

## CHAPTER 6

### DISCUSSION

This chapter discusses the results reported in Chapter 5 in an effort to provide answers to the four research questions of the current dissertation. Those results relating to the perception of English sound contrasts are discussed first, followed by an interpretation of the results on subjects' production of English segments. A brief general discussion section concludes this chapter.

#### **6.1. Auditory discrimination task**

It was described previously that in order to determine the effects of starting age of FL learning, exposure to the FL, L1, and gender on Catalan/Spanish NSs' perception of English sounds, an AX task examining both vowel and consonant contrasts was administered. Based on the results obtained (see Section 5.1), it can be seen that performance on the auditory discrimination task by the various learner groups differed primarily as a function of starting age and exposure to the TL. The following subsections then offer an interpretation of these findings, each subsection focusing on a given factor, and, by extension, on a different research question.

##### **6.1.1. Effect of onset age of FL learning**

The discrimination scores obtained on the AX task showed that generally older starters – i.e. 11- and 14-year-old and adult beginners – with 200 and 416 hours of exposure to the TL perceived English sound contrasts at significantly higher correct rates than younger learners matched for amount of exposure – i.e. 8-year-old beginners. Thus,

the finding of a late starting age advantage in the first stages<sup>107</sup> of L2 phonological acquisition in a formal learning context is consistent with Krashen et al.'s (1982) generalisation about older L2 learners' initial advantage over younger L2 learners in the TL performance in naturalistic settings. Moreover, the results obtained for the older learner groups of the present dissertation agree with actual findings of a late starting age advantage both in immersion settings (e.g. naturalistic study in Snow & Hoefnagel-Höhle, 1977/1982) and in FL learning contexts (e.g. García Lecumberri & Gallardo, 2003; and the laboratory study in Snow & Hoefnagel-Höhle, 1977/1982). Furthermore, the late starting age advantage was found to be short-lived, in line with Krashen et al.'s generalisations, as well as Snow and Hoefnagel-Höhle's (1977/1982, 1978/1982) findings. Particularly, when learners had benefited from long-term instruction in English – that is, learners with 726 hours of exposure to English – 8-year-old beginners were observed to have caught up with older starters. Additionally, 8-year-old starters even surpassed 11-year-old beginners after 726 hours of formal instruction in the FL, though not to the extent of discerning English sound contrasts at significantly higher rates.

If vowel discrimination and consonant discrimination are examined separately, it can be seen that, with regard to vowel discrimination, learner groups exhibited the same response patterns as in the overall answers to the AX task. That is, 11-year-old, 14-year-old, and adult beginners with 200 and 416 hours of instruction in English discerned vowel contrasts significantly better than 8-year-old beginners. Conversely, 8-year-old starters with 726 hours of exposure to English discriminated vowel contrasts nearly identically to 11-year-old starters matched for experience in English<sup>108</sup>.

Concerning consonant discrimination, it is in the perception of this type of sound contrast that 8-year-old beginners showed clearer signs of catching up with older learners. As stated in Section 5.1.5.2, 8-year-old starters with 416 hours of formal

<sup>107</sup> Based on Singleton's (1995) estimates of the differences in the quantity of input delivered in immersion settings vs. formal contexts, subjects in the present research with 416 hours might be considered to be still in their first stages of FL learning, though approaching the middle stages. By the same token, and in agreement with García Lecumberri and Gallardo (2003), learners with 726 hours of exposure might be regarded as being halfway through the process of FL learning or in the middle stages.

However, for the sake of brevity and of avoiding terminological confusion, in this discussion 200 hours of formal instruction will be also referred to as *short-term*, 416 hours of formal instruction as *mid-term*, and 726 hours of formal instruction as *long-term*. These equivalences are also based on the research design of the large study on the age factor from which subjects in this dissertation were drawn. In other words, the research design contemplated collecting data until subjects had reached a maximum of 726 hours of formal exposure to English, whereby 726 hours constituted the long-term stage of the large study on the age factor.

<sup>108</sup> Recall that there was no group formed by 14-year-old starters with 726 hours of exposure to the FL. In addition, very few subjects comprised the adult learner group with 726 hours of formal instruction, so their performance on the AX task could not be analysed statistically.

instruction (A2) discriminated consonants at slightly higher correct rates than 11-year-old starters with the same amount of exposure (B2); while they approached adults' (D2) correct discrimination scores (see Table 5.8 above). Besides, 8-year-old beginners' better discernment of consonant contrasts than that of 11-year-old beginners was more evident when they had received 726 hours of formal instruction in English ( $M = 2.85$  vs.  $2.60$ ), yet again younger learners' higher discrimination scores did not differ significantly from those of older learners.

The distinction between vowel and consonant sound discrimination in conjunction with starting age effects deserve further attention. As reported in Section 5.1.5, all age groups in the present research discriminated vowel contrasts at higher correct rates than consonant contrasts. This is in agreement with findings of FL formal learning contexts (García Lecumberri & Gallardo, 2003), whereas it runs counter to the predictions of the SLM. As noted earlier, the SLM hypothesises that perception (and production) of L2 consonant sounds can be mastered regardless of starting age of L2 learning. On the contrary, mastery of L2 vowels is hypothesised to be inversely correlated to onset age of L2 learning, in the event that L2 learning begins after the establishment of L1 phonetic categories. Rather than disconfirming the predictions of the SLM, though, the finding of Catalan/Spanish subjects' better discernment of vowel contrasts vs. consonant contrasts might be explained on the following grounds. First, as in García Lecumberri and Gallardo (2003), the learners in the present research were at the most in the mid stages of FL learning, which is in opposition to the subjects examined within the SLM framework – i.e. subjects who have reached their ultimate attainment in the TL. Thus, unlike research conducted within the SLM, the age effects observed in the present research refer to actual middle learning stages. Second, the nature of the sound contrasts included in the auditory discrimination task might have contributed to learners' better perception of vowel contrasts, irrespective of their starting age. As described in Section 4.2, initially all sound contrasts in the perceptual task had been expected to pose some difficulty for Catalan/Spanish learners of English. However, based on the findings of different degrees of perceptual difficulty the various age groups displayed (see Table 5.3), it was noted that the number of a priori high-difficulty sound contrasts for vowels and consonants was not balanced out. Specifically, consonant contrasts involved the problematic feature of voicing in word-final position in 60% of instances (3 pairs out of 5 consonant contrasts), whereas the most problematic feature involved in vowel contrasts – the tense/lax vowel distinction – constituted only 25% of instances (2 pairs out of 8 vowel contrasts). Thus,

consonant contrasts contained a larger number of sound contrasts of a high-difficulty nature, which seemingly resulted in all learners' better discernment of vowel contrasts rather than consonant contrasts.

A different pattern of results emerged when the different degrees of perceptual difficulty were considered. That is, learners tended to concur with the sound contrasts that presented higher and lesser degrees of perceptual difficulty, no matter their starting age of FL learning.

In particular, subjects in this study correctly discriminated /ɛ/-/æ/ in nearly 100% of instances, as shown in Table 5.3 above. Other vowel contrasts that were discriminated with little or no difficulty by all age groups were /ɪ/-/e/, /ɑ/-/ɒ/, and /ɒ/-/ʌ/, with the sole exception of 8-year-old beginners with 200 hours of formal instruction who exhibited a higher degree of difficulty (labelled as "some difficulty" in Table 5.3) discerning the three vowel contrasts. The successful (or near successful) discrimination of English /ɛ/-/æ/ corroborates previous findings of NNSs of English in L2 immersion settings, such as those Flege, Bohn, and Jang (1997) reported on native Spanish subjects, and Cebrian (2002a, 2002b, 2002c)<sup>109</sup> on Catalan learners of English. This result also agrees with studies examining native Catalan late starters with a higher amount of exposure in a formal setting (e.g. Rallo, 2003). In spite of this, results on /ɛ/-/æ/ might alternatively be interpreted as in Flege (1991a). That is, learners successfully identified /ɛ/-/æ/ not because they had previously formed a new phonetic category for these two sounds, but because they identified them with Spanish /e/ (or Catalan /ɛ/) and /a/ vowels, respectively. Further support for this interpretation appears in Flege, Munro, and Fox (1994), where Spanish learners of English were noted to have difficulty discriminating English /æ/ when contrasted with low back or mid central vowels, such as /ɑ/ and /ʌ/ (see also Section 2.3.2 above). The view that English /ɛ/-/æ/ might, in fact, fail to be discerned at high correct rates by Romance language NSs is evidenced by a recent study by Flege and MacKay (2004). In one of the experiments reported in Flege and MacKay (2004), native Italian learners of English with an average age of first exposure to the TL of 12 years and with a mean quantity of formal instruction in English of 9 years in their home

<sup>109</sup> Recall that Cebrian (2002a, 2002b, 2002c) did not examine /æ/.

country (and only a mean LOR of 3 months in Ottawa)<sup>110</sup> were found to have difficulty discriminating English /ɛ/-/æ/. In that case, the Italian subjects had identified both English /ɛ/ and /æ/ with Italian /ɛ/.

