

Abstract

The present paper reports on the findings and conclusions of an experimental study that investigated the perception of Scottish English /i/-/ɪ/ by Spanish speakers. The conclusions suggest that 1) L2 speakers can learn to identify and discriminate a new contrast, 2) these speakers may also learn to signal a new contrast using an acoustic cue that is secondary for the L1 perception of the same contrast, this could be evidence for “category formation” on the basis of non-native cue weighting, and 3) the L2 results showed to pattern in a way that may suggest stage-like development in the acquisition of new contrasts. Further/current research to find more evidence for these and other related ideas is presented towards the end of this paper.

Introduction

Part of the phonology of a language consists of sound distinctions that 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, first language (L1) speakers 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).

This paper reports the findings and discussion of an experimental study aiming to investigate the adult second language (L2) perception of new vowel contrasts, as well as the way in which such perception may develop. The paper includes possible answers to questions such as: 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 experimental study that will be reported tested the perception of the Scottish English /i/-/ɪ/ contrast by Scottish and Spanish speakers.

Before reporting the experiment and its results, it is worth mentioning some of the theoretical and experimental background (including the previous studies that tested the L2 perception of the same vowel contrast in Spanish speakers). The next section is a brief summary of the relevant, previous literature.

Theoretical background

A number of studies have been devoted to the perception of non-native sounds (see Strange 1995 and all articles in this volume). 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 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 briefly summarise and discuss 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).

The SLM is a mentalist model, 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.). The model suggests that TL sounds that are matched with different L1 sounds will be discriminated in a native-like fashion. However, there is no guarantee that they will be fully perceived in the way native speakers perceive them. The SLM 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: they are *perceptually equivalent*. However, the model also suggests that L2 learners may be able to perceive these non-phonemic features in terms of their experience with 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. In general, 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.

¹ Comments to this paper or any other general query should be addressed to p.r.escudero@reading.ac.uk. They are certainly very welcome.

On the other hand, the PAM assumes that new sounds will get assimilated, in one way or another, to the categories that already exist in the speakers' L1. Consequently, the model suggests degrees and types of assimilation of TL sounds to the L1 perceptual categories (e.g. they may get assimilated to a single category, two categories, etc). That is, 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. The model suggests that L2 speakers have already tuned their linguistic perceptual device to particular features so they would have difficulties in detecting the features that the TL sounds have that are not in their L1. Regarding learnability issues, the PAM predicts the degrees of discriminability but not the process involved in the acquisition of such discriminability.

From the two theories above mentioned, it may be inferred that L2 native-like perception is possible provided L2 speakers detect the sub-phonemic information or features in the TL sounds and *build up* new categories based on them. The models do not mention that L2 speakers should also detect the relative cue weighting of the information involved in the TL categories, neither do they explain what the developmental process for new sounds is. However, an extension of the models could address cue weighting. For instance, 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. In general, the acquisition of the right cue weighting include in a complete native-like perception would be extremely difficult (for some sounds more than for others), mainly because the input does not clearly manifest which cue is being preferred. Therefore, an extension and a better explanation of the above mentioned phenomenon (cue weighting) would seem necessary. Moreover, attempts to describe the developmental process of new contrasts ought to be considered.

The perception of tense/lax vowel contrasts

Studies have shown that the relevant acoustic information for the perception of tense/lax vowel contrasts is divided in two basic dimensions: quality and quantity differences. In terms of quantity, we can think of both the steady-state formant frequency of the vowel (mainly F1 and F2) and the transition of those formants from the start till the end of the vowel. On the other hand, quantity cues will be the vowel intrinsic duration as well as its duration depending on the linguistic context in which is produced (Nearey 1989).

For the experimental study that will be reported here, I chose a variety of English that has tense/lax contrasts in its vowel inventory namely Scottish Standard English (SSE). Within the tense/lax vowel contrasts that the variety has, the /i/-/ɪ/ one was chosen. In this variety /i/, as well as other vowels, is produced as a monophthong, as opposed to its production in General American or Southern British English in which it is produced as a diphthong. Consequently, it is suggested here that format transition (changes in F1 and F2 from the beginning till the end of the vowel) will not play a considerable role in signalling the distinction between Scottish /i/ and /ɪ/. Likewise, it is also argued that regardless of the linguistic context the vowels involved in this contrast will always be distinguished by L1 Scottish speakers. On that basis, intrinsic duration was thought as a relevant acoustic cue for signalling the contrast. Despite the possibility of finding that extrinsic duration does play a role in the perception of this tense/lax contrast, my principal purpose was to investigate the relative role of spectral steady-state cues and intrinsic duration cues only in the perception of tense/lax vowel contrasts. Therefore, Scottish English seemed to be methodologically appropriate to this purpose.

