Developmental patterns in the adult L2 acquisition of new contrasts: the acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers

M.Sc. dissertation
by Paola Escudero

Department of Theoretical and Applied Linguistics
University of Edinburgh

September 2000

Table of contents

Introduction

Part of the phonology of a language consists of sound distinctions that the speakers perceive and produce. These sound distinctions, in turn, are signalled by a number of sub-properties that integrate to constitute phonological contrasts. There is a many-to-one relation between the properties that signal a contrast and the contrast itself (Repp 1984). The information or sub-properties that signal contrasts are called phonetic cues. The speakers of every language have to learn what aspects of the phonetic signal function as cues and how to use each cue (Scobbie 1998).

Furthermore, the L1 speakers of a language assign relative amounts of perceptual attention to the different acoustic cues that signal a contrast. That is to say they have a particular weighting to such cues: some of them are primary and other secondary (Nittrouer 2000, Scobbie 1998). For instance, not only do L1 Scottish English speakers know that the principal cues to the tense/lax vowel contrast /i/-/ɪ/ are spectral and durational cues but also that spectral information is primary. This is manifested in their performance in identification tests that contained stimuli with a single acoustic cue (Escudero 2000a).

The main purpose of the present study is to investigate the adult L2 perception of new vowel contrasts as well as the way in which such perception may develop. For example, can L2 speakers learn to perceive a new contrast? If so, do they weight the acoustic information in the same way that L1 speakers do? And if not, could we think of the different patterns in their performance as being part of a sequential development? The studies that have compared the adult L2 perception of the vowel contrast /i/-/ɪ/ (Fox et al 1995, Bohn 1995 and Flege 1997) have arrived at different conclusions regarding the cue weighting and general perception of this contrast by native Spanish speakers. Furthermore, they do not suggest any mechanism for the development in the L2 perception of this group

of subjects. Consequently, I carried out a study that aimed to find answers to the L2 perception questions put forward in this paragraph as well as to provide more evidence that could avoid the contradictory nature of the claims made in previous studies. This new study tested the perception of the Scottish English /i/-/ɪ/ contrast by Spanish speakers. The motivation, findings and conclusions will be reported here.

The present paper consists of three chapters. Chapter 1 presents the relevant theoretical background, questions and hypotheses. The study and experimental design are presented in Chapter 2: this chapter includes a description of the subjects, methods, procedures, results and discussion. Finally, the conclusions and implications of the study as well as suggestions for follow-up research are stated in the last chapter.

Chapter 1 Theoretical background

In this chapter, the general trends, ideas and models that aim to describe and explain the L2 speech features will be presented. This literature review will include a consideration of previous studies that investigate the perception of English /i/-/ɪ/ contrast. I (Fox et al 1995, Bohn 1995 and Flege 1997) will critically review these studies and their arguments and weaknesses will be taken into account for the set of ongoing research questions and hypotheses that are tested in the present experimental study.

1. The adult L2 perception of sounds

A number of studies have been devoted to the perception of non-native sounds (see Strange 1995). There is an increasing interest in knowing the way in which L2 perception works and how different it may be from native perception. Most of the approaches to L2 perception suggest that the L1 background has a strong influence in the way the sounds of the *target language* (TL) are perceived, at the very least during the first stages of learning (Ingram and Park 1997).

Two of the most important and influential approaches to L2 vowel perception explain L2 speakers' behaviour mainly as a function of their previous, linguistic experience. These two approaches form two different but not incompatible theoretical models for non-native perception. I will summarise, criticise and test some of their claims. The models referred to here are the *Speech Learning Model* or *SLM* (Flege 1995) and the *Perceptual Assimilation Model* or *PAM* (Best 1995). Many of the current ideas in L2 speech research are reflected in these two models, hence the importance of reviewing their main arguments.

1.1. The Speech Learning Model

The SLM is an all-inclusive hypothesis about L2 speech, both perception and production. What is of interest for the present study is the model's claims regarding L2 perception. The model is a mentalist one, in which language-specific sounds are represented in long-term memory: the resulting representations that a speaker has of her L1 sounds constitute *phonetic categories* (e.g. /s/, /t/, /i/, etc.). It is possible to infer from the model that TL sounds that are matched with different L1 sounds will be discriminated in a native-like fashion. However, Flege also argues that there is no guarantee that they will be fully perceived in the way native speakers perceive them. The model, then, deals with the non-native perception of acoustic information, these ideas will be reported in the following section.

1.2. The Perceptual Assimilation Model

This model assumes that non-native perception is filtered by linguistic experience. That is, new sounds will get asssimilated, in one way or another, to the categories that already exist in the speakers' first language (L1). Consequently, the model suggests degrees and types of assimilation of TL sounds to the L1 perceptual categories. It also suggests the degree of discriminability for the sounds that form a contrast in the TL. In order words, it provides us with the different ways in which a sound can be assimilated and the results that this assimilation has for distinguishing contrastive TL sounds. Below are two charts that display these possibilities.

Degrees of assimilation of TL sounds to L1 perceptual categories:
a) Assimilated to an L1 category - A good exemplar of the L1 category - An acceptable but not ideal exemplar of the category - A notably deviant exemplar of the category b) Assimilated as uncategorisable speech sound c) Not assimilated to speech (non speech sound)

Table 1. Adapted from Best 1995

Assimilation and discrimination of L2 contrasts
1) Two-Category Assimilation Each L2 segment assimilates to a different L1 category. Excellent discrimination **2) Category-Goodness difference** Both segments are asssimilated to the same L1 category but they differ from the L1 "ideal", i.e. one is acceptable and the other one is deviant. Moderate - very good discrimination. **3) Single-category assimilation** Both L2 sounds are assimilated to the same L1 category, equally discrepant from the "ideal", i.e. both either equally acceptable or equally deviant. Poor discrimination. **4) Both uncategorisable** Both sounds fall within the L1 phonetic space (depending upon proximity to each other and to L1 sounds Poor - very good. **5) Uncategorised vs. categorised** One assimilated to an L1 category and the other falls in the phonetic space (it is considered speech) but it is not categorised Very good **6) Nonassimilable** (non-speech sounds) Both fall outside the speech domain. Good - very good

Table 2. Adapted from Best 1995.

From the tables above we coudl gather that the PAM predicts the level of difficulty that non-native sounds can have for L2 speakers according to the way these speakers first assimilate them to their L1 categories. Some of those types of assimilation are the ones that Spanish speakers of English could potentially have when first encountering the /i/-/ɪ/ Scottish English contrast (e.g. single-category, two-category, category-goodness difference, among others). However, these tables do not include suggestions for the way in which the acoustic or sub-phonemic information is perceived by L2 speakers. The PAM model does consider this phenomenon: its arguments and predictions with respect to it are presented in the next section.

2. The L2 perceptual attention and use of phonetic cues

What we saw above is what the models predict for the perception of non-native sounds in general. Let us turn, now, to what they suggest in terms of the perception of the acoustic/sub-phonemic information involved in the new sounds and contrasts.

2.1. L2 perception of acoustic information according to the SML

One of the main ideas in the SML model is the fact that sounds that are similar to L1 ones are less easily perceived in a native-like way than are new sounds. This, in turn, relates to the idea of *perceptual equivalence*: an L1 sound and an L2 one are merged into one category. In other words, a single phonetic category (from the L1) will be used to perceive sounds that are perceptually similar. Therefore, the theory predicts a degree of difficulty for acquiring L2 sounds. The closer a TL sound is to the closest L1 sound, the more difficult it is to have a new category for it.

This model explains the lack of ability to create new categories for *similar* sounds on the basis that L2 speakers are not able to perceive the *sub-segmental* (non-phonemic) features in which the L2 and L1 sounds differ. However, the model also suggests that L2 learners may, in time, be able to perceive these non-phonemic features, that is, in terms of their experience in the TL. L2 learners are, in principle, capable of forming new categories that resemble native-like ones, even for sounds that are similar to the L1 sounds. Furthermore, this model claims that it is only the perception of sub-phonemic features that can make L2 speakers form a new category for the sound and achieve complete native-like perception.

Nevertheless, the perception of the sub-phonemic information that distinguishes TL from L1 sounds may not be enough. It may turn out that the *relative* effect such sub-phonemic information has on the perception of contrasts is also important. This idea is supported by the fact that a phonological distinction is most likely to be signalled by a number of different acoustic cues (Repp 1984). However, these diverse acoustic information seems to have differential effects on the overall perception of phonemic contrasts (*cue weighting*).

From the discussion above, it seems that it is not enough to be able to perceive acoustic, sub-phonemic information conveyed in TL sounds to have native-like perception. L2 speakers also need to know what the relative effect of these cues is for native speakers to

be able to develop a new native-like category. The SLM model does not, however, seem to account for this further step. I will take up this point again later.

2.2. The PAM and the non-native perception of lower level phonetic information

The PAM model claims that what we perceive are the articulatory gestures (i.e. bilabial, alveolar, velar, uvular, dental stop, high, front, etc.) that the linguistic input manifests. There is no mental representation of sounds; what happens is that our linguistic, perceptual mechanisms get tuned to particular high level features that the articulatory gestures convey. That is, at the beginning young children listen to and detect every auditory feature (lower level ones), and later they learn to detect only high level features; i.e. those that signal contrasts (phonetic features) in their first language. This makes the task of perceiving L1 sounds easier, once the child is able to process the high level features that form the phonological system of its L1.

For non-native perception, the process is the inverse of this. Because these speakers have already tuned their linguistic perceptual device to particular high level features, they would have difficulty in detecting the features that the TL sounds have. That is, the TL may have high level features that signal contrasts that are actually low level ones for the L2 speakers. Consequently, the TL sounds will get assimilated to the L1 ones and we saw in what ways the process of assimilation could happen, according to this model.

After looking at what the models suggest about the perception of L2 sounds/contrasts in general, let us turn to their insights in the learnability and development of L2 sounds.

2.3. The Learning and development in the perception of new sounds

The SLM predicts varying degrees of difficulty for acquiring L2 sounds. The closer a TL sound is to the closest L1 sound, the more difficult it is to have a new category for it. Moreover, this model suggests that L2 learners are, in principle, capable of forming new categories that resemble native-like ones, even for sounds that are similar to the L1 sounds.

Likewise, the PAM predicts the degree of discriminability but not, at least not clearly, the way native-like perception is acquired. The fact that L2 speakers can discriminate non-native sounds on the basis of L1 categories does not account for the way in which the

learning of new categories happen. Best 1995 claims that the model allows for a learning process. However, she does not explain how this process works. It could be inferred that what L2 speakers have to do is learn to detect the low level gestures conveyed in TL sounds to be able to have new perceptual categories. They have to go from high level to low level gesture detection but how they do that is not explained.

On the other hand, we also know that the successful perception of non-native sounds should also include the cue weighting that native speakers of the TL have. Presumably, L2 speakers may also transfer the L1 weightings, especially when the TL sounds are similar to the L1 ones. However, if the dimension that is considered by the learner to be the strongest for L2 vowel contrasts cannot actually give information about a particular segment, then they may come to rely on other cues.

From the two theories, it may be inferred that L2 native-like perception is possible provided that L2 speakers detect sub-phonemic information in the TL sounds and *build up* new categories based on it. What is not mentioned is that they should also detect the relative cue weighting of the information involved in the TL catgories. Of course, this process would be extremely difficult (for some sounds more than for others), because it seems that the input does not clearly manifest which cue is being preferred. The PAM would be able to explain how cue weighting is acquired in terms of a new detection of high level features for the TL. Likewise, the supporters of the SLM may be able to cope with this phenomenon by suggesting that the phonetic categories are represented in long-term memory with their relative, acoustic weighting. Nevertheless, an extension and a better explanation of the above mentioned phenomenon (cue weighting) would seem to be necessary.

Some studies carried out within the SLM framework have investigated the relative effects that acoustic cues have in the perception of L1 and L2 contrasts. A good number of them (Fox et al 1995, Bohn 1995 and Flege et al 1997, among others) have looked into the perception of English tense/lax vowels by English and Spanish speakers (among other language groups). The findings and conclusions of these studies provide the specific background and motivation for the study that will be reported on in the next chapter.

Basically, the previous studies aimed to explain what the perception of non-native sounds was using English vowel segments as targets and Spanish speakers as the L2 group. From the careful analysis of these studies some problems and contradictions were found. Consequently, the aim here is to use the same contrast and subjects but with better methods to gather more reliable evidence for the L2 perception and development of new contrasts.

3. The L1 and L2 perception of tense/lax vowel contrasts

Before looking at the studies that investigate the L1 and L2 perception of English tense/lax vowels, the acoustic features of the segments that will be tested in this tudy are presented. Later, the studies that tested these segemnts (/i/-/ɪ/) in L1 and L2 perception will be discussed.