Likewise, the same argument (Flege, 1991a) might account for FL learners' high successful discrimination rates for the /ɪ/-/e/ contrast – i.e. English /ɪ/ and /e/ were likely to have been identified with Spanish/Catalan /i/ and Spanish /e/ and Catalan /ɛ/, respectively; thereby discriminating such contrast at high correct rates (see also Cebrian, 2002a). Similarly, the high correct discrimination scores obtained for /ɑ/-/ɒ/ and /ɒ/-/ʌ/ might also be interpreted in terms of Flege (1991a). However, it should be noted that there exist to date few studies that have examined contrasts involving low back and mid central vowels (e.g. Flege & MacKay, 2004; Flege et al., 1994; Rallo, 2005). Besides, and as previously mentioned, in those investigations perceptual difficulties have been reported when vowel contrasts involve /æ/-/ɑ/, /æ/-/a/, and /ɑ/-/ʌ/ comparisons (e.g. Flege et al., 1994), which were not included in the present auditory discrimination task.

At the other end of the scale, Spanish/Catalan learners of English showed a great deal of difficulty in discriminating the tense/lax vowel contrast /i/-/ɪ/, no matter their starting age of FL learning. In this case, the percent correct discrimination rate in all learner groups was below 65% (see Table 5.3), which might as well suggest that FL learners identified both /i/ and /ɪ/ as Spanish/Catalan /i/; thus consistent with many findings of Spanish and Catalan learners' nonnative-like perception of English /i/-/ɪ/ contrast in immersion settings (for Spanish, e.g., Escudero, 2002; Flege, 1991a; Flege, Bohn, & Jang, 1997; for Catalan, e.g., Cebrian 2002c, 2003). Another result related to the discrimination of tense/lax vowel contrasts has to do with both younger and older learners' somewhat lesser degree of difficulty in perceiving /ɪ/-/i/ in the minimal pair *still-steal* in the long-term, in contrast to high difficulty in distinguishing /i/-/ɪ/ in *seat-sit*. This apparent contradictory finding might be attributed to an effect of phonetic context. It has been previously observed (e.g. García Lecumberri & Cenoz, 1998) that Spanish learners of English perceive tense/lax vowel distinctions at better rates when presented in voiced consonant contexts than in unvoiced consonant environments. Consequently, the

<sup>110</sup> Note that the subjects examined in that experiment had more similarities to the Catalan/Spanish learners of English of the present dissertation, as opposed to most subject populations examined in Flege et al.'s research.

somewhat higher correct discrimination scores on English /ɪ-/i/ contrast in *still-steal*<sup>111</sup> agrees with this observation about phonetic context effects, rather than contradicting the scores obtained on the other vowel pair which involved the tense/lax vowel distinction.

As mentioned above, common to all age groups was the high degree of difficulty encountered in the perception of the consonant voicing distinction in word-final position, regardless of starting age. This finding further corroborates results of studies conducted on Spanish and Catalan learners of English in L2 naturalistic settings (e.g. Flege et al., 1992; Cebrian, 2000). On the whole, starting age did not have an effect on the degree of perceptual difficulty in the discrimination of /ʒ/-/dʒ/, being discerned with “some difficulty” as Table 5.3 displays. Last, all age groups in the long-term, except for adult starters, had little difficulty perceiving the /b/-/v/ contrast.

Finally, as an artefact of the design of the discrimination task, it could be ventured that some of the starting age effects reported above might have been partly confounded with effects of subjects’ chronological age at testing. In the first place, by administering an AX discrimination task it was originally thought that the task would involve the same cognitive load on the part of all learners. However, this “equality” in the cognitive load required to perform the AX task appears to have been overridden by the type of sound contrasts included. Thus, as noted in Section 5.1.5, AX discrimination tasks generally focus either on consonant or on vowel contrasts (cf. Snow & Hoefnagel-Höhle, 1977/1982, 1978/1982). Moreover, when examining one type of sound contrast, the task is restricted to one specific opposition (e.g. multiple instances of /i-/ɪ/), or if consisting of more than one block, each block contains one specific sound contrast. Only recently have studies incorporated a variety of sound contrasts into discrimination tasks (e.g. Flege & MacKay, 2004; Gallardo, García Lecumberri, & Cenoz, 2002; García Lecumberri & Gallardo, 2003). Thus, in relation to one of the experiments included in the Flege and MacKay (2004) study investigating native Italians’ perception of English /i-/ɪ/, /ɛ-/æ/, and /ɒ/-/ʌ/ contrasts, the authors indicate that

[i]n some categorial discrimination tests (e.g., Best et al., 2001) the trials testing each contrast of interest are presented in separate, counterbalanced blocks. In this

<sup>111</sup> It should be further noted that the [iɪ] sequence in English contains a distinct glide between /i/ and /ɪ/, making it two syllables in the pronunciation of some words, as is the case of *steal*. In contrast, the [ɪɪ] sequence in *still* lacks this epenthetic glide. Thus, the difference between /iɪ-/ɪɪ/ in *steal-still* is very noticeable, which might explain the good discrimination without the subjects actually hearing any vowel difference at all (I. MacKay, personal communication, June 1999, 2005).

study, the trials testing all nine contrasts [i/-ɪ/, ɪ/-ε/, e<sup>1</sup>/-ε/, e<sup>1</sup>/-i/, ε/-æ/, e<sup>1</sup>/-ɪ/, æ/-Λ/, ə/-Λ/, and ɒ/-Λ/] were presented in a single, randomized block to increase task difficulty and thus maximize the likelihood of observing significant between-group differences. (p. 28, footnote 2)

Based on all of the above, it can be inferred that by having included not only a variety of different vowel contrasts but also a variety of different consonant contrasts in the same discrimination task without separate blocks, the cognitive demands were indeed higher than expected. Moreover, whenever testing took place, learners with an AOL of 8 years were always younger than learners with AOLs of 11, 14, and 18+ years (see Table 4.3 for each subject group's mean chronological age at testing). Thus, it might be suggested that part of the late starting age advantage observed in the initial stages of FL learning to a certain extent might have been due to differences in subjects' cognitive maturation levels at testing. In turn, this would agree with findings of Spanish/Basque bilinguals learning English in a formal instruction setting, where late starting age has been found to be more beneficial than early starting age (García Lecumberri & Gallardo, 2003).

To sum up, onset age of FL learning did have an effect on the perception of English sound contrasts by Spanish/Catalan learners of English in a formal instruction setting. In agreement with much L2 acquisition research conducted in immersion settings and observations in the literature, in the initial states of FL learning a late starting age – in this case, 11, 14, and 18+ years – resulted in a better overall discernment of English sounds, as opposed to an early starting age – i.e. 8 years. Besides, consistent with the literature (cf. García Lecumberri & Gallardo, 2003), the late starting age advantage in the perception of English sound contrasts was short-lived, for 8-year-old beginners had already surpassed older starters in the long-term. However, neither in the short-term nor in the mid- or long-term did any learner group obtain overall correct discrimination scores within the NS range. This finding further corroborates the prediction of the SLM that when L2 learning commences after L1 category formation, native-like attainment of L2 phonological skills is not guaranteed. Besides, the SLM's hypothesis that age of first exposure to the TL will determine the degree to which learners perceive English sounds accurately – i.e. the earlier the starting age, the more accurately learners will perceive TL segments – is only partly supported by the findings of the present research, for the advantage of 8-year-old beginners in the long-term as to more accurate discrimination of

English sounds was small and nonsignificant. Finally, starting age failed to reveal significant effects on the different degrees of perceptual difficulty in the discrimination of English contrasts. In other words, all age groups tended to agree on which sounds contrasts were more difficult or easier to discern.