The findings of a first pilot (reported in Escudero 2000a) suggested that L1 Scottish speakers used both cues (spectral and durational) to signal Scottish /i/-/ɪ/. Furthermore, the results showed that the use of spectral cues was higher than the one for durational cues. That is, these speakers attended to spectral information comparatively more: spectral information was primary. These results were gathered by examining the speakers perception of vowels that presented one of the cues only: two continua of seven elements, manipulating one cue and keeping the other averaged for each continuum, were generated using the HLSyn version of the Klatt synthesiser.

Recall that the main interest here was in the L2 perception of new contrasts. Therefore, an immediate general research question is: what is the perception of /i/-/ɪ/ in speakers whose first language does not have this contrast? Spanish speakers of English fit this description and other researchers have previously chosen these language groups (target language, English and L1, Spanish) for examining the L2 perception of the same vowel contrast. Therefore, I will briefly discuss their findings and the findings of a second pilot study (also reported in Escudero 2000a) that retested the same L2 perception.

Previous studies addressing the perception of English /i/-/ɪ/ by Spanish speakers

Firstly, 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 serious difficulty in perceiving and producing English tense/lax vowels (e.g. /i/ and /ɪ/). They also claim that English speaker do use durational cues for the same vowel distinction.

On the other hand, Bohn 1995 assumes a single category assimilation to be the starting point for Spanish speakers' perception of /i/-/ɪ/ 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 the use of a *universal perceptual strategy* 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).

Finally, Flege et al 1997 used the same stimuli as Bohn 1995 did. According to these authors, the relative effect of the two cues in Spanish speakers is similar in percentages and that none of those results is significantly different from 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 statistical but observational. 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: more reliable evidence was needed. And this was the motivation behind a second pilot experiment (Escudero 2000a) set to retest the same L2 perception and compare it to the L1 one. The subjects were tested on the same identification test that had been used for L1 Scottish speakers (Escudero 2000a). The findings reported an L2 cue weighting in line with Bohn 1995 findings: the majority of the L2 subjects used durational cues as primary to signal the contrast. Furthermore, the findings suggested that not all L2 speakers behave in the same way but that they patterned in at least three sub-groups: a) L2 speakers who used durational cues exclusively, b) L2 speakers that used both cues but use durational cues much more than L1 speakers did, and c) L2 speakers that behaved native-like. Despite the revealing findings, the methods (subjects, materials and procedure) of this pilot were not free of problems. Consequently, the question about what cue weighting this L2 group has remained unanswered and further interesting questions emerged.

The research questions and hypotheses for the new experimental study will be now presented. They are the outcome of the questions and theoretical framework that motivated previous studies and they aim to find answers for some of the L2 behaviour that was not fully explained.

Research questions and hypothesis

Question 1: Can L2 speakers learn to identify and discriminate new contrasts?

Hypothesis 1: L2 speakers can learn to identify and discriminate a new contrast. They could perform very well in both discrimination and identification tests that present a vowel continuum with a manipulation of the two cues at the same time. The previous L2 pilot (Escudero 2000a) did not measure this perception. However, it is hypothesised here based on the theoretical framework (the SLM and the PAM) that L2 speakers are able to perceive new contrasts.

Question 2: do L2 speakers always use the same means that L1 speakers use when perceiving new contrasts?

Hypothesis 2: Previous studies (Bohn 1995, Flege et al 1997 and Escudero 2000a) suggest that L2 speakers may use means that are different from the ones L1 speakers use to perceive new contrasts. According to previous findings (Escudero 2000a), it is hypothesised here that some L2 speakers could perceive a contrast using an L1 secondary cue mainly or exclusively. For instance, Spanish speakers could perceive /i/-/ɪ/ using durational cues as the only or the primary cue for the new contrast.

Question 3: Are there more patterns than the contrast and non-contrast ones in the perception of new contrasts?

