Distributional training for discrimination of L2 contrasts: with or without explicit instruction?

MA Thesis

February 2007

Author: Margarita Gulian
0238384

1st Reader: Prof. P. Boersma
2nd Reader: Dr. P. Escudero

Research Master of Linguistics
Faculty of Humanities
University of Amsterdam

Table of contents

List of Figures …...………………………………………………………………… 4
List of Tables ………………………………………………………………………… 5
List of Appendixes ……………………………………………………………………… 5

1. Introduction ……………………….…………………………………........................ 6
   1.1. Introduction…………………………..………………………………............... 6
      1.1.1. Summary of research questions ……………………………………….. 7
      1.1.2. Organization of the thesis ……………………………………………... 7
2. L2 contrasts ………………………………………………………………………… 8
   2.1. L2 contrasts ……………………………………………………………………… 8
      2.1.1. Perception of L2 contrasts ……………………………………………. 8
      2.1.2. Two Potential Single Category contrasts ……………………………… 8
   2.2. Distributional Learning ……………………………………………………………… 13
      2.2.1. Different distributional learning studies ……………………………… 13
      2.2.2. Distributional learning of L2 contrasts ……………………………….. 14
   2.3. Other L2 perception studies ………………………………………………………… 17
      2.3.1. Transferability of L2 contrasts …………………………………………… 17
      2.3.2. Other training studies ………………………………………………… 18
      2.3.3. Other studies of L2 contrasts ………………………………….................. 19
   2.4. Degree of difficulty of the stimuli …………………………….............................. 21
      2.4.1. Synthetic vs. natural stimuli ……….………………….......................... 21
      2.4.2. In isolation vs. in context ………………………………….......................... 22
      2.4.3. Single speaker vs. speaker variability ……………………………........ 22
      2.4.4. Experimental stimuli striving semi-natural environment ………………… 23
   2.5. Summary of the previous studies and hypotheses and research questions of the present study ……………………….. 23

3. The experiment ……………………….. 26
   3.1. Participants ……………………….. 26
   3.2. Methodology ……………………….. 26
   3.3. Procedure ……………………….. 26
      3.3.1. Stimuli ………………………... 27
      3.3.2. Discrimination test ……………………….. 27
      3.3.3. Identification test ………………………... 28
      3.3.4. Training ………………………... 28
         3.3.4.1.Stimuli ………………………... 28
         3.3.4.2.Experimental group 1: Bimodal distribution, explicit instruction…30
         3.3.4.3.Experimental group 2: Bimodal distribution, no explicit instruction…………………………………………………… 30
         3.3.4.4.Experimental group 3: Monomodal distribution, explicit instruction……………………………………………………..30
         3.3.4.5.Experimental group 4: Monomodal distribution, no explicit instruction……………………….. 31
   3.4.Results…………………………………………………………………… 31
      3.4.1. Perception of the two contrasts …………………………….................. 31
      3.4.2. Effect of training……………………….. 36

4. Discussion ……………………………………………………………………… 38

4.1. Discussion of the hypotheses 38
4.1.1. Perception of /a/-/a/ is easier than /i/-/I/ 38
4.1.2. Bimodal training needs to last longer... 40
4.1.3. Instruction deteriorates perception.. 41
4.2.Summary and conclusion...43
4.3.Suggestions for further research... 44
4.4. Practical implications of the findings... 45
References... 46

List of Figures

Figure 1: Situation of the Dutch and Bulgarian /i/ and /a/ vowels in the vowel space, female productions are in bold; Bulgarian vowels are circled (Adank et al., 2004 and Tilkov, 1970)........10

Figure 2: Situation of the Dutch and Bulgarian /i/ and /a/ vowels in the vowel space, vowels in female production are in bold; Bulgarian vowels are circled (Escudero, 2006; Gulian, present)…12

Figure 3: Representation of the distribution of experimental tokens in Maye (2000)…………….15

Figure 4: Representation of the screen for the discrimination task………………………………. 27

Figure 5: Representation of the screen for the identification task for the /ɑ/-/a/ contrast…………28

Figure 6: Representation of the distribution of experimental training items…………………….. 28

Figure 7: First two formants of the two synthetic continua………………………………………. 29

Figure 8: Mean number of assimilation and correct responses in the pre-test …………………… 31

Figure 9: Mean number of correct responses for correct and different pairs in the discrimination test for the /I/-/i/ and /a/-/ɑ/ contrasts in the pre-test and in the post-test, where experimental groups are not differentiated……………………………………………………………………… 32

Figure 10: Mean number of correct responses in the identification test for the /I/-/i/ and /a/-/ɑ/ contrasts in the pre-test and in the post-test, by the four experimental groups…………………… 34

Figure 11: Mean number of correct responses for same and different pairs in the discrimination test for the /I/-/i/ and /a/-/ɑ/ contrasts in the pre-test and in the post-test, by the four experimental groups………………………………………………………………………………………… 35

List of Tables

Table 1: Mean frequencies of Dutch /I/, /i/ and Bulgarian /i/ and /e/ pronounced in context by female and male native speakers………………………………………………………………… 9

Table 2: Mean frequencies of Dutch /ɑ/, /a/ and Bulgarian /a/ and /o/ pronounced in context by female and male native speakers……………………………………………………………….. 11

Table 3: Mean frequencies of Dutch /I/, /i/ and Bulgarian /i/ and /e/ pronounced in context by female and male native speakers……………………………………………………………… 12

Table 4: Mean frequencies of Dutch /ɑ/, /a/ and Bulgarian /a/ and /o/ pronounced in context by female and male native speakers……………………………………………………………...12

Table 5: Mean formants and duration for synthetically formed Dutch /I/ and /i/ sounding like male voice……………………………………………………………………………………… 29

Table 6: Mean formants and duration for synthetically formed Dutch /ɑ/ and /a/ sounding like male voice…………………………………………………………………………………….29

List of Appendixes

Appendix 1: Table with the number of assimilations obtained in the identification test…………47

Appendix 2: Table with the correct responses obtained in the identification test……....................48

Appendix 3: Table of the correct responses obtained in the discrimination test ………………… 49

Chapter 1: Introduction

1.1. Introduction

L2 learners exhibit difficulty discriminating phonological contrasts that are not present in their native language. A particular non-native contrast which Best (1995) defines as Single Category contrast is produced when two contrasting L2 phonemes are equally deviant from their corresponding L1 phoneme. According to Best, Single Category contrasts are highly problematic for L2 learners. The question for language pedagogy is: what is an appropriate training for teaching L2 learners to discriminate such contrasts? In the input used for training, Maye and Gerken (2001) distinguish between bimodal and monomodal distributions of the phones in the input. In the bimodal distribution, listeners are provided with two distinct clusters of tokens along a certain continuum, where the end points of the continuum are presented more often than tokens at the centre; while in the monomodal distribution, the centre tokens are the ones represented more. According to Maye and Gerken, learners can infer new categories on the basis of the statistical distributions of phonetic tokens even without explicit instruction about the number of phonetic categories to be learned. This is how, if there is a bimodal distribution in the input a learner receives, he or she will infer that there are two distinct categories in the L2. Instead, when exposed to monomodal distribution, the learner will not learn to discriminate two contrasting phones.

In order to test Maye and Gerken's (2001) claim about distributional learning, a perceptual experiment was conducted. There it was tested how Bulgarian listeners perceived two contrasts, which were Dutch /i/-/I/ and Dutch /a/-ɑ/, the last one obtained from manipulated shortened Dutch /a:/ and lengthened Dutch /ɑ/. In addition, we trained these listeners to find out whether their Dutch vowel perception could improve. The vowels in the two Dutch contrasts differ from their single equivalents in Bulgarian in their formant values. In this way, it is tested whether two L2 vocalic contrasts differing in a nearly identical way from their L1 equivalents, are perceived comparably before and after training. In this way, besides revealing the effects of training, the results that will come out will shed light on whether two potential Single Category vocalic contrasts are perceived identically before and after training.

To test the effect of training and instruction as well as whether this effect is the same for the two vocalic contrasts, four experimental groups were formed. Thus, Bulgarian naïve listeners of Dutch, who have single /i/ and /a/ categories in their L1, were tested on their perception of Dutch /i/ and /I/ on the one hand, and on manipulated Dutch /ɑ/ and /a/, on the other. The first experimental group received bimodal distribution training without explicit instruction about the number of categories taught. The second experimental group was trained with stimuli distributed bimodally but with an explicit instruction of two categories being taught; in the third experimental group a monomodal distribution without instruction was given and finally in the fourth experimental group monomodal distribution training was accompanied with the instruction that two categories are taught. The first prediction which is made is that the two vocalic contrasts will be perceived identically and that the effect of training will be identical for both, since they are acoustically comparable. Following Maye and Gerken (2001), it can also be predicted that the performance of the participants in the two bimodal groups should be able to discriminate the target contrasts in contrast to the participants trained monomodally. At the same time, according to other training studies using explicit instruction (McClaskey et al., 1983) using instruction in training will help participants transfer their ability to perceive contrasts to other stimuli. Since in the present study the test stimuli differ from the training stimuli, on the basis of the evidence found in other training studies, the groups which receive instruction will be predicted to discriminate the contrasts better than the groups with no explicit instruction, which is in contrast to the claim by Maye and Gerken, mentioned above.

1.2. Summary of research questions

There are two main questions driving this thesis, namely are the two vocalic contrasts perceived identically before and after training? And, what is the effect of training? These two main questions can be divided into the more specific research questions stated below:

> *Research Question 1: Is the number of single-category assimilations of the two L2 contrasts comparable before training?*
>
> *Research Question 2: Is the perception of the two L2 contrasts comparable before training?*
>
> *Research Question 3: Is the effect of distributional learning identical on the perception of the two L2 contrasts?*
>
> *Research Question 4: Is the effect of instruction in training identical on the perception of the two L2 contrasts*
>
> *Research Question 5: Does bimodal training cause a significant improvement in the perception of L2 contrasts with respect to monomodal training?*
>
> *Research Question 6: Does training with explicit instruction cause a significant improvement in the perception of L2 contrasts?*
>
> *Research Question 7: What is the interaction between distributional training and instruction in the perception of L2 contrasts?*

1.3. Organization of the thesis

In Chapter 2, background information of both theoretical and empirical data underlying the present experiment will be given. First, attention will be paid to Best's (1995) definitions of non-native contrasts, followed by the concrete examples of non-native contrasts used in the present experiment. In the next subpart of Chapter 2, the findings of studies of distributional learning in different areas of language acquisition will be presented and after that attention will be paid to studies using distributional training for non-native contrasts and such discussing the transferability of non-native contrasts. Finally, the different aspects influencing the relative difficulty of the stimuli which form the non-native contrasts used in an experimental setting will be discussed. In this respect, three aspects will be discussed: type of stimuli (natural vs. synthetic); context of the stimuli (in isolation vs. in context); variation in the production of the stimuli (single vs. multiple speakers).

In Chapter 3, the experiment which was carried out to answer the above stated research questions will be described. Concretely, this experiment tests the effect of training for the perception of non-native contrasts and the difference between the perception of two comparable vocalic contrasts. This will be followed by a general discussion of the findings in the experiment.

Chapter 2. L2 contrasts

2.1. L2 contrasts

2.1.1. Perception of L2 contrasts

The Perceptual Assimilation Model (PAM) developed by Best (1995) is, as it is deduced by the name, a model concerning language perception which makes explicit predictions about assimilation and discrimination differences for diverse types of non-native contrasts. According to PAM, language users assimilate non-native speech sounds to native categories on the basis of their perceived gestural (articulatory-phonetic) similarities to native phones. The gestural similarities and dissimilarities referred to are, in their turn, based on the model of gestural phonology proposed by Browman & Goldstein (1986 (in Best, 1995) and refer to temporal and spatial properties (degree and location of constrictions) of the dynamic movements of the vocal tract articulators such as lips, jaw, tongue body, glottis etc. PAM as a language model makes predictions about similarities and dissimilarities between non-native segments and native constellations. These predictions are formulated in a list of assimilation patterns of non-native contrasts as for each pattern certain discrimination level is predicted.

The assimilation pattern formulated by PAM that is of biggest interest for this thesis is Single-Category Assimilation (SC) but I will also pay attention to two related assimilation patterns which are Category Goodness Difference (CG) and Two-Category Assimilation (Best, 1995).

Starting with SC, it refers to a non-native contrast in which both non-native sounds are assimilated to the same native category but are equally discrepant from the native ideal, in a way that they are both equally acceptable or both equally deviant. The discrimination of these two native sounds is predicted to be poor, though still above chance level. Examples of SC contrasts, are for instance English /l/-/r/, assimilated by Japanese as the Japanese phoneme /ɾ/; or the Hindi contrast /t/-/ʈ/ assimilated into /d/ by speakers of English.

The CG type is somehow similar to SC in the sense that here too, two non-native categories are assimilated to a single native category, but they differ in discrepancy from the native ideal phone; while one is acceptable, the other is deviant. It is predicted that the discrimination of the CG type will be moderate to very good, depending on the magnitude of difference in category goodness for each of the non-native sounds. Finally, the contrast the discrimination of which is predicted to be most successful is the TC type. In that case, each non-native segment is assimilated to a different native category.

As we see, different predictions are made about how successful the discrimination of the three types of contrasts should be, described by the three assimilation patterns. On the basis of these different predictions, the following hierarchy is formed: TC>CG>SC (Best, 1995). There are several studies testing this hierarchy, most of which are focused on the perception of non-native contrast by naïve listeners.

2.1.2. Two potential Single Category contrasts

As it was defined already in the previous paragraphs, SC refers to a phonological non-native contrast which is assimilated to a single native category, where the two L2 phones are equally deviant or equally similar to the native phone. A first potential case of SC in L2 Dutch is the vocalic distinction between /I/ and /i/. Since this contrast is absent in Bulgarian and the equivalent phoneme is neither identical to /I/ nor to /i/, SC assimilation is predicted. The second potential case of vocalic single category assimilation used in the present experiment is the one between shortened

Dutch /a:/ and lengthened Dutch /ɑ/, to which the phoneme /a/ in Bulgarian corresponds. Below the two SC contrasts will be described in more detail.

Although PAM is entirely a perceptual model, which predicts discrimination of non-native contrasts on the basis of their identification, here I will use the acoustic characteristics of vowels in order to define them as potential SC contrasts. By this it is expected acoustics to be a reliable predictor for perception, which will be tested in the present experiment as well.

Thus to compare both Dutch contrasts to their corresponding Bulgarian category, their characteristic formant values are used. In this way, the phonological analysis of vowels consists in the measurements of their first three formants. In order to be able to compare the phonological categories of two languages, it is preferable to record the phones of interest and analyse them in an identical way. Nevertheless, this is a difficult task for the present thesis, as it requires a study on its own. As compensation, here we will pay attention to two types of data sources for vowel analyses in Dutch and in Bulgarian. On the one hand, a comparison of the vowels in the two languages will be made on the basis of two separate studies of the two vocalic systems. While on the other hand, formant measurements will be used which derive of recordings by Escudero (2006) and of my own which are identical in the way they were recorded and in their formant analyses. The goal of this comparison between two sources of formant measurements is to reconcile the two and get a clearer image of the position of the Bulgarian vowels /i/ and /a/ with respect to the two Dutch contrasts.

Starting with the comparison between the Dutch and the Bulgarian vowels deriving form the literature, concretely the studies by Adank et al. (2004) for Dutch and Tilkov (1970) for Bulgarian will be used. For the Dutch data, the vowels were pronounced in the /sVs/ context by 20 speakers of standard Dutch, while with respect to the Bulgarian data, the formant values were obtained by several productions of the vowels in /sV/ context pronounced by a single female and male speakers of Bulgarian. In Tables 1 and 2, the correspondent formant values for the vowels of interest, plus two additional vowels, the Bulgarian /e/ and /o/ are presented.