3.1 Spanish and English vowels: acoustic information

Spanish has a five-vowel system formed by /i, e, a, o, u/ and General American English has 11 monophthongs, /i, ɪ, e, ɛ, æ, ɑ, ʌ, o, ɔ, ʊ, u/, most of which are actually *ingliding* and *outgliding* diphthongs. Phonologically, Spanish vowels only differ in spectral information (steady state formant frequencies). According to Fox et al 1995, Spanish speakers use only two dimensions, namely, high-low and front-back for distinguishing vowel contrasts. English speakers use three dimensions, the first of which includes durational information. According to this study, Spanish speakers do not use duration to signal vowel contrasts, whereas English speakers do. It is suggested that spectral and durational cues signal tense/lax English vowel contrasts (Strange 1989, Klatt 1976, Geigerich 1992).

An acoustic study (Bradlow 1995) of General American (GA) and Spanish vowels describes the spectral features for these vowel segments. The values tell us the relative distance of the spectral values for similar vowel segments in the two languages. However, it is worth bearing in mind that the spectral features of Scottish Standard English (SSE) are different from GA. The reason why the GA value sare presented here is because all the previous studies have tested the perception of this variety of English. The values for SSE /i/ and /ɪ/ are presented in the materials (for the present study) section.

Spanish (CVCV)			English (CVC)		
	F1 (s.d.)	F2 (s.d.)		F1 (s.d.)	F2 (s.d.)
/i/	286(6)	2147(131)	/i/	268(20)	2268(207)
/e/	458(42)	1814(131)	/ɪ/	463(34)	1995(199)
/a/	638(36)	1353(84)	/e/	430(45)	2200(168)
/o/	460(19)	1019(99)	/ɛ/	635(53)	1796(149)
			/æ/	777(81)	1738(177)
			/ɑ/	640(39)	1354(134)
			/ʌ/	780(83)	1244(145)
			/ɔ/	620(72)	1033(135)
			/o/	482(30)	1160(47)
			/ʊ/	481(36)	1331(161)
			/u/	326(26)	1238(160)

Table 3. Adapted from Bradlow 1995.

One interesting thing to note is the fact that English /ɪ/ and Spanish /e/ have very similar values for F1 and F2. Especially for F1 which, articulatory speaking, represents vowel height.

In the following study, we wanted to test Scottish Standard English to see if the L2 perception of it matched the one reported in the previous studies that tested tense/lax vowel contrasts. This variety of English has only 9 monophthongs, as opposed to the 11 of GA, (/i/, /ɪ/, /e/, /ɛ/, /a/, /ʌ/, /ɔ/, /o/, /u/). These vowels are straight monophthongs unlike the American vowels just discussed and that is the reason why steady state formants and intrinsic duration are the main and possibly only acoustic information involved in the /i/-/ɪ/ contrast. The acoustic values taken for this variety are shown in the methodology section of the experiment.

3.2. The L2 perception of tense/lax vowel contrasts

If one looks at the perception of the contrast /i/-/ɪ/ by Spanish speakers both the features that the sounds manifest and suggestions from previous studies seem to be of particular interest. Furthermore, this vowel contrast will help to explain the idea of cue weighting and the use of sub-phonemic information by L2 speakers.

Following the PAM's patterns of assimilation there are two possibilities of assimilation for the vowels that form the contrast (the order in which they are displayed here has no importance). These possibilities will be analysed following an acoustic, rather than an articulatory, approach.

	First Possibility		Second Possibility	
L1 English	/i/	/ɪ/	/i/	/ɪ/
L1 Spanish	/i/		/i/	/e/

Table 4. Possibilities of L2 assimilation to L1 categories according to PAM.

The first possibility suggests that the sounds may be assimilated to a single category in the L1. However, this assimilation, in turn, can also manifest a *category-goodness difference* pattern, as suggested by PAM. It may be that one of the sounds is a good exemplar of the category and the other is a deviant exemplar. If we consider the two dimensions involved in the English vowel contrast, both sounds are equally deviant. That is, English /i/ is longer than Spanish /i/, and English /ɪ/ has different spectral features.

Considering the above discussion, it is not clear why the two sounds are assimilated to the same L1 category. It seems that the L2 speakers may be using two different dimensions for assimilating the sounds to the same category. They use spectral information for English /i/ and durational information for /ɪ/. However, spectral information seems to have the strongest effect in these speakers' L1 vowel discrimination. Furthermore, it is worth mentioning again that Fox et al 1995 suggests that Spanish speakers do not have durational cues accessible. Whether the claim is that they do not have access to duration at all or that they have it as a sub-phonemic feature is not clear.

A different account for the same assimilation suggests that *orthography* may be the cause of SS assimilation of /ɪ/ to Spanish /i/ (Flege 1997). The claim is based on the fact that although the vowel in the words 'bit', 'ship', etc is pronounced as /ɪ/ in English, it is represented as an "i", which is a letter that always represents /i/ in Spanish. Young-Scholten 1995 presents a more detailed account of how phonological L2 acquisition can be biased by orthographical representations.

The second possibility has already been touched on. If the L2 learners were only using spectral information (the strongest cue in their L1) when assimilating TL segments to their L1, the most likely assimilation would be English /ɪ/ to Spanish /e/. It would be interesting to see what previous studies that investigated the same phenomenon suggest for the starting point in the perception of English /i/-/ɪ/ by Spanish speakers. Therefore, at this point we can turn to the findings of those studies in terms of the pattern of assimilation assumed and the use of the two acoustic dimensions (spectral and durational) for the L2 group when perceiving the /i/-/ɪ/ contrast.

3.3. Studies investigating the L1 and L2 perception of the contrast: findings and problems

It is worth mentioning again that Fox et al 1995 concluded that Spanish speakers of English do not have access to durational cues when perceiving vowel sounds. Furthermore, these authors suggest that this is the reason why these speakers have seriuos difficulty in perceiving and producing English tense/lax vowels (e.g. /i/ and /ɪ/). They also claim that English speaker do use this cue for the same vowel distinction.

Bohn 1995 assumes single category assimilation to be the starting point and that discrimination comes later. This author does not explain why he assumes this pattern of assimilation. However, his findings suggest that Spanish speakers rely on durational cues more than English speakers to discriminate the sounds. He concludes that this behaviour indicates an *interlanguage* or *general perceptual strategy.* He also suggests that when spectral information is not available (because of what he calls a *desensitisation* process) speakers will use duration as a universal, acoustic feature.

The problems with this study are that, 1) the stimuli did not vary in the same steps for both acoustic cues (eleven steps for spectral and three for durational cues): it may have been that Spanish speakers used durational cues because it was the dimension that varied the least. 2) The idea of desensitisation is not clearly explained. 3) It does not talk about what consequences further exposure to the TL could have (it does not have an experienced speaker group).

Flege et al 1997 seems to have used the same stimuli as Bohn did. According to these authros, the relative effect of the two cues in Spanish speakers is similar in percentages and that none of those results is significantly different to the results shown for English speakers. This contradicts Bohn 1995's conclusions. However, it is considered here that the fact that Spanish speakers weight the two cues similarly has a meaning, perhaps not a statistical but observational one. It may be that Bohn's findings are still correct. but this clearly needs further testing.

These three studies do not present conclusive evidence for the Spanish speakers' cue weighting of the /i/-/ɪ/ distinction. Some of the conclusions they draw seem to be contradictory. Consequently, there is a need for more reliable evidence. The research questions and the hypotheses will now be presented. They follow from the questions and theoretical framework that previous studies used and they try to find answers for some of the L2 behaviour that those previous studies did not fully explain.

4. Research questions and hypotheses

After presenting the theoretical background above, it is considered here that some questions remain unanswered. In this section, three of those questions (the most important ones) will be put forward after a further motivation. These questions will become the research questions that, in turn, motivate the experimental design that is reported in the following chapter. Likewise, answers for those three questions are proposed, these answers constitute the hypotheses are tested through the experiment.

4.1. The perception of phonetic information vs. its relative weighting

The two models for L2 speech perception, namely the SLM and the PAM, above mentioned above do suggest a possibility for the learning of new contrasts (Best 1995 and Flege 1995). For instance, Flege argues that learning to perceive a new sound/contrast is a by- product of the L2 listener's ability to perceive the phonetic information that the sound contains. That is, to this author this information is not accessible at the beginning mainly because it is non-contrastive in the L1. However, when it becomes available, the learners are likely to perceive new sounds or contrasts (Flege 1981). Likewise, Best suggests that there is a time in the L2 process where the single category may split and the learner acquires the perception of the contrast (Best 1995).

From what the above authors claim, we could gather that adult L2 learners could acquire a new contrast by listening to the phonetic cues involve in this contrast. This view suggests that the L2 acquisition process stops there, either because the learners are already native-like in the perception of the contrast or because they cannot go beyond this achievement. The first scenario seems to be the most favoured by the two models: L2 speakers managing to have two categories instead of one and being equally good as L1 speakers of the TL.

However, we also know now that all contrasts are signalled by more than one phonetic cue (normally by two most important ones). Furthermore, one of the cues may be primary or most important. Therefore, it might be the case that the L2 learners are able to perceive the contrast on the basis of a secondary cue, in which case they are not completely native-like perceivers. In fact, previous data (Escudero 2000b) shows that some L2 speakers consistently and exclusively use a secondary phonetic cue (secondary in terms of L1 perception) in their perception of a vowel contrast. L1 speakers of English use spectral information as a primary cue to the tense/lax vowel contrast /ɪ/-/i/, durational information is used but it is secondary (tested in Escudero 2000a and re-tested in Escudero 2000b). Spanish L2 speakers of English are shown to have a very good performance (deep slope and clear boundary in the middle of the continuum) for a durational continuum /ɪ/-/i/. However, they considered the two sounds to be the same or they were not sure about their decisions in a spectral continuum (Escudero 2000b). In this case, the L2 speakers are

making no use of the spectral information, which may mean that it is not accessible to them.

The analysis of further data (Escudero 2000b) also showed that some L2 speakers used durational cues more than native speakers do. That is, they are shown to have access to spectral information but still consider durational information as more or equally important. A third group, one with fewer subjects, had a perception that weighted the two cues in a native-like fashion.

The consequence of the previous discussion is that some questions still remain. The one below constitutes the first research question of the present paper. This question and hypothesis that aims to answer it are stated as follows:

Q1: Can all L2 speakers that correctly place exemplars of a new contrast in two categories be considered native-like perceivers?

H1: No, they cannot. Assuming that they started by having a single category and later managed to learn the new contrast by listening to phonetic information, they still have to detect all the relevant information and/or its use and relative weighting in L1 perception.

4.2. L2 Performance and development

Some studies have addressed the way in which L2 adult perception develops (cf. Flege 1995 and Gass 1984) and suggest developmental variables for some learner cases that were shown to be close to native performance. However, even though these and other studies may argue for a sequential development in L2 perception, they do not define or determine any specific phases/stages. Ingram and Park 1997 make an explicit reference to stages but make no attempt to suggest what the nature and features of these stages could be and/or what learners might be included (Ingram and Park 1997: Introduction).

It seems to be clear that most L2 learners do not perform in a consistent manner as they differ from L1 speakers in various ways. Great variability is one of the strongest and most typically non-native features in L2 acquisition. However, this variability does not seem to be completely chaotic. On the contrary, L2 speakers show trends in their performance and

development. Some evidence for these suggested trends will be presented here and at the end of the section a hypothesis for developmental stages for L2 perception will be stated.

As the above review of some studies (Fox et al 1995, Bohn 1995, Flege et al 1997) suggested, there are contradictory arguments concerning the L2 perception of English tense/lax vowels by Spanish speakers. However, it might be the case that these contradictions arise from the fact that the three studies tested L2 speakers in different points in their learning process. Fox et al 1995 tested Spanish speakers encountering the contrast for their first time and so he concluded that Spanish speakers did not listen to durational cues at all. On the other hand, Bohn tested more experienced learners and found out that they use durational information more than native speakers do. And finally, Flege et al 1997 found no significant difference between the tested Spanish group's cue reliance and the English L1 one for the same vowel contrast and claimed that this suggested that indeed there was not difference. Even though, they mention that there were different proficiency/exposure levels in the speakers (inexperienced, experienced or more experienced) the groups seemed to perform quite homogeneously and hence they did not suggest any sequential pattern/stage for the subjects in each study.

Clearly the reasons why these studies seem contradictory is because they are looking at subjects in different developmental stages and so they arrive at different conclusions for their performance. Therefore, the data gathered separately could be put together to constitute a developmental sequence. Then, it would be fair to try to test the existence of such a sequence having heterogeneous L2 subjects.

The data gathered from the previous study (Escudero 2000b) may constitute another piece of evidence that supports the attempt to find out whether there might be defined developmental stages in the L2 perception of new contrasts or not. Some of these data strongly supports the hypothesis in favour of an intermediate stage in the perception of English tense/lax vowels by Spanish speakers where the subjects use durational cues exclusively to signal such vowel contrast. The data also suggested that there could be multiple degrees in the reliance on the two main cues (spectral and durational) for this vowel contrast in the L2 speakers. That is, the intermediate stage could also be divided into at least two different patterns. A number of speakers had a performance that clustered

with these different trends. There was also a group that seemed to have a fully native-like performance. However, these results are inconclusive at present.

