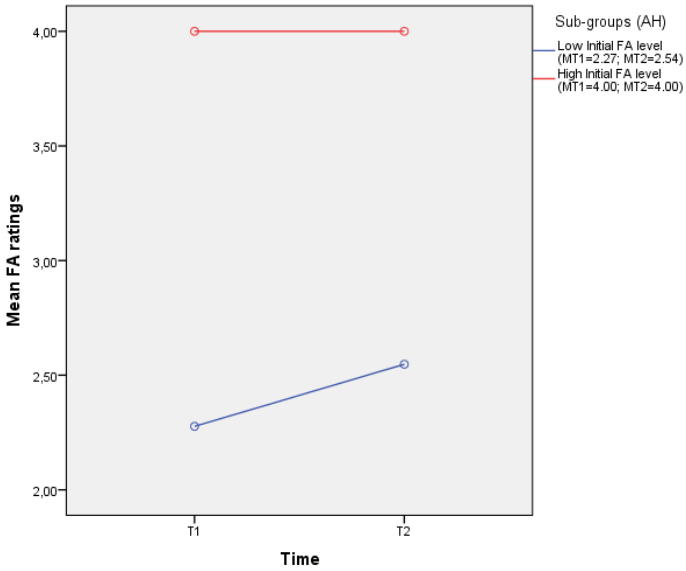


Figure 6.5: Mean FA ratings at T1 and T2
for High and Low initial FA level groups (AH)

A paired-samples t-test showed that there was significant improvement in FA between T1 and T2 for the Low-iFA-AH group [t(27)=-2.576, p=.016, η^2=.19]. The eta-squared value indicated a large effect size. In contrast, figure 6.5 illustrates that the mean FA ratings for the High-iFA-AH sub-group remained unchanged between T1 and T2. A Wilcoxon Signed

Rank-Test showed that there was not significant improvement between T1 and T2 for the High-iFA-AH group (p=1.000).

As was the case at T1, a Mann-Whitney U-test indicated that the FA ratings of the Low-iFA-AH and High-iFA-AH sub-groups continued to differ significantly at T2 (Z=-2.55, p=.011, r=.458) despite the significant improvement experienced by the Low-iFA-AH learners and the lack of changes in the scores of the High-iFA-AH learners (see table 6.12 and figure 6.6 below):

FA gains		
Group	**Mean**	**SD**
Low-iFA-AH (n=28)	.27	.556
High-iFA-AH (n=3)	.00	.794

Table 6.12: Mean FA gains for Low-iFA-AH and High-iFA-AH groups

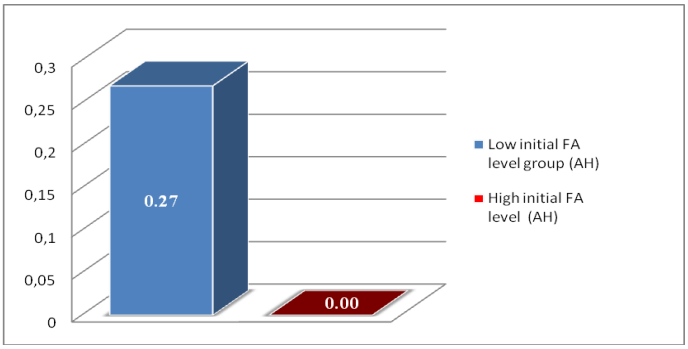


Figure 6.6: Mean FA gains for Low-iFA-AH and High-iFA-AH groups

As illustrated in Figure 6.6, data showed that the Low-iFA-AH group ($M_{FAgains}$=.27) obtained greater gains than the High-iFA-AH group ($M_{FAgains}$=.00). The difference between the gains in FA obtained by the Low-iFA-AH learners and the lack of gains for the High-iFA learners turned out to be non-significant (Z=-.503; p=.615).

An additional Chi-square analysis was conducted in order to assess whether there was an association between initial FA level and FA improvement. Table 6.13 illustrates the percentage of learners in each group who improved their FA score between T1 and T2:

	FA Improvement	
	no	**yes**
Low-iFA-AH (n=28)	11 39.3%	17 60.7%
High-iFA-AH (n=3)	1 33.3%	2 66.7%

Table 6.13: FA improvement of Low and High groups (AH)

The chi-square test statistic was not significant (p=.841), indicating that the proportion of High-iFA-AH learners who improved in FA was not significantly larger than the proportion of Low-iFA-AH learners.
We additionally considered whether the two sub-groups with the same initial FA level (i.e., low or high) had the same gains in the two different learning contexts considered (SA and AH). In other words, did SA participants with lower initial FA level have greater gains in FA than learners with lower initial FA level in the AH group? Likewise, did SA participants with higher initial FA level have greater gains in FA than participants with higher initial FA level in the AH group? These questions are addressed in the next sub-sections.

6.1.3.3 Gains in FA: a comparison between Low iFA-SA and Low-iFA-AH outcomes

Before analysing possible differences in amount of gains between the two low initial FA groups, we conducted an independent-samples t-test comparing the mean FA ratings of the Low-iFA-SA and Low-iFA-AH sub-groups at T1. In line with the results reported in section 6.1.1, results from this analysis showed that the two sub-groups already differed significantly at T1 [t(43)=2.144, p=.038, η2=.09], although the eta-squared indicated a small effect size [22].

Figure 6.7 illustrates the mean FA ratings assigned to each sub-group at the two testing times:

[22] An additional independent-samples t-test comparing the M_{FA} at T2 of the Low-iFA-SA and Low-iFA-AH group indicated that the difference in the M_{FA} at T2 between these two groups was also significant [t(43)=4.866, p<.001 η^2=.35].

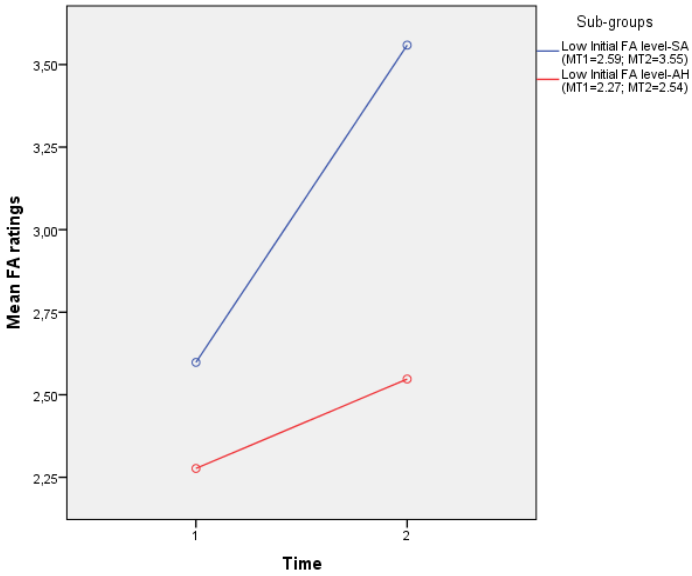


Figure 6.7: Mean FA ratings at T1 and T2
for Low initial FA level groups (SA and AH)

Subsequently, we explored FA gains for the two Low initial FA sub-groups. Table 6.14 shows the mean FA gains for each of these groups:

FA gains		
Group	**Mean**	**SD**
Low-iFA- SA (n=17)	0.96	0.71
Low-iFA-AH (n=28)	0.27	0.55

Table 6.14: Mean FA gains of low initial FA level groups (SA and AH)

As illustrated in Figure 6.8, the Low-iFA-SA group ($M_{FAgains}$=0.96) obtained greater gains than the Low-iFA-AH group ($M_{FAgains}$=0.27):

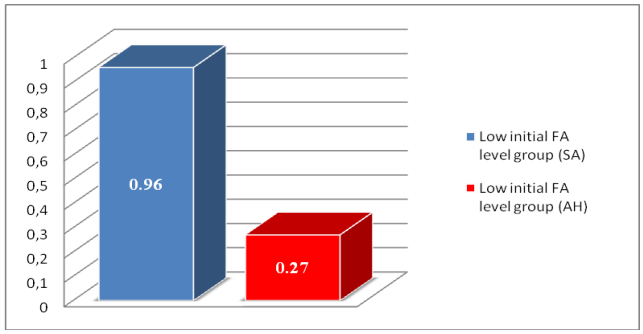


Figure 6.8: Mean FA gains for Low initial FA level groups (SA and AH)

An independent-samples t-test showed that there was a significant difference in FA gains between the two groups [$t(43)=3.629$, $p=.001$, $\eta^2=.23$]. The eta squared value reported a large effect size.

6.1.3.4 Gains in FA: a comparison between High-iFA-SA and High-iFA-AH outcomes

Non-parametric tests were conducted in order to address the impact of context on FA development in the case of the two High initial FA level sub-groups (SA and AH), although the small number of participants in the High-iFA-AH sub-group does not allow for generalizations and therefore results from analyses involving this group of participants must be taken cautiously. A Mann-Whitney U-test indicated that the High-iFA-SA learners and the High-iFA-AH learners did not differ significantly at T1 ($p=.413$)[23]. The fact that the difference between the two High-iFA sub-groups was not significant at T1, is worth mentioning since significant differences between the SA and AH groups, and between the Low initial FA subgroups (SA and AH) had been reported previously. Figure 6.9 illustrates the mean FA ratings assigned to each High initial FA sub-group at the two testing times:

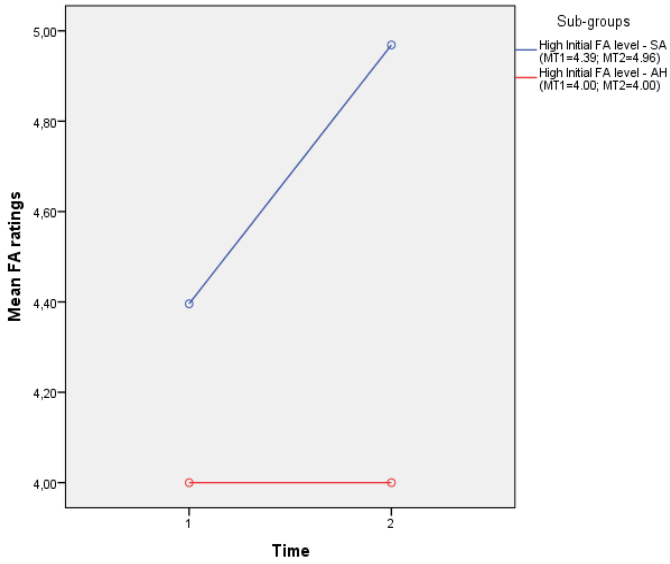


Figure 6.9: Mean FA ratings at T1 and T2
for High initial FA level groups (SA and AH)

[23] An additional Mann-Whitney U-test comparing the mean FA ratings of the High-iFA-SA and High-iFA-AH sub-groups at T2 indicated that the difference in FA between these two groups was not significant at T2 either ($p=.102$).

We explored FA gains for the two High initial FA sub-groups. Table 6.15 shows the mean FA gains for each of these groups:

FA gains		
Group	**Mean**	**SD**
H-iFA- SA (n=8)	.57	.92092
H-iFA-AH (n=3)	.00	.79495

Table 6.15: Mean FA gains for High initial FA level groups (SA and AH)

As illustrated in Figure 6.10, the High-iFA-SA group ($M_{FAgains}$=0.57) obtained greater gains than the High-iFA-AH group, for which scores actually remained unchanged between T1 and T2 ($M_{FAgains}$=0.00):

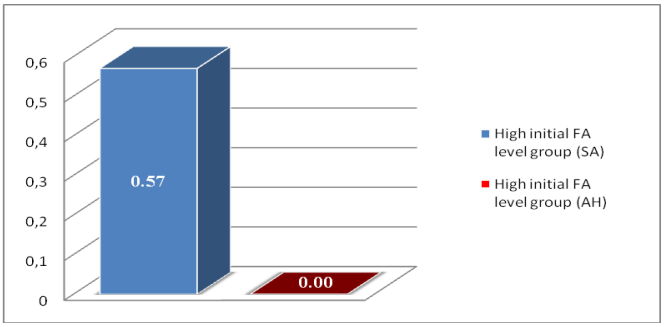


Figure 6.10: Mean FA gains for High initial FA level groups (SA and AH)

A Mann-Whitney U-test was run in order to check whether there was a significant difference between the gains obtained by the High-iFA-SA learners and the lack of gains for the High-iFA-AH learners. The results showed a non-significant difference in FA gains between the two groups (p=.356). As already mentioned, results from these analyses must be taken extremely cautiously given the small number of participants in the High-iFA-AH sub-group.

6.2 Changes in comprehensibility after SA and AH periods

In this section we examine changes in participants' comprehensibility after experiencing a SA or an AH learning context. We address RQ2 formulated below:

RQ2: Do participants improve in terms of comprehensibility after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is comprehensibility development different for participants with different initial comprehensibility level?

A total of 127 comprehensibility ratings[24] were obtained from the three groups of participants in the study: SA (n=25), AH (n=31), and NS (n=15). Inter-rater consistency was checked for these comprehensibility ratings by means of an intra-class correlation coefficient analysis. Results yielded a high Cronbach's Alpha for the listeners' ratings at T1 (α=.958) and the listeners' ratings at T2 (α=.939), revealing a high degree of agreement among the listeners at both rating times.

Normality was ascertained by means of Shapiro-Wilk tests (Larson-Hall, 2010:84). Preliminary analysis indicated that the comprehensibility scores for the two groups of non-native participants at the two testing times followed a normal distribution (statistic>.9; p>.05).

6.2.1 Longitudinal improvement in comprehensibility

After exploring the data and examining both inter-rater reliability and normality, we conducted the statistical analyses in order to answer the first research sub-question:

RQ 2a: Do participants improve their comprehensibility ratings after a 3-month SA period, or a 3-month FI period in an AH context?

A first exploration of the data suggested that listeners' ratings pointed towards some improvement in comprehensibility in the case of the SA group participants between Pre-test (*M*=4.04, *SD*=1.29) and Post-test (*M*=4.94, *SD*=0.87). Ratings assigned to the AH group were more similar between Pre-test (*M*=3.35, *SD*=1.17) and Post-test (*M*=3.88, *SD*=1.10). Table 6.16 shows the mean comprehensibility ratings of each group at the two data collection times:

[24] The listeners heard the speech samples produced by the SA group (n=25) at two data collection times (25x2=50 speech samples), the AH group (n=31) at two data collection times (31x2=62 speech samples), and a group of baseline native speakers (n=15), who produced the speech samples once (15 speech samples).

	Group	Mean	SD
Mean T1	SA (n=25)	4.04	1.29
	AH (n=31)	3.35	1.17
Mean T2	SA (n=25)	4.94	0.87
	AH (n=31)	3.88	1.10

Table 6.16: Mean comprehensibility ratings for SA group (Pre-test and Post-test) and AH group (Pre-test and Post-test)

One-sample t-tests were conducted in order to find out whether mean comprehensibility ratings at Pre-test and Post-test for non-native participants (SA and AH participants) were significantly lower than the native speakers' mean comprehensibility ratings (M=6.41). Results showed that non-native participants and native participants differed significantly in terms of comprehensibility at Pre-test [$t(24)$=-9.136, p<.001, d=2.45 for the SA group; $t(30)$=-14.457, p<.001, d=3.45 for the AH group], and they remained so at Post-test [$t(24)$=-8.362, p<.001, d=2.12 for the SA group; $t(30)$=-12.727, p<.001, d=3.01 for the AH group].

Mean comprehensibility ratings for the two groups of learners at both Pre-test and Post-test were submitted to a mixed-design analysis of variance (ANOVA) with *group* (SA and AH) as the between-subjects factor, and *time* (Pre-test and Post-test) as the within-subjects factor. Table 6.17 shows the results of the ANOVA evaluating changes in comprehensibility ratings over time:

Time	F(1)=30.28	p<.001	η2=.359
Group	F(1)=10.27	p=.002	η2=.160
Time x Group	F(1)=2.05	p=.158	η2=.037

Table 6.17: Results of mixed-design ANOVA with *group* (SA and AH) and *time* (Pre-test and Post-test) as factors

Results from this analysis yielded a significant main effect of *time* [F(1)=30.28; p<.001; η2=.35] and *group* [F(1)=10.27; p=.002; η2=.16]. The *time* x *group* interaction effect was not statistically significant [F(1)=2.05; p=.158; η2=.03].

Additional paired-samples t-tests showed that there was significant improvement in comprehensibility between T1 and T2 for the SA group [$t(24)$=-4.903, p<.001 , η^2= .50], and also for the AH group [$t(30)$=-2.935, p=.006 , η^2= .22]. The eta squared values indicated large time effects for both groups of learners, although the effect size was greater for the SA group. Figure 6.11 illustrates the mean comprehensibility ratings assigned to each group at the two testing times:

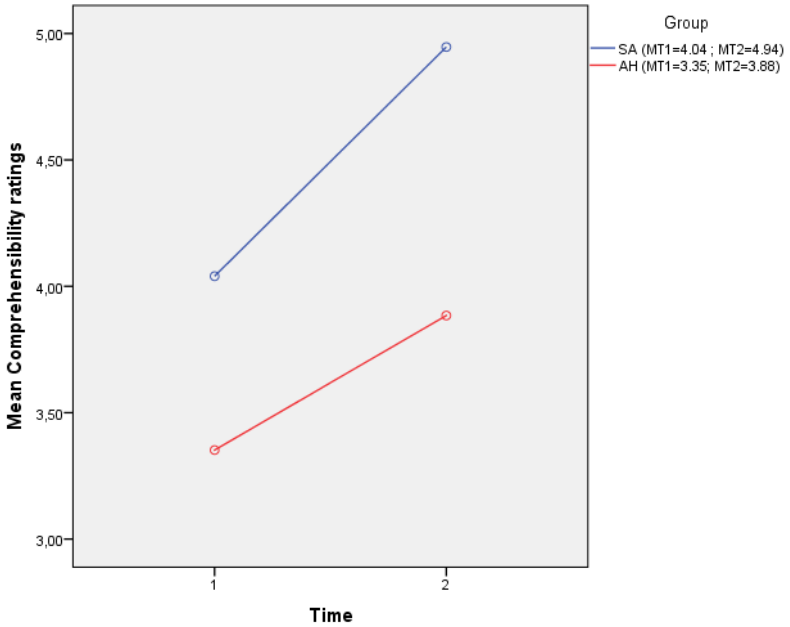


Figure 6.11: Mean comprehensibility ratings at T1 and T2 for SA and AH groups

Results from an independent-samples t-test comparing the comprehensibility ratings of the SA and AH groups at T2 showed a significant difference in comprehensibility ratings at T2 between the two groups [t(54)= 3.916, p<.001, η2= .22]. The eta squared value indicated a large effect size. As can be observed in Figure 6.11 above, comprehensibility ratings between the two groups already differed at T1. An independent-samples t-tests comparing the comprehensibility ratings of the two groups (SA and AH) indicated that the SA group and the AH group already differed significantly at T1 [$t(54)= 2.077$, $p=.043$, $\eta^2= .07$]. The eta squared showed a moderate effect size.

6.2.2 Gains in comprehensibility: a comparison between SA and AH outcomes

Another aim of this research was to analyse comprehensibility development over the course of the SA and AH periods, as formulated in sub-question 2b:

RQ2b: Is one learning context (SA vs. AH) more beneficial than the other? In other words, do SA participants obtain greater gains in comprehensibility development than the AH group?

In order to compare comprehensibility development over the course of the SA and FA periods, we analysed comprehensibility gains obtained by each group. A new variable: 'comprehensibility gains' was created, which indicated the amount of comprehensibility gains obtained by each learner. The value resulted from subtracting the mean ratings obtained at T1 from the mean ratings obtained at T2 (T2 minus T1). Subsequently, the mean of comprehensibility gains for each group was calculated. Table 6.18 shows these results:

Comprehensibility gains

Group	**Mean**	**SD**
SA (n=25)	0.90	0.924
AH (n=31)	0.53	1.009

Table 6.18: Mean comprehensibility gains for SA and AH groups

As illustrated in Figure 6.12, data showed that the SA group ($M_{Compreh\_gains}$=0.90) obtained greater gains than the AH group ($M_{Compreh\_gains}$ =0.53). In order to check whether this difference was statistically significant, an independent-samples t-test comparing $M_{Compreh\_gains}$ of the two groups (SA and AH) was conducted. This analysis revealed that the SA group did not obtain significantly greater gains than the AH group [$t(54)=1.432$, $p=.158$, $\eta^2=.03$]:

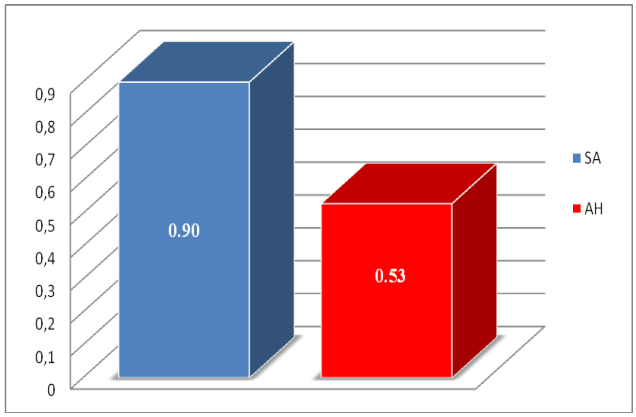


Figure 6.12: Mean comprehensibility gains of SA and AH groups

An additional chi-square analysis was conducted in order to assess whether there was an association between learning context and comprehensibility improvement. A new categorical variable was created, 'comprehensibility improvement', with two categories (yes/no). Learners were assigned to one of the two categories based on the presence of improvement in their comprehensibility score (FA gains>0) or lack thereof. Table 6.19 illustrates the percentage of learners in each learning

context (SA and AH) who improved their comprehensibility score between T1 and T2:

	Comprehensibility Improvement	
	no	**yes**
SA (n=25)	4 16.0%	21 84.0%
AH (n=31)	13 41.9%	18 58.1%

Table 6.19: Percentage of participants improving comprehensibility after SA or AH period

As shown in table 6.19, most of the SA participants improved their comprehensibility ratings (84% improvers versus 16% non-improvers), while the difference between the percentages of improvers and non-improvers in the case of the AH learners was smaller (58.1% versus 41.9%). The chi-square test statistic was significant [$\chi^2(1)=4.40$, $p=.036$], indicating that the proportion of SA learners who improved in comprehensibility (84%) was significantly larger than the proportion of AH learners (58.1%). Cramer's phi ($\varphi=-.280$, $p=.036$) suggested a low effect size. The odds ratio indicated that SA learners were 3.79 times more likely to improve their comprehensibility ratings than AH learners.

6.2.3 Effect of initial comprehensibility level

This sub-section explores the effect of initial comprehensibility level on comprehensibility development over time. We examined whether the participants' comprehensibility level at T1 conditioned comprehensibility development during the 3-month SA and AH periods. We addressed the following sub-question:

RQ2c: Is comprehensibility development different for participants with different initial comprehensibility level in each group (SA and AH)?

Correlations between comprehensibility scores at T1 and comprehensibility scores at T2 were conducted for each group. The results of this analysis are reported in Table 6.20:

	SA (n=25)	**AH** (n=31)
Pearson	.702	.610
Sig.	<.001	<.001

Table 6.20: Pearson correlations between comprehensibility level at Pre-test and comprehensibility level at Post-test for SA and AH groups

Pearson correlation coefficients indicated large significant positive correlations between comprehensibility scores at T1 and T2 in both groups, suggesting quite a strong relationship between the ratings obtained before and after the SA and AH periods: the higher the comprehensibility scores at pre-test, the higher the comprehensibility ratings at post-test in both contexts. Likewise, those participants with lower comprehensibility scores at T1 obtained lower comprehensibility scores at T2.

Additional correlations were conducted between comprehensibility scores at T1 and amount of gains in order to test whether there was a relationship between initial comprehensibility level and gains in comprehensibility. In other words, did participants with lower initial comprehensibility level obtain greater gains than those with higher initial comprehensibility level, or vice versa? Table 6.21 shows the results of this analysis:

	SA (n=25)	**AH** (n=31)
Pearson	-.739	-.499
Sig.	<.001	.004

Table 6.21: Pearson correlations between initial comprehensibility level and amount of gains for SA and AH groups

Pearson correlation coefficients yielded negative significant correlations between initial comprehensibility level and amount of gains in comprehensibility in both groups, indicating that the higher the initial comprehensibility level, the lower the gains; and the lower the initial comprehensibility level, the higher the gains. The relationship between the two variables was large in both groups, although it was larger in the case of the SA group.

In order to gain a better understanding of possible initial level effects on comprehensibility development and gains among participants, we examined the differences in comprehensibility development between participants with higher initial comprehensibility level and participants with lower initial comprehensibility level.

We operationalised initial comprehensibility level as a function of participants' comprehensibility scores at the beginning of the study (T1). We followed the same criterion for grouping participants that we had used in the FA analyses. We organised participants as follows in order to have two low initial comprehensibility groups (one for the SA group and the other one for the AH group), and two high initial comprehensibility level groups (one for the SA and the other for the AH group):

1. Low initial comprehensibility level group (Low-iC): participants obtaining a comprehensibility score <3.5 at Pre-test;

2. High initial comprehensibility level group (High-iC): participants obtaining comprehensibility score ≥3.5 at Pre-test.