### 6.1.2. Effect of exposure

The four main age groups in the study – groups A, B, C, and D – behaved differently as exposure to the FL increased. Firstly, 8-year-old beginners perceived English sound contrasts more accurately along with an increase in formal instruction. In all cases, younger learners obtained progressively higher correct scores for overall task, vowel contrast, and consonant contrast discrimination. Moreover, the differences in scores reached significance when younger learners had received 726 hours of formal instruction in English, as compared to the scores obtained with 200 hours and 416 hours of English experience. Thus, the effects of experience observed in 8-year-old beginners are in line with the predictions of the SLM.

Similar exposure effects were for the most part noted for adult beginners, in particular when exposure increased from 200 hours to 416 hours. In that case, though, the differences in discrimination scores did not reach significance. Besides, in comparison to 8-year-old beginners, the degree to which adults perceived English sound contrasts accurately as a result of exposure to the FL was less marked. Furthermore, when adults had received 726 hours of formal instruction, they tended to perceive English sound contrasts at lower correct rates than with lesser amounts of exposure, unlike younger starters. Nevertheless, the latter finding should be taken with caution, since the adult learner group in the long-term was composed of very few subjects.

On the other hand, 11-year-old and 14-year-old starters did not generally perceive English sounds more accurately as formal instruction in English increased. In fact, these two age groups displayed non-consistent exposure effects depending on whether they discriminated vowel or consonant contrasts. So, 11-year-old beginners perceived vowel contrasts at somewhat higher correct rates when formal instruction amounted to 416 hours and 726 hours, in opposition to 200 hours. By contrast, 11-year-old starters' perception of English consonant segments became less accurate when they had received



416 hours of instruction in English vs. 200 hours. Moreover, poorer performance along with an increase in exposure was also noted for distractor discrimination. As regards 14-year-old beginners, and contrary to 11-year-old beginners, 416 hours of formal instruction led to more accurate perception of consonant contrasts, but not to that of vowel contrasts. At this point only tentative explanations can be offered to account for learners' poorer performance along with an increase in English experience – especially noticeable in the mid-term of the present study. One explanation might lie in learners' motivation and (negative) attitude towards FL learning at that specific point in time, resulting in a diminishment in correct discrimination scores. Yet another plausible explanation has to do with NNS input that subjects might have been exposed to, instead of English NS input. This possibility will be further developed in the discussion of the accent studies conducted on subjects' production of English sounds.

Despite the fact that the use of nonparametric tests did not allow for the exploration of the interaction between the research variables in the study, it can be seen that there existed an interaction between onset age of FL learning and exposure to the TL in the case of 8-year-old beginners. Thus, exposure effects were consistently beneficial if FL learning had begun at the age of 8. However, the factor of exposure in English sound perception by 11- and 14-year-old starters, in addition to adult beginners, failed to reveal clear-cut results either in the direction of improvement or lack of beneficial effects.

What is more, even in those cases where an increase in English experience led to (significantly) more accurate perception of English segments, no learner group was able to discern English sound contrasts at native-like levels. Moreover, it was observed that a higher amount of formal exposure to English resulted in a better discernment of specific sound contrasts only, primarily /ɛ/-/æ/, /ɒ/-/ɑ/, and /ɪ/-/ε/ for vowel contrasts, and /b/-/v/ and /ʒ/-/dʒ/ for consonant contrasts (and mostly for 8-year-old beginners). Conversely, perception of English tense/lax vowel contrasts and consonant voicing in word-final position did not improve as a function of exposure to the FL.

Taken together, the differing exposure effects observed in this research are inconclusive and in line with findings that have been reported for FL formal learning settings where learners have limited exposure to the TL (e.g. Rallo, 2003, vs. García Lecumberri & Gallardo, 2003). In addition, the findings of the present study are consistent with García Lecumberri and Gallardo's (2003) observation that 6–7 years of

instruction in English in a formal setting are not sufficient to attain native-like performance in the FL phonology.

### 6.1.3. Effect of dominant L1(s)

The results obtained for the three language dominance groups – Catalan dominant, Spanish dominant, and Catalan/Spanish balanced bilinguals – revealed that they all perceived English sound contrasts very similarly. As further noted above, no consistent pattern of more accurate perception of English sounds emerged according to whether learners reported being Catalan or Spanish dominant speakers, or Catalan/Spanish balanced bilinguals. Thus, the phonetic features of the learners' L1, specifically those of Catalan, did not constitute an advantage when it came to discriminating English vowels and consonants, contrary to some observations (e.g. Coe, 1987) and to what the third research question in the present dissertation had hypothesised. This finding might be interpreted on the grounds that even when subjects claim to have Spanish or Catalan as their dominant L1, they are extensively exposed to both languages, and so are familiar with the phonology of both.

Although the AX task subjects performed could not test for their reliance on temporal cues rather than spectral cues in the perception of English tense/lax vowel contrasts, or for the generalisation of L1 neutralisation rule over English consonant voicing distinction in word-final position, the observed high degree of difficulty in learners' perception of English /i/-/ɪ/, /p/-/d/, /t/-/d/, and /s/-/z/ suggests that subjects indeed resorted to temporal cues and neutralisation rules in order to discern tense/lax vowel contrasts and consonant voicing contrasts, respectively. As a consequence, their perception of such sound contrasts was significantly poorer; further agreeing with findings of Romance language NSs learning English in both naturalistic and formal instruction contexts.

All in all, results for tense/lax vowel and consonant voicing contrasts, in conjunction with results for the remaining sound contrasts examined in the auditory discrimination task, were then consistent with previous findings of native Spanish and Catalan learners of English in both immersion and formal settings.

#### 6.1.4. Effect of gender

In agreement with previous studies, the comparison between male and female subjects in the discrimination scores obtained for English sound contrasts yielded inconclusive results. When the 281 subjects were considered together, it was found that female subjects discriminated English contrasts at slightly higher correct rates than male subjects ( $M = 15.43$  vs.  $15.12$ , respectively). However, this slight advantage did not hold when subjects were examined in their separate learner groups. For instance, as reported in Section 5.1.5.1, there was a great deal of variability between male and female subjects' scores for vowel contrasts, whereby no pattern emerged as to either one of the two gender subgroups obtaining consistently higher correct discrimination scores. Moreover, those instances where female participants discerned consonant contrasts better were subject to starting age of FL learning and exposure effects; hence extending previous inconclusive findings of gender effects in naturalistic settings to FL formal learning contexts.

#### 6.2. Imitation task – Study 1

Study 1 was conducted so as to look at the effects of the research variables on the FA ratings obtained for six English words (*jam, reading, red, speak, this, very*) and six respective segments (/æ, i, d, s, ɪ, v/) included in the imitation task as produced by FL learners and a control group of English NSs. It is worth noting again that at the time Study 1 was carried out, data for some learner groups still had to be collected. As a consequence, only learners groups A1, A2, B1, B2, B3, and D1 could be examined in detail and analysed statistically. Before discussing the effects of the research variables on the accent ratings subjects obtained on English words and segments, several general findings should be considered.

In the first place, it was observed that the six judges or listeners in the study assigned higher accent ratings (i.e. more foreign-accented) to words rather than segments. Moreover, words differed significantly from each other in their degree of FA. Similarly, accent ratings for vowel and consonant segments varied, vowels being rated as significantly less foreign-accented than consonants. In addition, those segments that were

judged to be more foreign-accented did not always correspond to the degree of FA assigned to the words in which they appeared. For instance, as reported in Sections 5.2.2.1 and 5.2.2.2, the word *red* was rated as significantly less foreign-accented than *very*, among other words; whereas its segment under examination – /d/ – was judged to have been produced with a significantly higher degree of FA than /v/ in *very*. That is, when rating words for degree of FA, judges often deemed features other than the segments subsequently rated for FA as more salient in the detection of FA on words. All of the above is in accordance with previous findings of FA research, whereby it has been stated that NS listeners weigh deviations from TL speech norms on a differing basis (e.g. Flege, 1981; Magen, 1998). Furthermore, the fact that words were considered to have been produced with a higher degree of FA than that of segments is in line with Munro et al.'s (1996) observation that FA is more readily detected in longer utterances (e.g. sentences) than in short stretches of speech (e.g. CV utterances) – in this case, words vs. individual segments.