First attention will be paid to the front vowel distinction and the phonological characteristics of the two phones in Dutch and their Bulgarian equivalent. In Dutch the two vowels forming the contrast are /i/ and /I/, which have a distinctive function, for instance, in the words *vis-vies* (fish-dirty), where in Bulgarian a single phonological category /i/ exists. Table 1 shows the correspondent formants of the Dutch front vowels, their Bulgarian equivalent and the Bulgarian vowel /e/ (which is of interest in order to see the position of Dutch /I/ with respect to it). The position of the Dutch and the Bulgarian vowels in the vowel space is represented in Figure 1.

Table 1: Mean frequencies of Dutch /I/, /i/ and Bulgarian /i/ and /e/ pronounced in context by female and male native speakers.

		Dutch /i/	Bulgarian /i/	Dutch /I/	Bulgarian /e/
female	F1	294 Hz	300 Hz	399Hz	500 Hz
	F2	2524	2750	2276	2300
	F3	2911	3300	2883	3000
male	F1	278	300	361	400
	F2	2162	2150	1919	1700
	F3	2665	2600	2536	2300

As shown in Table 1 and Figure 1, it is made clear what the position is of the Bulgarian /i/

formants with respect to the Dutch contrast between /i/ and /I/. Looking at the Bulgarian /i/, the following can be concluded for both female and male productions: Bulgarian /i/ appears in between Dutch /i/ and /I/, where it is closer to /i/, with higher F1 and F2 than its Dutch equivalents. This intermediate position of the Bulgarian /i/ with respect to the Dutch contrast, points that the Dutch /i/ and /I/ would form a potential case of SC assimilation for Bulgarian learners of Dutch.

It is interesting to mention that Dutch /I/ is located between Bulgarian /i/ and /e/. This fact is of importance for the experimental groups which receive explicit instructions, since during the instruction session the Dutch vowel /I/ is taught to the participants as 'sounding like between /i/ and /e/'.

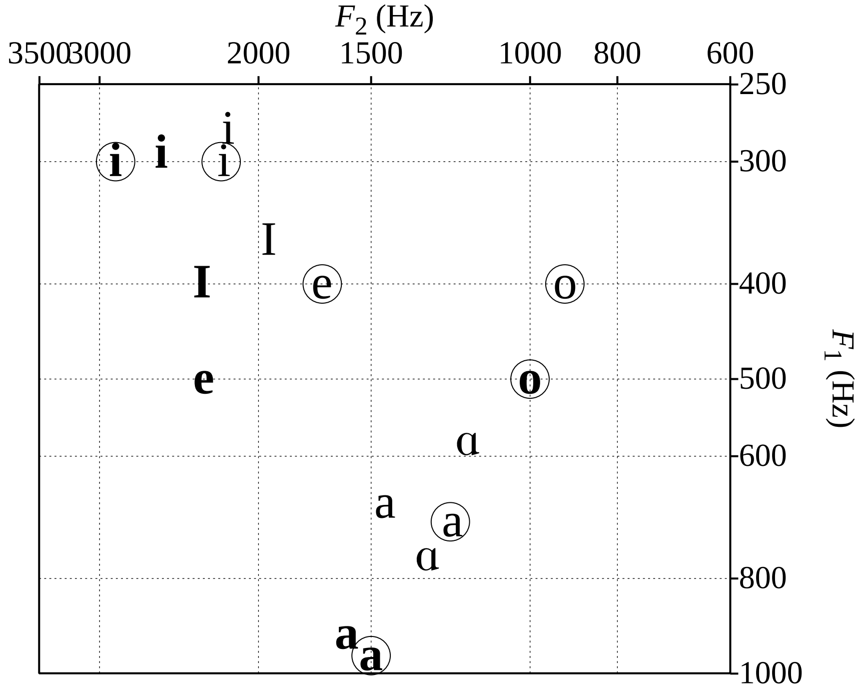

Figure 1: Situation of the Dutch and Bulgarian /i/ and /a/ vowels in the vowel space, female productions are in bold; Bulgarian vowels are circled (Adank et al., 2004 and Tilkov, 1970)

The second contrast used in the present experiment is an existing two-dimensional Dutch contrast, which was manipulated and reduced to a single dimension. Namely, the contrast between Dutch /a:/ and /ɑ/ is present both in frequency and duration. In order to test the perception of a non-native vocalic contrast comparable to this between /i/ and /I/, the duration dimension was eliminated by using a shortened /a:/ and a lengthened /ɑ/. The stimuli used in training were formed synthetically, so for forming them, mean duration was programmed for all stimuli in order to obtain equally long synthetic vowels, while for the test stimuli, natural /a/ and /ɑ/ production in /sVs/ context, were respectively shortened and if necessary lengthened so as to obtain equal duration of all stimuli. In order to be sure that the manipulated natural stimuli were still contrasting for native speakers, a native speaker of Dutch was asked to identify and rate a large number of them and among the correctly identified ones, the best rated were used in the experiment. In Table 2, as in Figure 1, the mean formant values for the Dutch and the Bulgarian /a/ phones and their position in the vowel space are presented. Here too, the values of Bulgarian /o/ are of interest in order to see the position of Dutch /ɑ/ with respect to it.

Table 2: Mean frequencies of Dutch /ɑ/, /a/ and Bulgarian /a/ and /o/ pronounced in context by female and male native speakers.

		Dutch /a/	Bulgarian /a/	Dutch /ɑ/	Bulgarian /o/
female	F1	912 Hz	959Hz	758 Hz	500 Hz[1]
	F2	1572	1500	1280	1000
	F3	2852	2900	2995	2700
male					
	F1	670	700	578	400
	F2	1425	1200	1172	900
	F3	2485	2100	2435	2200

Analysing the position of the Bulgarian central vowel with respect to the Dutch contrast, for both female and male production the following is true: it has higher F1 compared to both Dutch /a/ and /ɑ/, where it is closer to Dutch /a/; its F2 is placed in between the two Dutch phones, where it is still closer to Dutch /a/ in female production and quite in between the two in male production. Just like in the front vowels contrast, here too, the Bulgarian category seems to be located in between the two Dutch phones, where it approaches Dutch /a/, while the position of Dutch /ɑ/ is intermediate between Bulgarian /a/ and /o/. In this way, the Dutch /a/-/ɑ/ contrast as well represents a potential case of SC assimilation.

Up to this point, we have seen how the two Dutch contrasts and their Bulgarian correspondent vowels are located with respect to each other according to Adank et al. (2004) for Dutch and Tilkov (1970) for Bulgarian. In the next paragraphs a different source of formant analyses for a comparison of the languages in the needed contrasts will be used. Namely, identical recordings of the vowels in the two languages were used of which identical formant analyses were carried out. For Dutch the recordings made by Escudero (2006) of native Dutch speakers in /sVs/ context pronounced within the larger context of the type '*In sis en in sisse zit de i/ In sis and in sisse there is an i*.' The same type of recordings were made with Bulgarian native speakers, where the participants pronounced *'Vuv sis I vuv sise ima i'*. Out of these recordings, the vowels of twelve speakers for Dutch and six for Bulgarian were used, of which half were male and the other half were female productions and each speaker pronounced the vowels three times. In total, for each vowel 18 productions were obtained for Bulgarian and 36 for Dutch. Finally the first two and the first three formants for the front vowels were measured in an identical way for all vowels. Tables 3 and 4 show the formant values of the Bulgarian and Dutch vowels and Figure 3 demonstrate the position of these vowels in the vowel space according to the formant measurements of Escudero's (2006) recordings for Dutch and my own recordings for Bulgarian.

[1] The formant values for female producing the Bulgarian vowel /o/ in /sV/ context, Tilkov (1970) gives values identical to the values for male /o/. Since this is not so probable, values for /o/ in /ʃV/ context were used.

Table 3: Mean frequencies of Dutch /I/, /i/ and Bulgarian /i/ and /e/ pronounced in context by female and male native speakers.

		Dutch /i/	Bulgarian /i/	Dutch /I/	Bulgarian /e/
female	F1	336 Hz	370 Hz	449 Hz	506 Hz
	F2	2367	2415	2134	2077
	F3	2883	2911	2939	
male					
	F1	302	281	382	437
	F2	2208	2164	1964	1833
	F3	2766	2941	2631	

Table 4: Mean frequencies of Dutch /ɑ/, /a/ and Bulgarian /a/ and /o/ pronounced in context by female and male native speakers.

		Dutch /a/	Bulgarian /a/	Dutch /ɑ/	Bulgarian /o/
female	F1	735Hz	682Hz	653 Hz	506 Hz
	F2	1664	1653	1372	2077
male					
	F1	681	617	600	486
	F2	1482	1351	1289	990

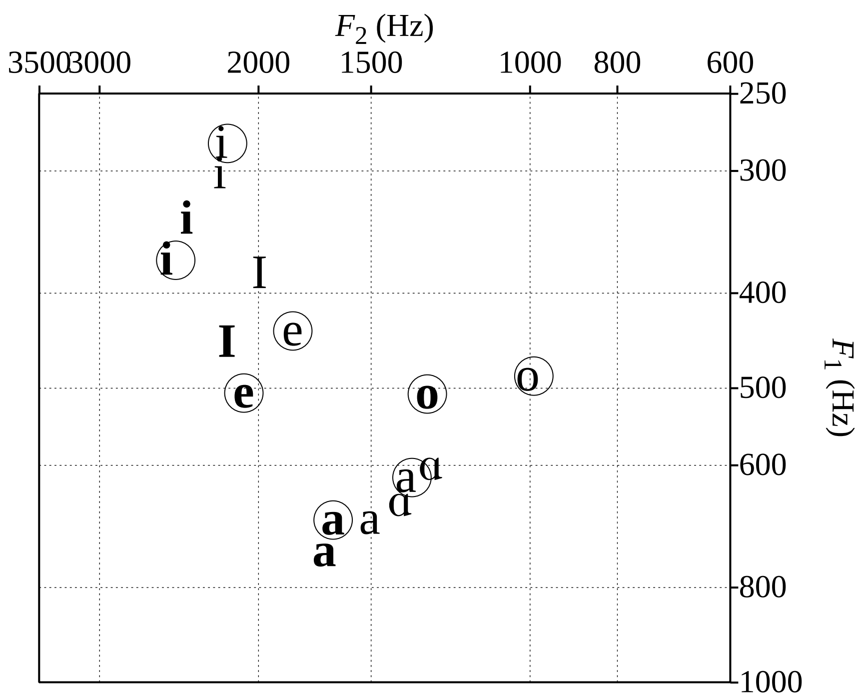

Figure 2: Situation of the Dutch and Bulgarian /i/ and /a/ vowels in the vowel space, vowels in female production are in bold; Bulgarian vowels are circled (Escudero, 2006; Gulian, present).

As shown by the formant values and as it can be seen in Figures 2, the comparison of the

Bulgarian and the Dutch vowels on the basis of Tilkov (1970) and Adank et al. (2004), respectively, does not drastically differ from the mutually comparable recordings of Escudero (2006) and of my own. The two comparisons do not differ considerably except for one vowel, which is the Bulgarian male /i/. While in the comparison between Tilkov and Adank et al. the male /i/ in Bulgarian was in between Dutch /i/ and /I/, where it approached /i/ significantly, here with lower F1 and identical F2 to Dutch /i/, the Bulgarian phone is close to Dutch /i/ but very distanced from Dutch /I/. Reconciliation between the two analyses would be that Dutch and Bulgarian /i/ are almost identical when produced by male, while for women the /i/ takes a rather intermediate position. If we turn to the different types of assimilations predicted by Best, while for female production Single Category assimilation will be predicted, for male production this will be Category Goodness Assimilation, which is predicted to be easier to L2 learners.

Up to this point, we have seen the two analyses of the Bulgarian vowels /i/ and /a/ with respect to their equivalent contrasts in Dutch. Firstly the comparison between the two vocalic systems based on existing literature with formant measurements led to the conclusion that both Dutch contrasts, /a/- /ɑ/ and /i/-I//, are potential cases of SC assimilation. Nevertheless, the second comparison based on Escudero’s and my own recordings and formant analyses on a limited number of productions led to a slightly different picture of the respective position of the vowels in the two languages. Namely, for male production the easier to perceive Category Goodness assimilation was predicted. In the perception test in the present experiment, the Bulgarian participants have to discriminate and identify both female and male productions. For this purpose, I will treat male and female productions of the front vowel contrast identically, both as potential SG contrasts. Nevertheless, I will keep in mind that the /i/-/I/ contrast may result easier than the /a/-/a/ contrast, given the male productions. If this is the case, in the statistical analysis a distinction will be found between the perception of the male and the female productions.

In the present section, it was predicted that the perception of the two Dutch contrasts by Bulgarians would lead to SC assimilation. This prediction was based on the acoustic description of the vowels in the two languages. Nevertheless, this is not the way in which Best (1995) determines the distances between the phonemes of different language systems. Rather, she would predict discrimination to be problematic, if a non-native contrast is identified as SC contrast. In other words, predictions for perception are based on perceptual, rather than acoustic characterisctics of vowels. In this study, it will be intended to find out weather the acoustic characteristics of non-native contrasts are in accordance with their perception. Besides, the effect of training on the perception of the two contrasts is tested.

Before moving to the very experiment, first attention will be paid to studies which relate to the present experiment on different ways. First of all, the findings of studies on distributional learning in different areas of language acquisition than speech perception will be described.

2.2. Distributional learning

2.2.1. Different distributional learning studies

In the present experiment distribution-based training is used for teaching the two non-native contrasts presented above. There are a number of studies which have already tested the effect of distribution-based training on the discrimination of L2 contrasts, but even before, the importance of distributional information for language learning was attested in other linguistic areas. From these studies it has been proven that humans are able to keep track of the frequency with which events occur (Hasher et al., 1987). In the following paragraphs some relevant findings of similar distributional learning studies will be presented.

Hasher et al. (1987) found evidence that adults were able to detect frequency even without explicit concentration on this aspect. Namely, they found out that the participants can reliably remember frequency of occurrence information about items they have been exposed to under truly incidental memory conditions. Subjects neither knew that the ultimate test task would concern item frequency nor that they had any reason to remember the items, since they were only doing a task in which pictures of objects are related with certain names (the Stroop task). Finally, it appeared that subjects who were made aware of something prior to the task for either a non-specific memory test or a frequency test were no better able to judge frequency than were subjects operating under truly incidental conditions.

Several other experiments have reported evidence that infants are able to learn statistical patterns in auditory input. Saffran, Aslin, and Newport (1996) investigated the segmentation of artificial speech by infants by training them with high and low probability syllable sequences. They tested the infants using the head-turning procedure on words identical as in the training, partly identical words (syllable sequences from the training but crossing a word boundary) and non-words. Infants listened longer to the partly familiar and the non-words, and in this way showed a novelty preference and thus were able to discriminate the familiar words from the other stimuli based on learning the transitional probabilities defining word boundaries.

These findings provide evidence for a functional statistical learning mechanism, available to infants, that segments speech into words on the basis of computation of input statistics, and gives rise to questions concerning the generality of this ability. Saffran, et al.(1999) found that both infants and adults detected transitional probabilities of non-linguistic tone sequences. The participants in this study received statistical information as the only boundary cue and they nevertheless succeeded at segmenting the tone stream. These findings indicate that statistical learning is not a purely linguistic mechanism.

Further evidence for generalizability comes from experiments by Aslin (2001) with non-human primates (cotton-top tamarins, a species of New World monkey). After exposure to the same set of auditory stimuli employed by Saffran et al. (1996), adult monkeys showed reliably greater interest in both non-words and part-words than in the familiar words, suggesting that they were able to extract the statistical information defining word boundaries in the artificial speech, in like manner to human infants. Taken together, this study demonstrates that humans are not the only primates who have the ability to keep track of statistical regularities.