Table 6.22 shows the participants' distribution following this grouping criterion:

	SA	**AH**
Low-iC group	n= 8	n = 17
High-iC group	n = 17	n = 14

Table 6.22: Distribution of SA and AH participants in Low and High initial comprehensibility level groups

The mean of comprehensibility ratings at T1 and T2 for each group are indicated in Table 6.23 below:

		Mean	**SD**
Low- iC-SA	T1	2.51	0.653
(n=8)	T2	4.23	0.461
High- iC-SA	T1	4.75	0.785
(n=17)	T2	5.27	0.829
Low- iC-AH	T1	2.44	0.507
(n=17)	T2	3.40	1.022
High- iC-AH	T1	4.45	0.711
(n=14)	T2	4.46	0.932

Table 6.23: Mean comprehensibility ratings for Low and High initial comprehensibility level groups (SA and AH) at Pre-test and Post-test

The normality of the distribution of scores at T1 for all sub-groups was assessed. The Shapiro-Wilk tests indicated that the comprehensibility scores for all initial comprehensibility level groups followed a normal distribution.

In the next sub-sections we present the results of the analyses examining the effect of initial comprehensibility level on comprehensibility development after the SA and AH periods. We compare the two low initial groups (and the two high initial level ones), in order to explore the impact of learning context on comprehensibility development in participants with different initial comprehensibility levels.

6.2.3.1 Initial comprehensibility level and SA outcomes

The first question that we examined was whether there were any differences in comprehensibility development between the two sub-groups in the SA group depending on their initial comprehensibility level. An independent-samples t-test comparing mean comprehensibility ratings at T1 between the Low-iC-SA and High-iC-SA groups was conducted in order to confirm that the two groups differed significantly at T1. Differences between these two groups were highly significant [$t(23)$=-7.019, p<.001, η^2= .68].

Subsequently, mean comprehensibility ratings for the two SA sub-groups at both Pre-test and Post-test were submitted to a mixed-design analysis of variance (ANOVA) with *group* (High-iC-SA and Low-iC-SA) as the between-subjects factor, and *time* (Pre-test and Post-test) as the within-subjects factor. Table 6.24 shows the results of the ANOVA evaluating changes in comprehensibility ratings over time:

Time	F(1)=50.37	p<.001	η2=.68
Group	F(1)=35.50	p<.001	η2=.60
Time x Group	F(1)=14.57	p=.001	η2=.38

Table 6.24: Results of mixed-design ANOVA with *group* (Low-iC-SA and High-iC-SA) and *time* (Pre-test and Post-test) as factors

Results from this analysis yielded a significant main effect of time [$F(1)$=50.73, p<.001, η^2=.68], group [$F(1)$=35.50, p<.001, η^2=.60], and *time* x *group* interaction [$F(1)$=14.57, p=.001, η^2=.38].

Additional paired-samples t-tests with testing time as the within-subjects factor showed that there was significant improvement between T1 and T2 for both the Low-iC-SA group [$t(7)$=--6.199, p<.001 , η^2=.84] and the High-iC-SA group [$t(16)$=-2.992, p=.009 , η^2=.35]. The eta-squared value indicated large effect size for both groups, although this was greater for the Low-iC-SA group. Figure 6.13 illustrates the mean comprehensibility ratings assigned to each sub-group at the two testing times:

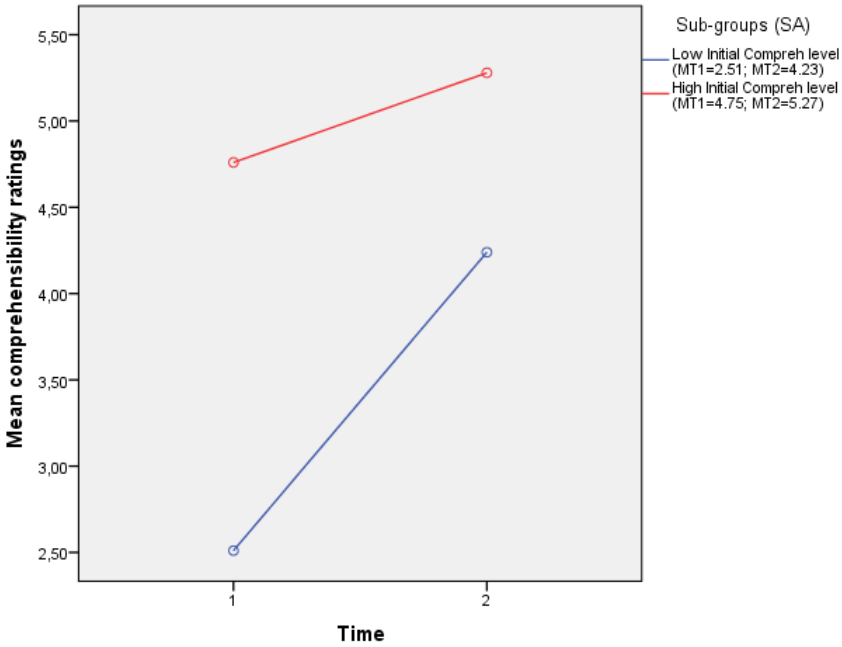


Figure 6.13: Mean comprehensibility ratings at T1 and T2
for High and Low initial comprehensibility level groups (SA)

As we have already reported at the beginning of this sub-section, results revealed a significant difference in comprehensibility ratings between the low and high level SA sub-groups at T1. An independent-samples t-test also showed a significant difference in comprehensibility ratings at T2 [$t(23)$=-3.291, p=.003, η^2=.32]. The eta squared indicated a large effect size, although this was smaller than the large effect size found at T1.

As indicated in the ANOVA (see Table 6.24 above), there was a *time* x *group* interaction, suggesting that the change in comprehensibility scores over time for the two sub-groups (Low-iC-SA and High-iC-SA) was not similar. In other words, time did not affect the two groups in a similar way. Figure 6.13 illustrates that the two groups did not experience a parallel increase in their comprehensibility ratings.

The analysis of comprehensibility gains obtained by each sub-group confirmed this. Table 6.25 below shows the mean of comprehensibility gains for each sub-group:

Comprehensibility gains		
Group	**Mean**	**SD**
Low-iC-SA (n=8)	1.72	0.78
High-iC-SA (n=17)	0.51	0.71

Table 6.25: Mean comprehensibility gains for Low-iC-SA and High-iC-SA groups

As illustrated in Figure 6.14 below, data showed that the Low-iC-SA group ($M_{Compreh\_gains}$ =1.72) obtained greater gains than the High-iC-SA group ($M_{Compreh\_gains}$=.51). An independent-samples t-test comparing $M_{Compreh\_gains}$ between the two groups revealed that the Low-iC-SA group obtained significantly greater gains than the High-iC-SA group [$t(23)=3.818$, $p=.001$, $\eta^2=.38$]:

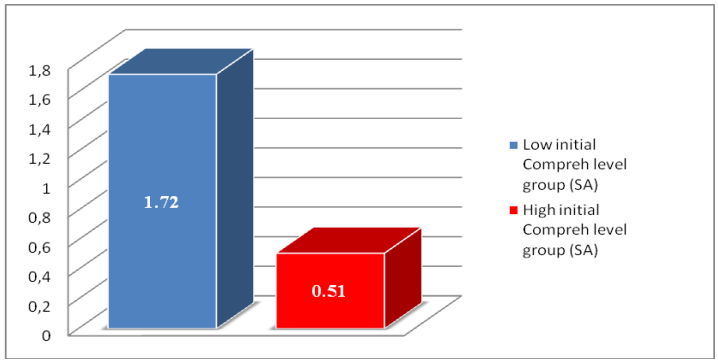


Figure 6.14: Mean comprehensibility gains for for Low-iC-SA and High-iC-SA

An additional Chi-square analysis was conducted in order to assess whether there was an association between initial level and comprehensibility improvement. Table 6.26 illustrates the percentage of learners in each group who improved their comprehensibility score between T1 and T2:

	Comprehensibility Improvement	
	no	**yes**
Low-iC-SA (n=8)	0 0%	8 100%
High-iC-SA (n=17)	4 23.5%	13 76.5%

Table 6.26: Comprehensibility improvement of Low and High groups (SA)

The proportion of High-iC-SA participants (76.5%) who improved in comprehensibility was not significantly larger than the proportion of Low-iC-SA participants (100%), as indicated by Fisher's exact test[25] (p=.269).

6.2.3.2 Initial comprehensibility level and AH outcomes

An independent-samples t-test confirmed that the comprehensibility ratings of the Low-iC-AH and High-iC-AH sub-groups differed significantly at T1[$t(29)=-9.15$, $p<.001$, $\eta^2=.74$].

Subsequently, mean comprehensibility ratings for each learner at both Pre-test and Post-test were submitted to a mixed-design analysis of variance (ANOVA) with *group* (High-iC-AH and Low-iC-AH) as the between-subjects factor, and *time* (Pre-test and Post-test) as the within-subjects factor. Table 6.27 shows the results of the ANOVA evaluating changes in comprehensibility ratings over time:

Time	F(1)=50.37	p<.001	η2=.687
Group	F(1)=38.83	p<.001	η2=.572
Time x Group	F(1)=14.57	p=.001	η2=.388

Table 6.27: Results of mixed-design ANOVA with *group* (Low-iC-AH and High-iC-AH) and *time* (Pre-test and Post-test) as factors

Results from this analysis yielded a significant main effect of time, group, and *time* x *group* interaction.

A paired-samples t-test showed that there was significant improvement in comprehensibility between T1 and T2 for the Low-iC-AH group [$t(16)=-4.569$, $p<.001$, $\eta^2=.56$]. The eta-squared value indicated a large effect size. In contrast, figure 6.15 illustrates that the mean comprehensibility ratings for the High-iC-AH sub-group remained unchanged between T1 and T2. A paired-samples t-test confirmed that there was no significant improvement between the T1 and T2 for the High-iC-AH group ($p=.963$):

[25] Fisher's Exact test was reported since 2 cells (50%) had expected frequency of less than 5. The minimum expected count was 1.28.

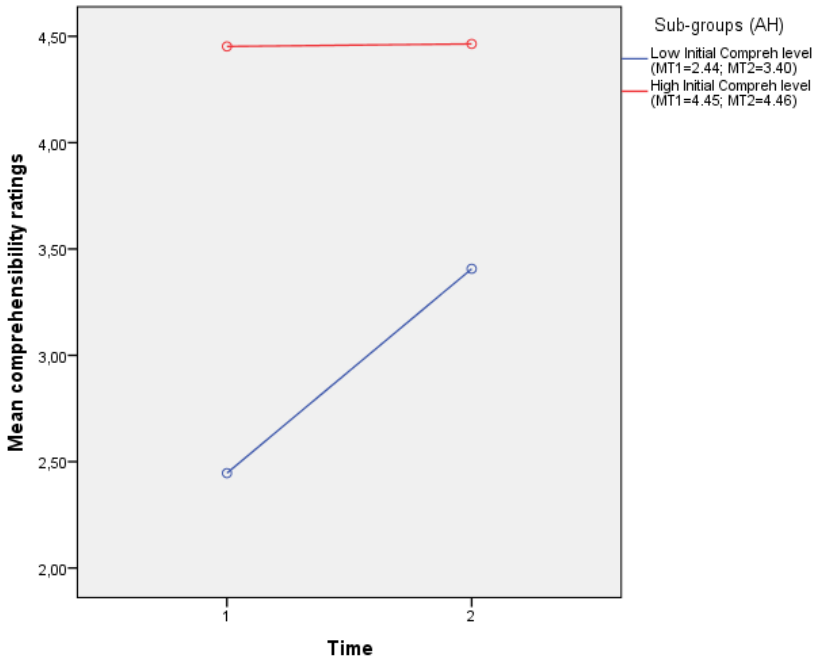


Figure 6.15: Mean comprehensibility ratings at T1 and T2 for High and Low initial FA level groups (AH)

As was the case at T1, an independent-samples t-test showed that the comprehensibility ratings of the Low-iC-AH and High-iC-AH sub-groups also differed significantly at T2 [$t(29)$=-2.98, p=.006, η^2=.23]. The eta squared revealed a large effect size, although this was smaller than the large effect size indicated at T1.

As indicated in the ANOVA (see Table 6.27 above) there was a *time* x *group* interaction, suggesting that the change in comprehensibility scores over time for the two sub-groups (Low-iC-AH and High-iC-AH) was not similar. We conducted an analysis of comprehensibility gains obtained by each group. Table 6.28 below shows the mean of comprehensibility gains for each group:

Comprehensibility gains		
Group	**Mean**	**SD**
Low-iC-AH (n=17)	0.96	0.86
High-iC-AH (n=14)	.01	0.94

Table 6.28: Mean comprehensibility gains for Low-iC-AH and High-iC-AH

As illustrated in Figure 6.16 below, the resulting data showed that the Low-iC-AH group ($M_{Compreh\_gains}$ =0.96) obtained greater gains than the High-iC-AH group ($M_{Compreh\_gains}$=.01). In order to check whether this

difference was statistically significant, an independent-samples t-test comparing $M_{Compreh\_gains}$ of the two groups was conducted. This analysis revealed that the Low-iC-AH group achieved significantly greater gains than the High-iC-AH group: $t(29)=2.91$, $p=.007$, $\eta^2=.22$:

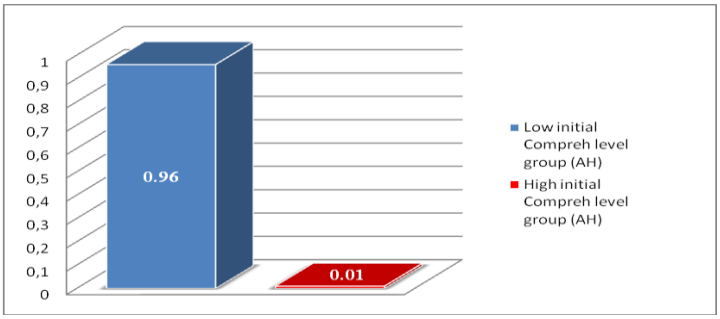


Figure 6.16: Mean comprehensibility gains for Low-iC-AH and High-iC-AH

An additional Chi-square analysis was conducted in order to assess whether there was an association between initial level and comprehensibility improvement. Table 6.29 illustrates the percentage of learners in each group who improved their comprehensibility score between T1 and T2:

	Comprehensibility Improvement	
	no	**yes**
Low-iC-AH (n=17)	4 23.5%	13 76.5%
High-iC-AH (n=14)	9 64.3%	5 35.7%

Table 6.29: Comprehensibility improvement of Low and High groups (AH)

The chi-square test statistic was significant [$\chi^2(1)=5.237$, $p=.022$], indicating that the proportion of Low-iC-AH participants who improved in comprehensibility was significantly larger than the proportion of High-iC-AH participants. Cramer's phi ($\varphi=-.411$, $p=.022$) suggested a moderate effect size. The odds ratio indicated that Low-iC-AH learners were 5.85 times more likely to improve their comprehensibility ratings than High-iC-AH learners.

We additionally considered whether the two sub-groups with the same initial comprehensibility level groups (i.e., low or high) had the same gains in the two different learning contexts considered (SA and AH). In other words, did SA participants with lower initial comprehensibility level have greater gains in comprehensibility than learners with lower initial comprehensibility level in the AH group? Likewise, did SA participants with higher initial comprehensibility level have greater gains in

comprehensibility than participants with higher initial comprehensibility level in the AH group? These questions are addressed in the next sub-sections.

6.2.3.3 Gains in comprehensibility: a comparison between Low iC-SA and Low-iC-AH outcomes

Before comparing the difference in amount of gains between the low initial comprehensibility groups, we ran an independent-samples t-test comparing the mean comprehensibility ratings of the Low-iC-SA and Low-iC-AH sub-groups at T1. In contrast with the results reported in section 6.2.1, results from this analysis showed that the two sub-groups did not differ significantly at T1 [$t(23)=.270$, $p=.790$, $\eta^2=.003$] [26].

Figure 6.17 illustrates the mean comprehensibility ratings assigned to each sub-group at the two testing times:

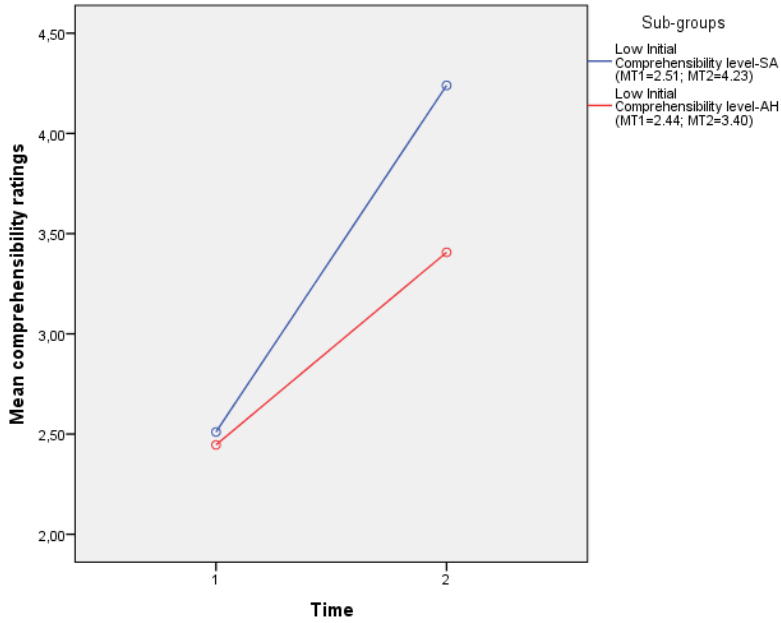


Figure 6.17: Mean comprehensibility ratings at T1 and T2 for Low initial comprehensibility level groups (SA and AH)

[26] An additional independent-samples t-test comparing the mean comprehensibility ratings of the Low-iC-SA and Low-iC-AH sub-groups at T2 indicated that the difference in the comprehensibility between these two groups was significant at T2 [$t(23)=2.804$, $p=.01$, $\eta^2=.25$].

Subsequently, we explored comprehensibility gains for the two Low initial comprehensibility level sub-groups. Table 6.30 shows the mean comprehensibility gains for each group:

Comprehensibility gains		
Group	**Mean**	**SD**
Low-iC-SA (n=8)	1.72	0.78
Low-iC-AH (n=17)	0.96	0.86

Table 6.30: Mean comprehensibility gains for Low initial comprehensibility level groups (SA and AH)

As illustrated in Figure 6.18, the Low-iC-SA group ($M_{Compreh\_gains}$=1.72) obtained greater gains than the Low-iC-AH group ($M_{Compreh\_gains}$=0.96):

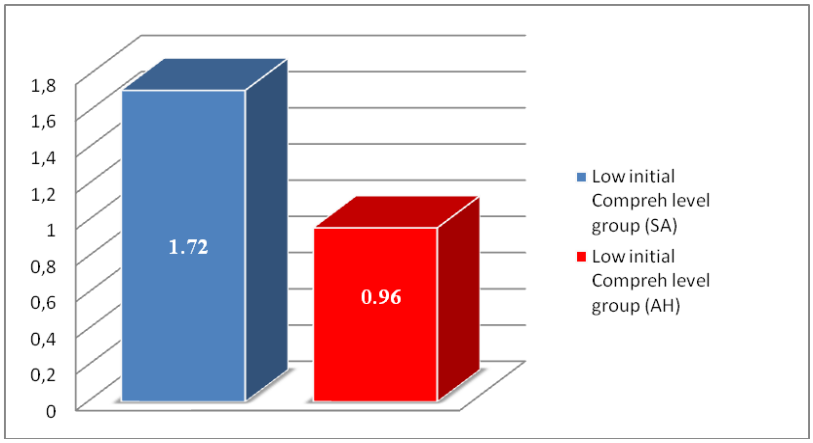


Figure 6.18: Mean comprehensibility gains for Low initial comprehensibility level groups (SA and AH)

An independent-samples t-test showed that there was a significant difference in comprehensibility gains between the two groups [$t(23)$= 2.123, p=.045 , η^2=.16]. The eta squared value reported a large effect size.

6.2.3.4 Gains in comprehensibility: a comparison between High iC-SA vs High iC-AH outcomes

Before comparing the difference in amount of gains between the high initial comprehensibility groups, we ran an independent-samples t-test comparing the mean comprehensibility ratings at T1 between the High-iC-SA and High-iC-AH group. Results from this analysis showed that the

differences in the $M_{Compreh}$ at T1 between these two sub-groups were not significant[27] [$t(29)=1.131$, $p=.267$, $\eta^2=.04$].

Figure 6.19 illustrates mean comprehensibility ratings assigned to each sub-group at the two testing times:

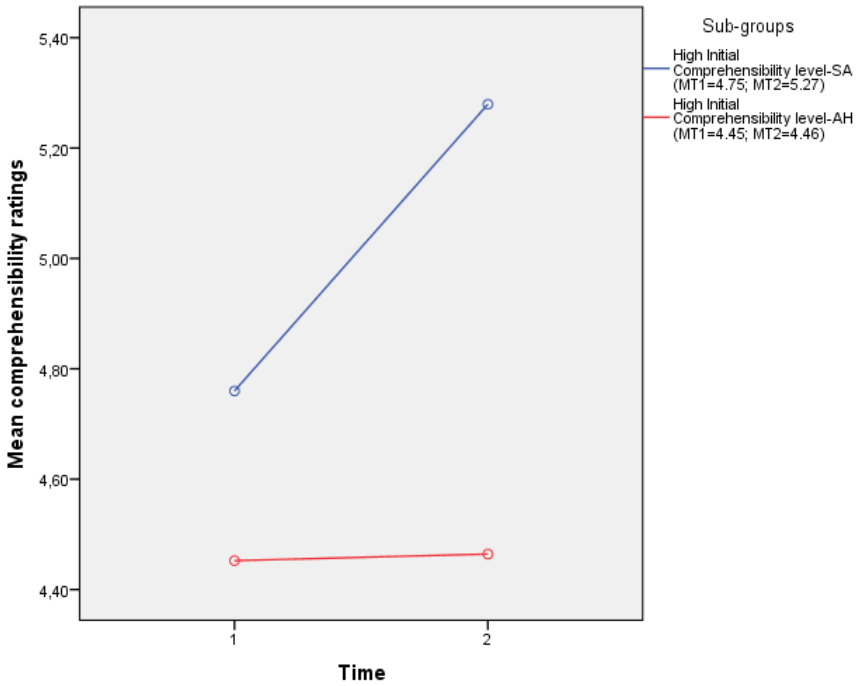


Figure 6.19: Mean comprehensibility ratings at T1 and T2
for High initial comprehensibility level groups (SA and AH)

Subsequently, we explored comprehensibility gains for the two High initial comprehensibility level sub-groups. Table 6.30 shows the mean comprehensibility gains for each group:

Comprehensibility gains

Group	Mean	SD
High-iC-SA (n=17)	0.51	0.71
High-iC-AH (n=14)	0.01	0.94

Table 6.31: Mean comprehensibility gains
for each high initial comprehensibility level group

[27] An additional independent-samples t-test comparing the mean FA ratings of the High-iC-SA and High-iC-AH sub-groups at T2 indicated that the difference in the comprehensibility between these two groups was significant at T2 [$t(29)=2.576$, $p=.015$, $\eta^2=.18$].

As illustrated in Figure 6.20, the High-iC-SA group ($M_{Compreh\_gains}$=0.51) obtained greater gains than the High-iC-AH group, for which scores actually remained unchanged between T1 and T2 ($M_{Compreh\_gains}$=0.01):

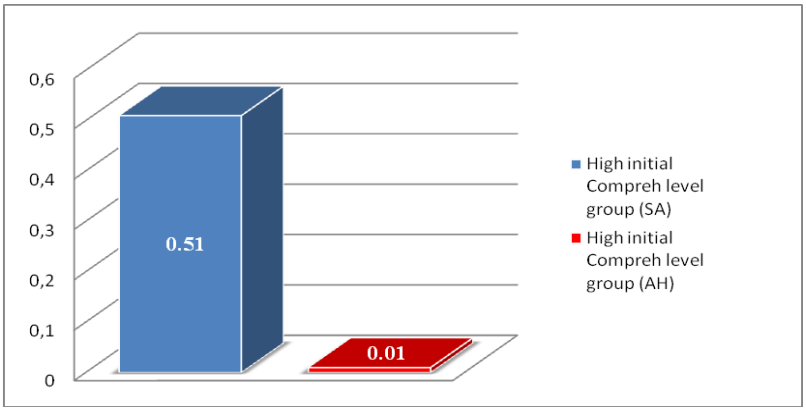


Figure 6.20: Mean comprehensibility gains for High initial comprehensibility level groups (SA and AH)

An independent-samples t-test was conducted in order to check whether there was a significant difference between the gains obtained by the High-iC-SA participants and the lack of gains for the High-iC-AH learners. The results showed a non-significant difference in comprehensibility gains between the two groups [$t(29)=1.701$, $p=.100$, $\eta^2=.09$].

6.3 Relationship between foreign accent and comprehensibility

In this section we will analyze the relationship between the participants' degree of FA and comprehensibility and thus address our final research question:

RQ3: To what extent are foreign accent and comprehensibility related speech dimensions?