A second general comment has to do with judges' performance. As seen above, inter-rater coefficients were acceptable, but not high enough to allow for a single mean accent rating. More specifically, it was reported that two judges (listeners 2 and 3) differed in the FA scores they assigned by rating subjects' word productions as being significantly more foreign-accented than did the remaining judges. This result further corroborates previous studies where listeners exhibited different degrees of sensitivity to foreign-accented speech (e.g. Flege, Bohn, & Jang, 1997; Flege et al., 1995a; Flege, Frieda, & Nozawa, 1997; Moyer, 1999, 2004; Munro et al., 1996). In spite of this, all judges carried out the rating task as expected in that they were able to identify the English control group successfully (and as presented in noise conditions). Therefore, in all instances NE foils' words and segments were judged to have been produced within the NS range. By contrast, Spanish and Catalan learners of English were always reported to have produced English words and segments with a moderate amount of FA. The following subsections then describe and interpret the extent to which the research variables influenced the foreign-accented ratings FL learners received. In addition, each subsection is concerned with the effects of a given research variable on accent ratings on both words and segments.

### 6.2.1. Effect of onset age of FL learning

As mentioned above, FL learner groups received significantly higher (i.e. more accented) FA scores on words and segments than the NE group. As for the effect of onset age of FL learning on accent ratings obtained for words, an early starting age resulted in a lesser degree of FA in the initial stages of FL learning. Specifically, 8-year-old beginners with 200 hours of formal instruction produced the six English words as less accented than 11-year-old beginners matched for amount of exposure. Furthermore, on occasion the differences in ratings between 8- and 11-year-old starters were significant.

However, this clear early starting age advantage did not hold when learners' experience in English had reached 416 hours. In that case, accent ratings varied as a function of judges and words. More precisely, half the judges considered the words as produced by 8-year-old starters with 416 hours of instruction as less foreign-accented than those of 11-year-old starters. On the other hand, the remaining half of listeners judged 11-year-old starters' word productions as less accented than 8-year-old beginners'. Specifically, younger learners received lower FA ratings on *red*, *speak*, and *very*; whereas older learners produced *jam*, *reading*, and *this* as less foreign-accented.

The observed age effects when learners had 416 hours of instruction are difficult to reconcile with findings of formal learning contexts, where a late starting age advantage in the initial and mid stages of FL learning has been reported on the production of English speech (e.g. García Lecumberri & Gallardo, 2003). Besides, in the case of learners with 200 hours of formal instruction, 8-year-old beginners' less accented word productions than 11-year-old beginners' are not consistent with the effects of starting age on the perception of English sound contrasts in the initial stages of FL learning discussed in Section 6.1.1 above. That is, while in the first stages of FL learning 11-year-old beginners perceived English sounds more accurately than 8-year-old beginners, 11-year-old beginners produced English words on a more poorly basis than 8-year-old beginners. Based on the latter finding alone, it might be suggested that at least in the initial stages of FL learning an earlier starting age in formal instruction settings is more beneficial for FL sound production, while a later starting age results in a better perception of FL segments. Up to a point, this could be taken as evidence that production precedes perception in FL learning in those cases where L2 learning started just after the formation of L1 phonetic categories – 8-year-old beginners – but not later in life – i.e. 11-year-old starters and

older beginners. However, as just indicated, when learners had benefited from 416 hours of formal instruction, age effects on the production of English words were not so well defined. A better alternative to accounting for 8-year-old beginners' less accented production of English words and segments in the initial stages of FL learning (and in contrast to their less accurate perception of FL sounds) lies in the characteristics of the imitation task itself. It was mentioned in Section 4.2 that the words subjects produced were repeated in isolation and immediately after hearing each target word as delivered by a taped model voice. In this case, the cognitive load was the same for all subjects irrespective of their starting age of FL learning and of their chronological age at testing. Nevertheless, one could hypothesise that the feature of direct imitation involved in the task was to the advantage of younger learners<sup>112</sup>, for a number of studies have shown that younger learners are better imitators than older learners, even when presented with completely unknown FL speech (e.g. Tahta et al., 1981b; cf. Loewenthal & Bull, 1984).

Evidence for inconclusive age effects surfaced in the examination of accent ratings on English vowel and consonant segments. With regard to vowels, the age groups in the study produced /i, ɪ, æ/ with varying degrees of FA. Thus, with 200 hours of instruction adult starters were reported to have produced /i, ɪ, æ/ as significantly less foreign-accented than both 8- and 11-year-old beginners. This finding would be consistent with studies showing an initial late starting age advantage in the phonology of the TL. However, an initial early starting age advantage was observed in the comparison between 8- and 11-year-old starters. Even then, the early starting age advantage was subject to specific segments. Thus, 8-year-old starters produced /i/ and /æ/ as less foreign-accented than 11-year-old starters, while they obtained higher FA ratings on /ɪ/ than did 11-year-old starters. Moreover, a reverse pattern emerged when learners had received 416 hours of formal exposure to the TL. Then, 8-year-old beginners produced /i/ and /æ/ as more foreign-accented and /ɪ/ as less foreign-accented than 11-year-old beginners. These non-systematic age differences in accent ratings for vowels are also difficult to explain on the basis of the distinction between new vs. similar FL sounds as the SLM predicts. Instead, they further corroborate inconclusive age effects on FL

---

<sup>112</sup> It was noted previously that earlier starters were always younger in chronological age than the other starting age groups whenever testing took place.

phonological learning in the long-term in instructed-classroom settings (e.g. García Lecumberri & Gallardo, 2003).

Additional support for inconclusive age effects is to be found in the accent ratings on consonant segments. First, with 200 hours of instruction, adult learners obtained lower accent scores on /s/ and /v/ than 8- and 11-year-old beginners. In turn, 8-year-old starters' accent scores on /s/ and /v/ were lower than those of 11-year-old beginners. By contrast, the latter produced /d/ as less foreign-accented than 8-year-old and adult beginners, differences in ratings even approaching significance. On the other hand, when exposure amounted to 416 hours, accent ratings were similar across learner groups, though 11-year-old starters obtained slightly less accented ratings on consonants than did 8-year-old starters.

It should be noted, as well, that the results on the production of consonants as a function of onset age of FL learning are in disagreement with those related to the perception of consonant segments in the current dissertation. At least in the initial stages of FL acquisition, older starters often produced English consonants at more accurate rates than younger beginners. As compared to the AX task results, older starters' slight superiority in consonant production, albeit not significant, might be interpreted in the following way: while younger learners are catching up with older learners in the perception of English consonants, they are still not going through the same process in the production of consonants. If that were the case, the present findings for consonant perception and production would be in line with the assumption of the SLM that perception leads production, in addition to the model's premise that more accurate perception of TL segments does not necessarily involve better production of TL sounds.

A final comment on the accent ratings on English segments deserves further attention. In turn, this might help account for inconclusive age effects on the production of English consonant segments, and, by extension, English vowels and words. As noted previously, /d/ was rated as significantly more foreign-accented than /v/ and /s/. In fact, all groups, including the NE group, obtained noticeably higher accent scores on /d/ (see Figure 5.38). One could hypothesise that the higher accent ratings resulted from English NS listeners' difficulty in discerning whether /d/ had been elided or unreleased, since it has been shown that under ideal listening conditions judges (trained phoneticians included) might find it difficult to distinguish instances of /d/ elision from those of

unreleased /d/ (e.g. Flege & Davidian, 1984). As a corollary, the fact that the recordings contained a considerable amount of background noise might have resulted in the acoustic cues that are still present in unreleased /d/ not being perceptible at all. Hence, listeners might have considered instances of unreleased /d/ as elisions, and thus the higher accent ratings assigned<sup>113</sup>. By the same token, the noise present in the recordings was likely to have played a part in judges' ratings on the remaining segments and English words, especially in the case of FL learners' productions. As mentioned earlier, noise did not appear to influence judges' performance on identifying the NE control group. By contrast, the inconclusive age differences in Study 1 in part might be thought to arise from an interaction between foreign-accented speech and noise. All this would be in line with preliminary findings of studies assessing the effect of noise on the perception of native- vs. nonnative-produced speech in immersion settings, such as Munro (1998), where noise has a greater (negative) effect on NS listeners' perception of foreign-accented speech than accent-free speech in the TL.