The follow question remains: is the ability of capturing distributional information also applicable to L2 speech perception? As mentioned in the beginning of this section, recently these effects of distributional information have been investigated too. In the following paragraphs more details about these studies will be given.

2.2.2. Distributional learning of L2 contrasts

In the first studies on distribution-based training in speech perception, participants received only acoustic information about the target language and no information about what distinguishes words in that language. This kind of training exploits the distributional tendencies in speech with respect to phonemic distinctions in the cases when different phonemes have distinct distributions on some acoustic dimensions. (Maye, 2000; Maye et al. 2002). Later on Hayes (2003) investigated, whether the interaction between distribution-based training and teaching of minimal pairs had a positive effect on perception. In the following paragraphs, these studies will be observed in further detail.

Starting with Maye's dissertation (2000), she trained native speakers of English either to discriminate or not to discriminate pre-voiced stops from voiceless unaspirated stops([d] vs. [t^{-h}]). These differences are subphonemic in English, since they do exist in English but never occur in the same environment. So, they could be discriminated under certain task conditions. Maye created a series of eight tokens on continua between prevoiced [d] and voiceless unaspirated [t^{-h}] where token 1 is the most [d]-like and token 8 the most [t^{-h}]-like. The different tokens appeared in three different vowel environments, which served as experimental tokens. One set of participants heard a bimodal distribution of these tokens, while another heard a monomodal distribution of the tokens during training (see Figure 3).

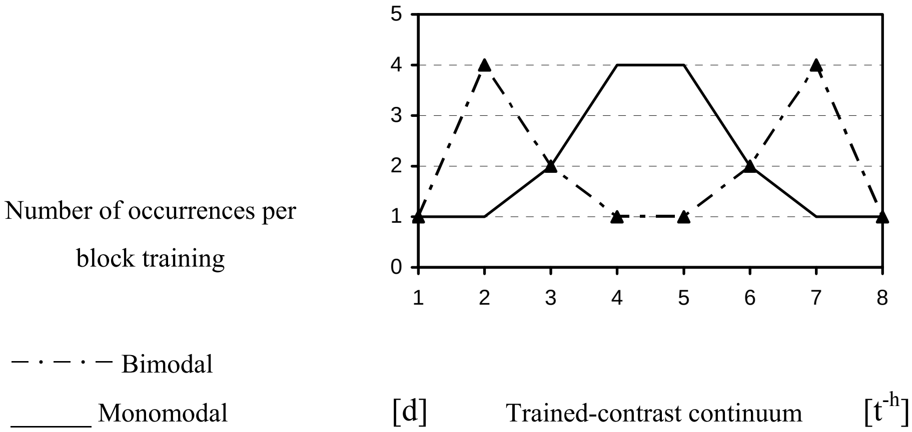

Figure 3: Representation of the distribution of experimental tokens in Maye (2000).

The bimodal group heard tokens 2 and 7 four times more often than the tokens in the middle of the continuum, whereas the monomodal group heard tokens 4 and 5 four times as often as the tokens at the endpoints of the continuum. Maye hypothesised that the bimodal training group provides listeners with two distinct clusters of phonemes along the continuum, while the monomodal training provides them with only one phoneme along the continuum. To distract the attention of the participants from the trained contrast, also filler items were included. After the training the participants were given a discrimination task in which they heard two monosyllabic words, which they had to characterise as same or different.

With respect to the test items, the same stimuli used during training were used and there were four types of test pairs: experimental-same; experimental-different; filler-same and filler-different pairs. It was found that participants trained with the bimodal distribution of tokens were better able to discriminate the two phonemes forming the contrast than participants trained in the monomodal distribution. The different groups were compared in their performance in the experimental-different pairs, since performance on all other stimuli was near-ceiling.

In a second experiment, the generalizability of the contrast was tested. For this purpose, two groups were trained on the [d]-[t^{-h}] contrast and tested on this contrast followed by a test on a comparable contrast between [g] and [k^{-h}]. These groups were respectively a bimodal training and a monomodal training group. Besides, there were two other groups in the same training conditions but these two groups were trained on the [g]-[k^{-h}] contrast and consequently tested first on the same and then on the [d]-[t^{-h}] contrast. Just like in Experiment 1, the results in Experiment 2, showed an effect of bimodal training above monomodal training. Nevertheless, the participants were not able to generalise the contrast they were trained on, since they could not transfer their knowledge to another similar contrast. This result indicates that the participants did not extract a phonological feature corresponding to the [g]-[k^{-h}] or [d]-[t^{-h}] contrasts. Maye interpreted these results as indicating that participants in the bimodal group were able to learn the phonemic status of the [d]-[t^{-h}] contrast based only on distribution-based information. Finally, it is of interest that the participants trained on the velar contrast scored lower overall on the trained contrast, in comparison with the group trained on the [d]-[t^{-h}] contrast.

As mentioned above, while Maye (2000) studies the effect of distributional input for the discrimination of non-native contrasts without explicit instruction of the words' meanings, Hayes (2003) compares the effect of distributional input with that of learning minimal pairs. Besides, the latter author is interested in the interaction of the two manners of training to discriminate a non-native contrast. Bellow, the two main experiments in Hayes (2003) will be described.

In Experiment 1, Hayes compared distribution-based evidence for learning to discriminate a non-native contrast to lexical contrast-based evidence. For this purpose, six different training groups and their performance on a discrimination test was compared. First, there were monomodal and bimidal training groups (replicated from Maye's study (2000). Groups 3 and 4 were planned to test the effect of lexical evidence, which were a contrast training group and a no-contrast training group. In the contrast group, items 2 and 7 from the continuum replicated from Maye (see Figure 3) were matched to two different meanings, while in the no-contrast group, items 2 and 7 were matched to the same meaning. The other two groups were rather control groups, which were: the near-endpoint training group, in which participants heard only tokens 2 and 7 but which were not linked to meaning since they saw no pictures and the no-training group, where the participants were tested without having received any training.

For the discrimination task, Hayes used the same stimuli as Maye (2000). In Hayes' study too a ceiling effect was found for all except for the experimental different pairs. For this purpose, the six different groups were compared in their performance in the experimental different pairs and it turned out that the bimodal group performed significantly better than the monomodal group, but that from all groups the contrast group showed the most accurate performance overall. Since an effect was found for both distribution-based evidence and lexical-based evidence, the question arose which of the two has a greater influence on learning the contrast. In order to be able to tease apart the effect of each of the two types of evidence the second experiment was carried out in which distribution based and lexical-contrast-based evidence were combined.

In Experiment 2, in order to compare the two types of evidence four experimental groups were formed. The groups could vary in two variables, first this was distribution, bimodal vs. monomodal and second, lexical contrast vs. no lexical contrast. In this way, the bimodal-contrast group obtained double evidence for a non-native contrast, the monomodal-no-contrast group obtained double evidence against such contrast, while the monomodal-contrast and bimodal-no-contrast groups received conflicting evidence. Looking at the results, it was found that the bimodal-contrast group performed significantly better than the monomodal-no-contrast group, while the comparisons between all other groups showed no significant differences. Hayes concludes that the results from Experiment 2 do not show evidence that either the distribution-based or the lexical-based evidence plays a stronger role in learning a novel contrast.

Finally, Hayes compares the groups from Experiments 1 and 2 with the intention to see which of the two types of evidence has a greater impact in the discrimination test. What calls to attention from this comparison is the lack of significant difference between both the bimodal and the contrast groups from Experiment 1 and the bimodal-contrast group from Experiment 2. And since neither Experiment 2, nor the combined results from Experiments 1 and 2 revealed which of the two types of evidence had a greater effect on the discrimination of a non-native contrast, the overall better performance of the contrast group in Experiment 1 was recalled. This makes Hayes conclude that the availability of lexical-based evidence results in more accurate discrimination of two new sounds than when obtaining distribution-based evidence. With other words, it is suggested that knowing the word meanings and the minimal pairs formed, leads a L2 learner to more accurate discrimination of a novel contrast than statistical evidence alone.

To sum up, three findings need to be outlined. First, both Maye (2000) and Hayes (2003) found that bimodal training of non-native contrasts leads to more accurate discrimination than monomodal training. Second, Hayes found that learners benefit more from an awareness of the meaning of words and of a minimal pair than from hearing a bimodal distribution of sounds without meaning, when learning a non-native contrast. Finally, it was found by Maye that although bimodal training led to better discrimination of non-native contrasts, it did not help L2 learners to transfer this ability to other novel contrasts of identical nature. In the next paragraphs, attention will be paid to other studies which investigate the aspect of transferability of non-native contrasts.

2.3. Other L2 perception studies

In the present subsection of the thesis a number of studies on L2 perception will be described, divided into three categories. First this will be studies on transferability, followed by other studies in which training was submitted and the third category will give an overview of studies concerning the perception of non-native contrasts, which do not make use of training.

2.3.1. Transferability of L2 contrasts

As we have seen, in her dissertation Maye (2000) investigates, among others, the subject of the transferability of non-native contrasts. There are more studies on the topic, which mainly differ in three aspects: the type of training administered; the type of contrast taught, and the kind of transferability investigated. In this way, as presented above, Maye compares the effect of a monomodally distributed vs. bimodally distributed input in training. Listeners are trained in a consonantal contrast and are tested whether they could transfer it to another comparable contrast (pre-voiced stops from voiceless unaspirited stops) but differing in place of articulation. Below the studies by Wang (1994) and McClaskey et al. (1983) on transferability will be observed in more detail, paying attention to the above stated three aspects in which they could differ.

The study by Wang (1994) differs in all three aspects from Maye (2000) since it investigated the learning of four different tones in Mandarin, used extensive training and tested whether the listeners can transfer it to different speakers. The four tones in Mandarin are mainly differentiated in F0 height and contour but since English is a language without contrastive tones, perception studies have shown that English speakers mainly rely on F0 height and are therefore not fully successful in tone discrimination (Wang, 1994). In her experiment Wang provides American English learners of Mandarin Chinese with explicit training of how to discriminate the four tones in that language. For that purpose the participants receive a two-week training programme, consisting of eight training sessions of 40 minutes each and feedback after every session. In the training, natural stimuli were used according to a “high-variability paradigm”, in which natural stimuli are presented in various phonetic contexts, produced by various talkers. This meant that six different speakers pronounced the four tones in different contexts in a way that different stimuli were available for the tests, the training and the generalization tests. In the generalization task, the ability to transfer the learned contrast to different speakers was tested. The results pointed to a significant improvement in distinguishing the different tones in the generalization test compared to the pretest. In other words, the explicit training helped the participants to learn to discriminate the four Mandarin tones and to transfer this ability when perceiving tones produced by new speakers.

After seeing Wang's (1994) study on transferability which differs in all the three aspects from the one by Maye (2000), a last attention will be paid to the study by McClaskey et al. (1983), which shares common aspects with both studies. Namely, comparable to Maye, here too a three-way consonant non-native contrast was synthetically formed and speakers of English were trained to discriminate the three phones. English divides stop consonants differing in VOT in two categories at each place of articulation (/b-p/, /t-d/, /k-g/), so for native speakers of English, a tripartite contrast is unusual. The participants in this experiment were tested with means of familiarization task and training with immediate feedback on a tripartite contrast (either labial or alveolar) and consequently their ability to transfer this contrast to a new place of articulation was tested, by testing labial-trained participants on the alveolar contrast and vice versa. As in Wang, here too the training was rather explicit, while the transferability tested is more of the kind of Maye, since generalization of the new contrast to another place of articulation is meant instead to other productions of the contrast. With respect to the results, McClaskey et al. reported that subjects who were able to identify the three voicing categories of the trained stimuli were consistently able to do so at the other place of articulation.

To sum up, it turns out that the participants in Maye's study (2000) who obtained implicit training were not able to transfer their ability to discriminate the new contrast to another place of articulation, while the participants in McClaskey et al. (1983) succeeded in a similar task after explicit training. Though testing a different type of transferability, Wang (1994) also reported a positive effect of explicit training on the ability to transfer a learned non-native contrast to productions of new speakers, which was much longer than the one used in the other two studies. If the findings of these three studies are considered comparable, two conclusions could be made. First, the discrimination of a contrast identical and produced by the same speakers as during training is easier to discriminate than new contrasts or contrasts produced by different speakers. Recall that in Maye, after obtaining a certain amount of bimodal training, the participants learned to discriminate a new contrast but not to transfer it to a new place of articulation. Second, by comparing the three studies, it turned out that explicit training is more successful than implicit training when transferability is tested.

2.3.2. Other training studies

Up to now, different training studies were described, first the distributional-based training studies, followed by studies concerning the topic of transferability. Almost all these studies differ in the type of training. Attention was already paid to the explicit vs. implicit character of the trainings presented up to this point. Another aspect in which training studies could differ is in the amount of training. With exception of Wang (1994), the rest of the described training studies used a single training session to teach the non-native target contrast. Instead, Wang, as it has become clear from the previous paragraphs, used a two-week training programme, which consisted of eight training sessions and each of these sessions lasted 40 minutes. Compared to the one-session training studies, Wang's eight-session training is a rather long-lasting one. In the following paragraphs other studies with more extended training will be mentioned.

García Pérez (2005) trained Spanish-speaking learners of English with a short residence in Canada to identify three vocalic contrasts in Canadian English /i-I/, /u-U/ and /a-A/, the last represented in the words *boss* vs. *strong*, according to another, production training study (García Pérez, 2006). With respect to the training, it is comparable to the type of training in Wang. Here too, the participants obtained explicit training, which in total lasted three weeks consisting of nine training sessions of 40 minutes. Besides, at the end of each session the participants made a test with feedback. Their perception was assessed with a pre and post-test design. Finally the results showed that there was a significant improvement in the identification of the contrasts /i-I/ and /u-U/ in the post-test and that the identification of /a-A/ improved too, even though the participants already had more than 70% correct responses on the pretest. As it becomes clear, as in Wang, here too the extended training with which the participants were provided resulted in a significant improvement in the perception of non-native contrasts.

The amount of training appears to be crucial for learning a new contrast, since there are as well studies which report the ability of animals to react to different phones. In this way, Kluender et al. (1998) trained birds, namely European starlings, to respond differently to the synthetically formed English /i/ and /I/ or the vowels corresponding to Swedish /y/ and /u/. This was attained by depriving the birds from food and training them to peck differently to vowels from the distribution of /i-I/ or /y-u/. For half of the birds pecks to /i/ or to /I/ were reinforced while for the other half theirs pecks for /y/ and /u/ were reinforced. The starlings received training to the point when they reacted to the trained vowel as expected, which resulted in 80 days of training, in which 5120 trials were made. Finally, the starlings were tested with novel stimuli, which were phonetic variants of the vowels by means of an "identification task" in which they perceived a pair of phonetic variants of the same vowel or a pair of phonetic variants of two different vowels. The results pointed significantly grater pecks per minute for different pairs than for same pairs.

These amazing findings are used by Kluender et al. (1998) to argue against innate predispositions of phonetic categories, replacing the innateness issue by the great role of experience in language acquisition. For the present thesis, what is of interest is the ability even of animals to become sensitive for phones which are new for them after a great amount of training. The conclusion that could be made here is that, besides the type of training, the amount of training is crucial for learning to discriminate a non-native contrast. Now that we have paid attention to a number of training studies below, other studies on the perception of non-native contrasts which do not make use of training will be presented.

2.3.3. Other studies of L2 contrasts

Here I will give an overview of a number of studies in which merely the perception of different non-native contrasts is tested and in which training is not provided. What is of interest for the present thesis is how, depending on the contrast and on the circumstances in which it was tested, different outcomes could be obtained. Therefore attention will be paid particularly to three studies which are Pallier et al. (1997), Werker and Tees (1984) and Bion et al. (2006).