We will present the results from correlations run between FA and comprehensibility scores at T1 and T2 for both the SA and AH participants. In addition, we will compare the aspects reported by the listeners as most influencing their FA and comprehensibility ratings in order to find out whether the same aspects affected their ratings of the two different speech dimensions. For this purpose, qualitative analyses of the data obtained from questionnaires and the rating task will be conducted.

6.3.1 Correlation between FA and comprehensibility

In order to explore the relationship between FA and comprehensibility, we first formulated research sub-question 3a below:

RQ 3a: To what extent do the degree of FA and comprehensibility correlate?

Correlations between FA scores and comprehensibility scores were performed for the two groups of participants (SA and AH) at the two testing times. The results of these analyses are reported in Table 6.32 below:

		SA (n=25)	**AH** (n=31)
T1	**Pearson**	.849	.789
	Sig.	<.001	<.001
T2	**Pearson**	.814	.741
	Sig.	<.001	<.001

Table 6.32: Pearson correlations between FA scores and comprehensibility scores at Pre-test and at Post-test for SA and AH groups

As can be seen, Pearson correlation coefficients yielded significant large positive correlations between FA and comprehensibility scores for the SA and the AH groups at the two testing times. This indicated that the higher the ratings in the FA scale, the higher the comprehensibility scores. In other words, the more native-like the accent, the greater the comprehensibility, whereas the more foreign-accented the learners' speech, the less comprehensible for our listeners.

We created a contingency table with initial FA level and initial comprehensibility level variables in order to explore the participants' distribution according to their FA and comprehensibility scores at Pre-test. Table 6.33 shows the distribution of participants as a function of their initial FA and comprehensibility level:

			Initial Compreh Level		
			Low	High	Total
Initial FA Level	Low	Raw number of participants	25	20	45
		% participants	44.6%	35.7%	80.3%
		% within InitialFALevel	55.6%	44.4%	100 %
		% within InitialComprehLevel	100%	64.5%	80.3%
	High	Raw number of participants	0	11	11
		% participants	0%	19.6%	19.6%
		% within InitialFALevel group	0%	100%	100%
		% within InitialComprehLevel	0%	35.5%	19.6%

Table 6.33: Participants' distribution according to their initial level of FA and comprehensibility (n=56)

Out of the 56 participants, 44.6% had a low initial FA level and a low initial comprehensibility level, whereas 35.7% showed a low initial FA level but a high comprehensibility level. A low initial FA score[28] was associated with low comprehensibility ratings in 55.6% of the cases, while 44.4% of the participants with low initial FA score obtained good results for comprehensibility. On the other hand, 100% of the participants with initial high comprehensibility scores also obtained high ratings in the FA scale. It is worth noting that none of the participants with high initial comprehensibility scores obtained low ratings for FA.

A similar contingency table was created displaying the distribution of participants as a function of their final FA level and final comprehensibility level (see Table 6.34 below):

			Final Compreh Level		
			Low	High	Total
Final FA Level	Low	Raw number of participants	10	24	34
		% participants	17.8%	42.8%	60.7%
		% within FinalFALevel	29.4%	70.6%	100 %
		% within FinalComprehLevel	100%	52.2%	60.7%
	High	Raw number of participants	0	22	22
		% participants	0%	39.2%	39.2%
		% within FinalFALevel	0%	100%	100%
		% within FinalComprehLevel	0%	47.8%	39.3%

Table 6.34: Participants' distribution according to their final level of FA and comprehensibility (n=56)

[28] As indicated in section 6.1.3 low initial FA scores referred to ratings <3.5 at the Pre-test (1 being = "heavy foreign accent").

As can be observed, out of the 56 participants in the study, 17.8% obtained low FA ratings and low comprehensibility ratings at Post-test, whereas 42.8% had low FA scores but were rated as highly comprehensible. Low scores in final FA level were associated with low comprehensibility scores in nearly 30% of the cases, while 70.6% of the participants with low final FA scores obtained good results for comprehensibility. Again, data confirmed that 100% of the participants with initial high comprehensibility level also obtained low ratings for FA and that none of the participants with low initial comprehensibility scores obtained a good result for FA.

We further tabulated the distribution of participants across the different FA and comprehensibility levels separately for the SA and AH groups so as to explore the relationship between participants' initial/final levels of FA and comprehensibility according to their SA/AH condition. Table 6.35 shows the InitialFALevel * InitialComprehLevel crosstabulations for the SA and AH participants:

			Initial comprehensibility level	
			Low	High
SA (n=25)	**Initial FA level**	Low	**8 (32%)**	**9 (36%)**
		High	**0 (0%)**	**8 (32%)**
AH (n=31)	**Initial FA level**	Low	**17 (54.8%)**	**11 (35.5%)**
		High	**0 (0%)**	**3 (9.7%)**

Table 6.35: SA and AH participants' distribution according to their initial level of FA and comprehensibility

As can be observed, the proportion of participants assigned to the low initial FA group and the low initial comprehensibility groups was higher in the case of the AH group (54.8%) than the SA group (32%). The proportion of participants receiving low FA scores in the rating experiment (i.e., those perceived to produce more foreign-accented speech), was 68% in the case of the SA group, and 90.3% for the AH. Moreover, the proportion of learners showing a high initial level in comprehensibility and a high initial FA level (more native-like accent) (high initial FA level group) was larger in the case of SA participants (32%) than in the case of AH learners (9.7%), indicating that there was a higher proportion of learners with better results in both dimensions at pre-test in the SA group.

Table 6.36 below illustrates the FinalFALevel*FinalComprehLevel crosstabulation for both groups of participants:

			Final comprehensibility level	
			Low	High
SA (n=25)	**Final FA level**	Low	**0 (0%)**	**8 (32%)**
		High	**0 (0%)**	**17 (68%)**
AH (n=31)	**Final FA level**	Low	**10 (32.25%)**	**16 (51.6%)**
		High	**0 (0%)**	**5 (16.1%)**

Table 6.36: SA and AH participants' distribution according to their final level of FA and comprehensibility

The proportion of AH participants in the low final FA and comprehensibility level groups decreased considerably from T1 (54.8%) to T2 (32.25%). In the case of SA participants, none of the learners were classified under this condition, since all the SA participants in the low final FA level group fell under the high final comprehensibility group (32%), a proportion that was very similar to the one at T1 (36%). On the other hand, the proportion of participants receiving low FA scores in the rating experiment (i.e "more foreign-accented speech") decreased from 68% at T1 to 32% in the case of SA group, but remained quite similar for AH participants, from 90.3% at T1 to 83.8% at T2.

It is worth remarking that a more native-like accent is always associated with good comprehensibility ratings. Results showed that none of the participants (either in the AH or the SA groups at pre-test or post-test) were assigned a high degree of FA and a low rating in comprehensibility. However, in order to confirm that a more native-like accent is linked to a positive effect on comprehensibility, we analysed teachers' reports on their FA and comprehensibility scores, which were collected throughout the rating experiment. The results from these qualitative analyses will be presented in the following sub-section.

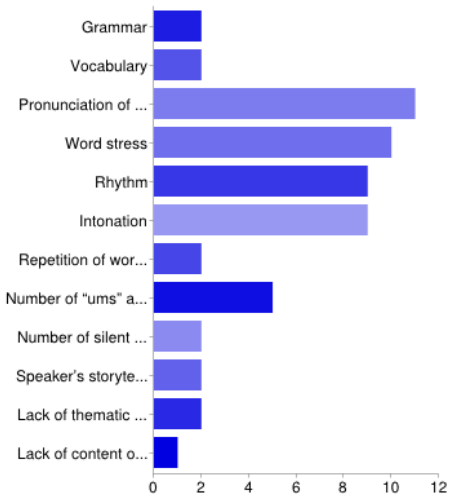
6.3.2 Aspects influencing FA and comprehensibility ratings

In order to gain a deeper insight into the relation between FA and comprehensibility, we explored the linguistic aspects which had influenced listeners' FA and comprehensibility ratings. Thus, we addressed the following research sub-question:

RQ3b: To what extent do non-native listeners report the same linguistic aspects as having influenced their ratings when evaluating FA and comprehensibility?

6.3.2.1 Aspects influencing FA ratings

In the online questionnaire that the listeners had to complete after the rating experiment (see Appendix 8), listeners were shown a list of twelve factors and were asked to select those that had most influenced their FA ratings. Table 6.37 shows the list with the twelve aspects and the raw number of listeners (together with the percentage) who selected each aspect:



Grammar	2	17%
Vocabulary	2	17%
Pronunc of vowels/consonants	**11**	**92%**
Word stress	**10**	**83%**
Rhythm	**9**	**75%**
Intonation	**9**	**75%**
Repetition of words or syllables	2	17%
Number of "ums" and "uhs"	5	42%
Number of silent pauses	2	17%
Speaker's storytelling ability	2	17%
Lack of thematic content	2	17%
Lack of content organization	1	8%

Table 6.37: Aspects affecting FA ratings
(summary of listeners' answers to question 1 in the online questionnaire)

Factors from the domain of phonology seemed to contribute the most to listeners' perception of foreign accent. Both segmental and suprasegmental aspects of speech were selected by most listeners. Pronunciation of individual sounds was selected by 92% of the listeners, followed by word stress (reported by 83% of the raters), and rhythm and intonation (75% each). The next aspect selected by most teachers was "the number of 'ums' and 'uhs'" (42%), but with a considerable lower percentage, however.

In addition, listeners were asked to indicate any other aspects which had affected their FA ratings. Table 6.38 reports all listeners' comments with respect to this question:

Listener	Comments
COGA	"Trying to imitate a native speaker".
ELME	"pronunciation of -ed for regular past verbs. -the "e" sound before the "s", as in "scared". -the "v" sound"
ESES	"It isn't an additional aspect but a remark. I paid attention to the correct pronunciation of initial and final consonants, which are so difficult for some students with Catalan and Spanish as their mother tongues".
MOLO	"In some cases the use of L1 words, if they interfere the pronunciation or rhythm".

Table 6.38: Aspects affecting FA ratings
(summary of listeners' answers to question 2 in the online questionnaire)

As can be seen, two listeners paid attention to aspects which are characteristic of Spanish learners' pronunciation in English (e.g., pronunciation of initial 's' in words such as 'scared'), and would be included under the 'pronunciation of vowels/consonants' category. Moreover, one listener referred to the effort of imitating native-like accent as a factor involved in her FA ratings, and another listener noted that the use of L1 lexical items had affected learners' pronunciation and, therefore, her ratings.

A final analysis was carried out using listeners' answers to question 5 of the questionnaire, which read as follows:

> *Of the aspects you indicated, please rank the top 3 that you felt most influenced your ACCENTEDNESS ratings. Write 1 (most important), 2 (second most important) and 3 (third most important) and the corresponding aspect next to each number.*

As illustrated in Figure 6.21 below, the three most selected aspects influencing listeners' FA ratings were pronunciation of individual sounds, intonation, and word stress and rhythm. Other aspects which were reported by the listeners' are shown in this figure (e.g. grammar, vocabulary, number of pauses and number of 'uhms' and 'uhs'):

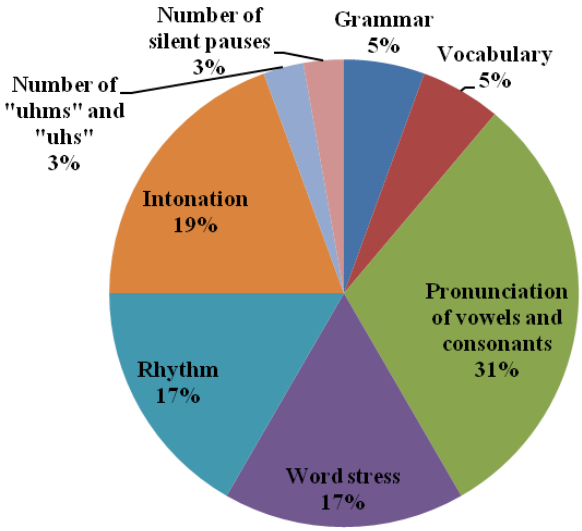
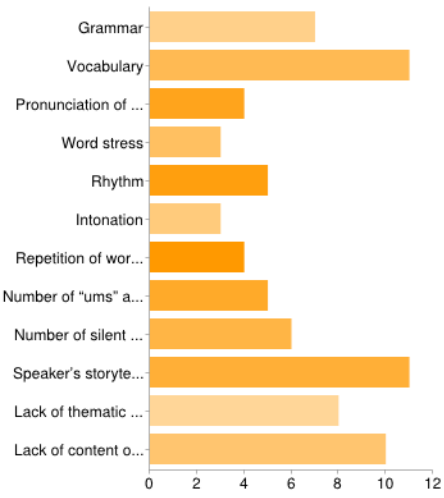


Figure 6.21Aspects affecting listeners' FA ratings most and their percentage

6.3.2.2 Aspects influencing comprehensibility ratings

As for the aspects influencing their comprehensibility scores, listeners were given the list of aspects in question 5 and asked to choose those which had most affected their comprehensibility ratings. A summary of the results is presented in Table 6.39 below:



Grammar	7	58%
Vocabulary	**11**	**92%**
Pronunciation of vowels/consonants	4	33%
Word stress	3	25%
Rhythm	5	42%
Intonation	3	25%
Repetition of words or syllables	4	33%
Number of "ums" and "uhs"	5	42%
Number of silent pauses	6	50%
Speaker's storytelling ability	**11**	**92%**
Lack of thematic content	8	67%
Lack of content organization	**10**	**83%**

Table 6.39: Aspects affecting comprehensibility ratings
(summary of listeners' answers to question 3 in the online questionnaire)

Unlike accent, comprehensibility was mostly associated with vocabulary and discourse (storytelling and content organization). More than 90% of the listeners selected 'vocabulary' and 'speaker's storytelling ability', and 83% bore in mind 'content organization' when assigning comprehensibility scores. Lack of thematic content was important for 67% of the listeners, and grammar influenced the ratings of 60% of the listeners.

In question 4 of the questionnaire, listeners were asked to type in any other aspects influencing their comprehensibility ratings. Table 6.40 below reports all listeners' comments:

Listener	Comments
ANSE	"The attitude of the speaker to make himself/herself understood. Some students tried really hard to communicate, whereas most of those who seemed to me native English speakers spoke fast and in a low voice".
ESES	"lack of connectors and correct use of reference (pronouns...), which is related to lack of content organization".
INCA	"Use of Spanish words maybe".
MOLO	"The use of L1 words in some cases which shows the lack of ability of the student to make the message be understood. Variety of specific vocabulary related to the topic provided".

Table 6.40: Aspects affecting comprehensibility ratings
(summary of listeners' answers to question 4 in the online questionnaire)

Seventy-five percent of these comments referred to vocabulary. Two listeners pointed out the use of L1 vocabulary as interfering in comprehensibility. The lack of vocabulary was stressed by one of the raters especially (MOLO). Interestingly, one of the listeners highlighted that speaker's attitude had also affected her comprehensibility ratings (ANSE).

Finally, we analysed listeners' answers to question 6 of the questionnaire, which read as follows:

> *Of the aspects you indicated, please rank the top 3 that you felt most influenced your COMPREHENSIBILITY ratings. Write 1 (most important), 2 (second most important) and 3 (third most important) and the corresponding aspect next to each number.*

The three most cited aspects were vocabulary, lack of content organization and speakers' storytelling ability (followed by grammar and pronunciation of individual sounds). Figure 6.22 below illustrates the results regarding question 6:

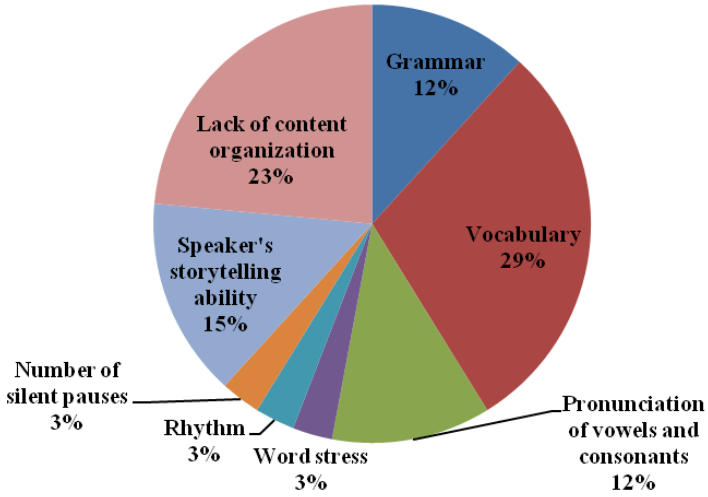


Figure 6.22: Aspects affecting listeners' comprehensibility ratings most and their percentage

As detailed in section 5.3.2, in order to delve into the comprehensibility construct, we asked listeners to type in their comments immediately after rating the comprehensibility of 20 speech samples scattered throughout the rating experiment. We obtained 688 comments about comprehensibility from the text entry boxes which were filled in by the listeners during the rating experiment. The listeners' descriptive comments were first classified as affecting comprehensibility positively or negatively. There were 426 negative comments and 262 positive comments. Then the comments were thematically coded, and re-coded to eliminate overlapping ones (e.g. 'L1 word', 'invented word' and 'wrong word' were combined under a 'vocabulary' category). Examples of each category are provided in Appendix 9.

We found that some listeners were more specific than others regarding their comments. Whereas some listeners made general comments about comprehensibility such as "grammar errors", other listeners specified in their reports the type of grammar errors they found in the participants' speech (e.g., "no subject", "wrong verb tense", etc.). Therefore, we recalculated the total amount of comments about comprehensibility assigning, for instance, the 'grammar' category in one listener report, independently of the number of comments made on this category.

After revising all listeners' comments, the total number of comments obtained was 596. Table 6.41 below shows all the categories obtained from the listeners' comments indicating whether they were considered in negative (N), positive (P) evaluations, or both (B). The initial number of

comments is provided together with the final number of comments obtained once double references to the same category made by the same listener were identified (number of comments deleted are indicated in brackets). The percentage of each category over the total number of final comments is indicated in the % column:

Category	Referred to in…	Initial number of comments	Final number of comments	%
Ambiguous[29]	B	10	10	1.67
Attitude	B	19 (-3)	16	2.68
Being a teacher	P	2	2	0.33
Communicative strategies	B	10	10	1.67
Content	B	37 (-1)	36	6.04
Discourse	B	61 (-12)	49	8.22
English Proficiency	N	1	1	0.16
Familiarity with the story	P	4	4	0.67
Fluency	**B**	97 (-11)	86	**14.42**
Grammar	**B**	99 (-12)	87	**14.59**
L1 familiarity	B	16 (-2)	14	2.34
L1 influence (general comment)	N	3	3	0.50
Listener's attitude	P	4	4	0.67
Low voice	N	3	3	0.50
Pronunciation	**B**	194 (-39)	155	**26**
Self-correction	P	4	4	0.67
Style	B	3 (-1)	2	0.33
Vocabulary	**B**	121 (-11)	110	**18.45**

Table 6.41: Frequency of coded categories for comprehensibility from teacher reports (initial raw number, final raw number and percentage of comments)

As can be observed, 26% of the comments referred to pronunciation (including segmental and supra-segmental aspects). Vocabulary was the second most frequent aspect considered by listeners in their comments (18.45%), followed by grammar (14.59%) and fluency (14.42%).

Further analyses explored whether the above-mentioned aspects were also taken into account in negative and positive comprehensibility ratings.

[29] There were a number of comments which were categorized as 'Ambiguous'. They were included in this category when it was not clear what the listeners were considering. For instance, for comments such as "I can't understand some words", it was not clear whether there was a pronunciation problem on the part of the speaker or if the speaker had invented a word which the listener could not understand (vocabulary). Given that we were not sure whether this was a comment referred to pronunciation or vocabulary, we assigned it to the 'Ambiguous' category.

Therefore, we examined the 426 negative comments and the 262 positive ones separately.

As for the comments describing aspects negatively influencing comprehensibility ratings, pronunciation, vocabulary and grammar were reported as the categories that most frequently affected listeners'scoring decisions. Twenty-six percent of the comments referred to segmental and supra-segmental aspects of participants' speech, 22.71% dealt with vocabulary items, and 19.11% with grammar. Fluency was mentioned in almost 15% of the comments. Table 6.42 below shows the results of this analysis:

Category	Initial number of comments	Final number of comments	%
Ambiguous	7	7	1.93
Attitude	10	8	2.21
Communicative strategies	1	1	0.27
Content	19	19	5.26
Discourse	20	19	5.26
English Proficiency	1	1	0.27
Fluency	62	53	**14.68**
Grammar	80	69	**19.11**
(No) L1 familiarity	3	1	0.27
L1 influence (general comment)	3	3	0.83
Low voice	1	3	0.83
Pronunciation	126	94	**26.03**
Style	2	1	0.27
Vocabulary	91	82	**22.71**
	426	361	

Table 6.42: Frequency of coded categories for negative comments on comprehensibility from teacher reports
(initial raw number, final raw number and percentage)

To have a better idea of the pronunciation features, we classified the comments according to the aspect of speech they were more specifically referring to. Table 6.43 shows this classification and the percentage of comments assigned to each pronunciation category:

Pronunciation aspect	%
Pronunciation of individual sounds and words	55.55
Foreign Accent	17.46
Intonation	11.11
Rhythm	3.96
Stress	2.3
Native-like pronunciation	4
Other[30]	5.55

Table 6.43: Pronunciation aspects reported by listeners as negatively influencing their comprehensibility ratings of participants' speech

Half of the comments regarding pronunciation problems referred to the pronunciation of individual sounds or words. It is worth remarking that specific reference was made to foreign-accented speech as an aspect affecting comprehensibility (17% of the comments referred to foreign accent). However, 'L1 interference' was mentioned when considering other pronunciation factors such as pronunciation of individual sounds or words, and intonation. Comments such as "Spanish intonation", "L1 influence on pronunciation" and "typical pronunciation mistake (from their L1)" were collected.

Having a native-like pronunciation was regarded as hindering comprehensibility to some extent by some of the listeners when rating native participants' speech. Comments such as those that follow were collected from listeners' evaluations:

> ELME: "after listening to so many recordings with the same type of syllable-timed speech, it was hard to readjust my ear to connected speech"
> INCA: "strong English native-like accent"

AS regards the comments referring to aspects positively affecting comprehensibility ratings, pronunciation was also considered the most influential aspect. As shown in Table 6.44 below, fluency, discourse and vocabulary were aspects reported in more than 10% of the comments:

Category	Initial number of comments	Final number of comments	%
Ambiguous	3	3	1.23
Attitude	9	8	3.29
Being a teacher	2	2	0.82

[30] Comments regarding pronunciation in general and speech clarity were categorized under the 'Other' category.

Communicative strategies	9	9	3.70
Content	18	17	7
Discourse	41	31	**12.75**
Familiarity with the story	4	4	1.64
Fluency	35	34	**14**
Grammar	19	19	7.81
L1 familiarity	15	13	5.34
Listener's attitude	4	4	1.64
Pronunciation	68	63	**26.33**
Self-correction	4	4	1.64
Style	1	1	0.41
Vocabulary	30	30	**12.34**
	262	242	

Table 6.44: Frequency of coded categories for positive comments on comprehensibility from teacher reports
(initial raw number, final raw number and percentage)

As done with the pronunciation comments negatively affecting participants' speech comprehensibility, we analysed listeners' reports on positive evaluations in further depth and found that listeners did not identify any particular aspects of pronunciation as positively affecting their comprehensibility ratings, but rather they referred to pronunciation in general. About 40% of the comments were similar to the following ones: "quite good pronunciation that facilitates comprehensibility", "pronunciation is OK", and "pronunciation is not that bad". Having a native-like pronunciation or imitating native-like pronunciation was the second most frequently cited aspect (22% of the comments). Moreover, general comments on accent were reported in 15% of the listeners' comments (e.g. "good accent"). Table 6.45 below shows the percentage of comments assigned to each pronunciation aspect reported by the listeners as positively affecting their comprehensibility ratings:

Pronunciation aspect	**%**
Pronunciation of individual sounds and words	2.94
Accent (general comment)	14.7
Intonation	10.29
Rhythm	7.35
Pronunciation (general positive comment)	41.17
Native or imitating native-like pronunciation	22.05
Being familiar with L1 accent	1.47

Table 6.45: Pronunciation aspects reported by listeners as positively influencing their comprehensibility ratings of participants' speech

While language aspects such as pronunciation, vocabulary, grammar or fluency were most frequently reported by listeners as affecting

comprehensibility, other aspects were mentioned which will be discussed in more detail in the following chapter. Reference to a native speaker model or the importance of native-like speech, L1 familiarity and the speaker's and listener's attitude were points made by the listeners which will also receive special attention in the next pages so as to provide further answers and comments in the context of English pronunciation teaching today.