### 6.2.2. Effect of exposure

As was the case of the perception of English sounds, formal exposure to English had a varying effect on learners' production of English words and segments. First, and unlike findings of the AX task in the present dissertation, 8-year-old beginners obtained more foreign-accented scores as their experience in English increased from 200 hours to 416 hours. Only on one occasion – i.e. ratings on /æ/ – did a gain in experience lead to the expected direction of less accented production of English segments. Conversely, and unlike the AX task too, an increase in exposure for the most part led to 11-year-old beginners' less accented word and segment productions, often yielding significant results. However, as was stated in relation to the AX task, in some instances halfway through their learning (416 hours) 11-year-old beginners produced words as more foreign-accented than when they had 200 hours and 726 hours, in addition to vowel /ɪ/ and consonant /d/. The latter finding then agrees with the results of the AX task and might be

---

<sup>113</sup> It should be remembered that listeners were specifically asked to rate /d/ elision in *red* as an instance of FA (see footnote 80 above).



tentatively attributed to learners' motivation and attitudes towards the learning of the FL phonology. However, a more likely explanation for the apparently negative effect of exposure on FL word and sound production lies in Flege's (1991b) "accented L2 input hypothesis" (see 2.3.1.1 above). A number of studies have illustrated that, in spite of an early start in L2 learning in immersion settings, L2 learners failed to attain native-like pronunciation in the TL as they had been exposed to NNS input (e.g. Flege, 1991b; Flege & Eefting, 1987; Flege & Fletcher, 1992). Similarly, in formal learning contexts early FL learners, in addition to late FL learners, have been reported to pronounce the TL with an inevitably degree of FA, based on the fact that the input they received was already accented (García Lecumberri & Gallardo, 2003). Thus, the exposure effects observed in the present dissertation might be interpreted in terms of the "accented L2 input hypothesis". In particular, 8- and 11-year-old beginners were likely to have been exposed to accented input in English through their school teachers<sup>114</sup>. This would explain 8-year-old beginners' more accented production of English sounds as a function of formal exposure to the FL, rather than disconfirming the predictions of the SLM concerning an increase in experience in the TL. The accented L2 input hypothesis is further supported by the finding of learners' failure to attain a native-like pronunciation in the FL in the long-term in those cases where an increase in exposure resulted in a more accurate production of English words and segments.

### 6.2.3. Effect of dominant L1(s)

Similar to the results obtained for the perception of English sound contrasts, the three language dominance groups did not differ in their FA scores on the production of English words and segments. What is more, there was no consistent pattern as to which language group produced English words and segments as systematically more foreign-accented (or less foreign-accented). It was suggested above that all learners being familiar with the phonology of both Catalan and Spanish might have influenced their perception of English sounds to the extent of discerning sound contrasts in a similar manner. Likewise, all learners' familiarity with Spanish and Catalan might have contributed to the lack of differences in accent ratings observed in Study 1 between

---

<sup>114</sup> See Naves and Muñoz (1999) for a detailed description of English language teachers' characteristics in Spanish school settings.

Spanish dominant, Catalan dominant, and Spanish/Catalan balanced bilinguals. In addition, the segments chosen for FA ratings were expected to have posed the same degree of difficulty for each one of the three language dominance groups.

Finally, it should be mentioned that judges 5's and 6's higher degree of familiarity with Spanish did not influence their ratings in a noticeable way, as compared to those listeners who were not familiar (or less familiar) with Spanish. This result further corroborates previous findings, where listeners' familiarity with the learners' L1 (and even foreign-accented speech) is not thought to play a significant role in accent detection (e.g. Flege & Fletcher, 1992; Flege, Frieda, et al., 1997).

#### **6.2.4. Effect of gender**

Unlike the perception of English sounds, consistent gender differences in the production of both English words and segments were evident in Study 1. More precisely, female subjects' pronunciation in the TL was always rated as less foreign-accented than that of male subjects, reaching significance levels in the words *red* and *speak* and for listeners 1, 3, and 6. This finding supports the popular belief that female subjects are better language learners – or, maybe better language imitators – and provides evidence that in formal language learning contexts female subjects produce English sounds at more accurate rates.

#### **6.3. Imitation task – Study 2**

The discussion of the results of Study 1 offered an insight into the influence that the four research variables exerted (or did not exert) on the production of English words and segments by Spanish and Catalan learners of English. It should be added that Study 1 was limited in the sense that at the time it was conducted subject data from several learner groups were still pending collection – in particular, data from 8-year-old beginners with 726 hours of instruction. In the second major study carried out on the data obtained by means of the imitation task – Study 2 – subjects belonging to all learner groups had already participated. However, in Study 2 group sizes were reduced, since

only those recordings that had a minimum S/N level of 10 dB were selected for further assessment. Consequently, data for C2, D2, and D3 could not be analysed statistically. Even then, Study 2 was more complete than Study 1 in that both the younger age groups – A and B – could be examined at three different points of exposure. Furthermore, the comparison between 8- and 11-year-old starters' performance in the TL phonology was of special interest, for the two age groups were representative of the new and former curricula in the Spanish educational system.

Based on the findings of Study 1, two methodological procedures were adopted on designing and implementing Study 2 in an attempt to obtain a higher degree of inter-rater agreement. In the first place, listeners were asked to only rate vowel segments, as Study 1 had shown that judges might have regarded the salience of several acoustic features on a non-systematic differing basis when rating words for degree of perceived global FA. Secondly, intra-rater consistency was examined by adding an extra set of repeated subjects' productions to a total of fourteen accent rating and vowel identification blocks. Yet the seven listeners who participated in Study 2 differed in their sensitivity to FA, as in Study 1. Particularly, judge 2 rated subjects' English vowel productions as significantly less foreign-accented than did the remaining judges. In contrast, judges 3 and 4 rated learners' productions as consistently more foreign-accented than judges 2, 5, 6, and 7. Rather than invalidating these judges' performance, this finding might be interpreted as in Flege, Bohn, et al. (1997). In that study, one of the three NS listeners recruited was found to have different internal category boundaries for English sounds from those of the remaining two listeners. Nonetheless, this difference did not affect that specific listener's identification of native-produced speech. Likewise, judges 2, 3, and 4 in Study 2 of the present dissertation rated the NE control group's vowel productions as native-like and later identified their vowel productions as intended in nearly all instances. Moreover, the high intra-rater coefficients reported added reliability to the seven judges' performance, as well as the high degree of inter-rater agreement that was often observed.

### 6.3.1. FA ratings on /i, ɪ, ε, æ, ɒ, u, ʌ/

In addition to the differences in accent ratings among judges seen above, overall accent scores on English /i, ɪ, ε, æ, ɒ, u, ʌ/ showed that the seven vowels were rated on a

varying basis. More precisely, /æ/ and /ɛ/ (*tests*) were considered to have a significant higher degree of FA than /i/ (*tea*) and /ɛ/ (*red*). At the other end, /ɪ/ and /ʊ/ were rated as less foreign-accented than /æ/, /u/ and /ʌ/, though not at significant levels. At first sight, the higher accent ratings for /æ/ and /ɛ/ (*tests*) and lower FA scores for /ɪ/ might come as a surprise based on the perceptual findings reported in Section 6.1.1 above. The following subsections aim to explain these unexpected results, together with the effects of the research variables of the present dissertation on the accent ratings on English vowels.

### 6.3.1.1. Effect of onset age of FL learning

Like the results of Study 1, the NE group obtained lower FA scores than learner groups on all vowel sounds and according to all judges, whereas all learner groups' vowels were judged to have been produced with a certain amount of FA.

Among FL learners, the observed age effects displayed a similar pattern to that of the first study, for judges differed as to which age group produced English vowels at more accented rates in the initial stages of FL learning. Thus, judges 1, 4, 6, and 7 considered 8-year-old beginners' vowel productions as being less foreign-accented than those of older starters with 200 hours of instruction, while judge 7 considered adults' production of English vowels as less foreign-accented than 8-year-old starters. Thus, the seven judges in Study 2 coincided with the six judges in Study 1 in exhibiting different degrees of sensitivity to FA as a function of onset age of FL learning. At the same time, this finding corroborated previous studies (e.g. Moyer, 1999; Munro et al., 1996). In spite of this, the majority of judges in the present study tended to rate 8-year-old beginners' vowel productions as less foreign-accented. Along this line, a tendency was observed for this younger age group to produce (tentatively classified) new FL sounds – /ɪ/, /æ/, /ʊ/, and /ʌ/ – with a lower degree of FA. This finding would disconfirm the initial late starting age advantage in the first stages of FL acquisition, at least concerning new TL sounds (note that older learners produced similar FL sounds at less accented rates: /i/ and /ɛ/ for 11-year-old beginners, and /u/ for adult starters). Besides, to a certain extent, this result would support the prediction of the SLM that, in the event that L2 learning starts

after the establishment of L1 phonetic categories, an earlier start in the acquisition of the TL will lead to a more accurate production of new TL segments.