Starting with Pallier et al. (1997), in their study the perception of a particular contrast by bilinguals with two different backgrounds was tested. Namely the participants were all Spanish-Catalan bilinguals, but while the first group was formed by Catalan raised bilinguals, the second was composed by Spanish raised speakers of Catalan. The two groups were tested in their perception of the Catalan contrast /e-ɛ/, which represents a case of Single Category Assimilation, given that in Spanish a single /e/ phone exists, which is different from the Catalan /ɛ/ and still more open than the Catalan /e/. To test their perception, the participants first had to classify synthetic tokens forming a /e-ɛ/ continuum either as /e/ or /ɛ/. Second, they were tested with a discrimination task, in which they had to decide whether they perceived two times the same phone or two different ones. The results in both tests, showed a significantly poorer performance of the Spanish raised participants, except for some individual near-native performances. The authors conclude that there is lack of behavioural plasticity in mid-late bilinguals and that early and extensive exposure to L2 is not sufficient to attain the ultimate phonological competence of native speakers.

In Werker and Tees' (1984) study the experimental settings were manipulated in which the perception of two non-native consonantal contrasts for speakers of English was tested. In particular, the two contrasts were first the Hindi contrast between voiceless unaspirated retroflex versus its dental variant /ʈa-ta/ and second the Thompson contrast between a glottalized velar and a glottalized uvular / ki-qi/. Here too, the participants were tested with a discrimination test in which they had to decide whether they perceived two same or two different phones. Two of the experimental groups varied in the between stimuli and between trial times, resulting in what we could call a quick and a slow group. In the quick group, within each trial the two stimuli were separated by 500 ms of silence and there was 2500 ms of silence between the different trials. In the slow group, the between stimuli time was 1500 ms and the between trial time was 3000 ms. The results showed that the performance of the quick group was significantly better for both the Hindi and the Thompson contrasts. This makes the authors conclude that under particular circumstances non-native contrasts could be discriminated and that there is a memory trace available after 500 ms which decays after 1500 ms and that this memory trace helps listeners to detect differences within phonemic categories.

The third study we will look at concerns the question of whether perception or production has a primary position by Bion et al. (2006). Here attention will be paid to the findings with respect to perception. The participants were Brazilian proficient learners of English and their perception of two front vowel English contrasts was tested (/i-I/ and /ɛ-æ/). For this purpose an oddity discrimination test was carried out, in which each trial contained three stimuli, which were either all identical or containing a single odd stimulus. The results pointed to a significantly more accurate performance of the /i-I/ contrast with respect to the /ɛ-æ/ contrast. Unfortunately, the study does not address the exact position of the English contrast with respect to the Brazilian Portuguese vowel system in a way that it remains unclear whether the contrasts are potential Category Goodness or Single Category assimilation contrasts within Best's Perceptual Assimilation Model (1995).

In the present subsection we have become familiar with three different perception studies. Though relatively heterogeneous, using their findings, a number of conclusions could be made. First looking at Pallier et al. (1997), their findings suggest that native-like perception of non-native contrasts is hardly attainable given that even mid-late bilinguals with a large amount of L2 input exhibit imperfect performance. The consequences of these findings is that if after training near-native perception is planned to be tested, the methodology should be considered critically, in other words, if near-native perception is assessed, the same success is not guaranteed in a natural environment. Besides, the effect of short term memory should be taken into account as it was pointed out by Werker and Tees (1984). This effect could have two consequences: first when an experiment entails both training and a perception test, it is better to use different stimuli for these two parts. Secondly, it is advisable to administer a perception test over a period of time in order to control for the effect of short term memory. Finally, the conclusion that could be made from Bion's et al. (2006, in press) findings is that two apparently identical non-native contrasts could be perceived significantly differently.

In the next section, particular attention will be paid to the different levels of transferability determining the relative difficulty of test stimuli. In a previous section transferability to a different contrast than the trained one was mentioned, but now, we will only look at transferability of the very contrast that was trained. Wang (1994) already tested transfer of trained contrasts, when placed in different contexts and produced by different speakers, which are one of the types of transferability that will be discussed further in the next section.

2.4. Degree of difficulty of test stimuli

Up to this point, some conclusions from perception experiments were made. One of these conclusions was the recommendation of using different stimuli in testing than in training. In this way it is controlled for short term memory and transferability. Another way to control for effects of short term memory is by testing perception over a certain period of time. Still, if an immediate perception test is administered, the possibility of using different stimuli than the ones used in training is also very adequate. As mentioned above, also by using different stimuli, it could be tested whether the learned contrast can be transferred to cases of phonological variation of the same contrast. In this way there are three degrees of transferability which could be tested, which are, whether synthetic stimuli are transferable to natural ones, whether stimuli pronounced in isolation could be transferred to stimuli in context and finally, in case the stimuli are natural, whether when pronounced by the same speaker, these could be transferred to different speakers. The different degrees of transferability, some of which have been discussed above, are directly related to the relative degree of difficulty of test stimuli, since each degree of transferability makes the stimuli used in a perception task more difficult. In the following paragraphs, attention will be paid to each of these degrees, which will be finally related to the stimuli used in the present experiment.

2.4.1. Synthetic vs. natural stimuli

As mentioned, one of the types of transfer is from synthetic stimuli to natural ones or vice versa. The question is whether the perception of natural stimuli is more difficult than perception of synthetically formed stimuli. A study which investigates this particular topic is the one by Mochizuki (1980) where proficient Japanese speakers of English and native English speakers were tested in their perception of English /r/ and /l/. The study is of interest as in its methodology the nature of the test stimuli and the context they are in were manipulated. First, Mochizuki investigates the perception of natural productions of /r/ and /l/ in CV context followed by productions of the consonants placed in different positions within a word. The results show accurate discrimination of the two consonants in CV context. With respect to the perception of /r/ and /l/ within a word, it turned out that some positions resulted easier than others. In this way the Japanese learners performed more accurately identifying both /r/ and /l/ in word final than in word initial position.

The perception test in which synthetic stimuli were used had a rather different character. The two groups of listeners had to label as /r/ or /l/ 11 stimuli original from a synthetic continuum between the two. From this test it turned out that the Japanese labelled the /l/ stimuli almost as accurately as the native speakers. However, they had problems with labelling the /r/ stimuli. This finding makes the author conclude that perceiving naturally produced non-native stimuli may result easier to L2 learners than the perception of synthetic stimuli. Besides, with respect to the perception of natural stimuli, Mochizuki concludes that when these are produced in a certain contexts, like word initial position, they could be more difficult to perceive.

What is of interest for the present thesis is that according to the findings by Mochizuki (1980), perception with natural stimuli after training with synthetic ones should not result difficult to L2 listeners. Still we should not forget that Mochizuki's study only gives evidence for consonant perception, while it might be the case that vowels are more "sensitive" to synthesis or in natural perception, compared to consonants. In any case, the fact that different type of stimuli are used in training and subsequently in testing, entails a degree of transfer which makes the perception more difficult. We will have this implication in mind, when discussing the stimuli for training and testing used in the present experiment.

2.4.2. In isolation vs. in context

Another type of transfer possible is from training with vowels in isolation to testing with vowels in context. A study by Kewley-Port (1995) gives an answer to the question of whether one of these two types of perception of vowels is more problematic. In this study, speakers of English were tested for the thresholds for the English vowel /I/. Namely, they were first trained on a 14-step synthesised /I/ in isolation and in a number of contexts. After that, they had to identify the synthetic vowels, determining their corresponding thresholds for each different position. The results showed that for some participants consonantal context made it more difficult to resolve the formant frequency than for vowels in isolation. A concrete result was that with respect to CVC context, manipulation of F2 degraded the discrimination of formant frequency more than manipulation of F1, where the last sort of manipulation did so especially in the /lIl/ context.

Another study where the difficulty in the perception of vowels in context is suggested is the study by Fry et al. (1962). Here too, synthetic phones are used as testing stimuli, namely, the English /I/, /ɛ/ and /æ/ were synthesised and the space in between was divided into 13 steps. In an identification task, the authors combined more or less prototypical exemplars of two of the target vowels in ABA or ABB type of grouping and in this way could compare the effect of context in vowels with non-speech sounds and consonants in context from a study by Eimas (1962, in Fry et al. 1962). They found out that the effect of context on vowels is much larger compared to consonants and non-speech sounds, advising that when vowels are trained, vowels in context should be used.

In summary, the two different studies point to the relative difficulty of the perception of vowels in context. As they point out, vowels in consonantal context appear to be more difficult to recognize than isolated vowels. For the present experiment this is of particular interest, since it suggests that when training with isolated vowels is given, testing perception with vowels in context would result more difficult. This difficulty is added to the difficulty entailed in the transfer from one type of stimuli to another.

2.4.3. Single speaker vs. speaker variability

The last type of transferability which entails on its turn another degree of difficulty is the transfer from single speaker to multiple speaker perception. As we have already seen in Wang (1994), there the intensive training makes the participants able to perceive the four tones in Mandarin produced by different speakers, though in the training session a single speaker production was used. In their study, it is assumed that perception of multiple speakers is more difficult than perception of a single speaker. This assumption is attested by Golginger's et al. study (1990) where they test the accuracy in recalling lists of spoken words produced by a single speaker vs. multiple speakers.

According to another study by Martin et al. (1989, in Goldinger, 1990), when recalling lists of spoken words, talker variability affects not only speech perception but memory processes as well, so that word lists produced by multiple talkers require greater processing resources for encoding and rehearsal in working memory than lists produced by a single talker. The findings in Golginger et al. provide support to Martini's et al. view. They found a strong relation between talker variability and rate of presentation of the word lists. Concretely, it turned out that recall of multiple talker lists was affected significantly more than single talker lists by changes in presentation rate. The authors concluded that processing talker specific characteristics requires a complex interplay of perceptual, attentional and memory processes.

What is of interest for the present experiment is the finding that the perception of stimuli produced by multiple speakers results more difficult than the production of single speaker. This would mean that transfer from training using stimuli of the same speaker to a test where multiple speakers' productions are used will entail a two-folded difficulty, first, from stimuli which are easier to perceive to more difficult ones and second because of the different nature of the stimuli in the two conditions.

2.4.4. Experimental stimuli striving semi-natural environment

In this subsection of the thesis, we have paid attention to three types of transfer of the stimuli used in training with respect to the stimuli used in testing, two of which appeared to result in an additional degree of difficulty. Namely, testing perception with different stimuli than the ones used in training will always be more difficult to the listener than when identical stimuli are used. Besides, this difficulty will be increased when either transfer from isolated vowels to vowels in context, or from single speaker to multiple speaker productions of the stimuli is tested. Still, from the studies observed above, it turned out that transfer from synthetic stimuli to natural ones will not result in an additional difficulty of the stimuli.

In the experiment reported in this thesis, which will be described in detail in the next chapter, the three types of transfer mentioned above are all combined and used to test the effect of training on the perception of Dutch vowels. In particular, the stimuli used in training are synthetic, isolated and sound as pronounced by a single speaker, while the vocalic stimuli in the perception tests were natural, produced in a consonantal context and by several different speakers. In this way a maximal difficulty of the test stimuli is reached.

These relatively difficult test stimuli are intended not only to test the ability of non-native listeners to transfer their knowledge from training in a three-folded way but also to test these listeners with stimuli which are closest to a natural environment. The goal of testing such three-folded transfer and perception in a semi-natural environment is to make conclusions about the universality of the training method. With other words, if an effect of training is found, we can only conclude that the object of training is learned, when a semi-natural methodology is used. Such a semi-natural testing environment is new to the training studies using distributional input. Recall the studies by Maye (2000) and Hayes (2003) where the effect of distributional learning is tested. Both Maye and Hayes use identical test stimuli and training stimuli and though the bimodally trained listeners showed better perception after training, it cannot be claimed that the improvement in the perception of non-native contrasts guarantees their successful perception in a real life environment.

How the semi-natural test stimuli were obtained will be described in detail in Chapter 3, where the experiment and its results will be outlined. Before turning to Chapter 3, a number of conclusions will be summarised on the basis of the different previous studies discussed up to this point. These will be followed by the formulation of the hypotheses and the concrete research questions which guide the present experiment.

2.5. Summary of the previous studies and hypotheses and research questions of the present study

In the present chapter an overview was given of distribution-based studies, training studies and studies on non-native contrasts. Each section of the chapter was meant to serve a different aim. This is how some studies are needed for concretising the hypotheses and research questions stated in Chapter 1, while other literature should help the selection of the experimental stimuli. This is how sections 2.1, 2.2. and 2.3. concern the two main question guiding the present experiment, while section 2.4. is meant to clarify the choice of test stimuli.

Recalling the two main questions in the thesis, the first is concerning the effect of training on vocalic contrasts, while the second refers to whether this effect is identical for two vocalic contrasts. In section 2.1, I opened the chapter with the model of perception of L2 contrasts by Best (1995). There too, on the ground of their acoustic values, two non-native contrasts were analysed as potential cases of SC contrasts within her model. The question is weather the acoustics characteristics of vowels can be good predictors for perception and weather they coincide with the perceptual predictions made in PAM. More concretely speaking, according to their acoustics two non-native contrasts were considered identical with respect to their Bulgarian equivalent, since the Bulgarian phoneme is situated in between the two Dutch phonemes forming the contrast in question. If a similar description is obtained according to the perception of non-native contrasts, the two non-native phonemes would be assimilated to the same imperfect native phoneme, which is a case of SC assimilation. Thus acoustically the Dutch contrasts /i/-/I/ and /a/-/ɑ/ are potentially cases of SC assimilation, and since acoustics and perception are narrowly linked, their outcomes will be expected to coincide. Besides; it is expected that when two non-native contrasts are acoustically comparable, they will be perceived identically before and after training. These expectations are formulated in the first three Hypotheses.

Hypothesis 1: Two acoustically comparable non-native contrasts are perceived identically.

Hypothesis 2: Effect of training is identical for the perception of two acoustically comparable non-native contrasts.

Hypothesis 3: Acoustic characteristics of non-native contrasts coincide with their perception according to Best's PAM.

In the following section 2.2, a number of distribution-based studies were discussed. First we have briefly seen studies which show us that it is inherent to both children and adults and even to other primates to keep track of statistically based information, followed by studies which prove that bimodal training of non-native contrasts leads to more accurate discrimination than monomodal training (Maye, 2000, Hayes, 2003). This evidence for the effect of bimodal training on the learning of non-native contrasts results in the formulation of Hypothesis 4.

Hypothesis 4: Training L2 contrasts with bimodally distributed stimuli helps their identification and discrimination in contrast to monomodal training.

In section 2.3. other training studies were described in which the effect of instruction takes a central place. Though successful, bimodal training appeared not to help L2 learners to transfer the ability to perceive non-native contrasts to novel contrasts of identical nature. Nevertheless, another training study (McClaskey et al., 1983) helped the participants transfer the perception of non-native contrasts after training with explicit instruction. Though testing different type of transferability, Wang (1994) too reported a positive effect of explicit and extensive training on the ability to transfer a learned non-native contrast to productions of new speakers. Besides, García Perez's (2005) study shows evidence that extensive and explicit training is successful for learning to perceive non-native contrasts, showing that the amount of training is crucial for learning to discriminate a non-native contrast. Based on the findings of these studies, Hypothesis 5 was formulated.

Hypothesis 5: Training L2 contrasts with explicit instruction helps L2 learners to transfer their ability to perceive these contrasts in different tokens than the ones they were trained on.

Related to the aspect of instruction in training which prepares L2 learners better for transferring their ability to perceive L2 contrasts, in section 2.4, attention was paid to three types of transfer in perception of contrasts which were analysed as potential degrees of difficulty of test stimuli. In this way, it appeared that testing transfer from isolated vowels to vowels in context and from single speaker's to multiple speakers' productions makes perception more difficult, while transfer from synthetic to natural stimuli does not cause an additional difficulty. Besides, it was argued that using test stimuli which are different from training can on the one hand control for short term memory and on the other hand they can guarantee the semi-natural environment of the tests. For this purpose, the test stimuli in the present experiment were composed containing all these three types of transfer, something we will turn to when presenting the experiment in detail.