From this preliminary evidence the question posed at the beginning of the section seems to be relevant as well as the hypothesis that aims to answer it:

Q2: Are there defined/specific stages in the development of adult L2 perception of new contrasts?

H2: Yes, there are. I hypothesised here that there are at least 3-4 different stages in the development of L2 new contrasts. They could be based on the patterns that were suggested for the L2 performance in the perception of the new contrast /i/-/ɪ/. That is, what could be called a stage zero (S0) can be represented as a time in the L2 contrast development when the learners cannot discriminate the contrast with all the acoustic information manipulated equally in a native-like fashion.

A second stage (S1) would correspond to a time in the L2 perception development when the learners are able to identify and discriminate a new contrast with all the acoustic information manipulated equally in the same way native speakers do. However, they do it exclusively on the basis of a secondary acoustic cue. That is, in the L2 perception of Scottish English /i/-/ɪ/, the Spanish speakers that are in this stage will discriminate the contrast native-like but have a cue weighting and reliance where durational information is not only primary but also the only mean they use to signal the contrast.

A third stage (S2) would have a developmental pattern that involves being able to perceive the contrast (native-like performance in a two-cue manipulated continuum) and have higher weighting of secondary information than the one manifested in L1 perception. In terms of the /i/-/ɪ/ contrast, the subjects that fit in this pattern will be able to perceive the vowel contrast in a native-like fashion, they will use both spectral and durational information for signalling the contrast but still weight durational information higher than the L1 Scottish English group.

The last stage (S3) would be characterised by a native-like perception of the contrast and a native-like cue weighting of the acoustic information involved. For the /i/-/ɪ/ contrast, the L2 subjects that manifest this pattern of development will be able to identify and discriminate the contrast in a two-cue manipulated continuum as well as use spectral information as primary for signalling the contrast. This stage may be called *the ultimate attainment* stage.

The stages proposed above seem to follow a sequential order. However, that may not be the case. We will need to test subjects longitudinally to gather evidence for the sequential nature of these stages. Due to time limits and methodological limitations a longitudinal study was not feasible. Therefore, a cross-sectional study including a highly heterogeneous group of L2 subjects was tested to attempt to find as much variability as possible. Despite these limitations, it was thought that some insights and evidence for the existence of these patterns/stages in the L2 development of new contrasts would be found through the experiment that will be described in the following chapter.

4.3. The starting point

The literature (Best 1995, Flege 1995, Bohn 1995, Fox et al 1995 and Flege et al 1997 among others) suggests that the L2 perception of contrasts that do not exist in the L1 is a single-category. That is, the L2 learners will first merge the sounds involved in the new contrast into one: they will assimilate the sounds to the one category that exists in the L1 (Best 1995).

However, most studies assume this starting point without giving clear evidence for it. They assume that the reason for a poor discrimination of new sounds is the assimilation to only one category (see Fox et al 1995, Bohn 1995, Flege 1995 and Flege et al 1997). It seems that there is a further need for more data that could evidence a single-category starting point in the non-native perception of a new contrast.

Therefore, the question of what the starting point is for adult L2 learners when encountering a new contrast is still relevant. The hypothesis of a single-category starting point will be further tested here. However, it will also be suggested that both the L1 and the input to which the learners are first exposed might influence this starting point.

Before the end of this section, I would like to restate and highlight the third research question and hypothesis for the present study:

Q3: What is the starting point in the adult perception of new contrasts?

H3: As suggested by previous studies, it is highly plausible that L2 adult learners start with the single category that exists in their L1. That is, they perceive the two sounds as being equal (as suggested by Flege's "perceptual equivalence" or Best's "single-category assimilation"). However, it may turn out that particular features in the L1 and the input to which the learners are exposed are the factors that may influence the starting point. That is, they could potentially assimilate the two sounds to two different categories in the L1. In the case of Spanish speakers' acquisition of /i/-/ɪ/ they may start by assimilating the sound to Spanish /i/ and /e/, respectively rather than merging them in the same category (Spanish /i/, most likely).

Chapter 2 The experiment

This chapter presents a cross-sectional and experimental study carried out to test the hypotheses posed in the previous section, namely, 1) L2 speakers may lexicalise a new phonological contrast on the basis of a non-native use of phonetic information, 2) there may be patterns/stages in the development of the acquisition of new contrasts and 3) the starting point for new contrasts depends on the acoustic information of both the new segments and the closest L1 segments: it is not always the case that new contrasts are assimilated to an L1 single category.

The subjects and experimental design are first presented. Following a detailed explanation of the tasks and the procedure, the findings and the discussion are presented.

1. Subjects

The subjects for this study were 50 in total, 20 of which were L1 speakers of Scottish English and the other 30 L2 speakers of English (L1 Spanish). In general, the subjects reported no hearing problems and accepted to participate in the study voluntarily.

The L1 speakers constituted the control group. They were 10 female and 10 male speakers, all from Scotland, who reported to have lived in Edinburgh for most of their lives. Their ages at the time of the study vary between 23-35 years.

The L2 subjects were the actual subjects of the study. They were 15 female and 15 male L2 speakers of English who, in all cases, started their L2 learning after the age of 12. They were visiting or living in Edinburgh when participating in the study. All of them were speakers of Spanish in the general sense of the word, no attempt to control for Spanish

varieties was made: some of them were from Spain (different regions within the peninsula) and some others from different countries in South America. Their ages vary between 18-58 years. They were middle and upper class students (undergraduate and postgraduate) and employees.

2. Methods

This section presents the methodology used for the study, which aimed to gather reliable data that would support the hypotheses. The section is organised so it can show the stimuli used and the expectations for the subjects' performance, as well as a detailed explanation of the experimental design, including the way it was presented to the subjects and what they were expected to do.

2.1. Materials

Two vowel sounds were synthesised using the values of the natural production of two L1 Scottish English speakers. The vowel sounds were representative exemplars of the vowels /i/ and /ɪ/ produced 10 times each by the Scottish English speakers.

The synthesis was carried out at the Linguistics microlab (University of Edinburgh) using the Sensyn version of the Klatt parameter synthesiser. In a previous experiment (Escudero 2000b), the HLSyn version of the same synthesiser had been used. The outcome of the two syntheses suggested that the sounds generated in Sensyn had a much better quality. That is, the synthesised versions of the vowel sounds /i/ and /ɪ/ used for the present experiment were much closer to the naturally produced ones. In this way, the quality problem that was noted in the materials previously used (Escudero 2000b) was solved. Likewise, the quantity problem, that is the fact that the sounds were too long and that this might have confused the listeners, was solved by using durational values that were two steps shorter than the ones used before (Escudero 2000b).

Figures 1 and 2 show the spectrograms of /i/ and /ɪ/, respectively, in their naturally produced and the synthesised versions. Table 5, after the spectrograms, show the F1 and F2 values (in Hertz) for the two sets of vowel sounds as well as the conversions made for the synthesis. The durational values are also presented there.

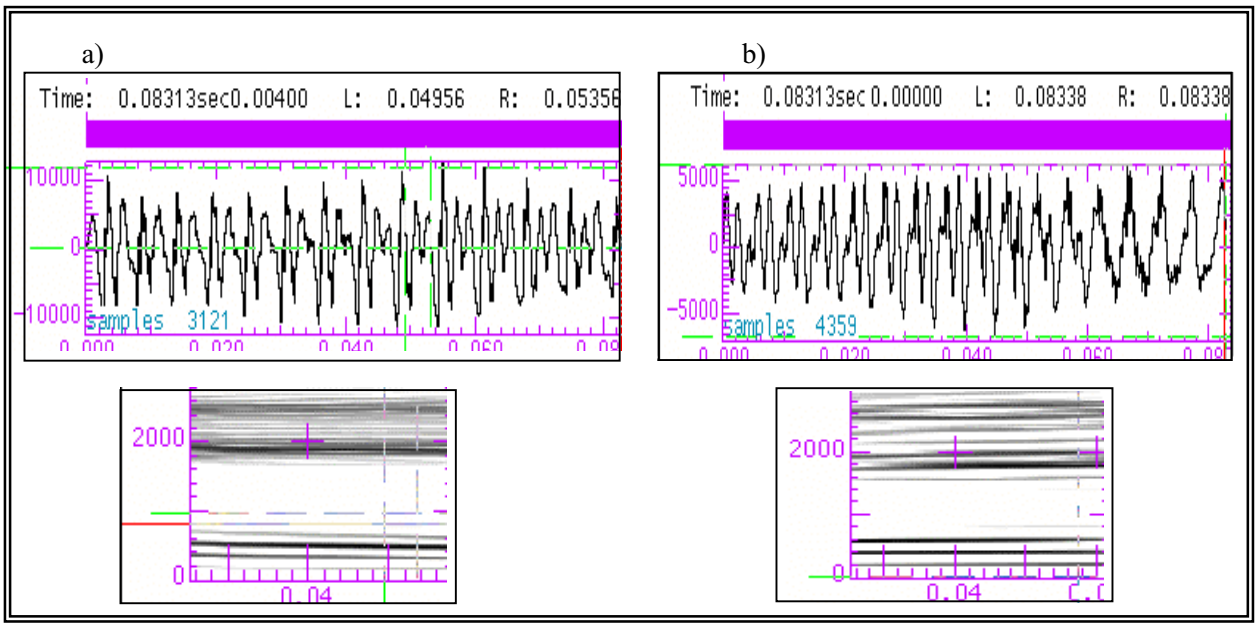

Fig. 1. Spectrum and spectrogram of the a) synthesised and b) naturally produced versions of the Scottish English vowel /ɪ/. The spectrograms represent the middle part of the sound[1]

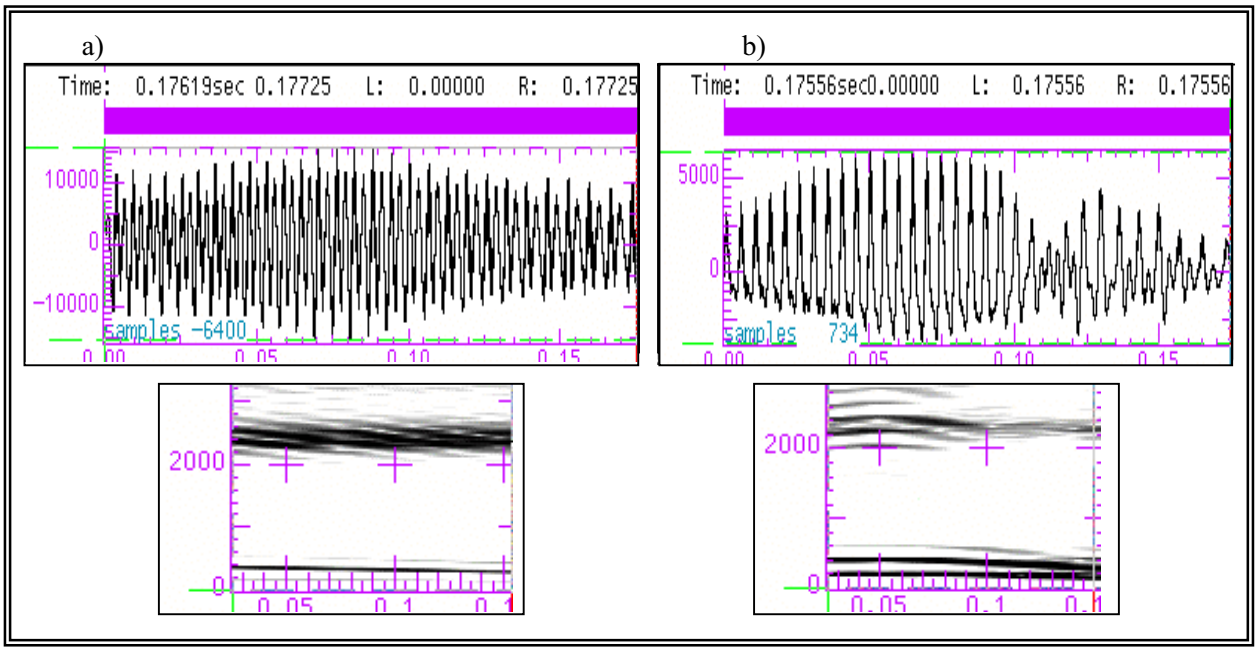

Fig. 2. Spectrum and spectrogram of the a) synthesised and b) naturally produced versions of the Scottish English vowel /i/.

[1] These pictures were created using the program Xwaves and Paintshop for their editing.

Sound	F1 (in Hertz)	F2 (in Hertz)	Duration (in msc.)	Sound version
/i/	343	2328	175	Natural production
	340	2320	177	Synthesis
/ɪ/	484	1890	82	Natural production
	480	1893	82	Synthesis

Table 5. F1, F2 and duration values for the natural produced and synthesised vowel sounds.

The two synthesised versions of prototypical productions of /i/ and /ɪ/ were used to make seven continua of six steps each (seven elements in each continua). An auditory scale rather than a physical one was used to get the values for the six steps in all seven continua. That is to say, the values in Hertz were converted to mels (auditory scale), then the values for seven elements were computed by linear interpolation, finally these values in mels were converted back to Hertz for synthesis (Sensyn can only accept values in Hertz). The equations used for such conversions are shown in Table 6 below. Likewise, seven durational values were computed, the increase in duration from stimulus one to stimulus seven was done through a logarithmic interpolation[2].