However, this prediction was not upheld when learners had received 416 hours of formal instruction, since judges differed greatly as to which age group produced English vowels with a lesser amount of FA. Therefore, as in Study 1, age differences in accent ratings for vowels were not consistent at this point of exposure. Nor were differences in accent scores maintained to the advantage of younger learners, as the result for /ʌ/ illustrated. In that case, /ʌ/ accent scores were noticeably higher and nearly approached significance in 8-year-old beginners, as compared to 11-year-old beginners. In turn, this would disconfirm the predictions of the SLM of an early starting age advantage in the production of new TL sounds.

Support for a late starting age advantage was found when learners had received 726 hours of formal instruction in English. As happened with the perceptual task, there were no significant differences in accent ratings for English vowels between 8- and 11-year-old beginners, although in this case 11-year-old beginners produced English segments at slightly less accented rates. This finding is in line with results from formal learning contexts where older starters in the long-term (as conceived in the present study) are somewhat better at pronouncing the TL (García Lecumberri & Gallardo, 2003), while disagreeing with findings of naturalistic settings (e.g. Snow & Hoefnagel-Höhle, 1978/1982).

Overall, onset age of FL learning had an effect on the production of English vowels, which was in the opposite direction of that reported for the perception of English sounds. That is, in the initial stages of FL learning younger beginners produced English vowels, particularly those segments that might be considered new FL sounds, with a lesser amount of FA. As was the case of the perception of English sounds, age differences became less noticeable when learners had received 416 hours of formal instruction in English. Finally, despite the lack of marked age differences, older starters (specifically 11-year-old beginners vs. 8-year-old beginners) produced English segments somewhat more accurately in the long-term, which does not agree with Krashen et al.'s generalisations or the predictions of the SLM. Even then, age differences in accent ratings were small and judges varied greatly in their consideration of which age group's production of English segments was less foreign-accented.

### 6.3.1.2. Effect of exposure

Parallel to the age effects reported in the previous subsection, the factor of exposure yielded mixed results. In the case of 8-year-old beginners, this age group followed the trends outlined in Study 1. That is, as exposure to the FL increased, their production of English vowels, both similar and new sounds, was rated as more foreign-accented. This is in contradiction with the predictions of the SLM and several studies (e.g. Flege, Bohn, & Jang, 1997), and even challenges this same group's performance on the AX task, where increments in the amount exposure led to their more accurate perception of English sounds.

As for 11-year-old beginners, it was difficult to determine the effect of exposure on their production of new vs. similar English sounds, since accent scores varied greatly according to each segment. Moreover, judges differed in their ratings as a function of exposure. Therefore, while judges 1, 2, 4, and 5 rated 11-year-old beginners' vowels as more foreign-accented along with an increase in formal instruction in English, judges 3, 6, and 7 considered that vowels had been produced with a lesser amount of FA as exposure to the TL increased.

In any event, neither 8-year-old nor 11-year-old beginners produced English vowel segments within the NS range. Besides, common to both age groups was the fact that halfway through their learning (i.e. 416 hours) their FA scores were higher, thereby corroborating findings of Study 1. All these results are not in line with the expected effect of exposure within the SLM framework, yet they concur with an alternative account the SLM puts forth in those cases where differences in the amount of exposure do not result in a less foreign-accented pronunciation of the TL segments by the learner group with a larger amount of experience in the L2, namely the accented L2 input hypothesis (see 6.2.2 above).

On the other hand, at a descriptive level exposure appeared to be more beneficial for older learners, i.e. 14-year-old and adult starters. More precisely, both age groups obtained progressively less accented scores – though not native-like – on the seven target vowels as they gained experience in English. In particular, adults' less accented vowel productions might be accounted for the fact that most of their teachers of English were

NSs of the TL. By extension, this would lend further support to the accented L2 input hypothesis.

### 6.3.1.3. Effect of dominant L1(s)

In agreement with previous findings of the present dissertation (see Sections 6.1.3 and 6.2.3), the three language dominance groups produced English /i, ɪ, ε, æ, ɒ, u, ʌ/ in a very similar fashion. As expected, the differences in accent scores were significant between NE foils and Catalan/Spanish learners of English. In addition, as was the case of Study 1, the judges' higher degree of familiarity with Spanish did not influence their ratings.

Even though Catalan dominant, Spanish dominant, and Catalan/Spanish balanced bilinguals did not differ significantly in their production of the target vowels under examination – all groups pronounced vowels with a moderate amount of FA – the accent ratings they obtained on /i, ɪ, ε, æ, ɒ, u, ʌ/ should be further considered. In Section 6.3.1, it was pointed out that /ε/ (*tests*) and /æ/ were rated as significantly more foreign-accented ( $M = 4.72$  and  $4.43$ , respectively) than the remaining vowel segments. Similarly, higher FA scores were assigned to /ɒ/ and /ʌ/ ( $M = 4.08$  and  $4.23$ , respectively). On the other hand, /i/ was rated as less foreign-accented ( $M = 3.79$ ). In general terms, the better production of /i/ and less accurate production of /æ/ and /ʌ/ are consistent with findings of Romance language learners of English (e.g. Flege et al., 1994; Fox et al., 1995). However, this is not the case of /ε/. But, instead of disconfirming previous studies, the higher accent ratings obtained for /ε/ (*tests*) might be understood as an effect of phonetic context or phonotactics, since /ε/ was produced in a word containing a three-consonant coda (non-existent in Spanish or Catalan). This interpretation is further supported if the accent ratings for /ε/ in *red* are examined. In that instance, /ε/ was judged to have been produced with a lower degree of FA ( $M = 3.68$ ).

Despite the fact that correlations between perception and production of FL sounds could not be performed in this dissertation, it would appear that the accent scores of Study 2 were not correlated with the findings reported for the perception of English

sound contrasts, in particular /ɛ/-/æ/ and /ɒ/-/ʌ/. Yet again two reasons might be suggested in order to corroborate, rather than disconfirm, previous findings. Firstly, the fact that learners discriminated these sound contrasts at high correct rates while producing them less accurately is in accordance with the assumption of the SLM that perception leads production. Secondly, and as will be seen in the misidentification patterns for the target vowel sounds (Section 6.3.2.3 below), the accent scores indicate that learners might have in fact identified each /ɛ/ and /æ/, and each /ɒ/ and /ʌ/, with two different L1 vowels. In the case of perception, this in turn resulted in /ɛ/-/æ/ and /ɒ/-/ʌ/ being discriminated successfully at very high rates, as suggested in Section 6.1.1 above. There it was further hypothesised that a different pattern of results might have emerged, had /æ/ and /ʌ/ been presented in the same contrasting pair (as shown below judges identified FL learners' incorrect production of both English /æ/ and /ʌ/ as Spanish /a/ at significant levels).

#### 6.3.1.4. Effect of gender

Last, findings of Study 2 relating to gender differences in the production of English vowels corroborated the results from Study 1 in that female subjects were deemed to be better language imitators than male subjects, as far as the production of English vowels was concerned. In other words, when FL learners were examined together, it was shown that female subjects produced vowel sounds as being significantly less foreign-accented than those of male subjects, according to judges 1, 3, 6, and 7. For the remaining listeners, differences in FA scores approached significance, particularly for vowels /i, ɛ, æ, ʌ/. On top of that, within each learner group female participants maintained their advantage over male participants in obtaining less foreign-accented scores on the seven target vowels, although within each learner group the differences in accent scores were nonsignificant.