Once the hypotheses which lay in the basis of the present thesis have been summarised, these can be split up in a number of concrete research questions which will have a leading function in the next chapter. In Chapter 3 the hypotheses will guide the presentation of the experimental setting on the one hand and the analysis of the results on the other. Before formulating the concrete research questions, first it should be mentioned that these are all related to the perception of two Dutch contrasts, namely /i/-/I/ and /a/-/ɑ/, by Bulgarian naïve listeners of Dutch.

Before starting with the formulation of the different research questions, recall that there are two main questions driving the present experiment. The first main questions concerns the perception of the two Dutch vocalic contrasts used in the present experiment, namely whether they are perceived on an identical manner by Bulgarian speakers. The second main question concerns the effect of the different types of training on the participants' perception of these contrasts in the present experiment.

While the first three hypotheses split up into Research Questions 1, 2, 3 and 4, which are related to the first main question. So, starting with the first set of research questions, they are meant to compare the two L2 contrasts.

> *Research Question 1: Is the number of assimilations of the two L2 contrasts comparable before training?*
>
> *Research Question 2: Is the perception of the two L2 contrasts comparable before training?*
>
> *Research Question 3: Is the effect of distributional learning identical on the perception of the two L2 contrasts?*
>
> *Research Question 4: Is the effect of instruction in training identical on the perception of the two L2 contrasts*

As mentioned, the second main question, aims to reveal an effect of training, which is reflected in the second set of research questions and from Hypotheses 3 and 4, Research Questions 5, 6 and 7 are born.

> *Research Question 5: Does bimodal training cause a significant improvement in the perception of L2 contrasts with respect to monomodal training?*
>
> *Research Question 6: Does training with explicit instruction cause a significant improvement in the perception of L2 contrasts?*

Research Question 7: What is the interaction between distributional training and instruction in the perception of L2 contrasts?

After the formulation of the concrete research questions, we can pass to the next chapter of the thesis where the experiment itself will be described. There it will be attempted to give an answer to the two sets of research questions.

Chapter 3: The experiment

3.1. Participants

In total, 40 Bulgarians took part in the experiment. They all had no knowledge of Dutch, so that they could be considered naïve listeners of Dutch. With respect to their particular origins, half of them were recruited in a small town in the North West of the country, while the other half was randomly recruited in the capital, Sofia. The different origins of the participants did not mean a difference in the dialect that they spoke. The participants' age varied between 16 and 60, where the mean age was 29. For a more detailed description, there were two 16 year-old participants, one 60 and one 50 year-old, while the rest varied between 20 and 40. About 80 percent of the participants had minimal or no knowledge of English, while the rest reported to be able to comprehend spoken English.

3.2. Methodology

As already mentioned in the present study two vocalic contrasts were tested, trained and afterwards tested again. For this purpose two blocks of tests were carried out. The first block concerned a /ɑ/-/a/ contrast, and the second block contained stimuli of a front vowel contrast. This order of the tests was motivated by the fact that the /ɑ/-/a/ contrast was experienced as easier by all the participants. Originally the idea was that half of the participants would perceive first the /ɑ/-/a/ contrast and then the /i/-/I/ contrast while the other half would be trained the other way around. Nevertheless, after the first half of the participants had all commented to experience the /ɑ/-/a/ contrast as much easier than the front vowel contrast, the same order was kept for the second half of the participants. A block of tests for the /ɑ/-/a/ contrast was formed by the following elements: identification pre-test, discrimination pre-test, training session, identification post-test and discrimination post-test; then the same sequence of elements for the /I/-/i/ contrast followed. The two blocks together lasted for half an hour. In the next paragraphs the separate tests will be described in more detail.

3.3. Procedure

In order to confirm the assumption that Bulgarian naïve listeners of Dutch are not able to perceive the two target contrasts in this study, namely /I/-/i/ and /ɑ/-/a/, a pre-test was carried out, which consisted of an identification and a discrimination tests. For each vocalic contrast, a separate block of tests was composed. After the training session all the participants had to perform in a post-test, identical to the pre-test. The post-test was intending to test the effect of training on the discrimination and identification of the two non-native contrasts. In contrast to the training where synthetic stimuli were used, in the tests various natural stimuli were used, which are described further below.

3.3.1. Stimuli

For the experiment recordings of Dutch vowels in context were used, which were already available before the experiment. Escudero (2006) recorded 20 native speakers of Dutch (10 female and 10 male speakers) in five different CVC contexts, forming a total of 300 vowels in contexts. Out of this large data 7 female and 7 male productions of each vowel in a sVs context were used, which were the following pronunciations: /sa:s/, /sɑs/, /sis/ and /sIs/. The stimuli with front vowels were used as they are, while the vowels /ɑ/ and /a/ had to be manipulated before being adequate for the experiment. Since Dutch /i/ and /I/ differ only in their frequency values, it was intended to compare this contrast to another frequency-based contrast and for this reason in the /a/-/ɑ/ contrast it was controlled for a possible effect of duration. In this way, while the Dutch /a:/ was shortened in order to be deprived of the contrasting factor duration, the Dutch /ɑ/ was made somewhat longer in order to reassure that its length will not play any role in its recognition. The manipulation of the natural stimuli was carried out using the program for phonetic analysis *Praat*. Concretely, the /a:/ usually with duration of 0.15 seconds, was shortened by extracting manually not consecutive periods, until its duration was not longer than 0.11 seconds, while for the /ɑ/, usually 0.09 seconds long, a number of periods were duplicated in order to last as long as the manipulated /a/, namely 0.11 seconds.

3.3.2. Discrimination test

Before starting with the actual discrimination test, the participants were allowed to practice with pairs of more clearly contrasting Dutch phones and the participants had to respond with 'same' if the pairs were considered identical and 'different' when non-identical pairs were perceived (see Figure 4 for a representation of the screen for this task). So they heard the /o/-/o/, /u/-/u/ pairs for same and /o/-/u/, /o/-/u/, /u/-/y/ for different. All the vowels appeared in a /sVs/ context, so the participants heard pairs like /sos/- /sus/ etc. The actual discrimination test consisted of ten pairs containing /I/-/I/, ten /i/-/i/ pairs and fifteen more pairs with different vowels /i/-/I/, for the front vowels, which were also in a /sVs/ context. The /ɑ/-/a/ contrast was tested with the same distribution: ten /ɑ/-/ɑ/ pairs; ten /a/-/a/ pairs and fifteen /ɑ/-/a/ pairs. So, per block, 35 randomly distributed pairs were presented to the participants and they had to choose whether they heard twice the same vowels or two different ones.

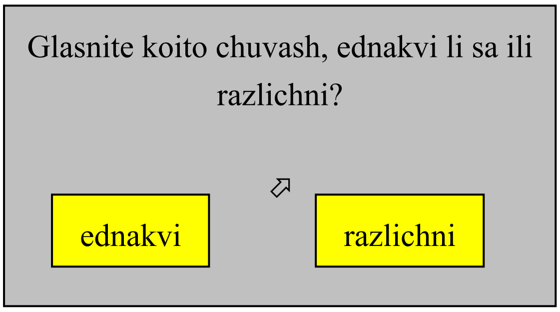

Figure 4: Representation of the screen for the discrimination task[2].

3.3.3. Identification test

Also in this test, there was a practice session. There, the Dutch /sus/, /sys/ and /sos/ had to be identified. For this purpose the participants had to decide whether the vowel in the monosyllabic

[2] The text on the screen appeared in Bulgarian, with meaning "Are the two vowels that you hear same or different?". Identically, in the identification test too, the text was in Bulgarian, meaning "Choose the vowel that you hear".

word was an /o/ or another sound different than /o/ (marked on the screen as ‘other’). For the actual identification task, the participants perceived ten monosyllabic Dutch words with each of the target vowels /ɑ/, /a/ and /i/, /I/. Here too, all the target phones appeared in a /sVs/ context. For instance for the /ɑ/-/a/ contrast, either /sas/ or /sɑs/ was perceived as input by the participants, which they had to characterise either as containing the Bulgarian vowel /a/ or ‘other’ vowel, which would mean a vowel different than /a/. A representation of the screen for the identification task is shown in Figure 5. The command and the ‘other’ button were written in Bulgarian but not in Cyrillic but rather with Latin characters, since it was not possible to use Cyrillic characters in Praat. But this was not a problem since every Bulgarian is able to read Latin characters too (street advertisements, e-mail communication). The participants could make a single choice and after having clicked, they perceived the next stimulus. The procedure was identical for the front vowel contrast.

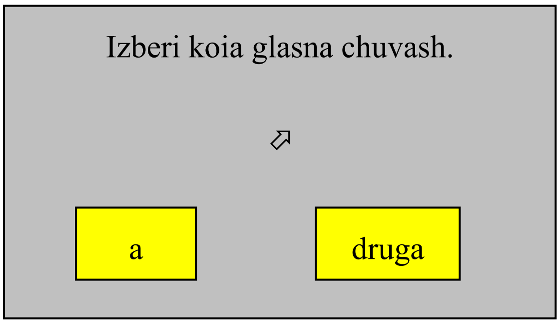

Figure 5: Representation of the screen for the identification task for the /ɑ/-/a/ contrast.

3.3.4. Training

3.3.4.1. Stimuli

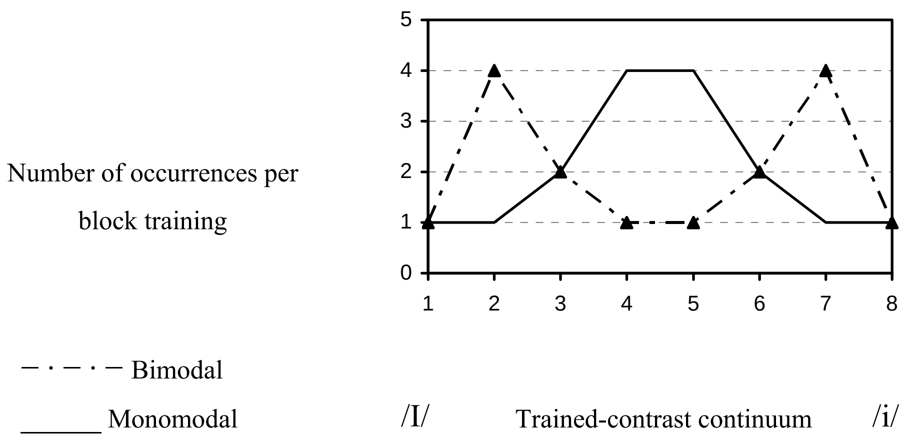

Figure 6: Representation of the distribution of experimental training items

The training stimuli consist of two continua between Dutch /I/ and /i/ on the one hand, and lengthened Dutch /ɑ/ and shortened Dutch /a/, on the other. These two continua were synthesized, having in mind the formant values and the mean duration of naturally produced /I/, /i/ /a:/ and /ɑ/. Taking the front vowels as example, the acoustic dimensions that were manipulated are the first two formants of the vowels. In this way, for a continuum from /I/ to /i/ F1 is decreasing, while F2 is increasing (Table 5, Figure 7). From the continuum, eight separate phones were picked out, from which the first had identical values to a prototypical /I/, while the eighth is identical to a prototypical /i/ (Figure 5). The formant values of /i/ and /I/ used as starting points to form the continuum were taken from mean formant values reported in Adank et al. (2004). These values were slightly adjusted according to the perception of a native speaker who was asked to select the best /i/ and /I/. With respect to the duration of the synthetic stimuli, this was kept constant at a value which was calculated as the mean duration of the two original vowels. For the /ɑ/-/a/ vowel contrast an identical procedure was carried out in order to obtain an /a/-/ɑ/ continuum (Table 6, Figure 7).

With respect to the distribution of the stimuli, the bimodal group heard tokens 2 and 7 four times more often than the tokens in the middle of the continuum, whereas the monomodal group heard tokens 4 and 5 four times as often as the tokens at the endpoints of the continuum. Both distributions formed one training block, while the entire training session consisted of eight training blocks and lasted ten minutes.

Table 5: Mean formants and duration for synthetically formed Dutch /I/ and /i/ sounding like male voice.

	/i/	/I/
F1	250 →	440
F2	2200 ←	1919
duration	0.09	0.09

Table 6: Mean formants and duration for synthetically formed Dutch /ɑ/ and /a/ sounding like male voice.

	/a/	/ɑ/
F1	839 ←	677
F2	1376 ←	1074
duration	0.139	0.139

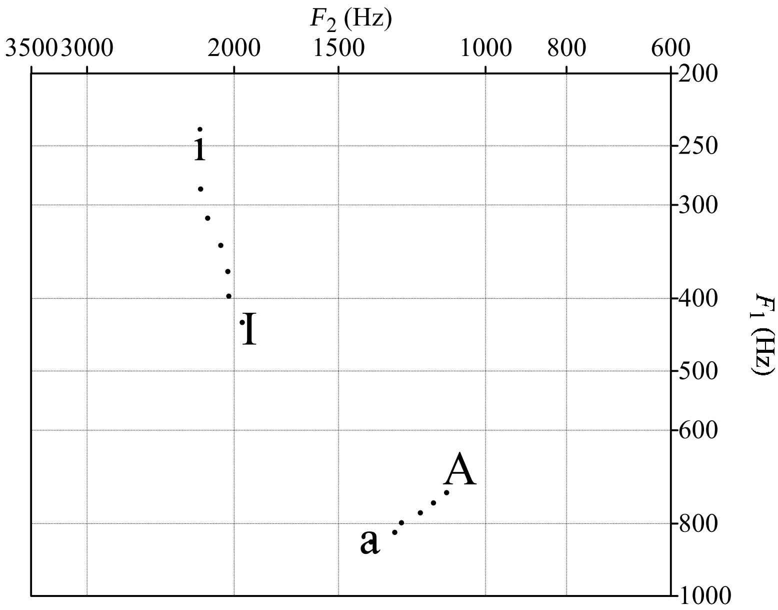

Figure 7: First two formants of the two synthetic continua.

3.3.4.2. Experimental group 1: bimodal distribution, explicit instruction

In this experimental group, the participants were trained in the two Dutch phones which were bimodally distributed and the training was accompanied by instruction. Namely, together with each item it appeared on the screen which of the two phones was presented. The symbols representing the /a/-/ɑ/ contrast were *a* and *A*, respectively and for /i/-/I/ this were just *i* and *I*. In this training condition, items near the two endpoints of the continuum (items 2 and 7) were presented four times as often as tokens in the centre of the continuum. See Figure 6 for a graphical representation of the bimodal distribution. This particular distribution formed one training block. In total, the participants perceived eight training blocks per training which lasted five minutes. This group perceived a complete training session of bimodally distributed vowels (first the /ɑ/-/a/ vowel contrast, followed by front vowels, each in 8 blocks), accompanied by instruction for each token.

Before the training session, the participants were informed that they will be perceiving two different foreign phones one of which sounds more like the Bulgarian equivalent (/a/ and /i/), while the other sounds like between two Bulgarian vowels (the /ɑ/ between /a/ and /o/; the /I/ between /i/ and /e/). Each time they heard an item form 1 to 4, the symbols ‘I’ or ‘A’ appeared on the screen; while each time they heard an item between 5 and 8, ‘i’ or ‘a’ appeared on the screen.

3.3.4.3. Experimental group 2: bimodal distribution, no explicit instruction.

The participants in this group heard experimental training stimuli in a bimodal distribution between /I/-/i/ and/ɑ/-/a/ (Figure 6). This group perceived a complete training session of bimodally distributed vowels, without being instructed about the number of phones that they perceived. In this way, the participants in this group obtained exclusively auditory training, by listening to the bimodally distributed items and not being told what they perceived. They were only informed in advance that they would hear a foreign language.