Conversions			
Hertz to mels	m	=	$(1000/\log 2) \log (f/1000 + 1)$
mels to Hertz	f	=	$1000 (10^{m/3322} - 1)$

m = mels
f = frequency in Hertz
1000/log 2= 3322

Table 6. Equations for both computing the values of the spectral elements and for the synthesis.

[2] That is, the durational values were increased in percentage, e.g. durational value 2 = durational value 1 * 1.1335 , value 3 = value 2 * 1.1335%, etc. This method was suggested by Dr. Alice Turk, it seems that differences in duration are perceived in this way (logarithmically).

The first continuum that was generated was the one with the prototypical /ɪ/ and /i/ (in that order) values. This continuum was called AC and it had seven elements (AC1, AC2, AC3, AC4, AC5, AC6 and AC7). AC1 and AC7 represented the endpoints /ɪ/ and /i/, respectively. Figure 3 shows a representation of the continuum.

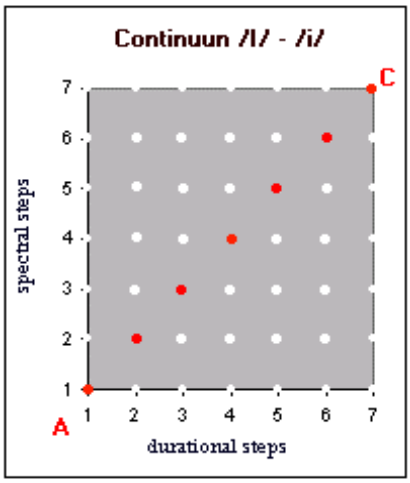

Fig. 3. Diagonal continuum with the two cues varied (contrast).

This continuum would be used in an identification test. Likewise, three elements or stimuli in the continuum would be used in a discrimination same/different test. Both tests would gather the perception of the vowel contrast and the place of the category boundary in both speaker groups.

Six other continua were generated in order to measure the subjects' reliance and weighting of durational and spectral cues. Therefore, the six continua were divided into two different sets. The aim behind the first set of continua was to measure the durational reliance. To do that, the spectral values were kept constant in three parts of the AC continuum and the duration is manipulated in the usual six steps. That is, the first durational continuum has a spectral value AC1 (the most /ɪ/-like spectral value), the second has an AC4 spectral value (a value in between /ɪ/ and /i/ for spectral information) and the third an AC7 spectral value (the most /i/-like). Figure 4 shows the three durational continua.

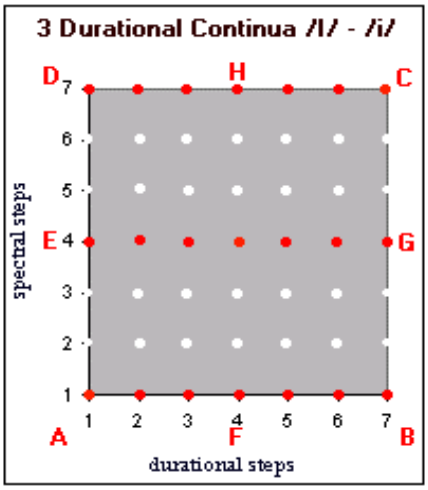

Fig.4. Durational continua.

If the subject relies on durational cues, she is expected to show very good to good performance, as represented by a clear category boundary (deep slope in the graph), in all three continua. However, it may be the case that some continua are performed better than others are. Predictions as to what continuum would be the highest in scores could not be made. A language group difference in scores is predicted.

The second set of continua to measure reliance and weighting of the cues is made up with stimuli that have a constant duration and involved a manipulation of spectral information. The first continuum has an AC1 durational value, the second an AC4 value and the third and AC7 one. Again, if the subject relies on spectral cues, his/her performance in all these three continua will be very good to good. Variation between continua and subjects is also expected. A performance difference between language groups is also predicted. Figure 5 shows the three spectral continua.

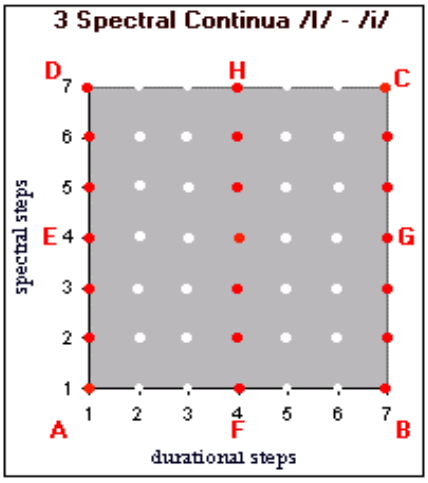

Fig.5. Spectral continua.

2.2. Procedure

The experiment was elaborated using Psyscope (Mac version), an experiment designing program. The subjects listened to all stimuli under comfortable hearing conditions. All subjects were tested in the Experiment Room at the Department of Theoretical and Applied Linguistics (University of Edinburgh), a soundproof room.

The experiment that all 50 subjects did was made up of three parts and had a total duration of 35 minutes. The first part was a discrimination same/different test. The subjects had to decide whether the sound pairs they heard were the same or different by pressing the button with the word “same” or the one with the word “different”. Together with the listening of the sounds they had a visual display in every trial, very much similar to what they had on the button box, that reminded them of what to press. The subjects were instructed both verbally and in writting. They were told to guess if they were not sure about the answer and to take as much time as they thought convenient to make a decision. Likewise, they knew that the next trial would not come up if they did not make a decision.

Figure 6 shows the display that matched the buttons that could be pressed as well as what they represented.

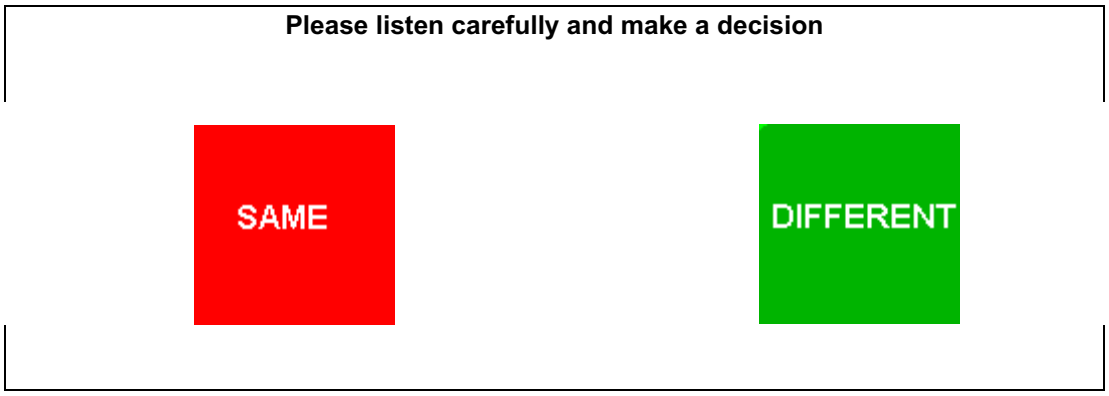

Fig. 6. Visual display for the Experiment first part.

For this first part the subjects were presented with 10 repetitions of 10 different pairs of sounds. The pairs were stimuli AC1, AC3, AC5 and AC7 from the AC continuum (continuum with /ɪ/ and /i/ prototypical values). These stimuli were paired in the following way to get the 10 pairs of the test: AC1-AC1, AC1-AC3, AC3-AC1, AC3-AC3, AC3-AC5, AC5-AC3, AC5-AC5, AC5-AC7, AC7-AC5, AC7-AC7 (following Pisoni 1973). The task had five blocks of 2 repetitions of the 10 pairs played in a randomised order. After each block the subjects had the opportunity to take a short break if wanted.

When the first part was finished, a message asking the subject to call the experimenter for further instructions appeared on the screen. That allowed the subjects to have a break before the second part. The second part of the experiment consisted of a forced identification test. The subjects had to decide whether the vowel sound they heard was represented in the picture of a “ship” or the picture of a “sheep” by pressing the button with one of the pictures. In this part, they also had a visual display in every trial with the pictures representing the ones on the button box with the possible choices. They also had verbal and written instructions and were told to guess if not sure. They were told to take

as much time as they thought convenient to make a decision and they knew that they would not have the next trial if they did not make a decision. Figure 7 shows the display they had.

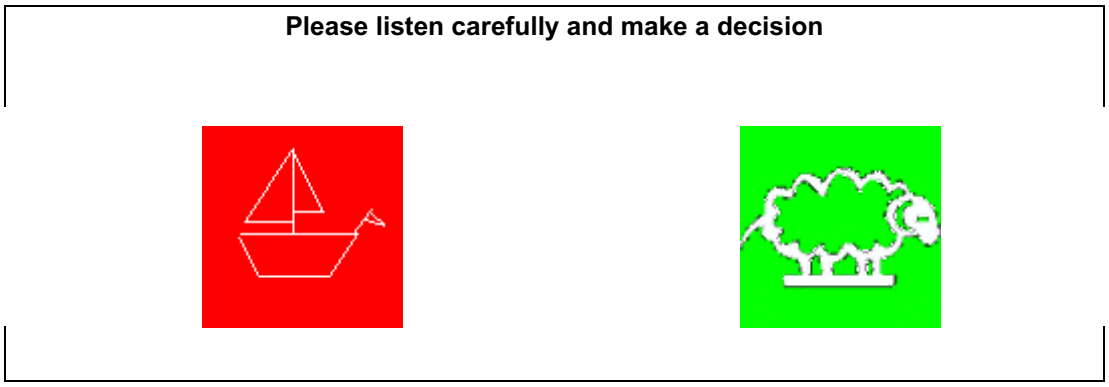

Fig. 7. Visual display for the experiment second and third parts.

In this second part, they were presented with 10 repetitions of the seven elements of the AC continuum. The test had five blocks and each block had two repetitions of the elements in the AC continuum played in a randomised order. After each block the subjects had the opportunity to take a short break. When the part was finished the subjects got a new message that asked them to call the experimenter.

The last part of the experiment was the longest one. It was, again, a forced identification test. The subjects took a break of 2 to 5 minutes before doing this last part. What they had to do, the visual display they had and the instructions were the same as the ones used for the previous part (the instructions were repeated). However, the experimenter encouraged the subjects to take breaks when the dialogue for doing it appeared. They were presented, this time, with 10 repetitions of 37 vowel sounds (the elements of the seven continua)[3]. The test had 10 blocks of 37 stimuli played in a randomised order. After the end of this part they had a short message thanking for their participation followed by verbal thanks and coffee and/or a bar of chocolate.

[3] The elements of the AC continuun were laso included to test the between-task stability of the diagonal contrast perception. The AC scores for both parts of the experiment were averaged for the anlysis. This methodology was approved by Dr. Mitsuhiko Ota.

It is worth mentioning that all subjects were asked for the names of the objects on the pictures ('sheep' and 'ship') before the identification tests. This was to find out whether they made a difference between the two objects or not. The experimenter never produced the words for neither of the pictures nor did she tell them that the words where different. For the L1 group, some of the subjects where not sure of what the object in the red picture was. They thought it was a "boat" or a "yacht", in that case the experimenter explained that it was something else until they produced the expected word ("ship"). All L1 subjects and 90% of the L2 subjects produced distinctive words for the two pictures. The majority of the L2 subjects appeared to produce "sheep" with a long [i] and "ship" with a short one. Likewise, the majority of these subjects, after producing the sounds, referred to the difference as being one of length ("the long and the short", they said). Of course, they could have thought that the vowels only differ in terms of length but still perceived and produced a quality difference. However, it is interesting to find out that none of the L1 subjects tested made a similar judgement.

3. Results

The results will be presented in three sub-sections[4]. The first one shows the comparison between the L1 and L2 group performance in the "diagonal" continuum: AC1-AC7, in both the discrimination and identification tests. The individual scores for this continuum in the two identification tests (part 2 and part 3 of the experiment) were averaged.

The second subsection shows the comparative cue weighting for the L1 and L2 groups. Only the subjects whose performance in the diagonal continuum, where the two cues were varied, was equal to the L1 performance were considered for this analysis. The groups' performance is shown for the two cues separately and the reliance and weighting of the cues are presented in different graphs.

A third section presents the individual cue weighting and reliance scores of those individuals that had a native-like performance in the AC continuum. It turned out that the

[4] Special thanks to Dr. Matthew Aylett who wrote the script that automatically counted the total scores of each subject in each of the continua as well as the total group scores.

21 one subjects performed in three different ways and do not show 21 completely different performances. Consequently, the individual results will be presented within the three patters that were found.

3.1. L1 vs. L2 perception of the vowel contrasts ("diagonal contrast")

The first analysis of the data from all fifty subjects showed that four of the L2 subjects perform the three parts of the experiment in a different manner than expected. These four subjects performance differ from the other 46. Basically, these subjects either change their decisions in terms of the buttons they chose from test to test, or they thought that the button with the picture of a "sheep" represented the lax vowel and the button with the 'ship', the tense one. At least two explanations for this unexpected performance will be suggested in the discussion section.