### 6.3.2. Vowel identifications for /i, ɪ, ε, æ, ɒ, u, ʌ/

If English vowels were rated for degree of perceived FA on a varying basis, so were they identified on differing grounds in the vowel identification task the seven listeners carried out. Irrespective of the possible effects of the research variables, when taken together, the percent correct identification scores obtained by subject groups agreed with findings of native Romance language speakers of English. Thus, /ε/ was identified as intended at higher correct rates than the remaining vowel segments, followed by /i/, /u/, and /ʌ/. In contrast, /ɪ/, /æ/, and /ɒ/ were identified at lower correct rates, resulting in significant differences in /ɪ/ identification scores vs. /i/, /ε/, and /ʌ/. Based on these global results, it can be said that identification scores differed considerably from FA ratings<sup>115</sup>. For instance, FL learners' /ε/ identification scores were near native-like, whereas accent ratings were far from being native-like – particularly in the word *tests*. These results, however, are in line with findings of Munro et al. (1996) who illustrated that a higher degree of FA on TL vowels on the part of native Italian speakers of English did not necessarily entail the target vowels having failed to have been correctly identified (produced) as intended.

#### 6.3.2.1. Effect of onset age of FL learning

The first age effect observed had to do with the differences in identification scores between the NE control group and FL learners. Overall, NE foils obtained significantly higher percent correct identification scores than Spanish and Catalan learners of English. To be exact, the significant results were located in /i/ (*tea*), /ɪ/, /æ/, and /u/. Differences in identification scores for /i/ (*speak*), /ε/, and /ɒ/ were nearly always significant or, instead, close to being significant. In contrast, the differences in identification scores for /ʌ/ between the NE group and FL learners failed to be significant, although the control group still obtained higher correct identification scores.

<sup>115</sup> Recall that correlations between accent ratings and vowel identification scores could not be performed, since listeners' ratings could not always be pooled into a single mean rating.

When FL learners were considered separately from NE foils, the effects of onset age of FL learning on the identification scores were more conclusive than those on the FA ratings. What is more, the starting age effects noted were in agreement with observations in L2 acquisition research of an initial late starting age advantage and findings of perception of English FL sounds reported in Section 6.1.1 above. Thus, although in the initial stages of FL learning the four age groups did not differ significantly in the correct identification scores for English vowels, older beginners obtained higher scores than younger learners. Moreover, in some instances differences approached significance in favour of older beginners – i.e. C1–D1 in /i/, A1–D1 in /u/, and B1–D1 in /æ/ and /ʌ/. Similarly, older learners (11-year-old beginners) with 416 hours of instruction produced English vowel sounds as intended at higher correct frequency rates than 8-year-old beginners. Finally, in the long-term the clear-cut late starting age advantage had disappeared. However, unlike younger starters' somewhat more accurate perception of English sound contrasts than that of 11-year-old beginners when formal instruction amounted to 726 hours, in the production of English vowels 8-year-old beginners did not always surpass 11-year-old beginners in the long-term. In fact, they were judged to have produced /i/ (*tea*), /ɪ/ (*this*), /æ/, and /ʌ/ as intended at lower frequency rates (vs. /i/ in *speak*, /ɒ/, and /u/, which were identified as intended at distinctly higher frequency rates). In that respect, it could be suggested that in the long-term 8-year-old starters were still in the process of catching up with 11-year-old beginners in the production of English segments, while in the perception of FL sounds younger learners had already caught up with older learners when they all had received 726 hours of formal instruction in English. In addition to agreeing with previous findings, the latter suggestion would corroborate the SLM's prediction that TL sound perception precedes production of TL segments, thereby extending it to a formal learning context.

#### 6.3.2.2. Effect of exposure

The correct identification scores obtained for all age groups as a function of exposure to the FL yielded inconclusive results. As was the case of the FA ratings on English vowels, together with the accent ratings on words and segments obtained in

Study 1, 8-year-old starters tended to produce English sounds – whether new or similar – as intended at lower correct frequency rates along with an increase in formal instruction. The vowel identification scores then further disconfirmed the beneficial exposure effects observed for younger learners on the attainment of more accurate perception of English sounds.

As for 11-year-old beginners, an increase in experience in English did not result in a significant better production of English sounds. On the whole, 11-year-old starters in the initial stages of FL learning and in the long-term (i.e. 200 and 726 hours of exposure) varied little in their vowel identification scores, parallel to their discrimination scores on English sound contrasts at these two points of exposure. Nevertheless, consistent with previous results reported on this age group's both perception and production of English sounds, halfway through their learning of the FL in the current research design (i.e. 416 hours of instruction), 11-year-old beginners tended to produce English /i/ (*speak*), /ɒ/, and /ʌ/ as intended at lower frequency rates. On top of that, 11-year-old beginners exhibited a higher degree of mixed exposure effects halfway through their learning. Therefore, and contrary to what has just been stated, they produced English /ɪ/, /æ/, and /u/ as intended at distinctly higher frequency rates when formal exposure to English amounted to 416 hours.

As mentioned in Section 5.3.2.3.2, the production of English vowels by 14-year-old and adult starters could only be examined at the descriptive level. Even then, the two older learner groups displayed the same pattern of inconclusive exposure effects as did 11-year-old beginners.

In summary, mixed exposure effects were obtained on all the age groups' vowel identification scores. Furthermore, in the long-term no starting age group produced the English target vowels as intended at native-like levels, although it should be mentioned that learners' /ɛ/ identification scores closely resembled those of English foils. Once again, all these findings might be interpreted in light of Flege's (1991b) accented L2 input hypothesis. Last, Catalan and Spanish NSs' failure to attain native-like vowel identification scores, in conjunction with their medium foreign-accented production of English vowels, as a function of exposure to English further corroborates the observation that more than 6–7 years of instruction in the FL are needed in a formal learning context in order to produce English sounds accurately (García Lecumberri & Gallardo, 2003).

### 6.3.2.3. Effect of dominant L1(s)

As reported in Section 5.3.2.2.3, the NE control group produced /i, ɪ, ε, æ, u/ as intended at significantly higher rates (i.e. native-like, in this case) than did the three language dominance learner groups. Concerning /ɒ/, the differences in scores approached significance in favour of English foils. Likewise, English NSs produced /ʌ/ as intended at higher frequency rates than Spanish and Catalan NSs, but not to the extent of yielding significant differences. The somewhat lower percent correct identification scores that NE foils obtained for /ʌ/ than for other vowels are in line with previous findings, such as Flege, MacKay, and Meador (1999). In that study, English NSs obtained a mean intelligibility score for /ʌ/ of 86%, in addition to one of 84% for /æ/, whereas for the remaining English vowels their mean intelligibility scores ranged from 91% to 99% (p. 2978, Table II).

Among FL learners, there were no differences between the three language groups in the identification scores for the seven English vowel segments under investigation. This finding provides further support for previous results of a lack of L1 dominance effects on the perception and production of English sounds, reported in the present dissertation.

For the most part, the degree to which native Spanish and Catalan subjects produced /i, ɪ, ε, æ, ɒ, u, ʌ/ as intended is in accordance with previous findings of Romance language learners of English in both immersion and formal instruction settings. In the same way, the misidentification patterns obtained agree with the most frequently reported substitutes for those seven English vowels. To be exact, FL learners produced /ε/ as intended at very high frequency rates, approaching native-like levels. In the event of mispronunciations, [e] was the only reported substitute. All this parallels findings of Spanish learners of English in naturalistic settings such as Flege, Bohn, et al. (1997). Also, in agreement with Flege, Bohn, et al. (1997) and Cebrian (2002c), subjects produced /i/ as intended in many instances – though at lower frequency rates than /ε/ – [ɪ] being one of the most common substitutes. Unlike previous studies, another reported substitute for /i/ was the pure tense vowel [i], which in fact was heard significantly more

often than its lax counterpart [ɪ]. As noted in Section 5.3.1.6, in the vowel identification task a specific distinction was made between English /i/ which shows more formant movement (i.e. it is diphthongised) and the pure tense vowel [i], whereas in previous studies (e.g. Flege, Bohn, et al., 1997) no such difference had been contemplated. All this might account for subjects' somewhat lower /i/ identification scores<sup>116</sup>, instead of any other effects such as NNS input they might have been exposed to. Likewise, both the identification scores for /u/ and its most frequent misidentification pattern – pure tense [u] – might be interpreted in the same manner.