3.3.4.4. Experimental group 3: monomodal distribution, explicit instruction

In the monomodal training group, the participants perceived items from the centre of the continuum, where items 4 and 5 were presented four times as often as items at the endpoints which are items 1 and 8 (see Figure 3). The tokens in this training session were accompanied by the corresponding symbol appearing on the screen (see 3.3.4.2.). This was a training condition with conflicting cues, since the distribution of the tokens suggested that the different tokens did not form a contrast, while the instruction suggested that there was a contrast.

3.3.4.5. Experimental group 4: monomodal distribution, no explicit instruction
In this training group, monomodally distributed items were perceived by the participants without further information about whether the tokens did or did not form a contrast.

3.4. Results

In this section, the results obtained from the identification and discrimination tests will be presented. The statistical analyses will be outlined, following the research questions formulated in section 2.5. Recall, that the research questions are divided into two main questions, each represented by a number of more specific research questions. The first main question is about how comparable are the two vocalic contrasts tested in the present experiment. The second concerns the effect of training on these two vocalic contrasts. The research questions will be answered one by one.

3.4.1. Perception of the two contrasts

As mentioned above, one of the two main research questions concerns the perception of the two vocalic contrasts used in the present experiment, namely, whether these are perceived in an identical way by Bulgarians. The perception of the /I/-/i/ and /a/-/ɑ/ contrasts is compared in two ways. First by comparing the number of assimilations for the two contrasts and second, by comparing the number of correct responses for the two contrasts obtained in the identification and the discrimination tests. In the statistical tests which will be carried out, we would like to find out the following: whether the number of assimilations is comparable for the two contrast (Research Question 1), whether the number of correct responses before training is identical for the two contrasts (Research Question2), and finally, whether the effect of training is identical for the two contrast (Research Questions 3 and 4). First of all, attention will be paid to Research Question 1.

> *Research Question 1: Is the number of assimilations of the two L2 contrasts comparable before training?*

Recall that the identification task consisted of perceiving a vowel, either Dutch /i/ or /I/, which had to be defined as Bulgarian /i/ or another vowel. Every identification of a Dutch vowel as Bulgarian /i/, which is every click on ‘i’ instead on ‘other’, was considered as an assimilation, independent of being correct or incorrect. In total the identification test consisted of 20 vowels, 10 productions of /i/ and 10 productions of /I/. For the /a/-/ɑ/ contrast, every identification as the vowel /a/ was considered as assimilation.

The participants’ perception of the two contrasts previous to training was compared with a paired t-test. Before training there was no difference between the participants, thus the perception of the two contrasts by the 40 naïve listeners of Dutch was compared. The paired t-test pointed out that the participants assimilated the /i/-/I/ contrast significantly more than the /a/-/ɑ/ contrast ($t(40) = -6.648$; $p<0.05$). This result points out that Bulgarian naïve listeners tend to perceive a categorical difference between Dutch /a/ and /ɑ/ while they hardly hear any difference between /i/ and /I/ (see Figure 8).

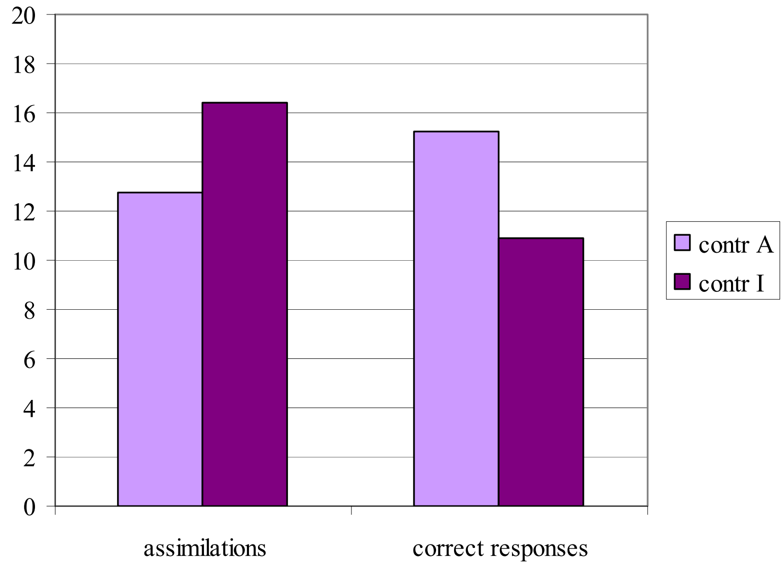

Figure 8: Mean number of assimilation and correct responses in the identification pre-test.

As shown in the bar chart, there are more assimilations made for the /I/-/i/ than for the /a/-/ɑ/ contrast. Namely, in the identification test the participants chose more often /i/ instead of 'other' than they did when performing the identification test for the /a/-/ɑ/ contrast. In other words, for more than half of the vowels that the participants heard, they did not perceive a difference between the Dutch vowels /i/ and /I/. Finally, the answer to Research Question 1 is negative, which is that the number of single category assimilations is not identical for the two contrasts previous to training, since in the pre-test the participants make significantly more assimilations of /i/ than of /a/. In the next paragraphs, I will continue with Research Question 2, repeated below, to see whether this difficulty in perceiving the difference between /i/ and /I/ also means that the identification and the discrimination of the Dutch vowels was less successful for the /I/-/i/ compared to the /a/-/ɑ/ contrast.

Research Question 2: Is the perception of the two L2 contrasts comparable before training?

In order to answer the second research question, which compares the performance in the different tests for the /I/-/i/ and the /a/-/ɑ/ contrasts previous to training, two other tests were carried out. On the one hand, a paired t-test analysed the number of correct responses in the identification pre test, while on the other hand, an ANOVA test was used to compare the correct responses in the discrimination pre test, given its multivariable nature.

Starting with the identification test, a paired t-test is used which compares the two contrasts in the number of correct responses that the 40 participants gave in the identification pre test. Here too, the outcomes for the two contrasts differ significantly, /a/-/ɑ/ appearing to be much easier than /i/-/I/ ($t(40) = 6.707$; $p<0.05$).

The analysis of the discrimination test is different than the previous tests, since here an ANOVA is used. An ANOVA was necessary because of the more complicated nature of the discrimination test itself, in which the test stimuli consisted of 20 same pairs and 15 different pairs, the 20 same pairs being 10 /i/-/i/ pairs and 10 /I/-/I/ pairs and the different pairs being mixed. In total the participants perceived 35 items to which they had to respond either with ‘same’ or ‘different’. In other studies, like in Hayes (2003), where a similar discrimination test was used, it turned out that participants tended to show better performance in the perception of same pairs with respect to different pairs. For this purpose, in the statistical analysis a distinction between the two types of stimuli is made. In the present result analyses, this results in a two-way ANOVA with two between subject factors (distribution: bimodal, monomodal; instruction: with instruction, without instruction), and two within subject variables (contrast: /a/-/ɑ/ contrast, /I/-/i/ contrast; pairs: same pairs, different pairs), where the dependant variable is the number of correct responses.

This analysis of variance revealed a main effect of contrast ($F(1, 36)$ is 58, $p < 0.05$). Looking at Figure 10, this effect can be distinguished. The significant effect of contrast shows that just like in the identification test, in the discrimination test too the /I/-/i/ contrast resulted the more difficult one.

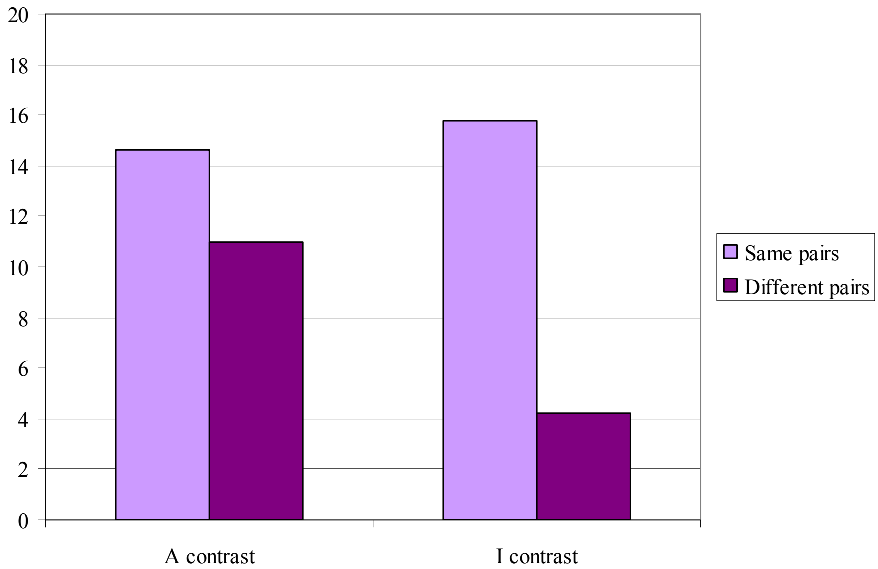

Figure 9: Mean number of correct responses in the discrimination pre test.

Finally, the results of the identification and those of the discrimination pre tests should be combined in order to give an answer to Research Question 2. The two tests revealed significantly better perception of the /a/-/ɑ/ contrast prior to training. This finding gives a negative answer to Research Question 2, namely that the two Dutch contrasts are not perceived identically by Bulgarian naïve listeners.

Up to this point I have treated the first two research questions which concern the perception of the two Dutch contrasts in the pre-test. The next two research questions, intend to find out whether the two contrasts differ after training, in other words, whether the effect of training is identical for the two contrasts. The effect of training is further split up into research questions 3 and 4.

> *Research Question 3: Is the effect of distributional learning identical on the perception of the two L2 contrasts?*

> *Research Question 4: Is the effect of instruction in training identical on the perception of the two L2 contrasts?*

In order to find out whether there is an identical effect of distributional learning and of instruction for the /I/-/i/ and /a/-/ɑ/ contrast, here too the correct responses first in the identification test and then in the discrimination test will be considered.

In order to analyse the responses in the identification test an ANOVA was carried out. Specificaly, this was a two-way ANOVA with two between subject factors (distribution: bimodal, monomodal; instruction: with instruction, without instruction), and two within subject variables (contrast: /a/-/ɑ/ contrast, /I/-/i/ contrast; time: pre-test; post-test), where the dependant variable was the number of correct responses.

The interesting result that the ANOVA shows is a significant interaction of time, distribution and instruction ($F(1, 36)$ is 5,5, $p< 0.05$). This means that at least two groups develop in a significantly different way after training. Looking at Figure 11, the interpretation of this effect can be that bimodal training without explicit instruction causes improvement in the perception of both contrasts. On the basis of the interaction pointed above, however, we can not know for sure that it is the bimodal training without instruction which improves the perception of the Dutch contrasts. This is why a post hoc test was carried out which compared the four experimental groups and pointed out a significant difference between the bimodal training group with and the bimodal training group without instruction ($p<0.05$). The conclusion is that in the identification test, an effect of bimodal training was found, while an effect of instruction was not found, since the group which received no instruction performed significantly better than the group with instruction.

With respect to the question whether the effect of bimodal training and the lack of effect of instruction is identical for the two non-native contrasts tested, it can be considered identical since the ANOVA mentioned before does not point out a significant difference between the two contrasts after training (the interaction between distributional learning, time and contrast not reaching significance ($F(1, 36)$ is 0,63, $p> 0.05$). Nevertheless there is a nearly significant effect of the interactive time and contrast, which suggests a trend that participants improve more in the /a/-/ɑ/ contrast than in their perception of /I/-/i/. However, because bimodal training without instruction causes improvement for both contrasts, the conclusion is that the effect of bimodal training and the lack of effect of instruction are identical for the two Dutch contrasts.

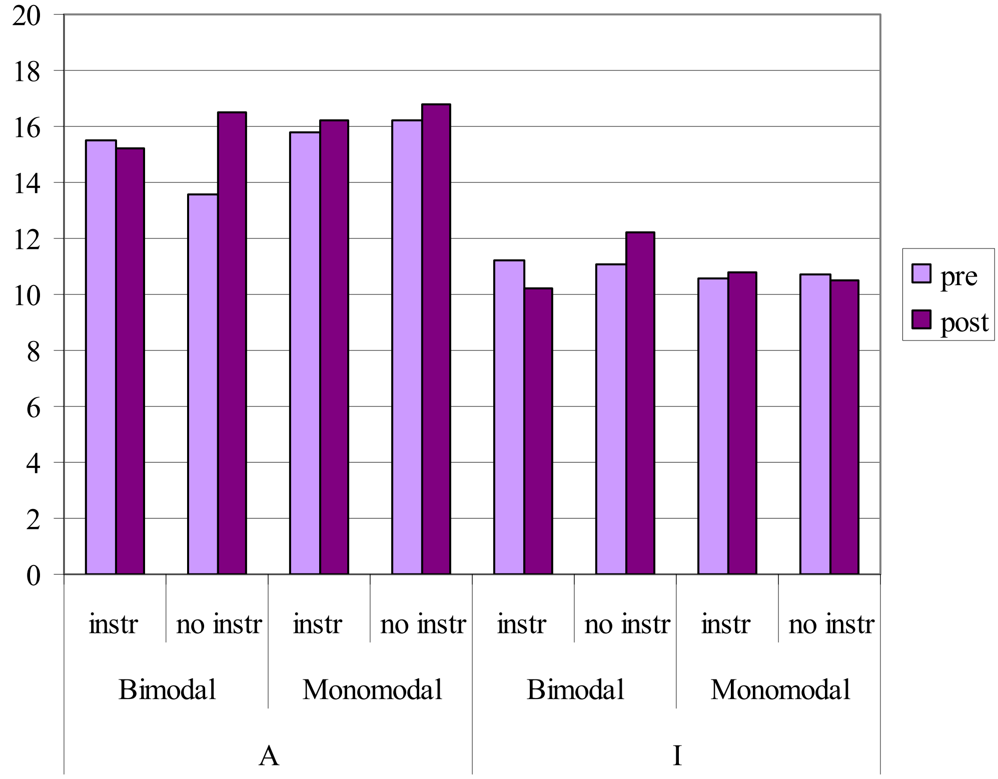

Figure 10: Mean number of correct responses in the identification test.

In order to answer Research Questions 3 and 4, as well another ANOVA which measures the results of the discrimination test is carried out. The three way ANOVA had two between subject factors (distribution: bimodal, monomodal; instruction: with instruction, without instruction), and three within subject variables (contrast: /a/-/ɑ/ contrast, /I/-/i/ contrast; time: pre-test; post-test; pairs: same pairs, different pairs), where the dependant variable is the number of correct responses. Here too, it is the combination of bimodal training and no instruction which results in improvement in the perception of both contrasts. Namely the interaction between time, distributional training and instruction is significant ($F(1, 36)$ is 6,5, $p< 0.05$). We can easily see this effect for both same pairs and different pairs and for both /I/-/i/ and /a/-/ɑ/ contrasts in Figure 12 but to be sure that the effect is caused by the combination of bimodal training and no instruction, here too a post hoc test is carried out. The results of the post hoc test are different for same and for different pairs. Namely, for different pairs bimodal training without instruction is significantly different than bimodal training with instruction ($p<0.05$), which is not attested for same pairs. The conclusion then is that the results of the discrimination test partly attests what was already suggested by the identification test, namely an effect of bimodal training, while effect of instruction is not found.

To sum up, the answer to Research Questions 3 and 4 is that there is an effect of bimodal training found, while effect of instruction is not found and this is valid for both Dutch contrasts tested.