Chapter 7

Discussion

In the previous chapter the results obtained in relation to our three research questions were presented. We analysed FA and comprehensibility in the oral productions of 25 Spanish learners of English before and after a 3-month SA programme, and 31 EFL Spanish learners receiving formal instruction in the AH learning context during the same period of time. The effect of time, learning context and initial FA and comprehensibility level were examined, as well as the aspects affecting FA and comprehensibility ratings assigned by a group of NNS listeners. The present chapter builds on the preceding one and considers the results gathered in relation to previous research reviewed in Part I of this dissertation. This chapter is organized around the three main research questions outlined in Chapter 4 and discusses the results obtained in the order in which they were presented in Chapter 6.

7.1 Changes in FA after SA and AH periods

Our first research question asked whether SA and AH participants improved in terms of FA over time, and compared FA development between the two groups of participants to determine whether one learning context was more beneficial than the other. The influence of initial FA level was also considered. We analysed the patterns of longitudinal improvement observed at the group level for participants with higher and

lower initial FA levels in both contexts (SA and AH). The first research question was formulated as follows:

> RQ1: Do participants improve in terms of foreign accent after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is foreign accent development different for participants with different initial foreign accent level?

The results presented in section 6.1 showed that SA participants improved significantly in terms of FA over the course a SA programme lasting three months. These results are in line with findings from some previous studies such as Højen (2003), who reported significant gains in perceived FA ratings after a SA period, but contrast with Avello et al. (2012), who found a slight but non-significant improvement in perceived FA after a 3-month period abroad.

Differences in initial proficiency level and previous L2 exposure can partly account for the difference between the results in Højen (2003), and our research, and those in the study by Avello et al. (2012). Participants in Højen's research had reported no (or limited) amount of exposure to conversational English before their SA experience, and participants in our study were adolescents in secondary education. Both groups of participants had a lower English language competence than the participants in Avello et al. (2012), who had to certify an upper-intermediate level of proficiency in English, equivalent to a B2 in the CEFR, in order to take part in the SA programme. Therefore, it may well have been that the lower initial proficiency in English of our adolescent learners favoured the significant improvement observed during their 3-month period abroad, something which is in accordance with results reported in other studies that have found greater gains in L2 oral production aspects for learners with lower initial proficiency levels than for rather advanced learners (Freed, 1995c; Juan-Garau, forthcoming; Lapkin et al., 1995; Llanes & Muñoz, 2009; Marriott, 1995; Regan, 1995; Valls-Ferrer, 2011).

DeKeyser (forthcoming) accounted for this general pattern of greater gains for lower level learners within a skill acquisition framework, according to which "all practice reaches a point of diminishing returns […] after which progress is still made for a very long time, but at a much slower pace than initially […]." He concluded therefore that the most advanced learners might need a stay abroad longer than just a few months in order to obtain measurable gains. In keeping with this view, DeKeyser (2007b), further noted that the more advanced learners may actually be

"the ones that are learning more in the long run", whereas the weaker learners would "make the quickest progress at the beginning" (pp. 211-212).

In line with the discussion in Llanes and Muñoz (2013), another factor that could explain the fact that the SA participants in our research improved their FA ratings could be the role of practice in L2 learning, as claimed by DeKeyser (2007b). Participants in our study may have enrolled in the SA programme with some previous declarative knowledge acquired in the course of the previous FI setting. Once abroad, after numerous hours of practice, they may have proceduralized their declarative knowledge and even automatized certain aspects of the L2. As Llanes and Muñoz (2013:17) suggested, "automatization could explain why on the post-test SA participants took less time to carry out the oral narrative task", something which could also be applied to the SA participants in our study (see Table 5.3).

As regards the AH participants in our study, significant improvement between T1 and T2 was also found for this group. Results indicated large time effects for both the SA and AH groups, although they were greater in the case of the SA participants. These results showing FA development in our AH participants between T1 and T2 are somehow optimistic, since previous studies involving Spanish adolescent learners of English in FI settings had reported no FA development, suggesting that amount of FI by itself did not decrease FA in learners' oral productions (García Lecumberri & Gallardo del Puerto, 2003; Fullana, 2006; Rallo & Juan-Garau, 2011). As described in Chapter 5, the AH learners in the present study confirmed that their EFL lessons responded to the common description of EFL lessons in Spain, where speaking and pronunciation are given less attention than aspects of grammar and vocabulary. However, AH participants in our study improved significantly in FA ratings over the FI period.

Despite the fact that the improvement in FA ratings for the AH group reached significance, in terms of amount of gains, the SA group obtained significantly greater gains in FA development than the AH group. This means that those students going abroad improved significantly more their FA ratings than those remaining AH. Our results confirmed findings in previous studies comparing SA and AH groups, which indicated a superiority of SA over AH settings regarding learners' L2 development in other oral dimensions, such as overall oral proficiency, accuracy, or fluency (Juan-Garau & Pérez-Vidal, 2007; Llanes, 2012; Llanes & Muñoz, 2013; Mora & Valls-Ferrer, 2012; Serrano, Llanes & Tragant, 2011; Trenchs-Parera, 2009). Unfortunately, no comparison can be made

with other studies examining FA in the oral speech produced by adolescent SA and AH learners, since to our knowledge ours is the first study comparing the effect of SA and AH on FA development in an adolescent population. Further studies are thus needed assessing FA development in similar pools of adolescent participants in order to confirm our results.

A more individual look at the data showed that up to 84% of the SA learners improved their FA ratings after their period abroad, whereas in the case of the AH learners the percentage of learners improving at post-test was lower (61.3%). This means that learners were 3.31 times more likely to obtain FA gains after a 3-month period in a SA setting than in a FI setting, and is consistent with the results for amount of FA gains. That is, not only did SA learners outperform AH learners in terms of mean FA gains, but the percentage of learners with improved FA ratings at post-test was also larger for the SA group than for the AH group. A Chi-square test revealed that the proportion of participants improving their accent after the SA period did not differ statistically from the proportion of AH participants within a 95% confidence interval (p=.061), although it was not far from reaching significance. It was significant within a 93.9% confidence interval. Actually, previous studies have referred to this p value within .06 and .10 as 'marginally significant' (Højen, 2003, Fullana, 2005; Valls-Ferrer, 2011). It is likely therefore that with an increased sample size in the present study the Chi-square test should have yielded a significant result.

Analyses comparing the two groups of participants showed a significant difference in FA ratings at Post-test, in the sense that FA ratings for the SA group were better than for the AH group. However, we cannot conclude that this difference is due to learning context effects, since the two groups of participants already differed significantly in FA ratings at Pre-test. As suggested in research analysing CLIL learning contexts (Alonso et al., 2008; Lagabaster and Sierra, 2009; Ruiz de Zarobe and Lasagabaster, 2010; San Isidro, 2010; Bruton, 2011), it seems that also in the case of SA participants there may be a kind of 'natural selection'. SA participants in our study were learners with higher L2 proficiency than the AH participants, at least in terms of FA, given the better mean FA ratings reported for this group at the two collection times. Further analyses comparing the performance of both groups of participants in the other language skills tested could help us to confirm this hypothesis. As illustrated in section 5.2.1, where the two groups are described, the learners who chose to go abroad were different from the learners who chose to stay at their home institutions. For instance, about half of the AH participants (54.8%) had not studied English abroad before, whereas only

one participant of the SA group (4%) had not. In addition, while most SA participants (80%) had already spent short periods of time in an English speaking country (half of them taking part in a two or three week summer course abroad), only 29% of the AH participants had been to an English speaking country for short periods of time and only 2 of them reported it had been for learning purposes. Thus, it seems that participants taking part in SA programmes may have a specific language learner profile different from those participants remaining at home (DeKeyser, forthcoming; Sanz, forthcoming).

A series of t-tests comparing FA ratings between native and non-native participants in our study showed that non-native participants (SA and AH) and native participants differed significantly in terms of FA both at Pre-test and at Post-test. This was something expected given the relatively low initial FA level of the L2 learners. As illustrated in Table 6.1, both groups of non-native speakers were far from native-like standards (NS M=6.71), mean FA ratings being farther from the native range for the AH group (M T1=2.44, M T2=2.68) than for the SA group (M T1=3.17, M T2=4.01). Effect sizes for the t-tests comparing NS and AH learners (η^2=6.08 for pre-test and η^2=5.36 for post-test) were likewise larger than those for the t-tests comparing NS and SA learners (η^2=3.43 for pre-test and η^2=2.44 for post-test). These results provided further evidence of the fact that SA group was better in terms of degree of accentedness.

As for the effect of initial FA level on FA development, participants in the two learner groups receiving better FA ratings at T1 also received better FA ratings at T2 and vice-versa. However, negative correlations between FA ratings at T1 and FA gains suggested a tendency according to which those participants with worse FA scores at T1 in both the SA and AH groups would be the ones obtaining higher gains, whereas the better the initial level regarding FA, the lower the gains, although these correlations did not reach significance in any of the groups. Further analyses were thus conducted to explore the role of initial FA level on gains in FA by grouping participants in the SA and AH groups as a function of their initial FA level (Low-iFA and High-iFA).

These group analyses, conducted separately for the SA and AH groups, revealed significant improvement between T1 and T2 for the SA participants with higher degree of FA at T1 (Low-iFA-SA group), but not for those SA participants who had been perceived as more native-like at T1 (High-iFA-SA group). The analysis comparing FA gains between the Low-iFA-SA and the High-iFA-SA sub-groups indicated that the Low-iFA-SA students obtained greater gains during SA than the High-iFA-SA students, although this difference in amount of gains between the two SA

groups was not significant. The same pattern of results was observed for the AH group, as participants in the Low-iFA-AH group (that is, those having a stronger FA at T1) improved significantly over the AH period, whereas results showed no significant improvement in the case of the High-iFA-AH participants (those with better FA ratings at T1), for whom FA scores remained unchanged between T1 and T2.[31] As was the case with the SA group, participants in the Low-iFA-AH sub-group obtained greater gains in the AH context than participants in the High-iFA-AH sub-group, although this difference was not significant. Thus, it seems that both SA and AH participants with worse FA at pre-test significantly improved their FA over time, whereas no significant improvement was found in the case of SA and AH learners with a more native-like accent at pre-test. This suggests some effect of initial FA level on FA development, in line with the above-mentioned trend of greater gains for lower level learners over more proficient learners (Juan-Garau & Pérez-Vidal, 2007; Llanes, 2012; Llanes & Muñoz, 2013; Mora & Valls-Ferrer, 2012; Serrano, Llanes & Tragant, 2011; Trenchs-Parera, 2009). This pattern of improvement was found both for the SA and AH contexts.

When comparing the effect of learning context on participants with high and low initial FA level, participants in the Low-iFA-SA group showed significantly greater gains than their peers in the Low-iFA-AH group. However, as it happened when comparing SA and AH groups as a whole, participants in both groups differed significantly in FA scores at T1. Thus, we cannot conclude that the significant difference in FA gains reported for the SA learners over the AH learners was only due to learning context effects, given that we were not comparing participants with the same pre-departure level. Therefore, other variables, such as the intrinsic differences in L2 proficiency between SA and AH learner groups commented above may have been at play.

Only when comparing participants in the High-iFA-SA and High-iFA-AH groups did we find that these groups did not differ significantly at T1. However, although the High-iFA-SA group obtained greater gains than the High-iFA-AH group, this difference was not significant. All in all, the reduced number of participants in the two high-level groups, particularly in the High-iFA-AH group, does not allow for a clear interpretation of these findings.

[31] Let us remember that conclusions involving analyses with the AH participants in the high initial FA sub-group should be taken cautiously given the low number of participants included in this group (n=3).

7.2 Changes in comprehensibility after SA and AH periods

Our second research question examined improvement in SA and AH participants' oral production in terms of comprehensibility, and compared comprehensibility development between the two groups of participants to determine whether one learning context was more beneficial than the other. The influence of initial comprehensibility level was also considered and longitudinal improvement for participants with higher and lower initial comprehensibility levels in both contexts (SA and AH) was analysed. Thus, we aimed at responding our second research question:

RQ2: Do participants improve in terms of comprehensibility after a 3-month SA period, or a 3-month period receiving formal instruction in an AH context? Is one learning context (SA vs. AH) more beneficial than the other? Is comprehensibility development different for participants with different initial comprehensibility level?

Results reported in section 6.2 indicated that the SA group improved significantly in comprehensibility between pre-test and post-test. These results confirmed previous findings showing the positive impact of SA periods in L2 learners' perceived oral comprehensibility (Allen & Herron, 2003; Martinsen, 2010). AH participants also showed significant improvement over the FI period. Large time effects were reported for both groups, although these were greater for the SA participants. Previous studies examining comprehensibility development in the oral productions of Spanish adolescent learners of English in a FI setting compared the effect of this learning context with that of a CLIL context. This was the case, for example, in Gallardo del Puerto et al. (2009), or Rallo and Juan-Garau (2011). In both studies the speech samples produced by SA participants were rated to be more comprehensible than the samples produced by their FI peers.

As regards comprehensibility gains, results showed that SA participants obtained greater gains than the AH group, although this difference did not reach significance. This contrasts with the results comparing FA gains, where SA participants were found to obtain significantly larger gains than AH participants. However, a more individual analysis of the data by means of a chi-square revealed that the proportion of SA learners who improved in comprehensibility (84%) was significantly larger than the proportion of AH learners (58.1%), as SA learners were 3.79 times more likely to improve their comprehensibility scores than AH learners. Taken together, these analyses suggest that the SA context was more beneficial

than the AH context in terms of comprehensibility development, since the percentage of learners improving their comprehensibility scores during SA was significantly larger than the percentage of learners improving their scores in the AH context, and SA learners were found to obtain larger mean comprehensibility gains than AH learners, although such improvement was not significant. These results are likewise in line with the findings reported for FA development.

As was the case with the FA ratings, data analyses indicated that the SA and AH groups differed significantly in comprehensibility ratings at T2, but we cannot refer to learning context as a decisive factor accounting for this difference since the two groups already differed significantly at T1 as far as comprehensibility is concerned. As discussed in the previous section, these differences between SA and AH learners could be the result of a kind of 'natural selection', by means of which the SA and the AH groups of participants would differ in terms of L2 learning profile, resulting in an imbalance between the two groups in terms of proficiency in the analysed dimensions of L2 speech.

As expected, and in accordance with the results for the FA ratings, analyses comparing native speakers' oral productions to those of the L2 learners indicated that non-native participants and native participants differed significantly in terms of comprehensibility both at Pre-test and at Post-test. Comprehensibility scores from the two non-native groups fell far below native-like scores (NS M=6.41). Scores from the AH participants were further away from the native range (M T2=3.35, M T2=3.88) than scores from the SA group (M T2=4.04, M T2=4.94). However, comprehensibility ratings from all groups of participants at both testing times were closer to native speakers' scores than FA ratings. Thus, we can conclude that both groups of participants approach NSs in terms of comprehensibility more than in terms of FA, and that SA participants get closer in both dimensions after the SA period, although not significantly in any case.

As regards the effect of initial comprehensibility level on comprehensibility development, analyses indicated a strong significant relationship between the ratings obtained at T1 and at T2 by both groups (SA and AH). In line with the FA ratings, this means that the higher the comprehensibility ratings at pre-departure time, the higher the comprehensibility ratings at post-test in both contexts. As for the relationship between gains in comprehensibility and initial comprehensibility level, significant negative correlations between comprehensibility ratings at T1 and gains indicated that participants with worse scores at T1 were the ones obtaining higher gains, and vice-versa.

In other words, the higher the initial level of comprehensibility, the lower the gains. Further analyses were conducted in order to examine the role of initial comprehensibility level on gains in comprehensibility by grouping participants in the SA and AH groups as a function of their initial comprehensibility level (Low-iC and High-iC).

These group analyses, conducted separately for the SA and AH group of participants, revealed significant improvement between T1 and T2 for the SA participants with low initial comprehensibility ratings, and also for those with high initial comprehensibility ratings. This contrasts with results regarding FA improvement for the same SA participants, as no significant improvement in FA was documented for those SA learners performing better at T1. As for the reason why a particular group of learners improved in one dimension but not in the other, we may conclude that FA and comprehensibility are to some extent separate dimensions so improvement or lack thereof in one does not imply the same in the other (Derwing & Munro, 2009; Munro & Derwing, 1999). High-level learners may improve in comprehensibility rather than FA, since the former may be more likely to be developed. However, further research on this question is needed.

The analysis comparing comprehensibility gains between the Low-iC-SA and the High-iC-SA sub-groups indicated that the Low-iC-SA learners obtained significantly greater gains than their peers in the High-iC-SA level group. The fact that learners with lower comprehensibility level at T1 obtained significantly greater gains after the SA period than learners with higher comprehensibility level confirmed findings from previous research which has concluded that learners with lower proficiency benefit more from SA periods in terms of oral production (Carroll, 1967; Freed, 1995c; Juan-Garau, forthcoming; Lapkin et al., 1995; Llanes & Muñoz, 2009; Marriott, 1995; Valls-Ferrer, 2011).

However, the proportion of learners who improved in comprehensibility was not significantly different between the two sub-groups (Low-iC-SA and High-iC-SA). It is worth highlighting that 100% of the SA learners assigned to the Low-iC-SA group improved their comprehensibility after the SA period. This finding is in line with previous results reporting that both SA participants with low initial comprehensibility level and high initial comprehensibility level improved their comprehensibility ratings over time.

In conclusion, all SA participants improved significantly in terms of comprehensibility regardless of their initial comprehensibility level. This contrasts with the results observed for the FA dimension, where only

participants in the low level group (i.e., showing worse FA at T1) significantly improved between T1 and T2. As far as comprehensibility gains are concerned, analyses revealed that there was a significant difference in development between the SA learners with a lower initial comprehensibility level and those performing better at pre-departure time, indicating that those learners with a lower comprehensibility level obtained more gains than learners with a higher comprehensibility level. This also contrasts with the results obtained for the FA dimension, where no significant difference in gains between the high-i-FA-SA and low-iFA-SA sub-groups was reported.

As for the AH group, participants in the initial low comprehensibility group improved significantly over the AH period, while no significant improvement was found in the case of participants performing better at T1. These results are in line with those reported for the FA dimension. Also in accordance with the results for the SA group, participants in the low-iC-AH sub-group obtained greater gains than participants in the high-iC-AH sub-group, with this difference being significant. This contrasts with results obtained for the FA dimension, where no significant difference was reported in terms of gains between the low-iFA-AH group and the high-iFA-AH one. Furthermore, the proportion of learners in the low-iC-AH sub-group who improved comprehensibility was significantly larger than the proportion of participants in the high-iC-AH sub-group.

When analysing the effect of learning context on participants of SA and AH groups with the same initial comprehensibility level (low and high), we found that participants with high initial comprehensibility level in the SA group and the AH group did not differ significantly in comprehensibility level at pre-test. More gains in comprehensibility were observed in the case of the high-iC-SA participants, although the difference in amount of gains between the high-iC SA and AH participant was not significant. As for the low initial comprehensibility sub-groups in the SA and AH contexts, they did not differ significantly at T1 either, and results showed that SA participants with low-iC level obtained significantly greater gains than their AH low-iC peers.

As previously discussed, SA participants obtained greater gains than the AH group in terms of comprehensibility, but this difference was not significant. However, when analyzing differences between participants according to their pre-departure comprehensibility level, results revealed that the SA context benefited those learners with lower initial comprehensibility level more than the AH context in terms of comprehensibility development. These results are in line with previous studies reporting the superiority of a SA period over an AH learning

context regarding development in other aspects of L2 oral competence (Juan-Garau & Pérez-Vidal, 2007; Llanes, 2012; Llanes & Muñoz, 2013; Mora & Valls-Ferrer, 2012; Serrano, Llanes & Tragant, 2011; Trenchs-Parera, 2009).

7.3 Relationship between foreign accent and comprehensibility

A final objective of our research was to examine the relationship between foreign accent and comprehensibility, as formulated in the third research question:

RQ3: To what extent are foreign accent and comprehensibility related speech dimensions?

The interest in exploring the relationship between FA and comprehensibility was twofold. First, we aimed at examining the correlation between these two speech dimensions based on the ratings provided by the listeners of our study (native-speakers of Spanish teaching English in secondary education). Second, we wanted to know about the aspects which listeners had considered when evaluating each speech dimension in the perception task. As seen in our literature review, comprehensibility has traditionally been claimed as the main goal for L2 learners. Therefore, we looked at this dimension in more depth by analyzing listeners' individual reports of their comprehensibility ratings for 20 of the speech samples during the rating task.

In order to find out about the relationship between foreign accent and comprehensibility, two types of analyses –quantitative and qualitative– were carried out. The former consisted of correlations between FA and comprehensibility scores at pre-test and post-test for the SA and AH participants. Results revealed significant large positive correlations between the two speech dimensions at the two testing times for both groups of participants, indicating that the more native-like the accent, the greater the comprehensibility, and vice-versa. These findings contrast with results reported by previous studies positing that heavily accented speech can often be perfectly intelligible (Derwing & Munro, 1997; Munro & Derwing, 1995a, 1999; Gallardo del Puerto, Gómez Lacabex & García Lecumberri, 2007; Hayes-Harb & Watzinger-Tharp, 2012). These results led us therefore to further examine whether listeners' FA and comprehensibility ratings were based on similar aspects of the learners' oral productions.

Subsequent analyses explored the distribution of participants according to their initial and final FA and comprehensibility ratings (see Tables 6.33 and 6.34). Almost half of the whole sample of participants (both SA and AH) were assigned low ratings both for FA and comprehensibility, that is, 44.6% these participants were perceived as both having a strong FA and being little comprehensible. Conversely, about 20% of the participants were given high scores for both dimensions, indicating that these learners were perceived as having a milder FA and being highly comprehensible. It is interesting to observe a kind of 'migration' of participants from the 'low FA-low comprehensibility' profile to the 'high FA-high comprehensibility' one after both the SA and AH periods: at T2 17.8% of the participants were assigned low ratings in both dimensions (instead of 44.6% at T1), and 39.2% were given high ratings in FA and comprehensibility (instead of 19.6% at T1).

Another interesting result from these descriptive analyses was the fact that none of the participants were assigned a high FA rating and a low comprehensibility score. In other words, participants who were assigned high comprehensibility scores were also given good ratings in FA.

Given these two results: (1) positive correlations between FA and comprehensibility (the better the ratings in FA, the higher the ratings in comprehensibility), and (2) more native-like accent always associated with good comprehensibility ratings, the interest in finding out about the aspects which influenced listeners' FA and comprehensibility ratings gained more relevance. Qualitative analyses were thus conducted of listeners' comments gathered in two different ways: from the questionnaires they completed after finishing the rating task, and from 20 reports which were typed in at the same time that they provided their ratings during the rating task.

In order to find out whether listeners had been influenced by the same aspects in the learners' oral productions when assigning their FA and comprehensibility ratings, we provided them with a list of twelve items in the final questionnaire (see Appendix 8) and asked them to select those which had particularly affected their ratings for each dimension. Concerning the aspects affecting their FA ratings, factors from the domain of phonology were selected by most listeners (see Table 6.37): pronunciation of vowels/consonants, word stress, rhythm and intonation were reported in this order as mainly affecting their FA assessment. This finding was confirmed when listeners were asked which three aspects had influenced their FA ratings most (see Figure 6.21). Pronunciation of vowels/consonants, intonation, word stress and rhythm were considered in this order. Actually, these phonological dimensions altogether represented

84% of the factors selected by the listeners as mostly influencing their FA ratings. These results confirmed the findings in Trofimovich and Isaacs (2012) indicating that accent was mainly associated with aspects of phonology (e.g., rhythm, segmental accuracy, and syllable structure).

Some listeners also mentioned other factors related to phonology and pronunciation which were not included in the list of 12 items provided. For instance, one of the listeners referred to 'imitation of native-like accent' as a factor affecting her FA ratings. However, a deeper analysis of the comments made by this listener revealed that this thought was not affecting the listener's FA ratings indeed, since the idea of imitation of FA accent was associated with comprehensibility, as can be read in her comments when rating comprehensibility of 20 of the speech samples:

> COGA: "The speaker tries to pronounce some words with a native-like accent, but it's even worse because it is quite difficult to understand those words".
> COGA: "mocks a native- like accent and slows down comprehensibility".

Another listener reported that the use of L1 words had affected learners' pronunciation in some cases:

> MOLO: "In some cases the use of L1 words, if they interfere the pronunciation or rhythm".

The same list of twelve items was provided to listeners in our study so that they chose the aspects that had affected their comprehensibility ratings. According to their comments, comprehensibility was mostly related to vocabulary and discourse aspects, such as storytelling and content organization (see Table 6.39). Lack of thematic content and grammar were also selected by more than half of the listeners. When asked to identify the three most influential aspects on their comprehensibility ratings, vocabulary (29%), lack of content organization (23%) and speaker's storytelling ability (15%) were reported in this order, followed by grammar (12%) and pronunciation of individual sounds (12%). On the other hand, it is worth noting that even though none of the phonological factors were as important individually for the comprehensibility ratings as in the case of the FA ratings, pronunciation of vowels and consonants represented 12% of the comments, word stress 3% and rhythm 3%. All these phonological factors taken together represented 18% of the comments related to comprehensibility ratings, a higher percentage than, for instance, speaker's storytelling ability, which was ranked third in the list presented above of the most influential factors affecting listeners' assessment of comprehensibility.