Hypothesis 3: There at least two more intermediate patterns: a) use of a secondary cue exclusively and b) use of a secondary cue as primary. It is suggested here that the patterns that are found in the L2 performance may constitute a stage-like development. The non-contrast pattern would be the first stage (S1), the a) pattern previously mentioned would be a second stage (S2), pattern b) would be a third stage (S3), and a native-like perception would be the last stage (S4). Note that stages 2 and 3 imply perceiving a contrast. Each one of the previous studies (Fox et al 1995, Bohn 1995, Flege et al 1997) have shown different patterns for the L2 speakers and arrive at general conclusions that may seem contradictory (as discussed above). These studies' data may constitute preliminary evidence and support for the formulation of a patterned/stage-like development.

These three questions and the hypothesis that aim to answer them were tested in an experimental study. The experimental design and findings will be reported and discussed in the following section.

The Experiment

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

The L1 speakers constituted the control group: 10 female and 10 male speakers 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: 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, some of them were from Spain (different regions within the peninsula) and some others from different countries in South America: no attempt to control for particular Spanish accents was considered. Their ages were between 18 and 58 years.

Materials

Two vowel sounds were synthesised based on the naturally produced vowels of two L1 Scottish English speakers. The synthesised vowel sounds had values that were represented in Scottish speakers /i/ and /ɪ/ productions (each sound was produced ten times by each of the 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 2000a), 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 2000a) was solved. Likewise, the quantity problem (the sounds had been too long and might have confused the listeners) was solved by using durational values that were two steps shorter than the previous ones (Escudero 2000a).

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 the synthesis (Sensyn can only accept values in Hertz).

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. The continuum was to be used in a forced identification and a discrimination same/different test. Both tests were to gather the perception of the vowel contrast and the place of the category boundary in both speaker groups.

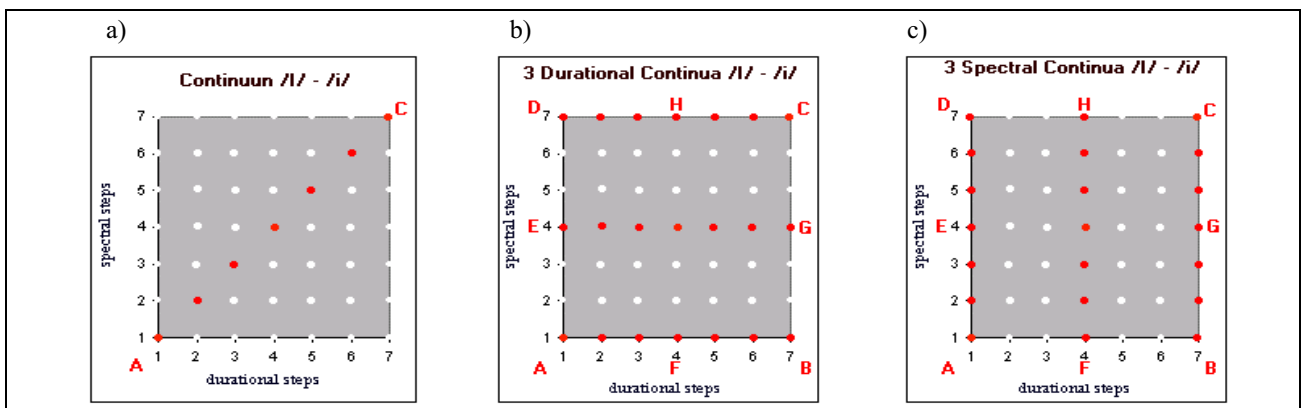


Figure 1. Stimuli used in the experiment divided in different types of continua.

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 was manipulated in the same 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). Consequently, if the subject relied on durational cues, she would be 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 performance in one or two continua would be better than in others. Predictions as to what continuum would be the highest in scores could not be made. A language group difference in scores was predicted.

A second set of continua was made up with stimuli that had a constant duration and involved a manipulation of spectral information. The first continuum had an AC1 durational value, the second an AC4 value and the third an AC7 one. Therefore, if the subject relied on spectral cues, her performance in all these three continua would be very good to good. Variation between continua and subjects was also expected. A performance difference between language groups was predicted. Figure 1 above shows a schematic representations of all the continua: a) the one manipulating the two cues at the same time or AC, b) the ones manipulating durational cues or AFB, EG and DHC and c) the ones manipulating spectral cues or AED, FH and BGC, respectively.

Procedure

The experiment was elaborated using Psycoscope: 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 had three parts and lasted 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 corresponding button. For all three parts of the experiment, the subjects 1) had a visual display in every trial, similar to what they had on the button box, which reminded them of what to press, 2) they were instructed both verbally and in writing, 3) 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 and 4) they knew that the next trial would not appear if they did not make a decision.