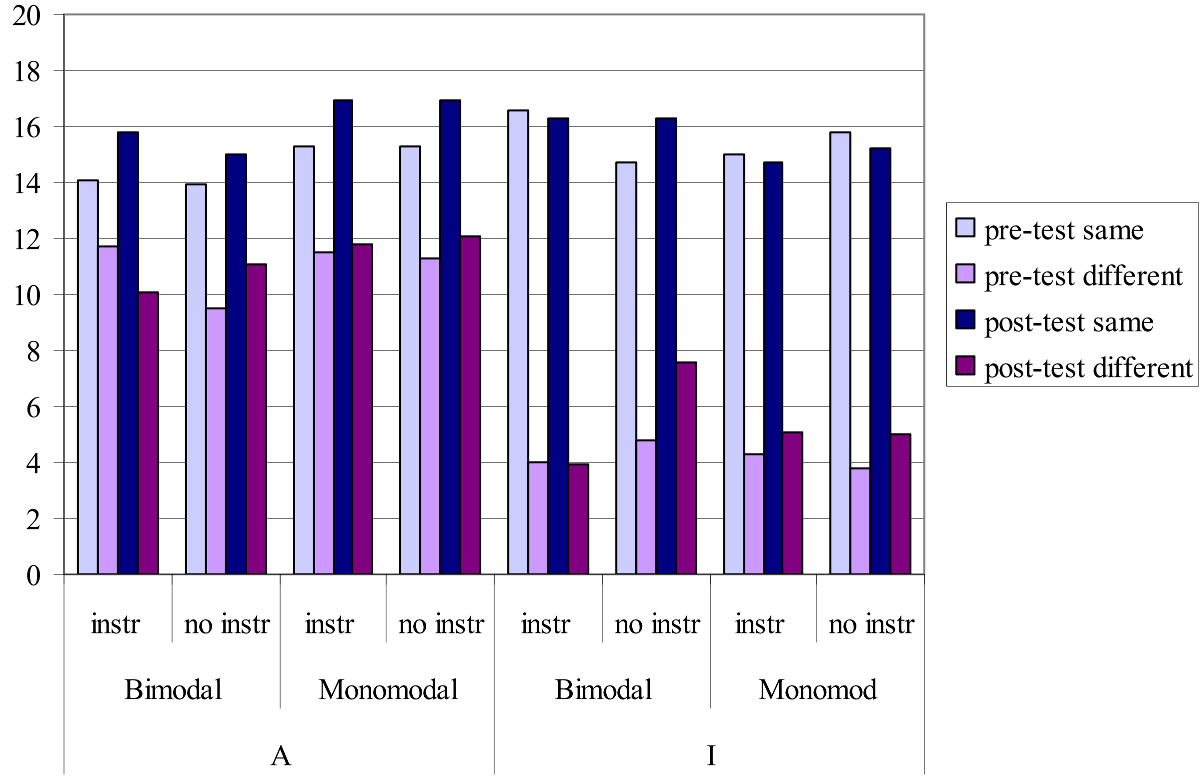

Figure 11: Mean number of correct responses in the discrimination test.

3.4.2. Effect of training

As pointed out earlier, the research questions to be answered with the present experiment were divided into two main topics, namely the comparison of the perception of two Dutch vocalic contrasts to Bulgarian naïve listeners and the effect of training on the perception of these. In the previous section, the first of these topics was treated. When comparing the two contrasts and especially while answering Research Questions 3 and 4 the topic of the effect of training was already partly touched. More specifically, it was concluded that there is an effect of bimodal distribution and lack of effect of instruction on the perception of both /I/-/i/ and /a/-/ɑ/ contrasts. Below, the second set of Research Questions, concerning the effect of training will be treated.

> *Research Question 5: Does bimodal training cause a significant improvement in the perception of L2 contrasts with respect to monomodal training?*

In order to compare the effect of bimodal training, the results from the identification and discrimination tests should be recalled. It are as well the same ANOVA's presented above which will be used to make the necessary comparisons. Starting with the identification test and the two-way ANOVA carried out to analyse the responses in that test, here too, the significant interaction between distribution, instruction and time is of importance. Again the post hoc test comparing the four experimental groups is carried out. As it already became clear, only bimodal training without instruction is effective. This means that if the effect of distributional training is questioned, bimodal training and monomodal training lacking instruction should be compared. The difference between the two groups turns to be insignificant when the responses for the two contrasts are pooled. Comparing the perception of the two contrasts separately, a nearly significant difference between the two groups is found for the perception of the /a/-/ɑ/ contrast (p=0.09). Thus the conclusion can be that there is a trend that bimodal training is more effective than monomodal training for the perception of the easier contrast of the two (see Figure 11).

To find out whether the results from the discrimination test coincide with what was found in the identification test, the tree-way ANOVA presented earlier is recalled. From this test, the significant interaction between time, distribution and instruction ($F(1, 36)$ is 6,5, $p< 0.05$), which was pointed out before is of interest. Again a post hoc test was carried out comparing the monomodal and the bimodal training groups lacking instruction. The results, however, turned out to be insignificant. That is, according to the discrimination test bimodal training did not appear to be more effective than monomodal training. Though the identification test revealed a tendency of bimodal training helping more than monomodal one for the perception of the /a/-/ɑ/ contrast, the more difficult discrimination test did not suggest such tendency. Next, Research Question 6 and 7 will be briefly treated.

> *Research Question 6: Does training with explicit instruction cause a significant improvement in the perception of L2 contrasts?*
>
> *Research Question 7: What is the interaction between distributional training and instruction in the perception of L2 contrasts?*

These two questions have been treated to a certain extent when comparing whether the effects of training are identical to the two Dutch contrasts tested. As already seen in section 3.4.1, there is an effect of bimodal training found by the results of the identification and the discrimination tests. But this effect is found only for bimodal training alone, since when accompanied by instruction, the perception of the two contrasts not only does not improve but even deteriorates perception(see Figures 11 and 12).

Continuing with Research Question 6, two findings, or better lack of findings have to be recalled. Namely, first the fact that bimodal training without instruction was significantly more successful than bimodal training with instruction, together with that monomodal training with instruction does not significantly differ from monomodal training without instruction. In other words, instruction deteriorates perception when combined with bimodal distribution and when combined to monomodal input, does not cause any change in the perception of non-native contrasts. First, it can be concluded that in the present experiment, no evidence was found for an effect of instruction. Second, it is expected that if a larger set of participants is tested, still the effect of bimodal training will be stronger than the effect of instruction.

Finally the question about the interaction between distributional training and instruction formulated in Research Question 7 could be answered simply by using the information already given about the effect of the two sorts of training separately. Namely, since the bimodal training without instruction caused significant better perception in comparison to bimodal training with explicit instruction, it is clear that the combination of the two types of training do not have a positive effect on one another. The lack of significant improvement or deterioration of the perception in the monomodal training with instruction reconfirms this lacking effect of interaction between the two training methods.

Now that all research questions have been answered, I will pass to the discussion, where the hypotheses will be discussed first. In the discussion of each hypothesis, an attempt will be made to clarify the findings of the experiment with respect to the different theoretical views presented in Chapter 2 of this thesis.

Chapter 4. Discussion

4.1. Discussion of the hypotheses

Chapter 2 of the present thesis consisted of a theoretical review on the basis of which the hypotheses were formulated. Each hypothesis was related to a certain part of the theoretical review and in this way, the first two hypotheses were related to Best's Perceptual Assimilation Model (1995). The third hypothesis used as base line the studies on distributional learning by Maye (2000) and Hayes (2003). Finally, the fourth hypothesis used the evidence from training studies in which explicit instruction was used like those by McClaskey et al. (1983), Wang et al. (1999) and García Perez (2005).

4.1.1. Perception of /a/-/a/ is easier than /i/-I/

Now that results from the experiment have been observed, the next step will be to treat the hypotheses one by one. To start, Hypothesis 1 will be the centre of my attention.

Hypothesis 1: Two acoustically comparable non-native contrasts are perceived identically.

According to the analysis in Chapter 2, the two contrasts, which perception was tested in the present experiment, were acoustically comparable. The non-native phonemes forming each contrast were not corresponding to their single equivalent in the L1 of the listeners. The L1 phoneme was rather in between the two non-native phonemes. If such position of the native phoneme with respect to a non-native contrast is experienced in perception, this will be a Single Category contrast according to Best's perceptual model. Later on, we will discuss the topic of weather the acoustics of non-native contrasts and the PAM come up with the same conclusions. Before that, however, we will see weather the comparable acoustics of two contrasts could say something about their perception as predicted.

In order to test Hypothesis 1, the perception of the Dutch contrasts /I/-/i/ and /a/-/ɑ/ was compared. Since it was expected that they would be perceived identically, the number of assimilations and the correct responses in the identification and discrimination tests were not expected to differ significantly for the two contrasts. Surprisingly, the opposite appeared to be the case, since the participants showed overall more difficulty perceiving the /I/-/i/ contrast as compared to the /a/-/ɑ/ contrast. With this finding Hypothesis 1 is rejected. Here the question rises why are the two contrasts perceived differently?

Giving an answer to that question, we might search explanation in the acoustic data of the vowels which were tested. Since we find that the acoustic predictions do not fit with the perception of contrasts, the first to doubt is the quality of the acoustic data that was used in the present study. Both the data provided by Tilkov (1970) and the recordings that I made myself of the Bulgarian vowels could be called unrepresentative because of the limited number of speakers that were recorded. Still, even if not representative enough, the two sources of formant values of the Bulgarian vowels have a very convincing aspect too. This is the fact that both sources of acoustic data suggest an intermediate position of the Bulgarian vowels with respect to the Dutch.

For a future study, it is first of all recommendable that the acoustic characteristics of as well non-native as native phonemes are obtained. In other words, the phonemes in question should be recorded and analysed in an identical way. Secondly, more non-native contrasts should be tested in order to obtain an idea about the predictive role of acoustics for perception. For this purpose, different types of contrasts should be compared. If the /I/-/i/ and /a/-/ɑ/ contrasts in acoustical terms would correspond to SC contrasts in perception, in the same way the other contrasts within PAM could be traduced into contrasts determined by their acoustic values. For instance, if the formant values of a native vowel are closer to one of the phonemes of a non-native contrast, then in acoustic terms this would be a Category Goodness contrast. Using the definitions of different types of contrasts in Best's PAM (1995), on the basis of their acoustics, several contrasts from different languages can be selected and their perception can be tested. Besides it is also interesting to compare different consonantal with different vocalic contrasts.

Hypothesis 2 relates to both the comparability of the two contrasts and to the effect of training. The question here is whether the effect of different types of training is the same for the two contrasts.

> *Hypothesis 2: Effect of training is identical for the perception of two acoustically comparable non-native contrasts.*

Though the perception of the /I/-/i/ and /a/-/ɑ/ contrasts appeared to be different, the effect of training on the two contrasts was identical. This is how Hypothesis 2 was confirmed. Thus the effect of training was identical for the two contrasts, independently of the fact that one of the contrasts was easier, but how do we interpret this finding? One would expect that more difficult contrasts would need more training in order to be perceived better. Still, this is not in discordance with the results of the experiment, since even after training the /I/-/i/ contrast remained the more difficult one (see Figure 11 and 12). An additional conclusion would be that though different in their difficulty, the two contrasts should be comparable to a certain extent. That is, if one of the two was much more difficult, the same amount of training would not have been enough for its perception to improve. Identically to the suggestion given in the discussion of Hypothesis 1, here too, it would be recommendable to compare different types of non-native contrasts and analyse the difference in the ease of perception. To recall the scores in the identification test for the easier /a/-/ɑ/ contrast, 16 correct responses out of a total of 20 items, was a relatively high score. However, if the contrast was indeed so easy, a ceiling effect would have been found. Since this is not the case, it could be concluded that both the /a/-/ɑ/ and the /i/-/I/ contrasts result relatively difficult for Bulgarian speakers.

The third hypothesis to discuss is related to Best's perceptual model and suggests that the acoustic characteristics of vowels are good predictors for perception and should therefore coincide with the perceptual predictions made in PAM.

> *Hypothesis 3: Acoustic characteristics of non-native contrasts coincide with their perception according to Best's PAM.*

Already in the discussion of Hypothesis 1, it became clear that in the present experiment the acoustics of non-native contrasts did not turn out to predict perception. Since the perception of one of the contrasts differed significantly from the other and this was not expected on the basis of their comparable acoustic nature, a problem appeared. Either the reliability of the acoustic data should be mistrust, or it should be doubted weather acoustics can predict perception. The answer to that question could be found in further research. The incompatibility between acoustics and perception has negative consequence for Hypothesis 3 too.

How would the /a/-/ɑ/ and the /i/-/I/ contrasts fit into PAM according to their perception? First of all, PAM will look at how these two contrasts were identified. Since the /i/-/I/ contrast appeared to be assimilated to Bulgarian /i/ in more than 80% of the cases, PAM would predict it to be a difficult to discriminate SC contrast. Nevertheless, looking at the /a/-/ɑ/ contrast which was assimilated to the Bulgarian /a/ in 60% of the cases, one already can predict better discrimination of this non-native contrast, defining it as Category Goodness contrast. According to this definition, one of the non-native phonemes would fit the native equivalent better, meaning that for instance the Dutch /a/ is perceived as similar to Bulgarian /a/, whereas Dutch /ɑ/ is perceived as deviant.

The predictions based on the identification of the two contrasts are found to be true for their discrimination. That is, looking at the responses obtained for the different pairs in the discrimination test, the /i/-/I/ contrast resulted the more difficult one. The same pairs, in contrast, give a rather unrepresentative image. Since the listeners did not appear to distinguish between the two Dutch /i/ phonemes, they tended to perceive pairs as identical. This overuse of same pairs in the discrimination test caused more correct responses for same pairs (see Figure 9). Summing up, according to PAM, the two non-native contrasts tested are perceptually different. While the /i/-/I/ contrast appeared to be a SC contrast, the /a/-/ɑ/ was a CG contrast. However, this is not supported by the acoustic characteristics of the two contrasts, since they do not point out that the Bulgarian phoneme /a/ is closer to Dutch /a/ than to /ɑ/. Instead, the acoustical data provided, pictured the two Bulgarian phonemes in between the non-native phonemes forming the two contrasts. This is how Hypothesis 3 is rejected and it is concluded that according to the results in the present experiment, their interpretation by PAM does not coincide with the acoustic values of the phonemes in the two language systems.

As already mentioned a couple of times, the reliability of the acoustic measurements of the Bulgarian vowels should be doubted. Besides, thought Best's model is entirely based on perception, one could argue that it should take acoustics into account too. By being a uniquely perceptual model, PAM misses a certain empirical strength. That is, that predicting certain perceptual performance, namely discrimination, on the ground of another one, identification, is not very chanllanging. Only by linking the empiricism of perception with this of acoustics, interesting predictions can be born.

Up to this point attention was paid to the perception of the two different non-native contrasts. The discussion of the next hypotheses will concentrate on the effect of training.

4.1.2. Bimodal training needs to last longer

The next to treat is Hypothesis 4, which predicts bimodal training to improve non-native contrast perception.

> *Hypothesis 4: Training L2 contrasts with bimodally distributed stimuli helps their identification and discrimination in contrast to monomodal training.*

As the results have pointed out clearly, an effect of bimodal training was only found with respect to bimodal training with instruction. Nevertheless, bimodal training without instruction did not show a significant improvement with respect to monomodal training without instruction. On the basis of this difference, Hypothesis 4 cannot be confirmed. Here, we need to recall the fact that in the identification test a nearly significant difference between the bimodal and the monomodal training for the perception of the /a/-/ɑ/ contrast was found. Still, this tendency was not confirmed by the findings of the discrimination test. This lack of significant effect of bimodal with respect to monomodal training is surprising since a positive effect of bimodal training was found by both Maye (2000) and Hayes (2003) in their dissertations.