These results are in line with Trofimovich and Isaacs' (2012) findings, which indicated that comprehensibility was mainly linked to grammatical accuracy and lexical richness. Actually, this claim was explicitly reported by one of our listeners:

> MOLO: "When having so many problems of vocabulary, sentence structures and fluency, intonation and pronunciation are not so important because the message is impossible to be understood".

As seen in the literature review, Trofimovich and Isaacs (2012) identified five speech aspects which differentiated between L2 learners at different comprehensibility levels: lexical richness and fluency distinguished between low-level learners, grammatical and discourse-level measures differentiated between high-level learners, and word stress errors discriminated between learners of all levels. Overall results in our study suggest that listeners regarded vocabulary and also discourse aspects (lack of content organization and speaker's storytelling ability) as the most important factors related to comprehensibility, factors which were also indicated in Trofimovich and Isaacs (2012) research. However, findings in our research do not support Trofimovich and Isaacs's conclusion pointing out that "four categories uniquely distinguished accent from comprehensibility, with all categories specific to the dimension of phonology (i.e., vowels and consonants, syllables, sounding native-like, and rhythm" (p. 912). As we have seen, pronunciation aspects (e.g., pronunciation of individual sounds) were also taken into account by the listeners in our study when assessing comprehensibility.

It is worth mentioning other aspects (different from the ones provided in the list) which some of the listeners reported as having influenced their comprehensibility ratings. Almost all comments referred to vocabulary, and some of them stressed the use of L1 items as hampering comprehensibility:

> INCA: "Use of Spanish words maybe"[32].
> MOLO: "The use of L1 words in some cases which shows the lack of ability of the student to make the message be understood"

The fact that these listeners considered the use and transfer of L1 words as negatively affecting comprehensibility did not necessarily lead us to

[32] L1 lexical interference was confirmed to negatively affect INCA's comprehensibility ratings, as she reported other comments throughout the perception task such as, "use of words translated from Spanish ("senior", from Spanish word "señor" -meaning "man").

refuse findings in previous studies suggesting a speech comprehensibility benefit in those situations where non-native speakers and listeners share the same L1 background (Gallardo del Puerto et al., 2009; Hayes-Harb et al, 2008). However, the analysis of all the comments gathered from listeners showed that L1 lexical interference was not positively considered in many instances:

> COGA: "The use of Spanish words makes it confusing"
> INCA: "usa vocabulario "traducido" de la lengua materna" (=He "translates" words from his L1).
> MAGU: "Clara influencia de la lengua materna. Adapta claramente vocabulario al inglés. Es dificil de comprender por el vocabulario" (=Clear influence of his L1. He adapts lexical items from his L1. It's difficult to understand because of the vocabulary").

So far, our results partly support those reported by recent research indicating that FA and comprehensibility are linked to different language aspects. On the one hand, we can conclude that aspects of phonology affected FA ratings more than aspects related to other domains such as grammar or vocabulary. On the other hand, although results confirmed that vocabulary and discourse factors, as well as grammar, were the main contributors to variation in comprehensibility ratings, reports from the listeners in our study suggested that factors related to pronunciation had also some influence on their assessment of comprehensibility.

As described in section 5.3.2, the listeners were also asked to type in the aspects of speech which they were taking into account when providing their comprehensibility ratings. Responses for 20 of the speech samples were analyzed. The analyses of these comments helped us to elucidate whether pronunciation was actually involved (or not) in listeners' comprehensibility ratings.

In Chapter 6, we presented the results from these reports in general, and we also examined the comments taking into account whether they referred to positive or negative evaluations of speech comprehensibility. When analysing all the comments provided by the listeners we found that 26% of the comments regarding their comprehensibility ratings considered aspects of pronunciation, 18.45% of the comments referred to vocabulary, 14.59% to grammar, and 14.42% to fluency. Therefore, a new distribution of the aspects affecting this speech dimension was obtained (compared to the 12-item classification from the final questionnaires presented above). Vocabulary was considered a key aspect for comprehensibility in many of the comments, but pronunciation was more frequently highlighted.

As presented in section 6.3.2.1, the comments from the reports were classified as affecting negatively or positively the listeners' ratings. With regard to the reasons hindering comprehensibility, 26% of the comments were related to pronunciation aspects, 22.71% associated with vocabulary, and 19% linked to grammar. Pronunciation was also regarded as the variable enhancing pronunciation most. Twenty-six percent of the comments arguing reasons for good comprehensibility had to do with pronunciation, 14% were related to fluency, 12.75% to discourse, and 12.34% were associated with vocabulary.

Therefore, according to these analyses, comprehensibility was related to pronunciation to a considerable degree. In order to gain a better understanding of the pronunciation aspects promoting (or hampering) comprehensibility, we classified the comments according to the pronunciation features which the listeners were particularly referring to (see Tables 6.43 and 6.45 in previous chapter). The top three pronunciation features which negatively influenced the listeners' comprehensibility ratings were pronunciation of individual sounds and words (55.5%), foreign accent (17.42%), and intonation (11.1%). As for the pronunciation aspects which enhanced comprehensibility, listeners cited pronunciation in general (41.1%), imitation of native-like pronunciation (22%), and degree of accentedness (14.7%), followed by intonation (10.29%).

Therefore, according to the reports on comprehensibility provided by the listeners while carrying out the perception task, pronunciation was found to be the most relevant aspect in their assessment of comprehensibility of L2 learners' speech. The fact that pronunciation was not ranked within the top three aspects affecting comprehensibility in the data obtained from the final questionnaires may respond to two possible reasons. First, since we presented the questions about the factors influencing FA and comprehensibility ratings in the same questionnaire, we could have implicitly motivated the distinction between the two constructs. Second, as already remarked, while it is true that specific pronunciation features (e.g., pronunciation of individual sounds, intonation, stress, etc.) did not greatly affect comprehensibility when considered separately, pronunciation aspects taken as a whole did have a considerable impact on the listeners' comments. On the other hand, we may consider the comments made by the listeners during the rating experiment as more reliable and ecologically valid than the comments collected at the end of the experiment, as the former were reported when listeners were actually rating the speech samples for comprehensibility.

According to these data, we can conclude that non-native listeners in our study took into account pronunciation aspects when assessing L2 learners' comprehensibility. First, strong positive correlations between FA and comprehensibility were found. Second, listeners in our study did pay attention to aspects related to accent or native-like pronunciation when providing their comprehensibility ratings, in contrast with previous research involving native listeners of English (Trofimovich and Isaacs, 2012). Further research is needed to gain a better understanding of the aspects affecting comprehensibility as reported by native and non-native listeners.

Last but not least, we would like to address two more issues presented in the theoretical background of this dissertation, taking into account the results obtained in this research: the conflation of the FA and comprehensibility dimensions, and the figure of the native-speaker interlocutor.

As pointed out when discussing some assessment criteria and scales in section 3.3, our listeners seemed to conflate the two dimensions of accentedness and comprehensibility. Although it seems that they have a clear idea of the two different constructs, their performance in the rating experiment provides contradicting results. Question 40 in the profile questionnaire, which the listeners had to complete before the rating experiment (see Appendix 4), reads as follows:

> "From your point of view, when assessing your students' oral productions, how important is each of the following dimensions? (Accent/Intelligibility-Comprehensibility/Fluency) Possible answers: Not important at all/Not so important/More or less important/Important/Very important".

Answers from the listeners revealed that when assessing oral productions, 33.3% of the listeners reported accent as 'not important at all' or 'not so important', and 50% considered it 'more or less important'. However, accent is a speech dimension primarily related to pronunciation aspects, and pronunciation was the aspect most frequently cited in the listeners' comments when rating the comprehensibility of 20 of the speech samples in the experiment. This suggests, therefore that although listeners in our study appear to believe that accent (and pronunciation) should not be the focus when assessing L2 learners' oral productions, they actually took into account this aspect to a great extent when rating the analyzed L2 learners' speech samples.

Another question related to accent and reflected in the listeners' comments about comprehensibility is the figure of the native speaker.

Listeners referred to the fact of sounding native-like (or not) on several occasions throughout their comments about comprehensibility. On some occasions native-like accent or pronunciation was explicitly regarded as positive:

> COGA: "tries to pronounce words in a native-like way"
> ESES: "use of Spanish words but with English-like pronunciation"
> INCA: "Intenta imitar la pronunciación de las "T"s en inglés"(=He tries to imitate the 't' sound in English"
> INCA: "Intenta pronunciar los sonidos como un nativo" (=She tries to pronounce words as a native-speaker)
> ROMA: "in certain parts of speech he shows a more English-like accent"

In other comments, it was clearly shown that listeners were assigning their ratings from a native-speaker perspective:

> LLCA: "A native speaker could understand him despite the grammar mistakes".
> INCA: "su acento es muy español, quizás para un nativo de habla inglesa sería muy dificil reconocer el verbo 'rob' tal y como es pronunciado en la grabación" (=His speech is clearly Spanish-accented. An English native speaker could find it hard to recognize some of the words as they are pronounced by the learners (e.g. the verb 'rob').
> MOLO: "Some grammar mistakes, but they could be assumed by a native speaker, but vocabulary problems make the message incomprehensible".

As far as native-like pronunciation is concerned, we also found some comments referring to a particular variety of English:

> ROMA: "His native-like pronunciation is quite evident and it makes comprehension easier, even though his accent is far from the standard British"

All these comments confirmed that still today many EFL teachers in Spain have an idealized view of the native speaker as a reference for L2 acquisition and assume a native-speaker interlocutor, and maybe a particular pronunciation/accent model. Actually, answers provided to questions 13, 15 and 16 in the profile questionnaires (see Appendix 4) showed that:

(1) 41.6% of the listeners reported that L2 learners should acquire a native-like pronunciation, while 41.6% was not sure.

(2) 66.6% of the listeners considered that having native-like or near native-like pronunciation in English was important.

(3) Only 25% of the listeners did not consider any geographical variety of English as the best accent to acquire.

Interesting definitions of what "good English pronunciation" is were also collected from listeners when answering question 14. Table 7.1 reports the listeners' comments:

ANSE	In our global world, the one that allows understanding each other.
COGA	Non international language
ELME	I would say intonation is one of the most important things in pronunciation (apart from individual sounds)
ESES	For me, "good English pronunciation" is that which can be easily identified and understood by a wide range of speakers from different areas. This should have the common features to all the varieties in pronunciation (with possible distinct or specific deviations).
ESMU	Comprehensible, "clean", well linked discourse, ...
HERO	Pronunciation in English is good when it is quite similar to the pronunciation of a native English speaker (native-like or near-native like).
INCA	A fluent one
LLCA	Any English variety is good English pronunciation, but it seems to me that American English is much more systematic and logical, as far as vowel contexts and sounds are concerned.
MAGU	The "good English pronunciation" is the one close to be native and without influence of the mother tongue.
MOLO	Any pronunciation of a word whose meaning could be decoded by the listener.
MJSE	That which allows them to communicate effectively.
ROMA	Clear, comprehensible and reflecting the main traits of a language (stress, rhythm and intonation).

Table 7.1: Answers from listeners to Q14:
In your opinion, what is "good English pronunciation"?

We can observe how some of the listeners clearly associated "good pronunciation" with "native-likeness" (see comments by COGA, HERO and MAGU). On the other hand, some listeners referred to pronunciation which is not related to any specific geographical variety and which allows communication between interlocutors (ANSE, ESES, MJSE).

Further research should be carried out to elucidate whether these findings reflecting the importance attached to native-like pronunciation in EFL teaching were specific to the particular group of non-native listeners in our study, or, on the contrary, whether these listeners are representative of what happens in EFL formal instruction learning contexts today. If this were the case, it would be important to consider the impact of such results

from research on actual teaching practice. For instance, even though several studies have outlined the importance of EIL and the idea of the mutual intelligibility between non-native speakers (Derwing and Munro, 2005; Jenkins, 2000; Kachru, 1992), results from the present study confirm that native-speakers are still a reference for many EFL teachers.

In addition, even though some studies have indicated that pronunciation training can help L2 speakers produce more intelligible speech (Derwing & Munro, 2009; Derwing, Munro & Wiebe, 1998), speaking in general and pronunciation in particular are not much the focus of attention in the FI context as reported by participants and listeners in our research (see section 5.2.1.2), which confirms previous results (Henderson, et al. 2012).

In line with Derwing and Munro (2009), we deem it appropriate to work on those aspects of pronunciation which may affect comprehensibility. As seen in section 3.1, Munro and Derwing (1999) suggested that prosodic errors seemed to affect intelligibility more than errors at the segmental level. Derwing, Munro and Wiebe (1998) compared the effects of two different types of pronunciation instruction: a programme focusing on segmental accuracy and a global programme addressing general speaking habits and prosodic factors. They concluded that only the L2 learners receiving global instruction in their study showed improvement in comprehensibility and fluency in their narratives. As already noted, the findings in Derwing et al. (1998) are at odds with the actual situation in the EFL classroom, where pronunciation practice, when available, tends to focus on error correction at the segmental level. Therefore, it seems that classroom practice remains very often far from the empirically-based findings in the research field.

Chapter 8

Conclusions and Future Research

In this dissertation we have examined the development of foreign accent and comprehensibility in the oral productions of 25 adolescent Spanish learners of English before and after a 3-month study abroad period, and 31 adolescent EFL Spanish learners receiving formal instruction in the AH learning context during the same period of time. Results showed that SA participants improved significantly between pre-test and post-test in both speech dimensions. Thus, our study confirms previous results reporting the beneficial impact of SA context on L2 learners' language development, in general, and on oral production, in particular.

In order to determine whether the SA learning context promoted more gains in FA and comprehensibility than in the at home formal instruction learning context, we also evaluated FA and comprehensibility progress made by a group of AH participants. Results revealed that the AH participants also improved significantly between pre-test and post-test in both dimensions. Subsequent analyses compared the amount of gains obtained by each group of learners in order to find out whether one learning context was more beneficial than the other. Results indicated that SA participants obtained significantly more gains in FA than the AH group. Thus, our findings suggest that the SA learning context is more beneficial than the AH setting in terms of improving those language aspects related to accent (i.e. pronunciation). As for comprehensibility, the SA context was more beneficial than the AH context in terms of comprehensibility development, since the percentage of learners

improving their comprehensibility scores during SA was significantly larger than the percentage of learners improving their scores in the AH context, and SA learners obtained larger comprehensibility gains than AH learners, although such improvement was not significant.

The lack of previous studies examining FA and comprehensibility in the oral speech produced by SA and AH L2 learners calls for further research exploring these speech dimensions in order to confirm ours. The adolescent population we have analysed is even more in demand of the attention of research which so far is scarce. In this respect, this situation reflects the important contribution we hope to be making with our study to the body of SA research.

As pointed out in Chapter 7, the comparisons made between the two groups of participants in our study should be taken cautiously since the L2 learners differed significantly in both FA and comprehensibility ratings at T1. This fact made us reflect about what seems a kind of 'natural selection' of the participants assigned to one group or the other (Bruton, 2011; Sanz, forthcoming). Studies comparing AH learners with participants in a different learning contexts (i.e. SA, CLIL) have usually recruited the learners in the AH institution who do not enroll in the SA or CLIL programme as the control group. Some of these studies have reported greater gains in particular language aspects for the SA or CLIL participants (Gallardo del Puerto et al., 2009; Llanes, 2012; Llanes & Muñoz, 2013; Rallo & Juan-Garau, 2011). However, given this possible 'natural selection' of participants when comparing two different learning contexts (i.e. SA vs AH), the best control group for context effect comparisons may not be the remaining group of students in the school(s) where a specific language learning programme is being carried out. As done in several studies edited by Juan-Garau and Salazar-Noguera (forthcoming), we may wish to find a control group which is more similar to the SA or CLIL group outside the research participants' institution, in a school where no CLIL or SA programmes are offered and a more comparable control group of participants can be obtained in terms of initial level. Taking this into account, the control group in our study may not have been the best one for learning context comparison purposes, since SA and AH learners already differed in FA and comprehensibility level at pre-test time. On the other hand, comparisons between the SA group and the AH group in our research may yield important conclusions for the school where these data were gathered, proving the ecological validity of our research.

In our analysis of longitudinal improvement we also examined the influence of initial level on FA and comprehensibility development. As in

previous research, we found that learners' initial level of FA and comprehensibility had an impact on the amount of progress they made. Significant improvement was found for SA and AH participants with heavy FA at pre-test (low initial FA level), but not for the SA and AH groups with a better initial level. Moreover, when comparing the SA low initial level group and the AH low initial group, the SA participants were found to obtain significantly greater gains in FA than the AH group. As for comprehensibility, participants in the SA programme improved significantly between the two testing times regardless of their initial level, but the low initial comprehensibility level group obtained significantly greater gains than the high initial level group. Unlike the SA group, only AH participants with a low initial level of comprehensibility improved significantly over time. Moreover, the low initial level group obtained significantly greater gains in comprehensibility than participants with higher initial scores. Regarding the learning context effect, SA participants with initially lower levels of FA and/or comprehensibility obtained significantly greater gains than their peers at home in both dimensions. These findings contribute to the body of SA research exploring oral production development and claiming that learners with lower proficiency benefit more from SA periods (Freed, 1995c; Juan-Garau, forthcoming; Lapkin et al., 1995; Llanes & Muñoz, 2009; Marriott, 1995; Regan, 1995; Valls-Ferrer, 2011).

In this research we were also able to examine FA and comprehensibility ratings assigned by English non-native instructors, who are frequently responsible for teaching EFL in the AH context in Spain, so as to determine the relationship between these speech dimensions. Contrary to previous findings (Derwing & Munro, 2009; Munro & Derwing, 1999), a strong correlation was found between FA and comprehensibility, indicating that those learners with better accent obtained higher comprehensibility ratings, and learners with heavier FA were also perceived as less comprehensible. Furthermore, we explored the aspects that listeners took into account when assessing FA and comprehensibility. Results showed that the FA dimension was mainly associated with pronunciation aspects, which also affected comprehensibility ratings assigned by the non-native listeners in our research. Confirming previous research (Trofimovich & Isaacs, 2012), aspects such as vocabulary and grammar were taken into account when rating L2 learners' speech comprehensibility but, contrary to previous findings in studies involving native listeners of English, pronunciation was the aspect that listeners heeded most when assigning comprehensibility scores.

It remains unclear whether the aspects reported by the group of non-native listeners of our study are specific to our groups of participants, or can be

generalized to learners from different L1 backgrounds. In addition it is important to validate our findings with English native and non-native listeners from other L1 backgrounds and profiles.

Further research should also take into account one of the limitations of this study. As already done in previous research examining other language dimensions (e.g. oral fluency), the design of our study did not include collecting data from the speakers' L1. It would be interesting to analyse L1 oral productions of the same participants for comprehensibility on the same task in order to find out which aspects of the L2 fluency are due to the command (or lack thereof) of particular second language aspects and which ones are attributable to more individual characteristics such as the learner's speaking style (as reported in some of the comments from our listeners) (Segalowitz, 2010; de Jong et al., 2012).

As for methodological issues, it is worth mentioning that the fact of using different data sources and time collections provided us with rich and complementary information which enabled us to better understand results and suggest possible explanations for our findings (i.e. questionnaires including different types of questions to be completed by the listeners after the rating experiment, and reports they had to type in while carrying out the perception task). The combination of quantitative and qualitative data analyses (i.e. mixed methods approach) has proved to be a useful and enriching way of interpreting data and drawing more reliable conclusions.

Finally, although it is widely accepted that the objective of L2 pronunciation teaching should be to help L2 learners be understandable for their interlocutors, classroom teachers have received little guidance on the pronunciation features on which they should focus during lessons (Derwing & Munro, 2009). The fact that pronunciation tends to suffer from neglect may not be due to teachers' lack of interest in the subject but rather to a feeling of doubt as to how to teach it. Lack of knowledge of phonetics and lack of formal preparation to teach pronunciation are two of the most cited problems, which have been corroborated in our research. The urgent need for specific pronunciation training for teachers in Spain has been called for frequently (Donovan, 2001; Levey 1999, 2001; Pavon, 2001; Pavon and Rosado, 2003). In this regard, it is worth highlighting the willingness reported by listeners to benefit from pronunciation training programmes and participation in studies like this one, which have provided them with food for thought.

We would like to conclude with what has been emphasized in the introduction of this dissertation. One of the main reasons and motivations for this research study was contributing to the relationship between

research and language teaching practice. Although further studies need to be conducted in order to confirm and generalise our findings, ours is a modest but ecologically valid contribution to empirical-based research aiming at exploring what really happens in classrooms with regard the teaching of pronunciation at least in the Spanish context.

References

Alonso, E., Grisaleña, J., & Campo, A. (2008). Plurilingual education in secondary schools: Analysis of results. *International CLIL Journal*, 1 (1), 36–49.

Allen, H.W. (2010). Language-learning motivation during short-term Study Abroad: An activity theory perspective. *Foreign Language Annals*, 43(1), 27–49.

Allen, H.W., & Herron, C. A. (2003). A mixed methodology investigation of the linguistic and affective outcomes of summer study abroad. *Foreign Language Annals*, 36, 370–384.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Cambridge University Press.

Anderson, J.R. (1992). Automaticity and the ACT∗ Theory. *American Journal of Psychology*, *105*, 165–180.

Anderson, J.R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmental, prosody, and syllable structure. Language Learning, 42, 529–555.

Asher, J. J. & García, R. (1969) The optimal age to learn a second language, *The Modern Language Journal*, 53(5), 334-341.

Avello, P. (2013). *L2 phonological development in speech production during Study Abroad.*¸PhD Thesis, Universitat Pompeu Fabra, Barcelona.

Avello, P., Mora, J.C., Pérez-Vidal, C. (2012) Perception of FA by non-native listeners in a study abroad context. *Research in language*, 10 (1), 63-78.

Avello, P., and Lara, A. R. (forthcoming) Phonological Development in L2 Speech Production during Study Abroad Programmes Differing in Length of Stay. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

References

Barquin, E. (2012). *Writing development in a Study Abroad Context*. PhD Thesis, Universitat Pompeu Fabra, Barcelona.

Beattie, J.; Valls-Ferrer, M., & Pérez-Vidal, C. (forthcoming). In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

Bent, T. & Bradlow, A.R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical society of America*, 114, 1600-1610.

Bohn, O. -S. & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131-158.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of vey advance later L2 learners. In D. Birdsong (Ed.), *Second language acquistiion and the critical period hypothesis* (pp. 133-159). Mahwah, NJ: Lawrence Erlbaum Associates.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters attain a native accent in a foreign language? A test of the Critical Period Hypothesis. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition: A critical look at the critical period hypothesis* (pp. 30-50). Clevedon, England: Multilingual Matters.

Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquision*, 19, 447-465.

Bradlow AR, Pisoni DB, Akahane-Yamada R, Tohkura Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, 2299–2310.

Brecht, R. D., & Robinson, J. L. (1993). Qualitative analysis of second language acquisition in study abroad: The ACTR/NFLC project. NFLC Occasional Papers, pp. 1–25. Washington, DC: National Foreign Language Center.

Brecht, R., Davidson, D., and Ginsberg, R. (1995). Predictors of foreign language gain during study abroad. In: Freed, B. (Ed.). *Second*

language acquisition in a study abroad context (pp. 37-66). Amsterdam: John Benjamins Publishing Company.

Brennan, E. M., & Brennan, J. S. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. Language and Speech, 24, 207-221.

Brown, A., Iwashita, N., & McNamara, T. (2005). An examination of rater orientations and test-taker performance on English for academic purposes speaking tasks (TOEFL Monograph 29). Princeton, NJ: Educational testing Service.

Bruton, A. (2011). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal and Moore (2010). *Applied Linguistics*, 32(2): 236-241.

Campbell, C. (1996). Socializing with teachers and prior language learning experience: A diary study. In K. bailey and D. Nunan (eds) *Voices from the language classroom* (pp. 201-223). Cambridge: Cambridge University Press.

Carroll, J. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, 1, 131_151.

Celce-Murcia, M., Brinton, D.M. & Goodwin, J.M. (1996). *Teaching pronunciation. A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Chiba, R., Matsuura, H., & Yamamoto, A. (1995). "Japanese attitudes towards English Accents". *World Englishes*, 14(1): 77-86.

Cohen L., Manion L., & Morrison, K. R. B. (2000). *Research methods in education* (5th ed.). London: Routledge.

Coleman, J. A. (1996). *Studying Languages: A Survey of British and European Students: The Proficiency, Background, Attitudes and Motivations of Students of Foreign Languages in the United Kingdom and Europe.* London: Centre for Information on Language Teaching and Research.

Coleman, J.A. (1998). Language learning and study abroad: the European perspective. *Frontiers*, 4 (1), 167-203.

Collentine, J. and Freed, B.F. (2004) Learning context and its effects on second language acquisition. SSLA, 26, 153-171. Cambridge University Press.

Collentine , J. (2004) The effects of learning contexts of morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26 (2), 227-248.

Collentine, J. (2009). Study abroad research: findings, implications, and future directions. In Long, M. and Doughty (Eds.), C. *The handbook of language teaching* (pp. 218-233). Malden, MA: Wiley-Blackwell.

Cubillos, J.H., Chieffo, L. & Fan, C. (2008) The impact of short-term study abroad programs on L2 listening comprehension skills. *Foreign Language Annals*, 41 (1), 157-185.