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 random order. After each block, and in all three parts of the experiment, the subjects had the opportunity to take a short break.

The second part of the experiment consisted of an 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 corresponding button. This time the subjects 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 (14 sounds in total) in the AC continuum played in a random order.

The last part was, again, an 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). 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.

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 were different.

Some of the L1 subjects were 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). None of the L1 subjects made such judgement.

Results

A first analysis of the data showed that four of the L2 subjects perform the three parts of the experiment in a different manner than expected: their results differed from the other 46. Basically, these subjects either change their decisions for 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.

In the analysis of the “diagonal continuum” AC, only the 26 L2 subjects that showed 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 20 L1 speakers had such distribution. Figure 2 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.

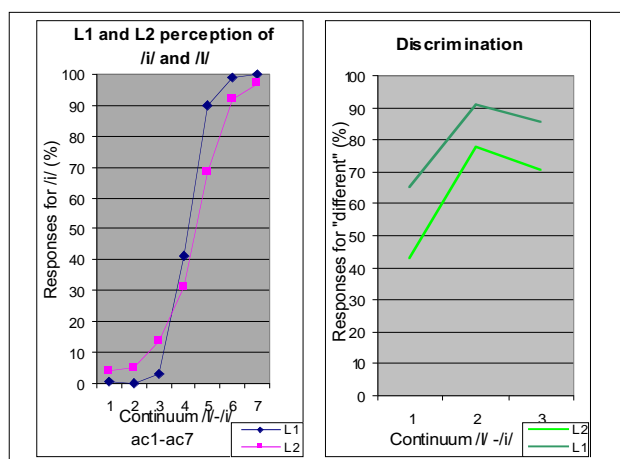


Figure 2. L1 and L2 identification and discrimination of the AC continuum (“diagonal” contrast)

The discrimination scores showed a peak in the category boundary for both language groups (L1 and L2). That, in line with traditional 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. From the same studies we know that vowel perception is not as categorical as consonant perception, the findings here confirm this.

For the identification results, the graphs show 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 light grey colour) was more gradual and not as categorical as the L1 one. After running a factor analysis test that had both the language group and the stimuli as factors, it was 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 all L2 subjects differ from the L1 ones or if it was only some of them that did. A Hierarchical Cluster Analysis (an statistical test) 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.

L1 vs. L2 cue weighting

The group score of 21 L2 subjects that proved to cluster with the L1 group was calculated, separately, for the spectral continua and the durational continua. The results showed very good performance, measured by a steep slope in the category boundary, for the L1 group in the spectral continua. The same group manifested a very poor performance in the durational continua. By way of contrast, the L2 group performed 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”)² 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 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)³. 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 calculate 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.

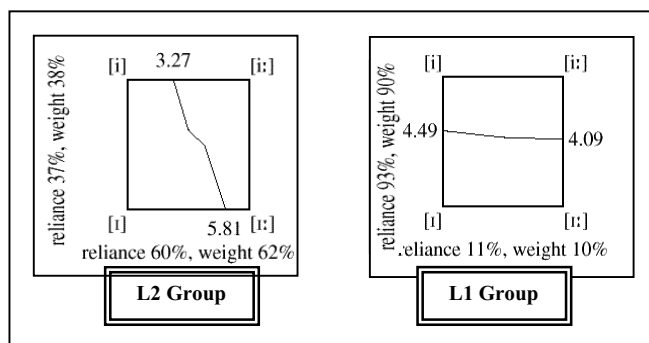


Figure 3. L1 and L2 cue reliance and cue weighting.

In these figures we can also see the category boundaries as represented by a line. The left or the bottom part from the line represents what was mainly reported as ‘ship’ and the right or top part what was reported as ‘sheep’⁴. From the data in Figure 3 above, it could be suggested that the L1 group used spectral cues much more than durational cues and that, conversely, the L2 group used 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

² Special thanks to Dr. Paul Boersma who wrote the script that generated the graphs in figures 2 and 3.

³ 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.

⁴ In this case, ‘mainly’ means more than 50% of the time.

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.

Patterns in the L2 individual cue weighting

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 highly 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.