For the analysis of the "diagonal continuum" AC, only the subjects that could show to have the native-like distribution of the sounds were considered: "sheep" representing /i/ and "ship" representing /ɪ/. It is worth mentioning again that all of the 20 L1 speakers showed such distribution. Therefore, fig. 8 shows the identification (a mean taken from the two times where the AC continuum was presented, that is, in part 2 and 3) and the discrimination scores for 20 L1 subjects and 26 L2 subjects.

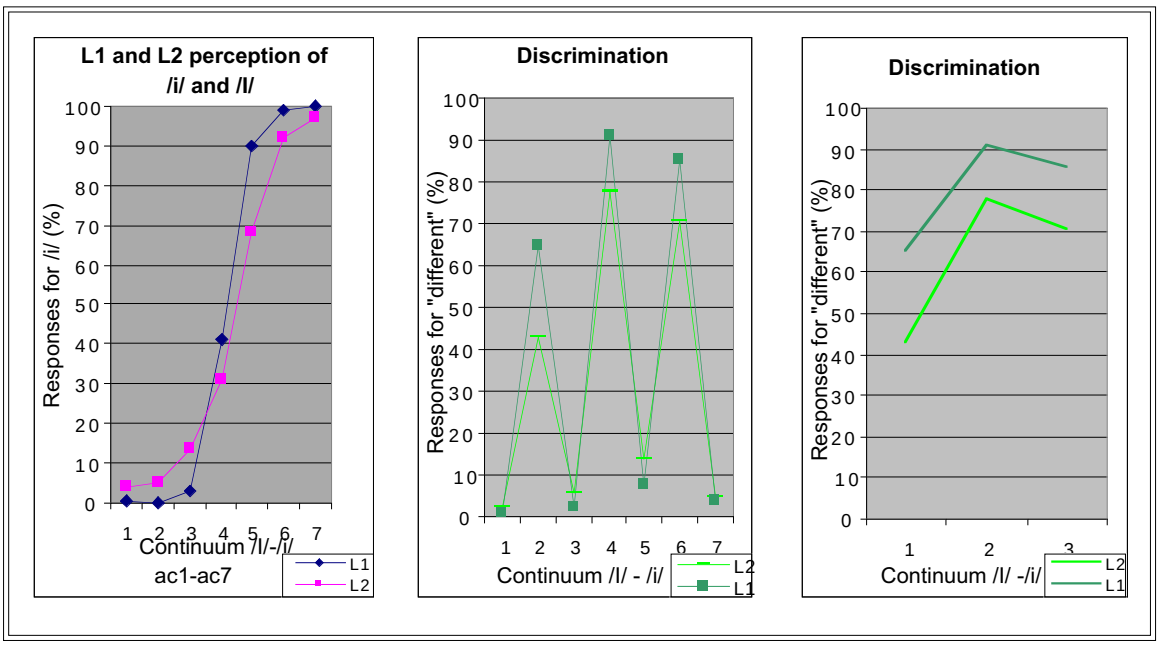

Fig. 8. L1 and L2 identification and discrimination of the AC continuum ("diagonal" contrast)

The discrimination scores showed a peak in the category boundary for both language groups (L1 and L2). That, in line with categorical perception studies (Pisoni 1973, Strange et al 1979, among others), suggests that both groups show a categorical perception for the two vowel segments tested here. We can see that vowel perception is not as categorical as consonant perception, this finding also agrees with previous studies on phonetic categorisation (Pisoni 1973). That is, rather than a single peak which is expected at the category boundary, three peaks are shown. As the picture shows, subjects are able to identify differences between the segments that are considered in the same vowel category, that phenomenon does not occur in consonant perception.

For the identification results, the graph shows that both groups have a category boundary placed in more or less the same part of the continuum. However, from pure observation it seemed that the L2 identification line (in pink colour) was more gradual than the L1 one. After running a factor analysis test that had both the language group and the stimuli in the continuum as factors, it was reliably found that the L1 and L2 performance was significantly different ($p < 0.025$).

However, from the group scores analysis it was not possible to decide whether it was the case that all the L2 subjects differ from the L1 ones or if it was only some of them that did. A Hierarchical cluster statistical analysis showed that 21 L2 subjects clustered with the L1 performance. That is, the node that minimally included the L1 subjects also included 21 of the 26 L2 subjects. The other 5 L2 subjects were found to be the ones that significantly differ from the L1 group. These 5 subjects were considered not to be able to perceive the stimuli in the diagonal contrast in a native-like fashion. The successful 21 L2 speakers were considered for the cue weighting and reliance analysis that is shown in the next subsection.

3.2 L1 vs. L2 cue weighting and reliance

Here the group score of the 21 subjects that proved to cluster with the L1 group was calculated, separately, for the spectral continua and the durational continua. Fig. 9 presents

graphs with the L1 and L2 group performances for the two acoustic cues for each of the three different continua. The L1 results are represented in blue and the L2 results in pink.

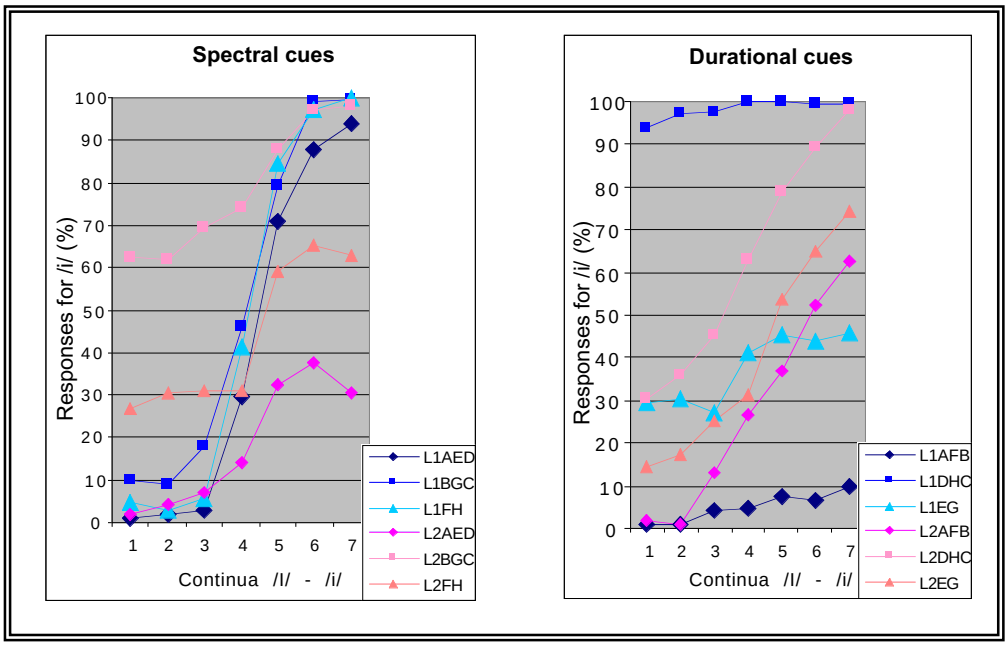

Fig. 9. L1 and L2 performance in the spectral and durational continua.

The graphs above show very good performance, measured by a steep slope in the category boundary, for the L1 group in the spectral continua. The same group shows a very pour performance in the durational continua. By way of contrast, the L2 group performs better in the durational continua than on the spectral continua. A univariate factor analysis between language and stimuli was run for each of the continua. The difference between the language groups showed to be significant for all the continua ($p < 0.05$).

The category boundary for the contrast, the cue reliance and the cue weighting was computed and plotted using Praat (“*a system for doing phonetics by computer”)*[5] statistics and graphs. The vertical percentages represent the spectral cues and the horizontal ones represent the durational cues. The cue reliance to signal the contrast was computed

[5] Special thanks to Dr. Paul Boersma who wrote the script that generated the graphs 10 to 13.

subtracting the score of the first stimulus from the seventh stimulus for each of the continua (cue effects according to Flege et al 1997). The scores for cue reliance do not sum up to 100% and were calculated using the end points of all the possible continua for the number of elements taken for spectral and durational cues (7 continua for each cue)[6]. The cue weighting or phonetic attention paid to each cue was computed by dividing each reliance value by the sum of the reliance values for the two cues. For example, if the reliance for durational cues were 90 and the spectral one 30 and we wanted to calculated the durational cue weighting, then we would have to divide 90 by 120 and the result would be 0.75, which would mean that the durational weighting is 75%. Fig. 10 shows the mean results for the cue weighting and reliance in each of the language groups.

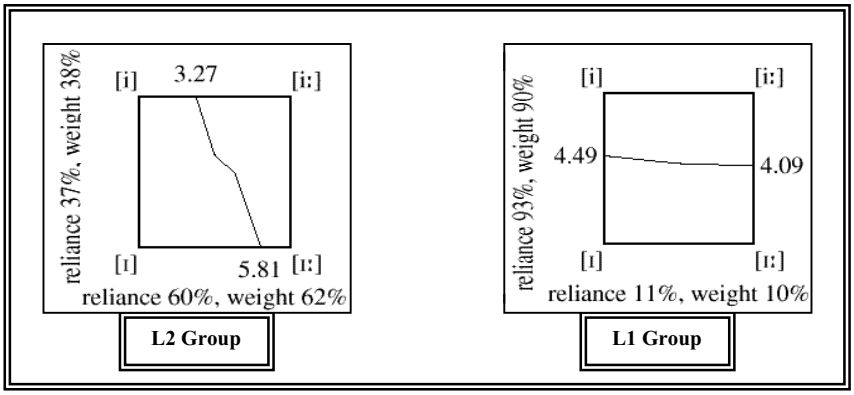

Fig. 10. L1 and L2 cue reliance and cue weighting.

In these figures we can also see the category boundaries as represented a line. These category boundaries are represented as connected line pieces in a two dimensional space:

[6] Not all the 14 continua were tested in this experiment but the end points of all of them were. These endpoints are the one that were used for the cue reliance computation. For example, the computation for durational reliance involved the following operation: ((AFB7 – AFB1) + (BG2-AED2) + (BG3 – AED3) + (EG7 – EG1) + (BG5 – AED5) + (BG6 – AED6) + (DHC7 – DHC1)) / 7 = durational reliance. The endpoints are all the stimuli in the edges of the "square" ABCD.

The left or the bottom part represents what was mainly reported as 'ship' and the right or top part what was reported as 'sheep'[7].

From the data in Fig. 10 above, it could already be suggested that the L1 group uses spectral cues almost exclusively and that the L2 group uses durational cues more than spectral ones. However, it seemed reasonable to think that not all the L2 subjects performed in the same way. However, we could not gather how different the individual behaviour was from the group scores. The analysis of the individual performance for the cue weighting was needed. The results were calculated using the same methodology used for the group scores. The findings are shown in the next section.

3.3 L2 individual cue weighting and reliance: three patterns

As expected, the individual analysis of the L2 group showed that there was variation in the performance of this group, unlike the L1 group that showed to be pretty homogeneous. However, the L2 individual results were not represented by 21 different performances: three main patterns were found.

The first pattern in the individual performance was found in seven L2 subjects. These subjects' performance has the same two features: 1) A very high reliance for durational cues and 2) negative reliance for spectral cues. The weighting of the cues could not be computed on the graph because of the negative values for spectral information. However, we could assume that the weighting is 100% for durational cues: some of the negative values computed for spectral reliance are so small that they might have been due to chance and so the actual spectral reliance may be zero. Fig. 11 shows the performance of the subjects that match this first pattern.

[7] In this case, 'mainly' means more than 50% of the time.