Dalton-Puffer, C., Kaltenboek, G., & Smit, U. (1997). "Learner attitudes and L2 pronunciation in Austria". *World Englishes*, 16(1): 115-28.

Dalton C. & B. Seidlhofer (1994). *Pronunciation*. Oxford: Oxford University Press.

Davidson, (2010) Study Abroad: When, How Long, and With What Results? New Data From the Russian Front. *Foreign Language Annals*, 43 (1), 6-26.

de Jong, N. H., Steinel, M.P., Florijn, A., Schoonen, R., & Hulstijn, j.H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5-34.

Dechert, H.W. (2005). Some critical remarks concerning Penfield's theory of second language acquisition. In D. singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition. A critical look at the Critical Period Hypothesis* (pp. 67-94). Clevedon: Multilingual Matters.

DeKeyser, R. (1991). Foreign language development during a semester abroad. In: B. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 104-119). Lexington, MA: D.C. Heath.

DeKeyser, R. (2007a) Introduction: Situating the concept of practice. In DeKeyser, R. (Ed.) *Practice in a second language. Perspectives*

from applied linguistics and cognitive psychology (pp. 1-18). New York: Cambridge University Press.

DeKeyser, R. (2007b) Study abroad as foreign language practice. In DeKeyser, R. (Ed.) *Practice in a second language. Perspectives from applied linguistics and cognitive psychology* (pp. 208-226). New York: Cambridge University Press.

DeKeyser, R. (2007c). Skill acquisition theory. In B. VanPatten, & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 97-114). Mahwah, NJ: Lawrence Erlbaum Associates.

DeKeyser, R. (forthcoming). Methodological Considerations about Research on Language Development during Study Abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins.

Derwing, T.M. (2008). Curriculum issues in teaching pronunciation to second language learners. In J.G. Hansen Edwards, & M.L.Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 347-369). Amsterdam: John Benjamins Publishing Company.

Derwing, T. & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19, 1-16.

Derwing, T. & Munro, M. J. (2005). Second language accent and pronunciation teaching: a research-based approach. TESOL Quarterly, 39 (3), 379 – 397.

Derwing, T. & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42 (4), 476 – 490.

Derwing, T. & Munro, M. J. (2011). Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39 (3), 379 – 397.

Derwing, T. & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: a 7-year study. *Language Learning*, 63 (2), 163-185.

References

Derwing, T., Munro, M.J., & Thomson, R.I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29, 359-380.

Derwing, T., Munro, M.J., & Wiebe, G. (1997). Pronunciation instruction for "fossilized" learners: Can it help? *Applied Language Learning*, 8, 217 – 235.

Derwing, T., Munro, M.J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, 393-410.

Derwing, T., Rossiter, M.J. & Munro, M.J. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development*, 23, 245-259.

Derwing, T., Thomson, R.I., & Munro, M.J. (2006). English pronunciation and fluency development in Mandarin and Slavic Speakers. *System*, 34, 183-193.

Dewey, D. (2004). A comparison of Reading development by learners of Japanese in intensive and domestic Immersion and Study Abroad contexts. *Studies in Second Language Acquisition*, 26, 303-327.

Dewey, D. (2008). Japanese vocabulary acquisition by learners in three contexts. *Frontiers, 15*, 127-148.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, 26 (2), 249-273.

Díaz-Campos, M. (2006). The effect of style in second language phonology: An analysis of segmental acquisition in study abroad and regular-classroom students. In C.A. Klee & T.L Face (eds), *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*. Somerville, MA: Cascadilla Proceedings Project, pp. 26-39.

Donovan, P.J. (2001). Making Pronunciation a Priority for EFL Teachers and Learners. In Levey, D., Losey, M.A. & González, M.A. (Eds.). *English Language Teaching Changing Perspectives in Context* (pp.245-249). Cádiz: Universidad de Cádiz (Servicio de Publicaciones).

References

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Douglas, D. (1994). Quantity and quality in speaking test performance. *Language Testing*, 11, 125-144.

Dufon, A. and E. Churchill (2006). *Language learners in Study Abroad contexts*. Clevedon: Multilingual Matters.

Dwyer, M. M. (2004). More is better: The impact of study abroad program duration. Frontiers: The Interdisciplinary Journal of Study Abroad, 10.

Dyson, P. (1988). *The year abroad. Report for the Central Bureau for Educational Visits and Exchanges*. Oxford: Oxford University Language Teaching Centre.

Fisher, L. and Evans, M. (2000). The school exchange visit: effects on attitudes and proficiency in language learning. Language Learning Journal, 22, 11-16.

Flege, J.E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76, 692-707.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America*, 84, 70-79.

Flege, J.E. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89, 395-411.

Flege, J.E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 229-273). Timonium, MD: York Press.

Flege, J. E. and Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91, 370-389.

Flege, J.E., Frieda, E.M., and Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2, *Journal of Phonetics*, 25, 169-186.

References

Flege,J.E. & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23 (4), 527-552.

Flege, J.E., Munro, M., & MacKay, I.R.A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.

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

Foster, P. (2009). Lexical diversity and native-like selection: The bonus of studying abroad. In B. Richards, M. Daller, D. Malvern, P. Meara, J. Milton and J.Treffers-Daller (eds) *Vocabulary Studies in First and Second Language Acquisition* (pp. 91-106). Basingstoke: Palgrave Macmillan.

Freed, B.F. (1995a). *Second language acquisition in a study abroad context*. Amsterdam: John Benjamins Publishing Company Publishing Company.

Freed, B. (1995b). "Language learning and Study Abroad". In Freed, B. (Ed.). (1995): *Second language acquisition in a study abroad context* (pp. 3-33). Amsterdam: John Benjamins Publishing Company.

Freed, B.F. (1995c). What makes us think that students who study abroad become fluent? In: Freed, B. (Ed.). Second language acquisition in a study abroad context (pp. 123-148). Amsterdam: John Benjamins Publishing Company.

Freed, B.F. (1998). An overview of issues and research in language learning in a study abroad setting. *Frontiers*: *The interdisciplinary journal of second language acquisition*, vol 4, 2, pp. 31-60.

Freed, B.F., Segalowitz, N. & Dewey, D. (2004). Context of learning and second language fluency in French. Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26 (2), 275-301. Cambridge University Press.

Freed, B., So, S., & Lazar, N. (2003). Language Learning Abroad: How Do Gains in Written Fluency Compare with Gains in Oral

Fluency in French as a Second Language? *ADFL BULLETIN, 34*(3), 34-40.

Fullana, N. (2005). *Age-related effects on the acquisition of a foreign language phonology in a formal setting*. PhD Thesis, Universitat de Barcelona, Barcelona.

Fullana, N. (2006). The development of English (FL) perception and production skills. starting age and exposure effects. in Muñoz, C. (ed.) *Age and the rate of foreign language learning.* Clevedon, England: Multilingual Matters, 41-64.

Gallardo del Puerto, F., García-Lecumberri, M.L. & Cenoz, J. (2005). Degree of foreign accent and age of onset in formal school instructions. *Proceedings of Phonetics Teaching and Learning Conference,* London, August 2005.

Gallardo del Puerto, F., Gómez Lacabex, E. & García Lecumberri, M.L. (2009). Testing the effectiveness of Content and Language Integrate Learning in foreign language contexts: the assessment of English pronunciation. In Ruiz de Zarobe, Y., Jiménez Catalán, R.M. (Eds.). *Content and Language Integrated Learning. Evidence from research in Europe.* (pp. 63-80). Clevedon: Multilingua Matters.

Gallardo del Puerto, F., Gómez Lacabex, E. & García Lecumberri, M.L. (2007). *Proceedings of Phonetics Teaching and Learning Conference,* London, August 2007.

García Lecumberri, M. L., Gallardo del Puerto, F. 2003. English FL sounds in school learners of different ages. In García-Mayo, M. P., García-Lecumberri, M.L. (eds.), *Age and the Acquisition of English as a Foreign Language* (pp. 115-135). Clevedon, England: Multilingual Matters.

Gardner, R.C. & Lambert, W.E. (1972). *Attitudes and Motivation in Second Language Learning*. Rowley, MA: Newbury House.

Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, 34, 65–89.

Gimson, A.C. (1994) *Gimson's pronunciation of English.* Revision by A. Cruttenden. London: Edward Arnold.

References

Griffen, T. (1991). A non-segmental approach to the teaching of pronunciation. In A. Brown (Ed.), *Teaching English pronunciation: a book of readings (pp. 17 – 190). London: Routledge.*

Hahn, L.D. (2004). Primary stress and intelligibility: research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38, 201-223.

Harding, L. (2008) . Accent and academic listening assessment: A study of test-taker perceptions. *Melbourne Papers in Language Testing,* 13, 1-33.

Harding, L. (2013). Pronunciation assessment. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics.* Hoboken, NJ: Wiley-Blackwell.

Hayes-Harb, R., Smith, B.L., Bent, T., & Bradlow, A.R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word- final voicing contrasts. *Journal of Phonetics,* 36, 664-679.

Hayes-Harb, R. & Watzinger-Tharp, J. (2012) Accent, intelligibility, and the role of the listener: perceptions of English-accented German by native German speakers. *Foreign Language Annals*, 45, 260-282.

Henderson A., Frost D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., … Curnick, L. (2012). The English Pronunciation Teaching in Europe Survey: selected results of a pilot study. *Research in language*, 10 (1), 5-26.

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

Howard, M. (2001). The effects of study abroad on the L2 learner's structural skills. *EUROSLA Yearbook*,1,. pp 123-141.

Howard, M. (2005). Second language acquisition in a study abroad context: A comparative investigation of the effects of study abroad and foreign language instuction on the L2 learner's grammatical development. In Alex Housen & Michel Pierrard (Eds.). *Investigations in instructed second language acquisition*, pp. 495-530. Berlin: Mouton de Gruyter.

References

Howard, M. (2009). Short-versus long-term effects of naturalistic exposure on the advanced instructed learner's L2 development: a case- study. In: Labeau, E., Myles, F. (Eds.), *The Advanced Learner Variety: The Case of French*. Peter Lang, Oxford, pp. 93e123.

Ife, A., Vives, G., & Meara, P. (2000). The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics*, 4(1), 55-84.

Ioup, G., Boustagi, E. El Tigi, M., & Moselle, M. (1994). Re-examining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16: 73-98.

Isaacs, T. (2010). Issues and arguments in the measurement of second language pronunciation. PhD Thesis, McGill University, Montreal.

Isaacs, T. & Thomson, R.I. (2013). Rater experience, rating scale length, and Judgments of L2 pronunciation: revisiting research conventions. *Language Assessment Quarterly* 10 (2), 135-159.

Isaacs, T. and Trofimovich, P. (2011). Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of second language speech. *Applied Psycholinguistics*, 32, 113-140.

Isaacs, T. & Trofimovich, P. (2012). "Deconstructing" comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34, 475-505.

Isabelli, C. L. (2001). Motivation and extended interaction in the study abroad context: Factors in the development of Spanish language accuracy and communication skills. *Dissertation Abstracts Internationa*l, 61(11), 43-62.

Isabelli, C. A. & Nishida, C. (2005). Development of the Spanish Subjunctive in a Nine month Study-abroad Setting. In D. Eddington (ed.) *Selected proceedings of the 6th conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 78- 91). Somerville, MA: Cascadilla Press.

References

Isabelli-García, C. A. (2003). Development of oral communication skills abroad. *Frontiers*, 9: 149-173.

Ismail, B., Morgan, M. & Hayes, K. (2006). Effect of Short Study Abroad Course on Student Openness to Diversity. *Journal of Food Science Education*, 1, 15-18.

Jenkins, J. (2000). *The phonology of English as an International language.* Oxford: Oxford University Press.

Juan-Garau, M. (forthcoming). Oral accuracy growth after formal instruction and study abroad: Onset level, contact factors and long-term effects. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

Juan-Garau, M. & Pérez-Vidal, C. (2007). The effect of context and contact on oral performance in students who go on a stay abroad. VIAL, 4, 117-134.

Juan-Garau, M. and Salazar-Noguera, S. Forthcoming. *Content-based learning in multilingual educational environments*. Berlin: Springer.

Juan-Garau, M., Salazar-Noguera, J. & Prieto-Arranz, J.I. (forthcoming) L3 English learners' lexico-grammatical and motivational development at home and abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

Jun, H. G. & Li, J. (2010). Factors in raters' perceptions of comprehension and accentedness. In J. Levis & K. LeVelle (Eds). *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, Iowa State University, Sept. 2009. (pp. 53-66), Ames, IA: Iowa State University.

Kachru, B.B. (1992). T*he other tongue: English across cultures*. Urbana; IL: University of Illinois press.

Kang, O. & Rubin, D.L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology,* 28, 441-456.

Kelly, G. (2000). *How to teach pronunciation*. London: Longman.

Kennedy, S., & Trofi movich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, 64, 459–489.

Kenworthy, J. (1987). *Teaching English pronunciation*. London: Longman.

Kinginger, C. (2004). 'Alice Doesn't Live Here Anymore: Foreign Language Learning as Identity (Re)construction'. In A. Pavlenko and A. Blackledge (eds). *Negotiation of identities in multilingual contexts*, pp. 219–242. Clevedon: Multilingual Matters.

Kinginger, C. (2008). Language learning in study abroad: Case studies of Americans in France. *Modern Language Journal Monograph,* Volume 1. Oxford: Blackwell

Kinginger, C. (2009). *Language learning and study abroad: a critical reading of research*. Basingstoke, UK: Palgrave MacMillan.

Kinginger, C. & Farrell, K. (2004). Assessing development of meta-pragmatic awareness in study abroad. *Frontiers: The Interdisciplinary ournal of Study Abroad*, 10, 19-42.

Knight, S. M. and Schmidt-Rinehart, B. C. (2002). Enhancing the Homestay: Study Abroad from the Host Family's Perspective. *Foreign Language Annals*, 35, 2, 190–201.

Lafford, B. (1995). Getting Into, Through and Out of a Situation: A Comparison of Communicative Strategies Used by Students Studying Spanish Abroad and 'At Home'. In B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 97-121). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Lafford, B. (2004): The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition* 26 (2), 201-225.

Lapkin, S., Hart, D., & Swain, M. (1995). A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay

in Quebec. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 67–94). Amsterdam: John Benjamins Publishing Company.

Larson-Hall (2010). *A guide to doing statistics in second language research using SPSS.* New York: Routledge.

Lasagabaster, D., & Sierra, J. M. (2009). Language attitudes in CLIL and traditional EFL classes. *International CLIL Research*, 1(2), 4-17.

Lenneberg, E.H. (1967). *Biological foundations of language*. New York: Wiley.

Lennon, P. (1990). Investigating fluency in EFL: a quantitative approach. *Language Learning*, 40 (3), 387-417.

Levey, D. (1999). Half Truths and White Lies: A Practical Pronunciation Guide for Spanish Speakers. In Harris, T. & Sanz, I. (Eds.) ELT: *Through the Looking Glass.*Granada: Greta. 215-226.

Levin, D.M (2001). Language learners' sociocultural interaction in a study abroad context. PD thesis, Inidana University. *Dissertation Abstracts International* – A 62 (2), 498.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39, 369-377.

Levis, J.M. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (Ed.), *Spoken English, TESOL and applied linguistics*: *Challenges for theory and practice* (pp,. 245 - 270). New York: Palgrave Macmillan.

Li, J. (2000). Cross-culture contact: A study of factors that contribute to culture shock on ESL students' adjustment in the English Language Institute at the University of Tennessee, 1999. PhD thesis, University of Tennessee. *Dissertation Abstracts International* – A 61 (2), 463-464.

Lima, E.F. (2011). Language and nonlanguage factors affecting nonnative undergraduate students' reactions to ITAs. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronuncaitionin Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 43 – 55), Ames, IA: Iowa State University.

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society*, 31, 419–441.

Lippi-Green, R. (1997). *English with an accent*: *Language, ideology, and discrimination n the United States.* New York: Routledge.

Llanes, À. (2010). *Children and adults learning English in a study abroad context*. PhD Thesis. Universitat de Barcelona, Barcelona.

Llanes, À. (2011). The many faces of study abroad: an update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism*, 3, 189 – 215.

Llanes, À. (2012). The short- and long-term effects of a short study abroad experience: The case of children. *System*, 40, 179-190.

Llanes, À. & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37 (3), 353-365.

Llanes, À. & Muñoz, C. (2013). Age effects in a study abroad context: children and adults studying English abroad and at home. *Language Learning* 63(1), pp. 63-90.

Llurda, E. (2002). "Una propuesta de reflexión sobre la pronunciación en la clase de lengua extranjera". Cots, J.M. & Nussbaum, L. *Pensar lo dicho*, Lleida: Milenio.

Long, M. H. (1981). Input, interaction and second language acquisition. *Annals of the New York Academy of Sciences*, 379, 259–278.

Long, M.H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.

Lord, G. (2010). The combined effects of immersion and instruction on second language pronunciation. Foreign Language Annals, 43(3), 488-503.

MacKay, I.R.A., Flege, J.E. & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, 27, 157-183.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26, 381–400.

Major, R. C. (2001). *Foreign accent*. Wiley Online Library.

Marriott, H. (1995). The acquisition of politeness patterns by exchange students in Japan. In: B.F. Freed (Ed.). *Second language acquisition in a study abroad context* (pp. 197-224). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Martinsen, R. (2010). Short-term study abroad: prediction changes in oral skills. *The Foreign Language Annals*, 43 (3), 504-530.

McKay, S.L. (2002). *Teaching English as an International Language: Rethinking Goals and Approaches.* Oxford: Oxford University Press.

Menard, S. (2002). *Longitudinal research*. Thousand Oaks, Calif.: Sage.

Milton, I., & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITI Review of Applied Linguistics*, 107, 17–34.

Mompeán, J. A. (2004). Options and Criteria for the Choice of an English Pronunciation Model in Spain. In J. Anderson et al. (Eds.) *Linguistic Perspectives from the Classroom: Language Teaching in a Multicultural Europe*. Lugo: Universidad de Santiago de Compostela. pp. 243-259.

Mompeán, Jose A. (2008). Consumers' Preferences and the Choice of English Pronunciation Models. In Monroy, R. & Sánchez, A. (Eds). *25 Years of Applied Linguistics in Spain: Milestones and Challeges*. University of Murcia, Spain, 959-964.

Mora, J.C. (2008) Learning context effects on the acquisition of a second language phonology. In C. Pérez-Vidal, M. Juan-Garau and A. Bel (Eds), *A portrait of the young in the new multilingual Spain*. Clevendon: Multilingual Matters, pp. 241-263.

Mora, J. C. (under review) Cognitive skills and individual differences in L2 speech perception. In Wright, C., Piske, T., and Young-Scholten, M. (eds.) *Mind Matters in SLA*. Bristol: Multilingual Matters.

Mora, J.C. and Fullana, N. (2007). Production and perception of English i:/-/ɪ/ and /æ/-/ʌ/ in a formal setting: Investigating the effects of experience and starting age. In J. Trouvain, & W. J. Barry (Eds.),

Proceeding of the 16th International Congress of Phonetic Sciences (pp. 1613-1616). Dudweiler: Pirrot.

Mora. J.C., Valls-Ferrer, M. (2012). Oral fluency, accuracy and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46 (4), 610 – 641.

Moyer, A. (1999) Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, 21, 81-108.

Moyer, A. (2004). *Age, accent, and experience in second language acquisition*. Clevedon, UK: Multilingual Matters Ltd.

Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition, 2*0(2), 139-154.

Munro, M.J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal*, 20 (2), 38 – 51.

Munro, M.J. (2008). Foreign accent and speech intelligibility. In J.G. Hansen Edwards, & M.L.Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 193-218). Amsterdam: John Benjamins Publishing Company.

Munro, M.J. (2013). Intelligibility. In Chapelle, C. (Ed.) *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell.

Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73–97.

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38 (3), 289-306.

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, 48, 159–182.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49, 285–310.

Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, 23, 451–468.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: an exploratory study. *System*, 34, 520-531.

Munro, M. J., & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, 44 (3), 316-327.

Munro, M. J., Derwing, T. M., & Burgess, C. S. (2003). The detection of foreign accent in backwards speech. In M-J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 535–538). Barcelona, Spain: Universitat Autònoma de Barcelona.

Munro, M.J., Derwing, T.M., & Holtby, A.K. (2012). Evaluating individual variability in foreign accent comprehension. In J. Levis & K. LeVelle (Eds.). Proceedings of the 3rd Pronunciation in *Second Language Learning and Teaching Conference*¸ Sept. 2011. (pp. 233-239). Ames, IA: Iowa State University.

Munro, M.J., Derwing, T.M, & Morton, S.L. (2006) The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28 (1), 111-131.

Munro. M.J., Flege J.E. & Mackay, I.R.A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics*, 17, 313-334.

Neufeld, G. G. (1979). Towards a theory of language learning ability. *Language Learning*, 29, 227-241.

Neufeld, G. G. (1980). On the adult's ability to acquire phonology. *TESOL Quarterly*, 14, 285-298.

Neufeld, G. G. (1988). Phonological asymmetry in second-language learning and performance. *Language learning*, 38, 531-559.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psychological Research* 5, 261-285.

Pallant, J. (2005). *SPSS Survival Manual*. Philadelphia: Open University Press.

Patkowski, M. (1990). Age and accent in a second language: a reply to James Emil Flege. *Applied Linguistics*, 11, 73-89.

Patkowski, M. (1994). The critical age hypothesis and interlanguage phonology. In M. Yavas (Ed.), *First and second language phonology* (pp. 205-221). San Diego: Singular.

Pavón, V. (2000). *La enseñanza de la pronunciación del inglés*. Granada: Granada Lingvistica.

Pavón, V. (2001). El Papel del Profesor en la Enseñanza de la Pronunciación. In Levey, D., Losey, M.A. & González, M.A. (Eds.). *English Language Teaching Changing Perspectives in Context* (pp. 289-300). Cádiz: Universidad de Cádiz (Servicio de Publicaciones).

Pavón, V. & Rosado, A. 2003. *Guía de Fonética y Fonología para Estudiantes de Filología Inglesa en el Umbral del Siglo XXI*. Granada: Comares.

Pennington, M.C., & Richards, J.C. (1986). Pronunciation revisited. *TESOL Quarterly*, 20 (2), 207-225.

Pennington, M.C., (1996). *Phonology in English language teaching.* London: Longman.

Pérez-Vidal, C. 2011. Language acquisition in three different contexts of learning: Formal instruction, stay abroad, and semi-immersion (CLIL).” In Y. Ruiz de Zarobe, J. M. Sierra and F. Gallardo del Puerto (Eds.). *Content and Foreign Language Integrated Learning. Contributions to Multilingualism in European Contexts* (pp. 103-127). Bern: Peter Lang.

Pérez-Vidal, C. (forthcoming). *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

Pérez-Vidal, C. & Barquin, E. (forthcoming). Comparing progress in academic writing after formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad*

and Formal Instruction Contexts. Amsterdam: John Benjamins Publishing Company.

Pérez-Vidal, C. & M. Juan-Garau (2009). The effect of study abroad on written performance. *Eurosla Yearbook* 9, 269-295.

Pérez-Vidal, C. & Juan-Garau, M. (2011). The effect of context and input conditions on oral and written development: A Study Abroad perspective. *International Review of Applied Linguistics in Language Teaching*. 49, 2, 157–185.

Pérez-Vidal, C., Juan-Garau, M., Mora, J.C. & Valls-Ferrer, M. (2012). Oral and written development in formal instruction and study abroad: differential effects of learning context. In C. Muñoz (Ed.), *Intensive exposure experiences in second language learning* (pp.213-233). Multilingual Matters.

Piske, T. (2007) Implications of James E. Flege's research for the foreign language classroom. In Bohn, O-S., Munro, M.J. (eds) *Language Experience in Second Language Learning* (pp. 301-314). Amsterdam: John Benjamins Publishing Company.

Piske, T., MacKay, I.R.A. & Flege, E.F. (2001) Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29 (2), 191-215.

Porter, D. & Garvin, S. (1989). "Attitudes to pronunciation in EFL". *Speak Out*!, 5, 8-15.

Purcell, E. T.& Suter, R. W. (1980) Predictors of pronunciation accuracy: a reexamination. *Language Learning*, 30 (2), 271-287.

Rallo , L. & Juan-Garau, M. (2011). Assessing FL pronunciation in a semi-immersion setting: the effects of CLIL instruction on Spanish-Catalan learners' perceived comprehensibility and accentedness. *Poznań Studies in Contemporary Linguistics PSiCL*, 47, 96 – 108.

Rees, J., & Kappler, J. (2008). Issues in the quantitative longitudinal measurement of second language progress in the study abroad context. In L. Ortega & H.Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 89-105). London: Routledge.

Regan, V. (1995). The acquisition of sociolingistic native speech norms: effects of a year abroad on second language learners of French. In:

B.F. Freed (Ed.). *Second language acquisition in a study abroad context* (pp. 245-267). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Regan, V. (2005). From speech community back to classroom: what variation analysis can tell us about the role of context in the acquisition of French as a foreign language. In: Dewaele, J.M. (Ed.), *Focus on French as a Foreign Language* (pp. 191-209). Clevedon: Multilingual Matters.