Consistent with previous findings (e.g. Flege, Bohn, et al., 1997), FL learners produced /ɪ/ as intended at noticeably lower frequency rates (about 50% of the time), while substituting the lax vowel for its pure tense and diphthongised counterparts, namely [i] and [ij]. Similarly, /æ/, /ɒ/, and /ʌ/ were identified as intended less often than other vowel sounds (range = 45% – 65%). Moreover, a wider variety of substitutions was observed for /æ/ and /ʌ/. Thus, both /æ/ and /ʌ/ were heard as instances of Spanish-like [a] and English [ɑ]. In addition, target /æ/ and /ʌ/ were misidentified with each other – [ʌ] and [æ], respectively. All in all, the findings agree with previous research (e.g. Flege, Bohn, et al., 1997; Flege, Munro, & Fox, 1994; Munro et al., 1996). Finally, FL learners' production of /ɒ/ was identified as intended the least often (in roughly 45% of instances), being heard instead as [ɔ], [ʊ], [u], [ʌ], and [ɑ]. The low identification scores for /ɒ/, together with [ʌ] as one of its most common misidentifications, are in sharp contrast to Catalan and Spanish NSs' successful discernment of sound contrasts involving /ɒ/ and /ʌ/ (but see native Italian subjects in Flege & MacKay, 2004). In this case, it might be hypothesised that to a certain extent NS listeners' own variety of English might have influenced their ratings, for Canadian English does not have an /ɒ/-/ɑ/ distinction and Canadian English /ɒ/ is somewhat lower (more open) in the vowel acoustic space than British English /ɒ/. Even if judges' variety of English had an effect on their accent ratings on /ɒ/, it seemed to be minimal as they did distinguish English NS vowel productions

---

<sup>116</sup> In fact, if both intended /i/ and pure tense [i] responses had been considered correct productions of English /i/, subjects' final identification scores for /i/ would have been higher than 70% (range = 70% – 100%), thereby fully corroborating previous findings.

from those of FL learners (in addition to the most common misidentification pattern being other than /a/, as would be expected based on the non-existent distinction /ɒ/-/ɑ/ in Canadian English).

#### 6.3.2.4. Effect of gender

As was the case of the FA ratings obtained in Studies 1 and 2, female subjects produced English vowels as intended at higher correct rates than male subjects, even yielding significant differences in the identification scores for /ɛ/ and /æ/ and approaching significance for /u/. Furthermore, female participants' superiority was consistent across all learner groups, but like the accent ratings, the differences were nonsignificant. Thus, the findings of identification scores fully confirmed the hypothesis that gender differences do exist at least in the pronunciation of the TL, thereby suggesting that female FL learners are likely to be better imitators of English sounds than are male FL learners.

### 6.4. General discussion

This chapter was aimed at discussing the effects of the factors of onset age of FL learning, exposure to the FL, L1, and gender on the results derived from the perception and imitation tasks that the subjects had performed. It is worth noting that in the interpretation of results a special emphasis was placed upon the comparison between 8- and 11-year-old beginners' perception and production of English sounds, since the introduction of English as an FL learning at the age of 8 (as opposed to the age of 11) in the school setting constituted one of the major changes from the former to new curriculum in the Spanish educational system.

Firstly, it was found that onset age of FL learning influenced Catalan and Spanish NSs' perception and production of English sounds in a differing manner. On the one hand, late starting age (11, 14, and 18+ years) was more favourable to the perception of TL sounds in the initial stages of FL learning, while in the long-term an early starting age (8 years) advantage over late starting age in English segment perception had become

evident. As stated before, these findings are in line with Krashen et al.'s (1982) generalisations and the predictions of the SLM, in addition to findings of research conducted in L2 naturalistic contexts.

On the other hand, starting age effects on the production of English sounds were not so clear-cut. When FA ratings on words, consonants, and vowels (both from Study 1 and Study 2) were considered, early starting age was more advantageous in the initial stages of FL learning. In contrast, late starting age proved to be more beneficial in pronouncing English words and segments as less foreign-accented in the long-term. In that case, the late starting age advantage observed after 726 hours of exposure to the FL corroborates findings of studies conducted in formal learning settings (García Lecumberri & Gallardo, 2003). However, several instances of late starting age and early starting age advantage in the accent ratings on words and segments were noted in the short-term and long-term, respectively, hence making age effects on the production of English speech inconclusive. Yet when vowel identification scores were examined (Study 2), starting age effects were more defined. In line with the initial late starting advantage reported for the perception of English sounds, older starters produced English vowels as intended at higher frequency rates in the initial stages of FL. However, unlike perceptual findings, in the long-term younger learners did not produce English target vowels as intended at consistently higher frequency rates than older learners.

Both the more conclusive effects of onset age of FL learning on perception of English sounds and the more contradictory or mixed starting age effects on the production of English sounds concurred with the observation that no age group perceived English sounds within the NS range. Finally, all of the above can be taken as evidence that perception leads production in the learning of an FL phonology, thereby extending the hypotheses of the SLM to an instructed-classroom setting.

Secondly, exposure to the FL (200, 416, and 726 hours) yielded somewhat inconclusive results. As was the case of onset age of FL learning, early starters benefited the most from an increase in instruction in the perception of English sounds, while gains in exposure did not lead to a noticeable improvement in older starters' perception of FL sounds. On the contrary, 8-year-old beginners did not produce English sounds as less foreign-accented or as intended as a function of exposure to the TL, while 11-year-old beginners often produced English sounds more accurately or with a lesser degree of FA, along with an increase in formal instruction in English. Nevertheless, learners' overall better performance in the TL phonology as a result of a longer exposure to the FL was

not marked and failed to be consistent for FL sounds tentatively classified as new vs. similar TL sounds. Moreover, a common pattern emerged halfway through their learning: with the exception of 8-year-old beginners' perception of FL sounds, both in the production and perception of English segments 8- and 11-year-old beginners' performance was distinctly poorer when they had received 416 hours of formal instruction. One explanation might lie in the bimodal or U-shaped pattern in performance that has been reported in L2 acquisition research. However, a more likely explanation that might account for this finding of mixed exposure effects has to do with the accented L2 input hypothesis.

As for the factor of dominant L1, Catalan and Spanish dominant speakers as well as Catalan/Spanish balanced bilinguals were found to perceive and produce English sounds very similarly. This was attributed to the fact that all learners were familiar with and exposed to the phonology of both languages on a daily basis. Additionally, the three language dominance groups' difficulty in discerning and producing English sounds such as the tense/lax vowel distinction and consonant voicing distinction in word-final position was in agreement with previous findings of Romance language speakers of English.

This study provided somewhat more conclusive evidence about gender effects on subjects' performance in the TL phonology, specifically in relation to the production of English sounds. In line with popular observations, in the short-, mid-, and long-term of the present study, female subjects consistently produced English segments with a lower degree of FA and as intended at higher correct rates than male subjects. In turn, these findings further suggested the hypothesis that female learners might be better imitators of FL sounds than male learners in a formal instruction setting.

Finally, two methodological issues that might have contributed to mixed starting age and exposure effects should be considered. The first aspect is concerned with the NS listeners who participated in Studies 1 and 2. As mentioned above, listeners were able to identify NE foils successfully, despite their varying degrees of sensitivity to foreign-accented speech. Moreover, the high intra-rater consistency coefficients obtained in Study 2 should be taken as indicative of listeners' valid and reliable accent judgements. However, up to a point the presence of background noise in the recordings was thought to have had a greater effect on judges' assessment of FL learners' accented word and segment productions. The second methodological issue is also related to NS listeners – particularly, the FA rating scale employed and the total number of participating judges in each experiment – in conjunction with FL learners' own characteristics. It was mentioned



earlier that 6–7 years of formal instruction in an instructed-classroom setting do not equal the amount of exposure that an L2 learner receives over 6–7 years in an L2 immersion context (García Lecumberri & Gallardo, 2003; Singleton, 1995). The subjects in the present dissertation had received a maximum amount of exposure of 7.5 years on average that, in fact, corresponded to a total of 726 hours of formal instruction in English. Thus, differences in exposure amounts between formal learning and naturalistic settings, together with subjects' starting ages of FL learning – all past the age at which L1 phonetic categories have already been established – suggest that the average 7.5 years of formal instruction in English had not been large enough to bring about noticeable differences (if any) in the pronunciation of the TL by native Catalan/Spanish learner groups. Consequently, the 9-point scale of FA that listeners were asked to use in both Study 1 and Study 2 might have been too fine-grained, in spite of previous research findings advocating the suitability of 9- and 11-point rating scales in the assessment of FA (e.g. Southwood & Flege, 1999). Alternatively, as Piske et al. (2001) put forward, it might be that a larger number of listeners is needed if differences in foreign-accented speech across very similar learner groups such as those of the current dissertation are to be detected.