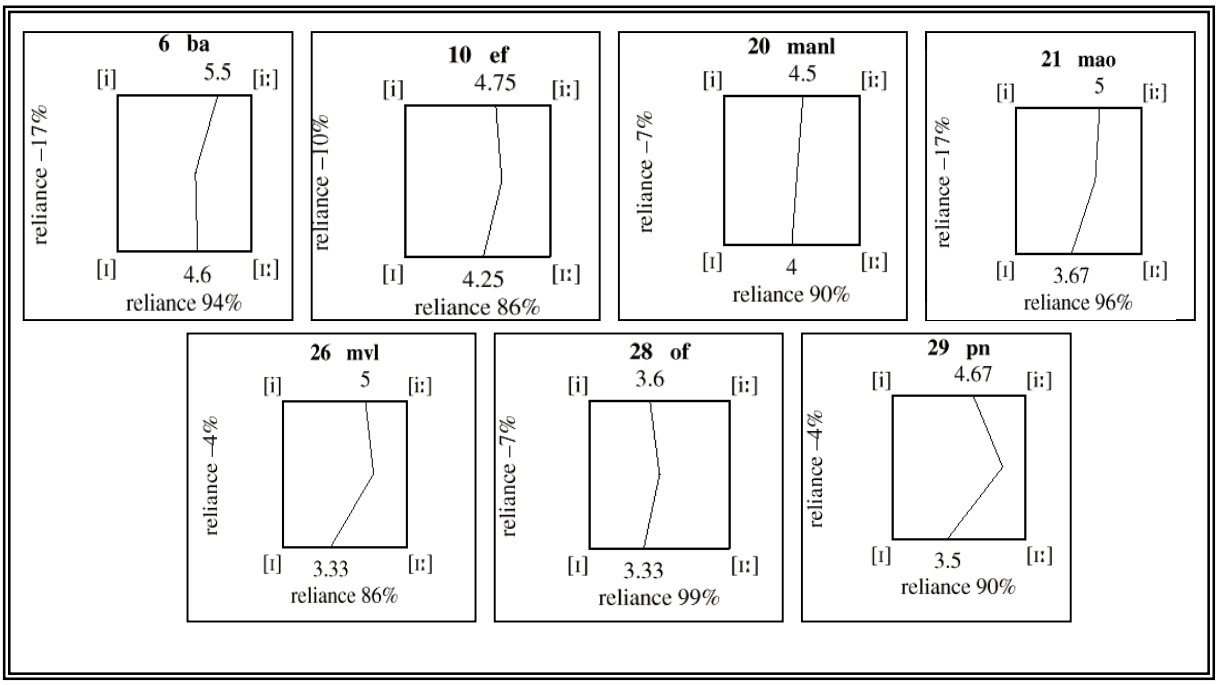

Fig. 11. L2 subjects matching pattern 1 (durational cues only). Spectral scores are on the vertical axis and durational scores on the vertical one.

The second pattern was found in the performance of 6 L2 subjects. This pattern presents the following two features: 1) A high durational reliance and weighting and 2) a low spectral reliance and weighting. The subjects within this pattern make use of spectral cues as well as durational cues but still consider durational information as primary. Fig. 12 shows the performance of these six subjects.

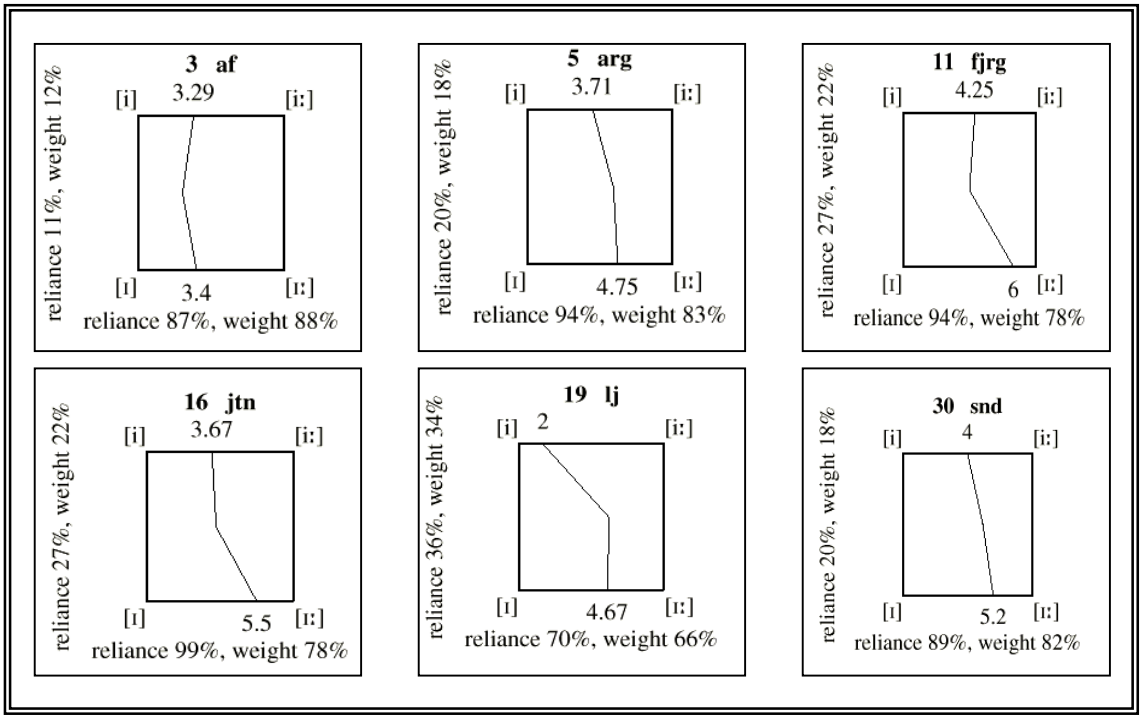

Fig. 12. L2 Subjects matching pattern 2 (durational information as primary).

The third pattern was thought to match the L1 group weighting and reliance. Eight L2 subjects showed to have such performance. The main feature for this pattern was a primary use of spectral information. Fig. 12 presents the scores of these 8 subjects as well as the L1 group performance as a mean of comparison.

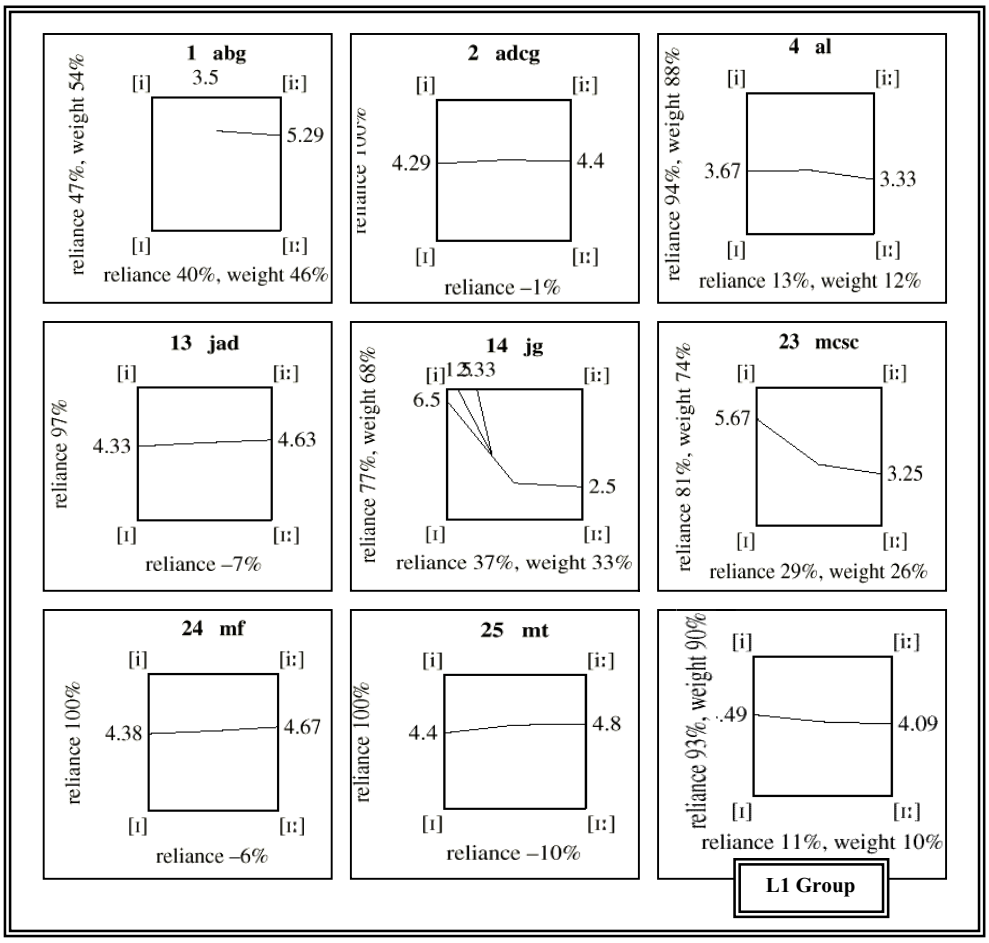

Fig. 13. 8 L2 subjects that match the L1 group pattern (spectral information as primary).

Having presented the results of the discrimination and identification as well as the cue weighting experiment, let us turn to the discussion of the general and particular findings. This will be done in the following section.

4. Discussion

This discussion will be organised in three subsections, each of them addressing one of the research questions that were put forward in line with the data and the findings gathered from the experiment.

4.1. The L2 perception of /i/-/ɪ/: auditory or phonetic strategies

For this first section, the possibility of having phonetic category formation (or a lexicalised phonological contrast) on the basis of different relative weighting of the acoustic cues involved is addressed. Likewise, the performance of the 4 L2 subjects that were removed from the data analysis is discussed.

4.1.1. The L2 perception of the contrast: discrimination and weighting

Recall that the data showed that 21 out of 30 L2 subjects had a native-like performance in the diagonal continuum AC. It was thought that this continuum would test the perception of the contrast per se and not the weighting of the cues because the acoustic information involved was varied at the same time in six steps. Out of those 21 subjects, 13 turned out to have a cue weighting that differed from the one the L1 speakers had. That is, they either used durational information exclusively to signal the contrast or they used the two cues (spectral and durational) but still weighted durational information as primary. On the basis of this data, it is claimed here that phonetic categories could be built up (and stored in the subject's grammar) with a non-native weighting of the acoustic information involved.

However, when this data was presented at *New Sounds 2000* (Escudero 2000c), James Flege (personal communication) suggested that there may be a different explanation for the L2 subjects behaviour in general. Flege claimed that the subjects might have made their decisions during the experiment using an auditory rather than a phonetic strategy or categorisation. That is, they could have categorised the vowel sounds (or one of them) as non-speech and as differ in length only. According to the PAM model, this sound will either be 1) "non-assimilable" or have an assimilation pattern of 2) "non-categorised vs. categorised", in both cases a good to very good discrimination is predicted. Flege suggested that there is no evidence from the performance on the continuum AC that could suggest that these L2 subjects have an independent phonetic category for /ɪ/. It is most

likely, on the basis of the acoustic feature of Spanish /i/ and Scottish English /i/, that the L2 group might have assimilated Scottish English /i/ to vowel /i/ in their L1 (Spanish). Therefore, the assimilation pattern would have been number 2. There is also the possibility of a goodness of category assimilation where the two English vowels are assimilated to Spanish /i/ but English /ɪ/ is a bad or less prototypical member of the category. In this case, we could not talk about phonetic category formation for English /ɪ/ in the L2 subjects.

Despite Flege's claims, there did seem to be some evidence of phonetic categorisation (or a lexical distinction). There are two main reasons why this is justified. First, if it were the case that the L2 subjects used length only to discriminate the sounds in the continuum, then they could have chosen either of the pictures to represent long or short. That is, the picture with the 'sheep' had the same probability to represent *short* as the picture for the 'ship' had. However, 26 out of 30 L2 subjects consistently decided that 'sheep' represented *long* and 'ship' *short*. It seems that 87% of the L2 group chose one of the two possibilities in the expected direction. This proves that the majority of the subjects had a lexical distinction and so it casts some doubt on the suggestion that they do not have a separate representation of the sounds.

Second, if it were the case that English /ɪ/ does not have a status separate from /i/ because it is only a non-prototypical member of the category, then it would be awkward to have the subjects producing a difference when asked to name the objects in the two pictures. Recall that the experimenter never told the subjects that the words for the pictures were or sounded different and that still more than 90% of them produced a clear difference between them. Consequently, the claim that they do not have a separate representation for the two sounds has less reliability mainly because the subjects showed to have a lexical distinction in the expected direction. The question that still remains is, would it be possible to produce a difference for two names that constitute a phonological minimal pair in the TL using two allophones of a single category? It might seem that the subjects can have separate lexical entries that minimally differ in terms of two members/allophones of the Spanish /i/ category. Although this does not seem to be likely, more experimentation to test if they do have a separate category for English /ɪ/ is certainly needed.

Therefore, despite some reasonable objections, it will be considered on the basis of the preliminary data gathered here that the L2 subjects that had a native-like performance in the AC continuum (21 subjects) do have a lexical distinction for the vowel sounds tested.

Likewise, the possibility will be considered that Spanish speakers of English could have a *covert contrast* (term suggested by Scobbie 1998 for children's production of L1 contrasts on the basis of a different cue weighting[8]) (see also Gerrits 2000 and Nittrouer 2000) when perceiving /i/ and /ɪ/. That involves the L2 perception of a contrast using different means than the ones used in adult L1 perception. On the other hand, this L2 perception could potentially match the production of the same contrast, in which case the L2 speakers' speech will become unintelligible for the L1 group. That is to say the L2 group will be perceiving and producing a covert contrast.

4.1.2. Different performance

Before the second hypothesis is discussed, the performance of the 4 subjects that do not manifest the expected behaviour has to be addressed. It is worth mentioning that none of the L1 subjects perform the tasks in an unexpected way, they all named and chose the expected picture for the vowel sounds. Two of these 4 subjects simply changed what button they thought represented the vowel sounds for the third part of the experiment. That is, they chose the 'sheep' button for the /i/-like sounds in one of the experiments and the 'ship' button for the other experiment. On the basis of this divergent performance, we cannot tell if they can perceive the contrast native-like or if they used auditory or different strategies to perceive the contrast. These subjects' decisions were not consistent enough for an attempt to describe their representation of the vowel sounds.