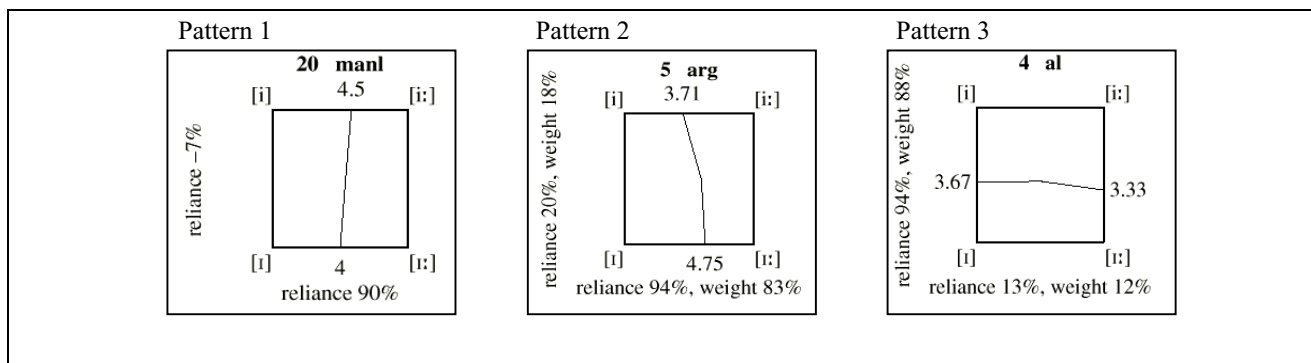


Figure 4. Patterns found in the L2 cue weighting for the /i/-/ɪ/ contrast.

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.

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.

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.

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

For this first section, the possibility of having phonetic category formation 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.

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. 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, James Flege in a personal communication suggested that there might be a different explanation for the L2 subjects' behaviour in general. Flege argued 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: /ɪ/) as non-speech and as differ in length only. Likewise, 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/. 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. 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 the 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. 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*.

Secondly, 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. 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, it is considered here that Spanish speakers of English may have a *covert contrast* (term suggested by Scobbie 1998 for children's production of L1 contrasts on the basis of a different cue weighting) (see also Gerrits 2000 and Nitttrouer 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.

Different performance

The performance of the 4 subjects that do not manifest the expected behaviour needs to be addressed. It is worth mentioning that none of the L1 subjects performed 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 three explanations for this performance. The first one relates to the auditory strategy suggested by Flege (see previous sub-section). 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 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 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.

Another interesting feature in 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.

L2 performance patterns and a possible stage-like development

The results of the experiment seem 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) manifested 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 patterns found in 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. Nevertheless, the patterns seem to be defined and distributed enough to consider them as possible 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.

Concluding remarks

The present study reported here intended to answer three research questions, namely, 1) Can L2 speakers learn to perceive new contrasts? 2) Do L2 speakers always use the same means that L1 speakers use when perceiving new contrasts? And 3) Are there developmental patterns/stages in the L2 perception of new contrasts?

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 are the conclusions specific to the research questions: they were drawn from the experiment findings and the discussion of them:

- 1) L2 learners can learn to perceive new contrasts: 21 out of 30 subjects showed to identify and discriminate /i/ and /ɪ/ in a vowel continuum (with the acoustic information manipulated at the same time) in the same way L1 speakers did.
- 2) 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. That is, 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.
- 3) 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.

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.

First, in order to confirm or disconfirm 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 least bad one

Second, 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 L2 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.

Third, 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. 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. This comparison would give very illuminating insights for the starting point on the acquisition of different English varieties.

Finally, 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 other the two abilities are ultimately attained. It would be expected that ultimate attainment in perception occurs before attaining the same level in production. Likewise, if the perception was completely native-like and production was not, then there could be explanations in line with the idea of *motor control problems*.

To sum up, the longitudinal studies that should be carried out ought to include a comparison between language varieties and between the perception and the production of vowel segments in L2 acquisition. Currently, there is some ongoing experimentation that investigates these issues (see Escudero 2000b and 2000c).

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.
- 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.
- 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 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. 2000b. *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*. Unpublished MSc Dissertation. The University of Edinburgh.
- Escudero, P.R. 2000c. *Input, L1 and universal strategies in the development of a new contrast: the relative weighting of acoustic information*. Paper presented at the 25th Boston University Conference on language Development. Boston University.
- 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., 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.
- 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.
- 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. 2000. *Learning to apprehend 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.
- Pisoni, D. 1973. 'Auditory and phonetic memory codes in the discrimination of consonants and vowels'. *Perception and Psychophysics* 13, 2:253-260.
- 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.
- Strange, W. (ed.) 1995. *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Timonium, MD: York Press.
- 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.
- 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.