Rivers, W. R. (1998). Is being there enough? The effects of homestay placements on language gain during study abroad. *Foreign Language Annals*, 31, 492–499.

Roach, P. (1991). *English Phonetics and Phonology*. 2nd ed. Cambridge: Cambridge University Press.

Rossiter, M. (2009). Perceptions of L2 fluency by native and non-native speaker of English. *The Canadian Modern Language Review*, 65 (3), 395-412.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English speaking teaching assistants. *Research in Higher Education*, 33, 511–531.

Ruiz de Zarobe, Y. & Lasagabaster, D. (2010). CLIL in a bilingual community: The Basque autonomous community. In D. Lasagabaster and Y. Ruiz de Zarobe (eds): *CLIL in Spain: Implementation, results and teacher training*. Cambridge Scholars Publishing, pp. 12-29.

San Isidro, X. (2010). An insight into Galician CLIL: Provision and results. In D. Lasagabaster and Y. Ruiz de Zarobe (eds): *CLIL in Spain: Implementation, results and teacher training*. Cambridge Scholars Publishing, pp. 55-78.

Sanz , C. (forthcoming). Contributions of study abroad research to our understanding of SLA processes and outcomes: The SALA project, an appraisal. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, 54(3), 525-582.

Sasaki, M. (2007). Effects of study-abroad experiences on EFL writers: A multiple data analysis. Modern Language Journal, 91(4), 602-620.

Sasaki, M. (2009). Changes in English as a foreign language students' writing over 3.5 years: A sociocognitive account. In R. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 49–76). Clevedon, UK: Multilingual Matters.

Sasaki, M. (2011). Effects of Varying Lengths of Study-Abroad Experiences on Japanese EFL Students' L2 Writing Ability and Motivation: A Longitudinal Study. *TESOL Quarterly, 45*(1), 81-105.

Schmidt-Rinehart, B.C. and Knight, S.M. (2004). The homestay component of study abroad: Three perspectives. *Foreign Language Annals,* 37 (2), 254-262.

Schumann, J.H. (1976). Social distance as a factor in second language acquisition. *Language Learning*, 26 (1), 135-143.

Schumann, J.H. (1997). *The Neurobiology of Affect in Language*. Oxford: Blackwell.

Scovel, T. (1969). Foreign accents, language acquisition and cerebral dominance. *Language Learning*, 19, 245-253.

Scovel, T. (1988). A time to speak: A psycholinguistic inquiry into the critical period for human speech. Cambridge, MA: Newbury.

Segalowitz, N. (2010). *Cognitive Basis of Second Language Fluency*. New York: Routledge.

Segalowitz, N., & Freed, B. (2004) Context, contact, and cognition in oral fluency in acquisition. Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26 (2), 173-199.

Segalowitz, N., Freed, B., Collentine, J., Lafford, B., Lazar, N. & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning contexts: study abroad and the domestic classroom. *Frontiers*,10, pp. 1-18.

Seliger, H., Krashen, S. and Ladefoged, P. (1975). Maturational constraints in the acquisition of second language accent. *Language Sciences*, 36, 20-22.

Serrano, R.; Llanes, À.; Tragant, E. (2011). Analyzing the Effect of Context of Second Language Learning: Domestic Intensive and Semi-Intensive Courses vs. Study Abroad in Europe, *System*, 39:2, p. 133-143

Serrano, R., Tragant, E., Llanes, À. (2012) A longitudinal analysis of the effects of one year abroad. *The Canadian Modern Language Review*, 68 (2), 138-163.

Siegal, M. (1995). Individual differences and study abroad: Women learning Japanese in Japan. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 225-244). Amsterdam: John Benjamins Publishing Company.

Simões, A. (1996) Phonetics in second language acquistion: An acoustic study of fluency in adult learners of Spanish. *Hispania*, 79, 87-95.

Smith, U., & C.Dalton. (2000). "Motivational patterns in advanced EFL pronunciation learners". *International Review of Applied Linguistics,* 38: 229-46.

Suter, R.W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26, 233-253.

Swain, M. (1985). Bilingual education for the English-speaking Canadian. In J.Atlantis & J. Staczek (Eds.), *Perspectives on Bilingualism and Bilingual Education* (pp. 385–398). Washington, DC: Georgetown University Press.

Thompson, I. (1991) Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning*, 41, 177-204.

Thomson, R.I. (2013) Accent reduction. In C.A. Chapell (Ed.), *The encyclopedia of applied linguistics*. Hoboken, NJ: Wiley-Blackwell.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics,* 17(1), 84-119.

Trenchs-Parera (2009). Effects of formal instruction and a stay abroad on the acquisition of native-like oral fluency. *The Canadian Modern Language Review*, 65, 3, pp. 365-393.

Trenchs-Parera, M. & Juan-Garau, M. (forthcoming) A Longitudinal study of learners' motivation and beliefs in at home and study abroad contexts. In C. Pérez-Vidal (Ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts*. Amsterdam: John Benjamins Publishing Company.

Trofimovich, P. & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1-30.

Trofimovich, P. & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15, 905-916.

Turner, C.E., & Upshur, J.A. (2002). Rating scale derived from student samples: Effects of the scale marker and the student sample on scale content and student scores. *TESOL Quarterly*, 36, 49-70.

Twombly, S. B. (1995). Piropos and friendships: Gender and culture clash in study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 1.

Tyler, A. (1992). Discourse structure and the perception of incoherence in international teaching assistants' spoken discourse. *TESOL Quarterly*, 26, 713-729.

Valls-Ferrer, M. (2011). *The development of oral fluency and rhythm during a study abroad period*. PhD Thesis. Universitat Pompeu Fabra.

Varonis, E., & Gass, S. (1982). The comprehensibility of nonnative speech. *Studies in Second Language Acquisition*, 4, 114-136.

Werker, J., & Polka, 1. (1993). Developmental changes in speech perception: New challenges and new directions. *Journal of Phonetics*, 21,83-101.

Wilkinson, S. (1998). On the nature of immersion during study abroad: Some participant perspectives. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 4.

Wingstedt, M. & Schulman, R. (1984). Comprehension of foreign accents. In W. Dressler (Ed), *Phonologica 1984: Proceedings of the 5th*

international phonology meeting (pp. 339-344). Cambridge: Cambridge University Press.

Wrembel, M. (2010). Third language phonological acquisition: An investigation into metalinguistic awareness and foreign language accent in L3. *Paper presented at the 20th International Conference of the European Second Language Association – Eurosla*, University of Modena and Reggio Emilia, Italy.

Yager, K (1998). Learning Spanish in Mexico: The effect of informal contact and student attitudes on language gain. *Hispania*, 81, 4, 898 – 913.

Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69-84.

Appendices

Appendix 1: Linguistic profile questionnaire (SA and AH participants)

Appendix 2: Online questionnaire for SA participants
(to be completed after their SA period)

Appendix 3: Teaching/learning practice in their English lessons at school (Questionnaire for SA and AH participants)

Appendix 4: Online questionnaire for EFL teachers in Spain

Appendix 5: Oral narrative (picture story): *A bank robbery*

Appendix 6: Listeners' experiment on Moodle

Appendix 7: Instructions for listeners

Appendix 8: Summarizing your listening / rating experience (to be completed after the rating experiment)

Appendix 9: Examples of listeners' comments about comprehensibility

Appendix 1: Linguistic profile questionnaire (SA and AH participants)

SALA Project
UNIVERSITAT POMPEU FABRA

Universitat de les Illes Balears

CUESTIONARIO DE PERFIL LINGÜÍSTICO

El objetivo de este cuestionario es conocer las lenguas que hablas y cómo has aprendido inglés:

Nombre y apellidos:	Curso:
Fecha de nacimiento:	Lugar de nacimiento:

1) ¿Desde qué año vives en Valencia? .....

2) ¿Qué lengua consideras tu lengua materna?
☐ castellano ☐ valenciano ☐ otras (especifícalas).....

3) Lengua en la que habitualmente te habla tu:
padre/tutor: ☐ castellano ☐ valenciano ☐ otras: .....
madre/tutora: ☐ castellano ☐ valenciano ☐ otras: .....

4) ¿Qué lengua/s hablas...?

en casa con tu madre/tutora:	☐ castellano	☐ valenciano	☐ otras: .....
en casa con tu padre/tutor	☐ castellano	☐ valenciano	☐ otras: .....
con tus hermanos/as:	☐ castellano	☐ valenciano	☐ otras: .....
con los amigos/as:	☐ castellano	☐ valenciano	☐ otras: .....
en el colegio:	☐ castellano	☐ valenciano	☐inglés ☐otras: .....

5) ¿Estudias alguna lengua extranjera, sin contar el inglés, que no hables en casa?
☐ Sí ☐ No En caso afirmativo, ¿cuál/es?..... ....

6) ¿Has estado en algun país de habla inglesa? ☐ Sí ☐ No
En caso afirmativo,
a. ¿Dónde? .....
b. ¿En qué fechas / año? .....
c. Tiempo total pasado en países de lengua inglesa:
☐ 0 a 15 días ☐ 15 días a 1 mes
☐ de 1 mes a 6 meses ☐ más de 6 meses

d. El motivo de esta estancia era:
☐ aprender la lengua inglesa ☐ de vacaciones ☐ otros: .....

7) ¿Por qué te has decidido a hacer una estancia en un país de habla inglesa el primer trimestre del curso que viene? ¿Dónde vas a ir?

.....

.....

SALA Project
UNIVERSITAT POMPEU FABRA

Universitat de les Illes Balears

8) ¿Es la lengua inglesa la primera lengua extranjera que has aprendido en el colegio?
☐ Sí
☐ No, he aprendido..... como primera lengua extranjera en el colegio.

9) ¿En qué curso empezaste a estudiar inglés?
Educación Infantil: ☐ 3 años ☐ 4 años ☐ 5 años
Educación Primaria: ☐ 1o ☐ 2o ☐ 3o ☐ 4o ☐ 5o

10) ¿Has estudiado alguna vez alguna asignatura de contenido (Ciancias Sociales / Educación Física / Tecnología, etc.) en inglés? ☐ Sí ☐ No En caso afirmativo, ¿cuál/es? .....

11) El curso pasado, ¿qué nota sacaste en…?
castellano..... valenciano.....
inglés..... asignatura de contenido impartida en inglés (si hiciste) .....

12) ¿Has estudiado alguna vez inglés fuera del colegio (academias de idiomas, clases particulares, cursos de verano, etc.)? ☐ Sí ☐No
En caso afirmativo:
a. Años de estudio en academias o clases particulares de inglés: ☐ 1 ☐ 2 ☐ 3 ☐4 ☐5 o más
b. Número de cursos intensivos de verano de inglés realizados: ☐ 1 ☐ 2 ☐ 3 ☐4 ☐5 o más
c. Durada total de los cursos de verano de inglés realizados:
☐ 15 días ☐ 1 mes ☐ 2 meses ☐ 3 meses o más

13) ¿Este curso, has hecho algún curso de inglés fuera del colegio? ☐ Sí ☐ No
En caso afirmativo, ¿cuántas horas semanales de clase de inglés tenías fuera del colegio?
☐ 1 ☐ 2 ☐ 3 ☐4 ☐5 o más

14) Otras actividades: (señala la casilla que corresponda):

	Siempre / casi siempre	**A veces**	**Casi nunca**	**Nunca**
¿Ves películas en inglés?				
¿Escuchas música en inglés?				
¿Lees en inglés?				
¿Consultas páginas web en inglés?				
¿Te comunicas por Internet con alguien en inglés?				
¿Hablas con alguien en inglés?				

¡Gracias por tu colaboración!

.

Appendix 2: Online questionnaire for SA participants (to be completed after SA)

Cuestionario a alumnos de ESO que han realizado una estancia en el extranjero durante el curso académico.

Responde a las siguientes preguntas relacionadas con la estancia en el extranjero que realizaste.

* Necessari

I. INFORMACIÓN PERSONAL

1. Nombre y apellidos: *

2. Sexo: *
Seleccione una opción

3. Fecha de nacimiento (dd/mm/aaaa): *

4. Lugar de nacimiento: *
Indica ciudad/pueblo y país.

5. Lengua en la que hablas habitualmente con tu madre: *

6. Lengua en la que hablas habitualmente con tu padre: *

7. Lengua en la que hablas habitualmente con tus hermanos/as *
Si no tienes hermanos, indícalo.

II. INFORMACIÓN GENERAL SOBRE LA ESTANCIA

8. Fecha aproximada de la estancia: *
Indica el período de tiempo aproximado en el que realizaste la estancia (dd/mm/aaaa-dd/mm/aaaa).

9. Curso escolar que cursabas en el momento de realizar la estancia: *
(por ejemplo, 2º de ESO).

10. Lugar de la estancia: *
Indica el nombre del colegio y la localidad (y el país) donde el colegio se encontraba.

11. Alojamiento durante la estancia:
Escoge la opción que mejor describa el tipo de alojamiento durante tu estancia.

- ○ Durante mi estancia, me alojé en la residencia del colegio.
- ○ Durante mi estancia, me alojé en casa de una familia nativa (o que hablaba en inglés), y con la que yo hablaba en inglés.
- ○ Durante mi estancia, me alojé en casa de una familia donde hablaba español.

12. Si ninguna de las respuestas anteriores describe tu situación de alojamiento, explica tu situación a continuación.

13. Si te alojaste en residencia, escoge la opción que mejor describa tu situación durante la estancia.

- Hablaba siempre/casi siempre en inglés con alumnos nativos.
- Hablaba siempre/casi siempre en inglés con alumnos nativos y/o de otras nacionalidades.
- Hablaba siempre/casi siempre en inglés con alumnos de otras nacionalidades.

14. Si a lo largo de tu estancia hablabas con alumnos de otras nacionalidades, indica de dónde eran.

III. APRENDIZAJE DURANTE LA ESTANCIA

15. Escoge la opción que mejor describa el tipo de aprendizaje que seguiste durante tu estancia en el extranjero: *

- Asistía a clases para alumnos extranjeros como yo. Las clases eran en inglés.
- Asistía a las clases de los alumnos nativos que van a ese colegio.
- Una combinación de las dos opciones anteriores, según la asignatura.

16. Describe brevemente las asignaturas que cursaste durante la estancia. Inidca en qué lengua se impartía cada asignatura y el número de horas semanales de clase. *

IV. LENGUAS DE COMUNICACIÓN FUERA DEL COLEGIO

En esta sección queremos que nos informes sobre tu contacto con el inglés y otras lenguas fuera del horario escolar.

17. ¿Realizabas alguna actividad extraescolar? Si es así, explícanos qué actividad(-es) hacías, cuánto tiempo dedicabas semanalmente, y en qué idioma te comunicabas. *
Si no realizabas ninguna actividad extraescolar, responde "no" a esta pregunta.

18. ¿Utilizabas el inglés como lengua de comunicación fuera del colegio? ¿En qué situaciones? ¿Con quién? *
Si no utilizabas el inglés fuera del colegio, responde "no" a esta pregunta.

19. ¿Utilizabas el inglés para hablar por teléfono, correo electrónico, sms, redes sociales, etc? *
Explícanos brevemente cuándo utilizabas el inglés en medios virtuales. Si no utilizabas el inglés en estos casos, responde "no" a esta pregunta.

20. Por el contrario, es probable que durante la estancia te comunicaras en otra lengua distinta al inglés. Si fue así, explica brevemente las situaciones en las que te comunicabas en otra lengua diferente al inglés (con quién, en qué lengua, etc). *
(por ejemplo: "utilizaba el español para hablar por Skype con mi familia en España cada día; hablaba en español con otros alumnos españoles que conocí, etc."). Si sólo utilizaste el inglés, responde "sólo utilizaba el inglés".

21. En resumen, si piensas en un día normal durante tu estancia, ¿podrías indicar qué porcentaje aproximado del tiempo utlizabas el inglés y qué porcentaje aproximado del tiempo te comunicabas en tu lengua materna u otra? *
Para hacer tu valoración, piensa en todas las actividades que realizabas normalmente durante un día (colegio incluído).

V. VALORACIÓN DE TU APRENDIZAJE DEL INGLÉS DESPUÉS DE LA ESTANCIA

22. Después de tu estancia en el extranjero, ¿cuánto crees que has mejorado en los siguientes aspectos de tu inglés? *
Escoge una de las cuatro opciones para cada aspecto.

	Nada	Poco	Bastante	Mucho
Reading	○	○	○	○
Listening	○	○	○	○
Writing	○	○	○	○
Speaking	○	○	○	○
Grammar	○	○	○	○
Vocabulary	○	○	○	○
Pronunciation	○	○	○	○

23. Después de tu estancia en el extranjero, ¿cuánto crees que has mejorado en los siguientes aspectos de tu "speaking"? *

Escoge una de las cuatro opciones para cada aspecto.

	Nada	Poco	Bastante	Mucho
Pronunciación en inglés.	○	○	○	○
Comprensión de mi discurso (es decir, que entiendan cuando hablo).	○	○	○	○
Fluidez (es decir, ritmo/velocidad cuando hablo).	○	○	○	○
Corrección (gramática, vocabulario).	○	○	○	○

24. Desde que regresaste de tu estancia en el extranjero hasta ahora, ¿cuánto crees que has mejorado en los siguientes aspectos de tu inglés? *

Escoge una de las cuatro opciones para cada aspecto.

	Nada	Poco	Bastante	Mucho
Reading	○	○	○	○
Listening	○	○	○	○
Writing	○	○	○	○
Speaking	○	○	○	○
Grammar	○	○	○	○
Vocabulary	○	○	○	○
Pronunciation	○	○	○	○

25. Desde que regresaste de tu estancia en el extranjero hasta ahora, ¿cuánto crees que has mejorado en los siguientes aspectos de tu "speaking"? *

Escoge una de las cuatro opciones para cada aspecto.

	Nada	Poco	Bastante	Mucho
Pronunciación en inglés.	○	○	○	○
Comprensión de mi discurso (es decir, que entiendan cuando hablo).	○	○	○	○
Fluidez (es decir, ritmo/velocidad cuando hablo).	○	○	○	○
Corrección (gramática, vocabulario).	○	○	○	○

VI. PREPARACIÓN, SEGUIMIENTO Y VALORACIÓN DE LA ESTANCIA

26. ¿Recibiste algún tipo de preparación en el colegio antes de irte a tu estancia? *

Escoge una opción

Si recibiste algún tipo de preparación, explícanos brevemente en qué consistió.

27. Durante la estancia en el extranjero durante parte del curso académico, ¿tuviste algún tipo de acompañamiento/seguimiento desde tu colegio en España? *

(por ejemplo, contacto con vuestro profesor de inglés en España, con el director del colegio, etc).

Escoge una opción

Si tuviste algún tipo de acompañamiento/seguimiento, explícanos brevemente en qué consistió.

28. Después de la estancia en el extranjero durante parte del curso académico, ¿hablaste sobre esta experiencia con alguien en tu colegio en España? *

(por ejemplo, valoraste la experiencia con tu profesor de inglés en España, con el director del colegio, etc).

Escoge una opción

Si valoraste con alguien tu estancia después de regresar, explícanos brevemente en qué consistió.

29. En cuanto a tu inglés, ¿qué has hecho para conservar todo lo que mejoraste durante tu estancia en el extranjero? *

MUCHAS GRACIAS POR TU COLABORACIÓN

Envia

Appendix 3: Teaching/Learning practice in their English lessons AH (Questionnaire for SA and AH participants)

Nombre y apellidos: ……..…… **Curso:** ……...

CUESTIONARIO SOBRE LA ENSEÑANZA-APRENDIZAJE DEL INGLÉS LE EN EL AULA DE SECUNDARIA

1. ¿Te gusta aprender inglés?
   - ❑ Sí. ¿Por qué? .....
   - ❑ No. ¿Por qué? .....

2. ¿Crees que es importante aprender inglés?
   - ❑ Sí. ¿Por qué? .....
   - ❑ No. ¿Por qué? .....

3. A la hora de aprender inglés, crees que trabajar los siguientes aspectos es...

	Nada Importante	Poco importante	Importante	Muy importante
Comprensión escrita (reading)				
Comprensión oral (listening)				
Producción escrita (writing)				
Producción oral (speaking)				
Gramática (grammar)				
Vocabulario (vocabulary)				
Pronunciación (pronunciation)				

4. ¿Qué aspectos de tu inglés crees que deberías / te gustaría mejorar? (Indica los 3 que creas más urgentes)
   - ❑ Comprensión escrita (reading)
   - ❑ Comprensión oral (listening)
   - ❑ Producción escrita (writing)
   - ❑ Producción oral (speaking)
   - ❑ Gramática (grammar)
   - ❑ Vocabulario (vocabulary)
   - ❑ Pronunciación (pronunciation)

5. ¿Qué importancia le da tu profesor/a de inglés a los siguientes aspectos dentro de las clases?

	Nada importante	Poco importante	Importante	Muy importante
Comprensión escrita (reading)				
Comprensión oral (listening)				
Producción escrita (writing)				
Producción oral (speaking)				
Gramática (grammar)				
Vocabulario (vocabulary)				
Pronunciación (pronunciation)				

6. ¿A lo largo del trimestre qué tiempo se dedica a los siguientes aspectos?

	Cada sesión o casi todas	A veces	Casi nunca	Nunca
Comprensión escrita				
Comprensión oral				
Expresión escrita				
Expresión oral				
Gramática				
Vocabulario				
Pronunciación				

7. ¿Habla en inglés el profesor durante las clases? (señala una de las siguientes opciones)

- ❑ El profesor/a siempre habla en inglés con los alumnos
- ❑ El profesor/a casi siempre habla en inglés con los alumnos
- ❑ El profesor/a casi nunca habla en inglés con los alumnos
- ❑ El profesor/a nunca habla en inglés con los alumnos

8. Mi interacción con el profesor/a: (señala una de las siguientes opciones)

- ❑ En clase siempre/casi siempre hablo en inglés con el profesor/a
- ❑ En clase a veces hablo en inglés con el profesor/a
- ❑ En clase alguna vez hablo en inglés con el profesor/a
- ❑ En clase no hablo en inglés con el profesor/a. Me dirijo a él/ella en castellano/catalán

9. Mi interacción con mis compañeros/as: (señala una de las siguientes opciones)

- ❑ En clase siempre/casi siempre hablo en inglés con el resto de compañeros/as
- ❑ En clase a veces hablo en inglés con el resto de compañeros/as
- ❑ En clase alguna vez hablo en inglés con el resto de compañeros/as
- ❑ En clase no hablo en inglés con el resto de compañeros/as

10. Cuando se trabaja la pronunciación en tus clases... (señala una de las siguientes opciones)

- ❑ dedicamos sesiones completas a practicar algún aspecto de pronunciación
- ❑ dedicamos un rato de la sesión trabajando algún ejercicio del libro de texto
- ❑ dedicamos un rato de la sesión trabajando algún ejercicio que trae nuestro profesor/a
- ❑ no hacemos ejercicios concretos. El profesor/a corrige nuestra pronunciación cuando nos equivocamos
- ❑ no trabajamos la pronunciación

11. ¿Te gusta hablar en inglés? (señala una de las siguientes opciones)

- ❑ Me gusta hablar en inglés. Intento aprovechar toda oportunidad para practicarlo
- ❑ No me importa tener que expresarme en inglés
- ❑ Hablar en inglés me da vergüenza. Cuando hablo inglés me siento ridículo/a, es como si no fuera yo
- ❑ No me gusta hablar en inglés. Motivos: .....

12. ¿Qué importancia le das a la buena pronunciación? (señala una de las siguientes opciones)

- ❑ Es importante hablar inglés con "buena pronunciación" para que nos entiendan
- ❑ No hace falta hablar inglés con "buena pronunciación" para que nos entiendan

13. ¿Intentas imitar la pronunciación nativa? (señala una de las siguientes opciones)

- ❑ Intento imitar el modelo de pronunciación inglesa cuando hablo
- ❑ Cuando hablo en inglés lo hago sin imitar el acento / pronunciación inglesa

14. ¿Corrige tu profesor/a la pronunciación de los alumnos/as en clase? (señala una de las siguientes opciones)

- ❑ El profesor siempre corrige los errores de pronunciación
- ❑ El profesor casi siempre corrige los errores de pronunciación

- ❑ El profesor casi nunca corrige los errores de pronunciación
- ❑ El profesor nunca corrige los errores de pronunciación

15. ¿Te gusta que te corrijan la pronunciación? (señala una de las siguientes opciones)

- ❑ Me gusta que me corrijan si cometo errores en mi pronunciación
- ❑ No me importa si no me corrigen en pronunciación
- ❑ No me gusta que me corrijan si cometo errores en mi pronunciación

16. ¿Crees que puedes mejorar tu pronunciación en inglés? (señala una de las siguientes opciones)

- ❑ Con práctica, estoy seguro de que mejoraré mi pronunciación en inglés
- ❑ No creo que llegue a tener una "buena pronunciación" aunque practique mucho

17. ¿Crees que necesitas mejorar tu pronunciación? (señala una de las siguientes opciones)

- ❑ Creo que mi pronunciación en inglés es lo suficientemente clara. Me entenderán si hablo en inglés.
- ❑ Creo que he de trabajar mucho la pronunciación. No creo que me entiendan si hablo en inglés.
- ❑ Me cuesta la pronunciación en inglés, pero no creo que haga falta trabajarla. Llegaré a entenderme con la gente si hablo en inglés, aunque mi pronunciación no sea del todo correcta.