The other two subjects of this group consistently thought that the 'ship' button represented a high vowel and the button 'sheep' a lower one. There are at least two explanations for this performance. The first one relates to the auditory strategy suggested by Flege. That is, they may have used the length of the vowels to make their decisions and chose 'ship' to represent long and 'sheep' short. However, when looking at their cue weighting data, it was found out that their decisions were made mostly on the basis of spectral information

[8] This proposed behaviour and the developement of the name suggested for it will be further developed in a forthcoming paper by Scobbie et al (personal communication, see reference section).

and when they used durational cues (which was very rarely) they opted for 'sheep' when the sound was long. That is, they were using the right distribution for duration. Another explanation for their performance can be a pure confusion in terms of buttons, which it is considered unlikely given the many times they see the display on the screen. These two explanations are in a sense rather weak.

A third, and more convincing, explanation for these two subjects' performance comes from the possible use of an 'orthographic strategy' (as suggested in Flege 1997 and mentioned in the review section above) and a two-category assimilation to the Spanish vowel sounds /e/ and /i/. That is, they recalled the spelling version of the pictures and matched the vowel sounds they heard to the closest Spanish vowel sounds. Spanish has a very transparent orthography in which the "i" as in the word "ship" always represents an /i/ sound and "e" as in the word "sheep" always represents an /e/ sound (and that is why the vowel for "sheep" could be long, it has a double "e"). However, we can not be sure of what these subjects actually did and what their category representation is on the basis of this experiment. Nevertheless, if they were, in fact, using the spelling of the words to cope with the task, their performance could be seen as evidence for an even stronger effect of literacy in L2 phonology (Young-Scholten 1995). Even though we tried to avoid having this effect by using pictures, it was found in the performance of these two subjects. Clearly, the effect of literacy in the L2 acquisition of phonology deserves much more attention than L2 research has been paying to it.

Another interesting feature of the previous two subjects' performance is their assimilation of Scottish English /ɪ/ to Spanish /e/. Previous studies have never entertained this possibility but it seems that this performance could suggest a two-category assimilation rather than a single-category one. This point will be further addressed when discussing the starting point for the perception of English /i/-/ɪ/ in section 3.3.

4.2. L2 performance patterns and a possible stage-like development

If we recall the results of the experiment we are able to suggest the possibility for stage-like development in the L2 subjects. After the first analysis of the data, there were some L2 subjects that performed poorly in the perception of the AC (diagonal) continuum. According to the approach followed here, these subjects have not completely acquired a

native-like perception of the contrast, as tested by that diagonal continuum. Five subjects were found to fit this pattern.

Likewise, seven subjects of the group that had a native-like performance in the AC continuum (21 L2 speakers) manifest a second pattern: native-like perception of the diagonal contrast + exclusive use of durational cues. A third pattern was found to be shared by six L2 subjects: native-like perception of the diagonal contrast + use of durational and spectral cues + primary reliance and weighting of durational cues. Finally, eight L2 subjects showed to have a native-like performance in both the diagonal contrast and in the cue reliance and weighting.

Even though the study did not gather longitudinal data, it is plausible to think of the well distributed patterns found among the L2 group as being part of a developmental sequence (Wode 1976, Werker et al 1981, Polka et al 1996). Of course, we could not know on the basis of these data what the order for this sequence would be. However, the patterns seem to be defined enough to consider them evidence for the possibility of stages/phases in the learning of new contrasts.

Here a rather impressionistic and not very elaborated notion of stage/phase is taken: stages are patterns found in the development of a particular knowledge/performance that remain stable for a period of time before more learning takes place. This notion could include the opportunity of going into a next stage once some kind of learning has happened. Again, from the data collected here it cannot be said which of the intermediate stages (S1 or S2) appears first in a L2 developmental sequence. Nevertheless, we could still think of them as being some of the steps L2 learners may take from first encountering a new contrast to ultimately attain the perception of it.

4.3. The starting point and the ultimate attainment of new contrasts

Among the patterns that were found in the analysis of the L2 data, the ultimate attainment pattern/stage was the one with the most L2 subjects. That is, taken all the patterns/stages separately, 26.6 % of the subjects manifested the fourth pattern, S3 (as opposed to 23.3% for S1, 20 % for S2, 16.6 % for S0 and 13.3% for the 'other strategies' performance). This

finding seems to contradict the suggestion of previous studies that the acquisition of a new contrast takes time and that it is difficult (Fox et al 1995, Bohn 1995, Flege et al 1997).
The basis for saying that Spanish speakers will have problems in discriminating the English /i/-/ɪ/ contrast is the belief that these subjects will assimilate these segments to a single category in their L1. Therefore, it is argued by the same studies that the learning of those segments as being separate entities will constitute a long process. However, some of the L2 subjects that turned out to perform in a native-like fashion had not been exposed to the TL for a long time.

The analysis of the linguistic background data from the 8 subjects included in the ultimate attainment stage revealed that 3 of them had had less than five years of exposure to the TL. Furthermore, these five years of learning took place in their native towns where their L1 was the dominant language and the TL was just a foreign one. By the time these subjects were tested, they had only been in Edinburgh for two weeks and their use of Spanish (their L1) in Edinburgh largely outperform the use of the TL. Therefore it is not likely that they could have attained a native-like level in the perception of the contrast had they had a single-category as a starting point.

From the S3 subjects that had had more than 5 years of exposure to the TL, only 3 had lived in a country were the TL was spoken for more than 2 years. The other two had visited English speaking countries but not spent more than a week in them. However, of these two subjects, one uses the TL language much more than the L1, whereas the other reported to do the opposite. Among the 3 subjects that had spent the most time in English speaking countries, two had been living in Edinburgh for more than 4 years; the third subject had only spent months in Edinburgh. One of these 2 subjects reported to use the L1 much more than the other.

From these data, we can confirm that the ultimate attainment stage subjects have a very diverse L2 background. This difference in the time and type of exposure to the TL could cast doubt on whether they are in a S3 in the learning of a new contrast. That is, the L2 subjects with the least exposure may have assimilated the sounds they heard to two categories in their L1 during the task. That might be confirmed by the fact that none of these subjects produced a difference between the objects represented in the two pictures

('sheep' and 'ship'). Moreover, one of them manifested some problems in remembering what the words for the objects were[9]. Therefore, from this analysis no strong evidence for their ultimate attainment of the contrast could be found. Regarding the other 5 subjects' performance, they could have also assimilated the segments involved in the contrast to two categories rather than one when they first encountered the vowel sounds.

After the discussion above, it could be suggested that the acoustic properties of /i/-/ɪ/ in Scottish Standard English may well cause a two category assimilation (to Spanish /i/ and /e/, respectively) by Spanish speakers. That is, there may be no lexical distinction between the L1 and the L2 categories for this contrast in these L2 subjects. However, from the data gathered in this study we could not reliably argue about the nature of the S3 subjects' perception. Data from the perception of these same subjects in three segments of Scottish Standard English should be analysed. That is, the analysis of the L2 perception of the Scottish three-way contrast /i/-/ɪ/-/e/ could provide evidence for the category formation of /ɪ/ and for the starting point in the perception of Scottish /i/-/ɪ/ by Spanish speakers.

[9] This subject rememebered what the words were in the end but still did not produce any difference between them.

Chapter 3 Conclusions, implications and further research

The present study intended to find answers for three research questions, namely, 1) is it possible to have a phonological category distinction (be able to identify and discriminate a phonological contrast) of a new contrast on the basis of a non-native use of the phonetic information? 2) are there developmental patterns/stages in the L2 perception of new contrasts?, and 3) what is the starting point in the acquisition of new contrasts? This chapter presents the concluding answers proposed after the analysis and discussion of the findings. Likewise, two implications, one theoretical and the other applied, for the field of second language acquisition are suggested in a separate section.

Despite the existence of evidence that supports the conclusions here, there is a clear need for further experimental studies that could provide stronger and more reliable evidence for the conclusions here. Consequently, suggestions for possible follow-up research are made in the third and final section.

1. Conclusions

As a general conclusion, and in line with Bohn 1995's claim, this study has shown that the majority of the L2 speakers (considering both intermediate patterns) listen to durational cues more than the L1 speakers do when perceiving tense/lax vowel contrasts (e.g. /i/-/ɪ/). Furthermore, this different cue weighting was manifested in two different patterns. That is to say, it may be argued that the development of the adult L2 perception of new contrasts has more than two instances (i.e. the classical non-contrast/contrast view) from first exposure until ultimate attainment.

The following conclusions directly address the questions and hypotheses that were put forward at the end of the review chapter according to the results of this study and the discussion of them. The first conclusion addresses the existence of a different cue weighting in L2 speakers that can identify and discriminate a vowel continuum with a manipulation of the two cues involved. The second conclusion suggests the possibility of a stage-like development in the L2 perception of new contrasts. Finally, the last conclusion addresses the idea of ultimate attainment by L2 learners, in line with the possible starting points for the L2 perception of new contrasts.

1) As far as the data that were gathered in this study is concerned, evidence of L2 phonetic category formation on the basis of non-native cue weighting was found. 62% of the subjects that could identify and discriminate the English contrast in a native-like fashion (as tested by a hierarchical cluster analysis test) did turn out to be using a different cue weighting/reliance of the acoustic information involved. These subjects used durational cues exclusively or primarily to signal the contrast rather than spectral cues, which constitute the primary information for the L1 perception.

2) L2 individual performance did not turn out to be completely heterogeneous. The data showed four well-defined and well-distributed patterns in the performance of 26 out of 30 L2 subjects. It is argued here that those patterns may reflect stages in the development of a new vowel contrast. However, further data to test the sequential nature as well as the "true" existence of these stages is needed.

3) The third pattern found in the L2 performance ("ultimate attainment") casts some doubt on the assumed starting point (S0) for the English contrast /i/-/ɪ/ in the L2 group. Therefore, it is concluded here that the starting point for new contrasts depends on the particular acoustic features of the English variety the L2 subjects are first exposed to together with the features of their L1. From the data gathered in this study, we can also suggest that the starting point for the perception of Scottish English /i/ and /ɪ/ by some Spanish speakers may be a two-category assimilation to Spanish /i/ and /e/, respectively. However, evidence from the perception of other Scottish English vowel segments by the

same L2 group is clearly needed. Likewise, a comparison between the acquisition of the same segments in two different English varieties may turn out to be illuminating.

2. Implications

The conclusions above stated have some implications both for the idea of individual variation in the acquisition of a second language and for the possibility of facilitating the ultimate attainment of new contrasts. The following two sections will briefly address these two topics.

2.1. Individual variation in L2 acquisition

The conclusions for the present study suggest that even though there may be a great deal of variation between L2 speakers, these subjects could pattern in terms of their performance in the perception of new contrasts. Recall here that 30 different patterns were not found in the L2 performance. On the contrary, only 5-6 patterns were found. This may contradict in some way the idea posed by some L2 researchers that every L2 speaker is different (summarised in the New Sounds 2000 discussion session)[10].

Because of the findings here, it would be interesting to know how much more variability is found in L2 as opposed to L1 development of phonological contrasts. It may be that the variability between the individuals of each acquisition group is equal and the difference is based on the fact that the L2 developmental patterns last for longer and may never change. That is, the individual differences are equal but as groups they differ on what their final developmental stage is.

The idea of developmental patterns for both L1 and L2 perception implies that not every subject will manifest an identical learning process. However, this variability is predicted in the same way for both L1 and L2 perception.

The fact that the developmental patterns found in this L2 performance are shared by a number of individuals and the possibility for the longer duration of them in the learning

[10] The proceeding including the discussion of this conference will be published in January 2001.

process may suggest that L2 acquisition could be illuminating in terms of a more general picture. That is, investigating the development of L2 perception might turn out to give interesting and helpful insights to our knowledge of human language perception.

2.2. Speech training

From the conclusions above, there seems to be a possibility for the patterns found in the L2 performance to constitute steps in the learning process of new contrasts. Therefore, it may be feasible to devise a training program that could facilitate the advancement of L2 speakers based on such developmental patterns.

Some studies (e.g. Flege 1988, Bradlow et al 1996) have suggested that L2 speakers could be trained to perceive and produce new contrasts. From conclusion 2 here, it may be possible to train learners to perceive vowel contrasts by facilitating their detection of the acoustic information involved. Furthermore, the possibility for defined developmental stages gives a stronger motivation for us to think such a training program might be successful. Traditionally, speech training was given in formal instruction contexts through the repetitive production of pairs of words that differ only in terms of the target segments (e.g. ship vs. sheep). However, this type of training seemed to be successful for some learners but not for the majority of them.

Here we consider that the ability to discriminate a new contrast should precede the complete acquisition of a phonological contrast. Therefore, detecting acoustic information in steps may turn out to be most successful for both ultimate attainments in the perception and the production of the new sounds. That is, it may be the case that the traditional and previous training was not successful because 1) it did not start with perception and 2) it trained learners in the last step of a developmental process. The second problem involves presenting the subjects with input that was well above their possibilities of processing it.