18. ¿Crees que es malo que se note tu acento español/catalán cuando hablas en inglés? (señala una de las siguientes opciones)

- ❑ Creo que es no es malo que se note mi acento español/catalán si hablo en inglés. No importa si reconocen de qué país procedo.
- ❑ Creo que si hablo en inglés es mejor que no se note mi acento español/catalán.
- ❑ Me da igual si se nota o no mi acento español/catalán cuando hablo en inglés.

GRACIAS POR TU COLABORACIÓN

Appendix 4: Online questionnaire for EFL teachers in Spain

QUESTIONNAIRE FOR TEACHERS OF ENGLISH AS A FOREIGN LANGUAGE (EFL) IN SPAIN

* Necessari

Please enter your identification code. Your identification code will allow us to identify you and all your questionnaire information without using your real name. Your ID consists of the first two letters of your name and surname.
Ex. Julia Ascaso (JUAS)

Please enter your email if you allow us to contact you for further information after completing this questionnaire.

I. PARTICIPANT INFORMATION

Gender *
- ○ Male
- ○ Female

Age *

Mother tongue (-s) *

School location *

Levels that you teach *

Years of English as a Foreign Language (EFL) teaching experience *

Please list your qualifications *
(e.g. BA, DipTEFL, MA TESOL, PhD, etc.)

II. LINGUISTIC PROFILE

1. How would you rate your knowledge of phonetics/phonology in your mother tongue? Please rate from 1 to 5 (1= "extremely poor", and 5 = "excellent"). *

	1	2	3	4	5	
extremely poor	○	○	○	○	○	excellent

2. How would you rate your knowledge of phonetics/phonology in English? Please rate from 1 to 5 (1= "extremely poor", and 5 = "excellent"). *

	1	2	3	4	5	
extremely poor	○	○	○	○	○	excellent

3. How would you rate your own pronunciation in English? Please rate from 1 to 5 (1= "extremely poor", and 5 = "excellent"). *

	1	2	3	4	5	
extremely poor	○	○	○	○	○	excellent

If you think your own pronunciation is not excellent (=5), please give reasons for your answer:
This can be done using bullet points.

4. How familiar are you with the following varieties of English? (i.e. you can understand in both written and spoken discourse) *

NOTE: "English as an International Language" refers to the English language as a global means of communication, not related to any specific geographical variety.

	Familiar with	More or less familiar with	Not familiar with
British "RP"	○	○	○
General American English	○	○	○
English as an International Language	○	○	○
Canadian English	○	○	○
Irish Englisn	○	○	○
Scottish English	○	○	○
Welsh English	○	○	○
Australian English	○	○	○
New Zealand English	○	○	○
South African English	○	○	○

If you are familiar with another variety of English, please tell us which one:

5. Which variety of English do you usually speak / write? *

Select one

If you usually speak/ write using another variety of English, please tell us which one:

6. Are you proud of your English accent? *

○ Yes
○ Somewhat
○ No

Please give reasons for your answer: *

This can be done using bullet points.

7. How good is your level of English? Please rate your language abilities and other aspects of your competence in English. *

	poor	acceptable	good	excellent
Reading	○	○	○	○
Listening	○	○	○	○
Writing	○	○	○	○
Speaking	○	○	○	○
Grammar	○	○	○	○
Vocabulary	○	○	○	○
Pronunciation of individual sounds and words	○	○	○	○
Word stress	○	○	○	○
Rhythm	○	○	○	○
Intonation	○	○	○	○
Cultural information	○	○	○	○

III. TEACHER TRAINING

8. Have you ever had any specific training in English pronunciation? *

○ Yes
○ No

9. Have you ever had any specific training in how to teach English pronunciation? *

○ Yes
○ No

10. Which specific training in EFL teaching would you like to receive? *

IV. VIEWS, ATTITUDES AND BELIEFS ABOUT EFL TEACHING

11. In your opinion, how important is to practise/work on the following aspects in order to have a good command of the English language? *

	not important at all	not so important	more or less important	important	very important
Reading	○	○	○	○	○
Listening	○	○	○	○	○
Writing	○	○	○	○	○
Speaking	○	○	○	○	○
Grammar	○	○	○	○	○
Vocabulary	○	○	○	○	○
Pronunciation	○	○	○	○	○
Cultural information	○	○	○	○	○

12. How do you feel teaching English pronunciation? *

Possible answers: confident, insecure, very comfortable, a bit frightened, really motivated, untrained, well-prepared, etc.

13. From your point of view, which pronunciation accent would be best for your learners? You may choose more than one option. *

NOTE: "English as an International Language" refers to the English language as a global means of communication, not related to any specific geographical variety.

- ☐ British "RP"
- ☐ General American English
- ☐ English as an International Language
- ☐ Canadian English
- ☐ Irish English
- ☐ Scottish English
- ☐ Welsh English
- ☐ Australian English
- ☐ New Zealand English
- ☐ South African English
- ☐ Other

If you have chosen "other" , please tell us which one:

14. In your opinion, what is "good English pronunciation"? *

15. Generally speaking, how important do you think having native-like (or near-native like) pronunciation in English is? Please rate from 1 to 5 (1= "not important at all", and 5="extremely important"). *

	1	2	3	4	5	
not important at all	◯	◯	◯	◯	◯	extremely important

16.Do you think your learners should acquire a native-like pronunciation? *

- ◯ Yes
- ◯ No
- ◯ Not sure

Please give reasons for your answer.

17. To what extent do you feel your students aspire to sound native-like or near native-like in English? Please rate from 1 to 5 (1= "do not aspire to this at all", and 5="aspire to this 100%"). *

Indicate how your students might answer this question.

	1	2	3	4	5	
do not aspire to this at all	◯	◯	◯	◯	◯	aspire to this 100%

18. Think about your students speaking in English. To what extent do the following aspects influence the comprehensibility of their oral productions? *

	not influencing at all	not influencing	somehow influencing	influencing	influencing a lot
Pronunciation of individual sounds (vowels and consonants)	○	○	○	○	○
Accent	○	○	○	○	○
Word stress	○	○	○	○	○
Rhythm	○	○	○	○	○
Intonation	○	○	○	○	○
Grammar	○	○	○	○	○
Vocabulary	○	○	○	○	○
Fluency	○	○	○	○	○
Your being an EFL teacher	○	○	○	○	○
Your familiarity with the student's mother tongue	○	○	○	○	○
Your knowledge of the context the student is referring to	○	○	○	○	○

19. For your students, how important is the practice and / or knowledge of the following linguistic areas? *

Indicate how your students might answer this question.

	Not important at all	Not important	More or less important	Important	Very important
Reading	○	○	○	○	○
Listening	○	○	○	○	○
Writing	○	○	○	○	○
Speaking	○	○	○	○	○
Grammar	○	○	○	○	○
Vocabulary	○	○	○	○	○
Pronunciation	○	○	○	○	○
Cultural aspects	○	○	○	○	○

V. YOUR OWN DAILY CLASSROOM PRACTICE

20. How much English do you speak during the lessons? *

Select one

21. How much English do your students speak during the lessons? *

Select one

22. Do your students speak to you in English? *

- ○ Yes, all of them
- ○ Yes, most of them
- ○ Some of them
- ○ Only a few
- ○ One or two in each group
- ○ No

23. Consider your teaching throughout one term. How often do your students practice the following aspects? *

	Never	Rarely	Sometimes	Frequently	Always
Reading	○	○	○	○	○
Listening	○	○	○	○	○
Writing	○	○	○	○	○
Speaking	○	○	○	○	○
Grammar	○	○	○	○	○
Vocabulary	○	○	○	○	○
Pronunciation	○	○	○	○	○
Cultural aspects	○	○	○	○	○

24. When you work on pronunciation with your students, what do you usually do? *

If you don't work pronunciation with your students, please write "nothing".

25. Do you teach your students how to RECOGNIZE phonetic symbols? *

Select one

26. Do you teach your students how to WRITE phonetic symbols? *

Select one

27. Regarding the course book you use with your students, does it have a specific pronunciation section in every unit? *

Select one

28. If there is a pronunciation section in every unit, do you work this section with your students?

- ◯ Yes, every lesson
- ◯ Yes, in some of the lessons
- ◯ Rarely
- ◯ No, we don't

29. How often do you use the following resources for pronunciation practice? *

	Never	Rarely	Sometimes	Frequently	Always
Other course books or materials	◯	◯	◯	◯	◯
Portable sound player, speakers, etc.	◯	◯	◯	◯	◯
Separate language lab	◯	◯	◯	◯	◯

30. Do you think the amount of time devoted to pronunciation in your lessons is sufficient? *

- ◯ Yes
- ◯ No
- ◯ Not sure

31. Is teaching English pronunciation an easy task? *

Please select one of the options below.

- ◯ Yes, it is easy.
- ◯ No, it is not easy.
- ◯ 50-50.

32. In your opinion, what makes the teaching/practise of pronunciation in your daily lessons difficult? *

If you can't think of any problems when teaching English pronunciation, please write "no difficulty". You can answer by using bullet points.

33. For RECEPTIVE work (listening, reading), which variety(-ies) of English do you use in your classes? *

You may choose more than one option below.

- ☐ British RP
- ☐ General American English
- ☐ English as an International Language
- ☐ Canadian English
- ☐ Irish English
- ☐ Scottish English
- ☐ Welsh English
- ☐ Australian English
- ☐ New Zealand English
- ☐ South African English
- ☐ Other

Please say which other variety(-ies):

34. For PRODUCTIVE work (speaking, writing), which variety(-ies) of English do you use in your classes? *

You may choose more than one option below.

- ☐ British RP
- ☐ General American English
- ☐ English as an International Language
- ☐ Canadian English
- ☐ Irish English
- ☐ Scottish English
- ☐ Welsh English
- ☐ Australian English
- ☐ New Zealand English
- ☐ South African English
- ☐ Other

Please say which other variety(-ies):

VI. EVALUATION OF PRONUNCIATION

35. Do you evaluate your learners' pronunciation at any point during the course? *
Please select all answers that apply.

- [x] At the beginning of the course.
- [ ] During the course.
- [ ] At the end of the course.
- [ ] I don't evaluate my learners' pronunciation.

36. Is evaluation linked to an established scale, for example, a national or international scale? *
e.g. CEF, IELTS, TOEFL, ACTEFL, etc.

- ○ Yes
- ○ No

If yes, please say which scale is used as a reference for your evaluations

37. Which types of task do you use for assessment of pronunciation? *
Please select all answers that apply.

- [ ] Written work, e.g. transcription into symbols or letters.
- [ ] Oral performances, e.g. dialogues, presentations, etc.
- [ ] Individual oral exams.
- [ ] Oral exams in pairs or groups.
- [ ] Listening and questions, e.g. multiple choice, short answer, note-taking, etc.
- [ ] Reading aloud (with or without preparation time).
- [ ] Other
- [ ] I don't evaluate my learners' pronunciation.

Please say which other type of task you use for the assessment of pronunciation.

38. To what extent do you provide immediate or delayed feedback on your learners' speaking regarding English pronunciation? *

- ○ Whenever I listen to any type of pronunciation error, I immediately tell my student.
- ○ I immediately correct the learner's pronunciation, if this is affecting communication.
- ○ I usually refer to pronunciation mistakes once they finish speaking.
- ○ I don't usually correct my students' pronunciation.

Further comments:

39. What type of corrections regarding pronunciation do you usually make? *

Please select all answers that apply.

- ☐ Segmental errors (e.g. pronunciation of vowels or consonants).
- ☐ Word stress
- ☐ Rhythm
- ☐ Intonation
- ☐ I don't usually correct my students' pronunciation errors.

40. From your point of view, when assessing your students' oral productions, how important is each of the following dimensions? *

	Not important at all	Not so important	More or less important	Important	Very important
Accent	○	○	○	○	○
Intelligibility / comprehensibility	○	○	○	○	○
Fluency	○	○	○	○	○

41. Please identify which aspects you bear in mind when assessing your students' oral productions regarding pronunciation. *

Please select all answers that apply.

- ☐ Pronunciation of sounds (e.g. vowels, consonants)
- ☐ Word stress
- ☐ Rhythm
- ☐ Intonation
- ☐ Accent
- ☐ Intelligibility / comprehensibility
- ☐ Fluency

VII. STUDY ABROAD PROGRAMMES

All teachers must complete question number 42.

The rest of the questions in this section (questions 43-46) are to be completed only by those teachers working at schools where one-term study abroad programmes are offered to their students.

42. In which aspects do you think students improve most during a study abroad programme? *

Please select all aspects that apply.

- ☐ Reading
- ☐ Writing
- ☐ Listening
- ☐ Speaking
- ☐ Grammar
- ☐ Vocabulary
- ☐ Pronunciation
- ☐ Cultural information

Further comments:

43. As for the students in your school who spend one academic term abroad, to what extent do they show any development in the following aspects?

	I don't know	No progress at all	Little progress	Some progress	A lot of progress
Reading	○	○	○	○	○
Writing	○	○	○	○	○
Listening	○	○	○	○	○
Speaking	○	○	○	○	○
Grammar	○	○	○	○	○
Vocabulary	○	○	○	○	○
Pronunciation	○	○	○	○	○
Cultural information	○	○	○	○	○

Further comments:

44. Does your school prepare the students for their term abroad?

- ◯ Yes
- ◯ No
- ◯ I don't know

If your school prepares the students for their term abroad, what does this preparation consist in?

45. Does your school monitor the students somehow DURING their term abroad?

- ◯ Yes
- ◯ No
- ◯ I don't know

If your school monitors the students during their term abroad, what does this monitoring consist in?

46. AFTER their term abroad, do the students provide the school with any feedback about their experience?

- ◯ Yes
- ◯ No
- ◯ I don't know

If your the students give some type of feedback after their term abroad, what does this feedback consist in?

End of the questionnaire - THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE.

Envia

Appendix 5: Oral narrative (picture story): *A bank robbery*

A BANK ROBBERY

You have witnessed a bank robbery. Report what happened to the police.

1

2

5

6

Appendix 6: Listeners' rating experiment on Moodle

WELCOME TO THIS EXPERIMENT!

We are very grateful for your contribution to our study dealing with the development of language competence by EFL secondary education learners. More specifically, our research focuses on the area of pronunciation in EFL learning in Spain, and how it is affected by different contexts of learning. This study is part of a PhD dissertation.

Your contribution is central to the nature of this research, as we approach the study of pronunciation from the point of view of listeners/judges, in this case, you. We want to know about your perception of EFL learners' accent and comprehensibility when they speak in English.

Moreover, we are interested in learning about the factors that you find most salient when listening to and scoring foreign language comprehensibility.

The estimated duration of the whole procedure is 3 hours. As you know, this is an online activity which lets you organize yourself at your convenience. Please read the instructions document that you will find in the Procedure section of this page.

Do not hesitate to contact me at carmen.delrio@upf.edu if you have any questions about the procedure.

THANK YOU FOR PARTICIPATING IN THIS STUDY

CONTEXT

Both EFL learners and native speakers of English were asked to describe a six-frame narrative story about a bank robbery. Click on the following pdf file to become familiar with the story of the narrative.

Oral narrative: A bank robbery

PROCEDURE

Click on the following pdf file and read the instructions for the rating and assessment tasks you will be asked to complete.

Fòrum de notícies

Instructions

PRACTICE

Now, let's try a few practice items... Remember the definitions below to guide your judgements:

ACCENTEDNESS: How different you think the speaker sounds from a native speaker of English (1 = *heavy foreign accent,* and 7 = *native-like accent).*

COMPREHENSIBILITY: How easy or difficult the speech sample is to understand (1 = *extremely difficult to understand,* and 7 = *extremely easy to understand).*

For every practice speech sample, **type in the aspects of speech that you found most striking and you took into account when rating comprehensibility**. T

? PRACTICE (4 samples)

Please, contact me if you have any questions after completing the practice samples. If not, you can start the rating task now . Click on PART 1 to start the experiment.

RATING AND ASSESSMENT EXPERIMENT

Now, let's start... Remember the definitions below to guide your judgements:

ACCENTEDNESS: How different you think the speaker sounds from a native speaker of English (1 = *heavy foreign accent,* and 7 = *native-like accent).*

COMPREHENSIBILITY: How easy or difficult the speech sample is to understand (1 = *extremely difficult to understand,* and 7 = *extremely easy to understand).*

Remember that **in each part you will find 1 or 2 speech samples** where you will be asked to **type in the aspects of speech that you found most striking and you took into account when rating comprehensibility**.

Try to use the whole scale over the course of the experiment.

- PART 1 (8 samples)
- PART 2 (9 samples)
- PART 3 (8 samples)
- PART 4 (9 samples)
- PART 5 (9 samples)
- PART 6 (8 samples)
- PART 7 (9 samples)
- PART 8 (9 samples)
- PART 9 (8 samples)
- PART 10 (9 samples)
- PART 11 (9 samples)
- PART 12 (8 samples)
- PART 13 (9 samples)
- PART 14 (9 samples)
- PART 15 (8 samples)

Please fill in the following short FINAL QUESTIONNAIRE. It will take you only a few minutes to complete it.

THIS IS THE END OF THE EXPERIMENT.

THANK YOU FOR PARTICIPATING IN OUR STUDY.

Appendix 7: Instructions for listeners

You will listen to 129 short speech samples from both foreign language learners and native speakers of English describing part of the six-frame narrative story *A bank robbery*. All of the speech samples are taken from the middle part and/or the end of the story.

We are interested in knowing to what extent learners' pronunciation sounds more or less native like, and to what extent it is more or less comprehensible. Moreover, we would like to know about the factors that you find most salient when listening to and scoring foreign language comprehensibility. For this purpose, we will ask you to type in your impressions regarding comprehensibility in 20 speech samples (out of the 129). You will find these samples scattered throughout the experiment.
Your task is as follows:

1. **RATING ACCENTEDNESS**

We are interested in knowing **how different you think the speaker sounds from a native speaker of English, if at all**. Please use the 7-point scale provided below each sample in order to assign your ratings for accentedness: 1 = *heavy foreign accent* and 7 = *native-like accent*.

2. **RATING COMPREHENSIBILITY**

We would like you to assess **how easy or difficult the sample is to understand**. Can you understand it without paying much extra attention, or do you have to work hard at it? Please use the 7-point scale provided below each sample in order to assign your ratings for comprehensibility: 1 = *extremely difficult to understand,* and 7 = *extremely easy to understand*.

3. **TYPING IN YOUR IMPRESSIONS FOR COMPREHENSIBILITY**

In 20 speech samples you will find a text box under the rating scale for comprehensibility. Please type in the aspects of speech that you find most striking and you take into account when rating this dimension. Feel free to use English, Spanish, and/or Catalan. This can be done using bullet points.

We are interested in the aspects of the speech that most influenced your rating. Please tell us about:

- any aspects related to the speaking style that were taken into account during your scoring decision for comprehensibility.
- any aspects that you noticed or were distracted by although you decided not to take them into account in your scoring.

- Also tell us if you experienced a dilemma when assigning a score. Please describe this and how you arrived at your rating decision.

At times you may feel that your thoughts about the speech are not relevant to our study. Rest assured that they probably are. Try to include all your thoughts during the rating process. Report what comes into your head as you reflect on the speech and the score you are assigning for comprehensibility.

4. FINAL QUESTIONNAIRE

After completing all the rating tasks, you will be asked to summarise your listening experience by answering a short questionnaire.

SOME FURTHER TIPS:

- It is your decision whether you rate accentedness or comprehensibility first. This will probably depend on which scoring decision you arrive at most quickly. Do the quickest one first.
- Please listen to the whole sample before making your decisions. You may listen to the speech sample as many times as needed.
- We recommend that you either **click on the "play" button** (not on the "Àudio MP3" label) OR **open the audio file by clicking on the right mouse button and selecting the "open link on a new tab" option**. This way, you'll be able to score and type in your impressions while you are listening to the recording.
- Try to use the whole scale (1-7) over the course of the experiment.
- There are four samples for practice.
- If you have any questions regarding the procedure, please contact me at carmen.delrio@upf.edu before starting the rating task.

PROCEDURE:

The 129 speech samples have been organised in 15 parts. **In each part, you will find one or two samples which require typing in your impressions after rating comprehensibility.**

As you know, this is an online exercise which lets you organize yourself at your convenience. There is one restriction, though. Once you start whichever part of the experiment, you must carry it out until the end. You can have a break or stop the experiment after finishing any part.

Here you have the list of sections in the experiment:

0. PRACTICE (4 samples)
1. PART 1 (8 samples)
2. PART 2 (9 samples)
3. PART 3 (8 samples)
4. PART 4 (9 samples)
5. PART 5 (9 samples)
6. PART 6 (8 samples)
7. PART 7 (9 samples)
8. PART 8 (9 samples)
9. PART 9 (8 samples)
10. PART 10 (9 samples)
11. PART 11 (9 samples)
12. PART 12 (8 samples)
13. PART 13 (9 samples)
14. PART 14 (9 samples)
15. PART 15 (8 samples)

FINAL QUESTIONNAIRE (to be completed after the rating task).

Thank you for participating in this study.

Appendix 8: Summarizing your listening/rating experience (to be completed after the rating experiment)

SUMMARIZING YOUR LISTENING / RATING EXPERIENCE

* Necessari

Please enter you identification code. Your identification code will allow us to identify you and all your questionnaire information without using your real name. You ID consists of the first two letters of you name and surname. *

Ex. Julia Ascaso (JUAS)

1. Please indicate which aspects of the speech you think negatively affected your ratings for ACCENTEDNESS: *

- ☐ Grammar
- ☐ Vocabulary
- ☐ Pronunciation of vowels and consonants
- ☐ Word stress
- ☐ Rhythm
- ☐ Intonation
- ☐ Repetition of words or syllables
- ☐ Number of "ums" and "uhs"
- ☐ Number of silent pauses
- ☐ Speaker's storytelling ability
- ☐ Lack of thematic content
- ☐ Lack of content organization

2. Are there other important aspects which affected ACCENTEDNESS and do not appear in the list above? Please tell us:

3. Please indicate which aspects of the speech you think negatively affected your ratings for COMPREHENSIBILITY: *

- ☐ Grammar
- ☐ Vocabulary
- ☐ Pronunciation of vowels and consonants
- ☐ Word stress
- ☐ Rhythm
- ☐ Intonation
- ☐ Repetition of words or syllables
- ☐ Number of "ums" and "uhs"
- ☐ Number of silent pauses
- ☐ Speaker's storytelling ability
- ☐ Lack of thematic content
- ☐ Lack of content organization

4. Are there other important aspects which affected COMPREHENSIBILITY and do not appear in the list above? Please tell us:

5. Of the aspects you indicated, please rank order the top 3 that you felt most influenced on your ACCENTEDNESS ratings. Write 1 (most important), 2 (second most important) and 3 (third most important) and the corresponding aspect next to each number: *

(For instance: 1 = vocabulary; 2: = intonation; 3 = speaker's storytelling ability).

6. Of the aspects you indicated, please rank order the top 3 that you felt most influenced on your COMPREHENSIBILITY ratings. Write 1 (most important), 2 (second most important) and 3 (third most important) and the corresponding aspect next to each number: *

(For instance: 1 = vocabulary; 2: = intonation; 3 = speaker's storytelling ability).

7. If you have any other comments about your listening and rating experience, please share them with us:

THANK YOU FOR PARTICIPATING IN OUR STUDY

Envia

Appendix 9: Examples of listeners' comments about comprehensibility

Category	Example from listeners
Ambiguous	"Clear" "I couln't understand some of the words"
Attitude	"He's trying to do his best" "It does't seem the person is making an effort to communicate"
Being a teacher	"The student says some words as they are 'written', I know students usually have this problema (scared)"
Communicative strategies	"The speaker explains 'gun' by giving a definition (very good strategy to make the listener understand the message)" "Lack of strategies to explain words that the speaker doesn't know in English"
Content	"Uses some imaginary details which makes it easier to follow" "Not much information"
Discourse	"Connectors help to follow the message (and… later…)" "Incoherent content"
English Proficiency	"Low level of English"
Familiarity with the story	"I can understand because I know the story"
Fluency	"Fluent speech" "Lack of fluecy (long pauses, repetition, false starts"
Grammar	"Good grammar to make the message straightforward and clear" "Wrong verb forms"
(No) L1 familiarity	"He made up words but I could understand because they were a transfer from their L1, which we share" "Muy difícil de entender desde el punto de vista de un no nativo español" (= Hard to understand for a non-native speaker of Spanish)
L1 influence (general comment)	"Presence of Spanish at all levels"
Listener's attitude	"If I listen carefully, I understand what he's saying"
Low voice	"Her voice is very low"
Pronunciation	"Quite good pronunciation" "Pronunciation of some key words (thieve, hands up, shoot)"
Self-correction	"Self-correction (arm... pistol)"
Style	"Easy-going speech" "Poca naturalidad" (= Sound unnatural)
Vocabulary	"Correct use of vocabulary" "He makes up words"