Consequently, the testing of the developmental pattern/stage of the learner as well as a training equivalent to the next step that should be attained could prove to have much better results. Since the findings here suggest that L2 subjects tend to manifest few rather than many patterns, it is likely that the learners would constitute groups and that the training would be collective rather than individual.

3. Further research

In this section, some suggestions for further testing the hypotheses and conclusions of this study are presented. The new studies that are proposed would also confirm or disconfirm the possible implications that were discussed in the previous section.

3.1. L2 category formation and the L1 and L2 perception of other vowel segments

In order to further confirm the fact that the L2 subjects have a different phonetic category for /ɪ/, it is necessary to test their perception of English /e/. Likewise, experiments that do not involve forced choices are needed. The L2 subjects in this study might have chosen one of the alternatives because it was the less bad one. That would mean that they do not have a different category but that because it is a very bad member of a category it could be placed somewhere else, this other place not necessarily being the TL category.

Among the experiments that could be designed, we could have Spanish speakers perform categorisation tasks in English and in Spanish. They would have to press the button with the vowel that represents best the sound they hear in both languages separately. The sounds they would hear would always be naturally produced English sounds (isolated vowels and words containing the vowels).

These experiments would be carried out, if possible, trying to test some or all the L2 subjects that were tested before. The L1 perception of this vowel segment will also be tested. Furthermore, a test for goodness of fit in the category would also be included within the series of experiments.

3.2. Longitudinal studies

The findings of this study could not be seen as evidence for a sequential change in the perceptual development of new contrasts. Therefore, longitudinal studies of at least a year in length seem to be needed. These studies would aim to investigate the sequence of the development as well as to find out how stable the patterns found in this study could be. The similarities and variation between the L2 subjects would also be analysed. The best of the situations would involve testing the subjects once a week. It would also be very interesting to compare this L2 development to the L1 development in the perception of the same segments. We would expect to find a much faster change in L1 perception and

perhaps the non-finding of some of the intermediate stages that could be found in L2 perception.

3.3. Factors for the starting point (S0)

A comparison between target varieties could investigate whether the starting point for subjects with the same linguistic background and English as a TL is the same regardless of what English the L2 group is exposed to.

It would be interesting to investigate the acquisition of less used English varieties. That is, most L2 acoustic studies have focused on the acquisition of General American English by different language groups and from their findings they have made general conclusions about the acquisition of English. The other two varieties that could be tested are Standard British English as spoken in the south of England and Standard Scottish English as spoken in Edinburgh. The reason being that there is a very large amount of L2 speakers of both varieties that tend to be established in the TL environment.

This comparison would give very illuminating insights for the starting point on the acquisition of different English varieties. Of course, it would also provide more information about the acoustic features of the vowel segments in each variety and the implications for the L1 and L2 acquisition processes.

3.4. The relation between the perception and the production of new sounds

It would certainly be of interest for the general L2 acquisition of new contrasts to compare the perceptual performance with its production. That is, it would seem obvious to try to find out whether the developmental changes that happen in perception occur in the same way in production and which of the two abilities is ultimately attained first. It would be expected that ultimate attainment in L1 and L2 perception happens before attaining the same level in production. Likewise, if the perception were completely native-like and production were not, then there could be explanations in line with the idea of *motor control problems.*

To sum up, the longitudinal studies that would be carried out should include a comparison between language varieties and between the perception and the production of vowel segments in both L1 and L2 acquisition.

References

Best, C.T. 1995. 'A direct realist view of cross-language speech perception'. In Strange, W. (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues.* Baltimore: York Press. pp. 171-203.

Best, C.T., and Strange, W. 1992. 'Effects of phonological and phonetic features on cross-language perception of approximants'. *Journal of Phonetics* 20: 305-30.

Best, C., Faber, A. and Levitt, A. 1996. 'Perceptual assimilation of non-native contrasts to the American English vowel system'. *Journal of the Acoustical Society of America* 99, 2602.

Bohn, O.-S. 1995. 'Cross Language Speech Perception in Adults First Language Transfer Doesn't Tell It All'. In Strange, W (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues.* Baltimore: York Press. pp. 279-304.

Bohn, O.-S. and Flege, J. E. 1990. 'Interlingual identification and the role of foreign language experience in L2 vowel perception'. *Applied Psycholinguistics* 11, 303-328.

Bradlow, A.R. 1995. 'A comparative acoustic study of English and Spanish vowels'. *Journal of the Acoustical Society of America* 97, 3: 1916-1925.

Bradlow, A., Pisoni, D., Yamada, R., & Tohkuda, Y. 1996. 'Training Japanese speakers to identify English /r/ and /l/: Some effects of perceptual learning on speech production'. *Journal of the Acoustical Society of America* 101, 2299-2310.

Cohen J.D., MacWhinney B., Flatt M., and Provost J. 1993. 'PsyScope: A new graphic interactive environment for designing psychology experiments'. *Behavioral Research Methods, Instruments, and Computers* 25, 2: 257-271.

Escudero, P.R. 2000a. *'The perception of Scottish Standard English: The role of acoustic cues in the /i/-/ɪ/ disitnction'.* Unpublished manuscript submitted as a project for the Advanced Phonetics option course, MSc in Applied Linguistics 1999-200. The University of Edinburgh.

Escudero, P.R. 2000b.*'The perception of English vowels by Spanish speakers: Spectral and temporal effects in the perception of the /i/-/ɪ/ contrast'.* Unpublished manuscript submitted as a project for the Second Language Acquisition option course, MSc in Applied Linguistics 1999-2000. The University of Edinburgh.

Escudero, P.R. 2000c. *'The perception of English vowel contrasts: Acoustic cue reliance in the development of L2 perception'.* Paper presented at the Fourth International Symposium on the Acquisition of Second Language Speech, New Sounds 2000. University of Amsterdam, The Netherlands.

Flege, J. E. 1987. 'The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalent classification. *Journal of Phonetics* 15, 47-65.

Flege, J. E. 1988. 'Using visual information to train foreign language vowel production'. *Language Learning* 38, 365-407.

Flege, J.E. 1991. 'Perception and production: The relevance of phonetic input to L2 phonological learning. In Huebner, T. & Ferguson, C. A. (eds.). *Crosscurrents in second language acquisition and linguistic theories.* Amsterdam: John Benjamins. pp. 249-264 (speech perception).

Flege, J.E. 1992. 'Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America* 91,1: 370-389.

Flege, J.E. 1995. 'Second language speech theory, findings and problems'. In Strange, W (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues.* Baltimore: York Press. pp. 233-277.

Flege, J. E., Munro, M. J. and Fox, R.A. 1994. 'Auditory and categorical effects on cross-language vowel perception'. *Journal of the Acoustical Society of America* 95,6: 3623-3641.

Flege, J.E, Bohn, O. –S, and Jang, S. 1997. 'Effects of experience on non-native speakers' production and perception of English vowels'. *Journal of Phonetics* 25: 437-470.

Fox, R.A., Flege, J.E., and Munro, M.J. 1995. 'The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis'. *Journal of the Acoustical Society of America* 97, 4: 2540-2550.

Gass, S. 1984. 'Development of speech perception and production abilities in adult second language learners'. *Applied Psycholinguistics* 5: 51-74.

Gerrits, E. 2000. *The perceptual weighting of acoustic cues by Dutch children.* Paper presented at the VIIIth Meeting of the International Clinical Phonetics and Linguistics Association, Queen Margaret University College, Edinburgh.

Gottfried, T. and Beddor, P. 1988. Perception of temporal and spectral information in French vowels. *Language and Speech* 31, 57-75.

Giegerich, H. 1992. *English Phonology.* Cambridge: Cambridge University Press.

Hillenbrand, J., Getty, L., Clark, M. and Wheeler, K. 1995. 'Acoustic characteristics of American English vowels'. *Journal of the Acoustical Society of America* 97: 3099-3111.

Ingram, J.C.L, and Park, S. –G. 1997. 'Cross-language vowel perception and production by Japanese and Korean learners of English'. *Journal of Phonetics* 25: 343-370.

Jenkins, J. and Yeni-Komshian, G. 1995. 'Cross-language speech perception: Perspective and promise'. In Strange, W (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues.* Baltimore: York Press. pp. 463-479.

Klatt, D. H. 1976. 'Linguistic uses of segmental duration in English: Acoustic and perceptual evidence'. *Journal of the Acoustical Society of America* 59, 5: 1208-1221.

Maxwell, E.M. 1981. 'The use of acoustic phonetics in phonological analysis'. *Journal of the National Student Speech Language Hearing Association* 9, 20-37.

Nakai, S. 1999. *The perceptual distance between vowels: The effects of prototypicality and extremity.* Unpublished PHD Thesis, University of Edinburgh.

Nearey, T. M. 1989. 'Static, dynamic and relational properties in vowel perception'. *Journal of the Acoustical Society of America* 85, 5: 2088-2113.

Nittrouer, S., Manning, C., and Meyer, G. 1993. 'The perceptual weighting of acoustic cues changes with linguistic experience'. *Journal of the Acoustical Society of America,* 94: 1865.

Nittrouer, S. 2000. *Learning to apprenhed phonetic structure from the speech signal: the hows and whys.* Paper presented at the VIIIth Meeting of the International Clinical Phonetics and Linguistics Association, Queen Margaret University College, Edinburgh.

Ohala, J.J. 1974. 'Phonetic explanation in phonology'. In Bruck, A., Fox, R. and La Galy, M. (eds.) *Papers from the Parasession on Natural Phonology.* Chicago: Chicago Linguistic Society.

Pisoni, D. 1973. 'Auditory and phonetic memory codes in the discrimination of consonants and vowels'. *Perception and Psychophysics* 13, 2:253-260.

Polka, L. 1991. 'Cross-language speech perception in adults: Phonemic, phonetic and acoustic factors'. *Journal of the Acoustical Society of America* 89, 2961-2977.

Polka, L. 1995. 'Linguistic influence in adult perception of non-native vowel contrasts'. *Journal of the Acoustical Society of America* 97, 1286-1296.

Polka, L. & Werker, J. F. 1994. 'Developmental changes in the perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20: 421-435.

Repp, B. H. 1984. 'Categorical Perception: Issues, Methods, Findings. In Lass, N. (ed.). 1984. *Speech and Language. Advances in Basic Research and Practice.* Volume 10. pp. 243-335.

Scobbie, J. 1998. 'Interactions between the Acquisition of Phonetics and Phonology'. In Gruber, K., Higgins, D., Olsen, K. and Wysochi, T. (eds.). *Papers from the 34th Annual Meeting of the Chicago Linguistic Society, II.* Chicago: Chicago Linguistic Society.

Scobbie, J., Gibbon, F., Hardcastle, W. and Fletcher, P. to appear. 'Covert contrast as a stage in the acquisition of phonetics and phonology'. In Broe, M. and Pierrehumbert, J. (eds.) *Papers in laboratory phonology V: language acquisition and the lexicon.* Cambridge: Cambridge University Press.

Stoel-Gammon, C., Buder, E. and Kehoe, M. 1995. 'Acquisition of vowel duration: a comparison of Swedish and English'. In Elenius, K. and Branderud, P. (eds.). *Proceedings of the International Congress of Phonetic Sciences 95.* Stockholm: Arne Stombergs.

Strange, W., Edman, T. R., and Jenkins, J. J. 1979. 'Acoustic and Phonological Factors in Vowel Identification'. *Journal of Experimental Psychology: Human Perception and Performance* 5, 4: 643-656.

Strange, W. 1995. 'Phonetics of Second Language Acquisition: Past, present, future'. In Elenius, K. and Branderud, P. (eds.). *Proceedings of the International Congress of Phonetic Sciences 95.* Stockholm: Arne Stombergs.

Strange, W. (ed.) 1995. *Speech Perception and Linguistic Experience: Issues in Cross-language Research.* Timonium, MD: York Press.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S., Nishi, K. and Jenkins, J. 1998. 'Perceptual assimilation of American English vowels by Japanese listeners'. *Journal of Phonetics* 26: 311-344.

Tyler, A., Edwards, M. and Saxman, J. 1990. 'Acoustic validation of phonological knowledge and its relationship to treatment'. *Journal of Speech and Hearing Disorders* 55, 251-261.

Weismer, G. 1984. 'Acoustic analysis strategies for the refinement of phonological analysis'. In Elbert, M., Dinnsen, D. and Weismer, G. (eds.) *Phonological theory and the misarticulating child.* ASHA Monographs Volume 22. Rockville, MD: ASHA. pp. 30-52.

Werker, J.F., Gilbert, J.V.H., Humhrey, K. and Tees, R.C. 1981. 'Developmental aspects of cross-language speech perception. *Child Development* 52: 349-355.

Wode, H. 1976. Developmental sequences in naturalistic second language acquisition. *Working papers in bilingualism* 11: 1-31.

Young-Scholten, M. 1995. 'The negative effects of 'positive' evidence on L2 phonology'. In Eubank, L., Selinker, L., and Sharwood Smith, M. (eds.) *The Current State of Interlanguage*. Amsterdam: John Benjamins Publishing Co. pp. 107-121.