The studies by Maye (2000) and Hayes (2003), however, do differ in one respect from the present experiment, namely in that here natural test stimuli were used. These stimuli were produced by different speakers and in consonantal context and therefore resulted more difficult to the participants than the synthetic training stimuli. More specifically, here a transfer from more simple to more difficult productions was tested. The aspect of transfer of the trained stimuli is also studied by Maye but in a slightly different way. In her dissertation she tests the transfer from training with a certain contrast to testing with another contrast of the same type but with a different place of articulation. When testing transfer, she did not find an effect of bimodal training on the perception of the new contrast. Here, something similar is found, namely, the difficulty to transfer the perception of non-native contrasts to natural productions. Nevertheless, since the near significant results of the identification test suggested that bimodal training could be effective with respect to monomodal training, if longer training would be given, this effect might be reached.

Another aspect of the test stimuli which should be mentioned is their natural character. As already pointed out, they were natural stimuli, pronounced by multiple male and female speakers and appeared in consonantal context. This is in contrast to Maye's and Hayes' dissertations, where both during training and in the test, synthetic stimuli were used. The advantage of using natural stimuli pronounced by multiple speakers is that a rather real-life experimental setting is created. Thus, if after a longer training it turns out that bimodal distribution helps the participants perceive this type of real-life stimuli in which they were not trained, conclusions can be made about the effects of training on the perception outside the experimental setting.

There is, nevertheless, one more requirement which can make the eventual effects of training even more generalizable to real-life situations, namely, by testing the perception of the trained contrast after a certain amount of time. Only if the bimodally trained participants would show an improved perception after some time, could it be definitely concluded that bimodal training is of use to the L2 learner even outside the experimental setting.

To sum up, here Hypothesis 4 was rejected. Nevertheless, it is expected that if distributional training lasts more than ten minutes, bimodal training will prove its effect with respect to monomodal training. Still, if conclusions about perception outside the experimental settings are aimed to be made, two experimental requirements are formulated. First, participants should be tested with real-life stimuli and second their perception needs to be tested after a certain amount of time.

4.1.3. Instruction deteriorates perception

Moving to Hypothesis 5, there it is predicted that explicit instruction would have a positive effect on the perception of non-native contrasts. And though distributional learning has not been combined with explicit instruction before, here it was hypothesised that when combining two effective training methods, the result would be positive.

> *Hypothesis 5: Training L2 contrasts with explicit instruction helps L2 learners to transfer their ability to perceive these contrasts in different tokens than the ones they were trained on.*

As mentioned, the expectation in the present study was that the participants who received bimodal training accompanied by instruction would exhibit significantly better perception of the non-native contrasts than the other experimental groups. Nevertheless, the results pointed exactly the opposite, the bimodal training with instruction being the one causing perception to deteriorate significantly with respect to the bimodal training without instruction. This unexpected result leads to the rejection of Hypothesis 4. Three possible explanations could be given in order to clarify the rejection of this hypothesis.

The first attempt to explain the unexpected results is by interpreting the combination of the two training methods as confusing. Recalling the idea behind bimodal and monomodal training, they are composed by eight different stimuli, each representing a different grade in a continuum between the two prototypical phonemes forming the contrast. In bimodal training the most representative exemplars of the two phonemes (items 2 and 7) are presented most often but still the remaining six stimuli forming the continuum are presented too (see section 3.3.4.1.) Though the stimuli which the participants perceive during training do not have a bipolar character, the instruction confronts them with a bipolar contrast. That is, the instruction tells the participants that the stimuli that they perceive are examples of either one or another foreign phoneme. Thus they perceive four slightly differently produced vowels and are told that this is /i/ in a language they don't know and four other stimuli are told to be /I/, a foreign vowel sounding like between Bulgarian /i/ and /e/. And since we are dealing with a continuum, some of the stimuli which are instructed as /i/, items 3 and 4, are sounding very similar to vowels presented as /I/, items 5 and 6 (see Figure 6, in section 3.3.4.1. for a clearer image of the continuum). Finally, this mismatch between the bipolar instruction and the multiple stimuli training might have resulted confusing to the participants. The conclusion is that if the effect of instruction in training is to be investigated, bipolar training has to be given to the participants, using only items 1 and 8 (the prototypical exemplars) or 2 and 7 (the most representative exemplars) of the continuum in Figure 6.

The second possible explanation of the opposite than expected effect of the combination between bimodal training and instruction could be complementary to the previous explanation. Here, the reason will be searched in the unconscious versus conscious nature of the two training methods. On the one hand, bimodal training with its bimodal distribution of the most representative items (items 2 and 7 being represented most often) suggests that the listeners discover the bipolar pattern of the presented stimuli at an unconscious level. On the other hand, training with explicit instruction implies a conscious understating and perception of the fact that two non-native phonemes form a contrast in a foreign language. So it might have been the case, that by making the participants perceive the training consciously, they became conscious too about the miscellaneous nature of stimuli and as a result the two training methods lost their effect.

The third possible explanation is still in a narrow relation with the conscious nature of instruction training. Given its concsious character, training with instruction requires more time to be successful, just as it is pointed out by the studies of Wang (1994) and García-Pérez (2005). Here the question rises, how could we explain the successful training with instruction given in a shorter time reported by McClaskey et al. (1983)? A possible explanation could be that the contrasts tested in that study and the transfer from one contrast to another are much easier than the contrasts tested in the present experiment and the ones tested in Maye (2000). To remind McClaskey's study, there speakers of English were trained and tested in their perception of a tripartite voicing contrast between /b-p/, /t-d/, /k-g/. If we analyse these contrasts, it might be that adding a third phoneme to a two-way contrast results easier to perceive than acquiring a contrast out of a single phoneme in the L1. In this way the findings by McClaskey could be left out, given the too simplified stimuli. In that case, it is expected that longer instruction training could be effective for the perception of SC contrasts

The conclusion of the discussion of Hypothesis 5 and its rejection is that not only if instruction training will be given with bipolar stimuli, thus separated from distributional learning, but as well if it will take longer, it will be expected to have a positive effect on perception.

After having discussed the four hypotheses, there is one more aspect of the findings which needs to be observed. Already in the discussion of Hypothesis 2, the relatively low responses were mentioned. Though the /a/-/ɑ/ contrast appeared to be easer for the participants, the responses in the perception tests were quite low. This is how, for the identification test, the participants scored between 13.5 and 16.5 correct responses out of 20. This was much lower for the /I/-/i/ contrast, with responses between 10 and 12 out of 20 (see Figure 11). The discrimination test and its responses showed a comparable pattern. A reasonable question would be, why did the participants score so poorly in the perception tests for both contrasts? First of all, the low results point to the difficulty of the contrasts. Even though bimodal training with instruction was effective, the mean scores did not pass the border of 16.5 for the /a/-/ɑ/ contrast. Here it is important to say that a ten-minute training was probably just too short to help the participants really learn the contrasts. Finally, a last reason for the low responses scores obtained, which we should not forget is the relative difficulty of the test stimuli. As presented in section 2.4, the test stimuli were made more difficult than the ones used in training, in a way that they had a more real-life character. Namely, the test stimuli were natural, produced by multiple speakers from both sexes. It is expected that if the participants were tested with the synthetic stimuli which were used in training, they would show better performance overall.

To sum up, in the discussion of the relatively low results obtained in the present study, the following was concluded. First, that the two contrasts tested were difficult to perceive, second, that the training was too short and finally that the test stimuli were difficult as well, just as planned already in the experimental design.

Already in the discussion of the four different hypotheses, conclusions were made in the light of future research. In the following section these conclusions will be summarised and suggestions will be given for improvement of the research of non-native contrasts, distributional training and training with instruction.

4.2. Summary and conclusion

In the present thesis the perception of two Dutch non-native contrasts by Bulgarian speakers was compared. Besides, the effect of training on the perception of these contrasts by Bulgarian naïve listeners was tested. On the one hand the effect of distributional learning on perception was tested on the other hand, the effect of instruction, combined with distributional training on perception was of interest.

The thesis was built up in the following way. First, in Chapter 2, a review was made of the literature which is of importance for the experiment. In this way attention was paid to Best's (1995) Perceptual Assimilation Model followed by the presentation of the two contrasts tested, the Dutch contrasts /I/-/i/ and /a/-/ɑ/. Further on, studies of distributional learning were presented and attention was paid to studies using instruction in training and the transferability of non-native contrasts. In this respects, different aspects of transferability were discussed. In Chapter 3, the experiment was described, after which the results were analysed. The analysis of the results pointed out that the perception of the two contrasts was not identical, the /I/-/i/ contrast being more difficult than the /a/-/ɑ/ contrast. Further on, there was only a trend found of bimodal training being more effective than monomodal training, which would be confirmed with longer training.

Finally, evidence was shown that instruction in bimodal training deteriorates perception. As it was suggested by the discussion, the reason of the negative effect of instruction might lie in the combination between distributional training and instruction. Summarizing the main conclusions once again, these are:

- **The /I/-/i/ contrast is more difficult than the /a/-/ɑ/ contrast**

- **Bimodal training should last longer to be effective**

- **Instruction in combination with bimodal training deteriorates perception**

4.3. Suggestions for further research

Just as in the entire thesis, here too, attention will be paid first to suggestions for future studies on non-native contrasts, followed by suggestions for training studies. The Perceptual Assimilation Model defined by Best (1995) is very useful, since it should be able to predict the difficulty of non-native contrasts to language learners, which is of importance for language education. Nevertheless, in order to reassure its predictive power perception should as well be related to acoustics. Once that acoustics have been introduced to the model, it is also recommendable that not only the perception of different types of contrasts within PAM has to be compared, but as well different contrasts within each type. This is how a perception experiment has to be carried out where at least two cases of Category Goodness and two other cases of Single Category contrasts will need to be compared. If this comparison proves the difference in difficulty between the two types of contrasts within PAM, its predictive power will be strengthened.

For the further research of distributional learning two recommendations can be given. Firstly, distributional training should last longer in order to be effective. In the present experiment the complete training sessions lasted ten minutes. In future research, the training could take three times longer, each training session being interrupted by a break. Breaks are obligatory, given the monotonous character of the training, during which uncountable many phones are perceived. Secondly, it is recommendable that bimodal and monomodal training are not accompanied by other training methods like instruction. The conclusion of the present experiment is that distributional training can better be given alone. This is what is basically done in the study by Maye (2000), the greatest difference with the present experiment being that Maye tests the participants' perception with synthetic instead of natural stimuli.

If it is intended to test the effect of explicit instruction in the training of non-native contrasts, instruction training should as well be given alone. When training with instruction is given, the contrasts which are taught can be presented with prototypical productions of the phonemes. Besides if instruction training is given, it might be of use to control for time. This would mean that a comparison could be made between groups which receive different amounts of training in order to see if time is indeed of big importance for instruction training.

Moreover, in order to improve the target of the present study, which would be to compare the effect of distributional training with this of instruction, the two training methods should be compared separately. In this way, a group receiving bimodal training can be compared to a group which receives training with prototypes and explicit instruction about the non-native contrasts. The two groups will be comparable only if the amount of training is comparable too.

With respect to the research of transfer in perception, on the basis of the present experiment and its findings a useful advice can be given too. It will be of interest to investigate the difficulty of different sorts of transfer. For instance, the intuitive conclusion was made that transfer to a different production of a learned contrast is easier than transfer from one contrast to a different one. Besides, it was concluded on the basis of other studies, that the transfer from a single speaker to multiple speakers as well as transfer form vowels in isolation to vowels in context is more difficult. In order to define the degrees of difficulty of these different types of transfer, they could be compared in a separate study where this topic can be treated in more depth.

4.4. Practical consequences of the findings

After the fully theoretical discussion and the theoretical suggestions for further research, a few practical aspects of the findings in the present experiment can be mentioned. First of all, there is a consequence of the findings which are related to language education. Here it was found that after longer training bimodal training might be effective for the perception of non-native contrasts in a way that it could be used in language education. Instead, here evidence was not found supporting the effect of instruction in training. In section 4.1. a number of possible explanations of this finding were given. Still, before the effectiveness of distributional training and instruction are compared, distributional training should be considered more effective. So, when teaching a non-native contrast in the classroom, it is recommendable to train the L2 learners bimodally even before having presented the contrast to them. Once the learners have obtained enough training, the teacher could help them discover the contrast themselves. This is not a usual scenario of a classroom setting, however. There usually instruction is used which is accompanied by limited phonological support, which does not seem very effective.

Another practical aspect of the present study concerns the potential Bulgarian L2 learners of Dutch. After Bulgaria was accepted in the Schengen space in 2001, the stream of Bulgarians in the European countries and more concretely to the Netherlands increased notably, given that since then Bulgarians could travel visa-free through Europe. Since January 2007, the Balkan state became a member of the European Union, which permits even greater migration of the Bulgarian citizens. As a consequence, the number of Bulgarians learning Dutch might increase even further. This prediction makes it realistic that more teachers of Dutch as a second language will be confronted with the difficulty of Bulgarian learners to perceive and produce Dutch contrasts like /I/-/i/ and /a:/-/ɑ/. And even if the group of Bulgarian immigrants in the Netherlands remains relatively small, there are other language groups which experience these Dutch contrasts as problematic.

References

Abramson, A. S., Lisker, L., (1970). Discriminability along the voicing continuum. Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences.* Prague: Academia.

Adank, P., van Hout, R., Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Dutch. *Journal of the Acoustic Society of America, 116, 3*, 1729-1738.

Baroni, M. (2000). Using distributional information to discover morphemes. Paper presented at the Annual Meeting of the Linguistic Society of America, Chicago.

Best, C. T., McRoberts, G. W., Sithole, M. N., (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology, 14*, 345-360.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*. 171-203. Timoniun, MD: York Press.

Bion, R. A. H., Escudero, P., Rauber, A. S., Baptista, B. O., (in press).Category formation and the role of spectral quality in the perception and production of of English front vowels.

Boersma, P., Escudero, P., Hayes, R., (2003). Learning abstract phonologicqal from auditory phonetic categories: an integrated model of acquisition of language specific sound categories. *Proceedings in the 15th International Congress of Phonetic Sciences*, 1013-1016. Barcelona: Causal Productions.

Escudero, P. (2006). An acoustic description of Dutch vowels. Unpublished paper, University of Amsterdam.

Flege, J. E., (1995). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics, 16, 4*, 425-442.

Frieda, E. M., Walley, A. C., Flege, J. E., Sloane, M. E., (2000). Adults' perception and Production of the English vowel /i/. *Journal of Speech, Language and Hearing Research, 43*, 129-143.

Fry, D. B., Abramson, A. S., Elimas, P. D., Liberman, A. M., (1962). The identification and discrimination of synthetic vowels. *Language and Speech, 5*, 171-189.

García Pérez, G. M. (2005). Perception of English vowels by native speakers of Spanish in a regular classroom setting. *Revista Virtual de Estudos da Linguagem, 3, 5*. (http:// revelhp.cjb.net/)

García Pérez, G. M. (2006). Production the starting link. *Revista Virtual de Estudos da Linguagem, 4, 7*. (http://paginas.terra.com.br/educacao/revel/)

Gulian, M. (2006).Vowel vs. consonant Single-Category assimilation. Term paper for the class Speech Perception and Production, University of Amsterdam.

Hasher, L., Zacks, R., T., Rose, K.C., Sanft, H. (1987). Truly incidental ecnoding of freqency information. *American Journal of Psychology: General, 108*, 356-388.

Hayes, R. (2003). *How are second language phoneme contrasts learned?* Unpublished doctoral dissertation. University of Arizona, Tuscon.

Kewley-Port, D., (1995). Thresholds for formant frequency discrimination of vowels in consonantal contexts. *Journal of the Acoustic Society of America, 97,* 3139-3146.

Kingston, J., Bartels, C., et al. (1996). Effects of training procedures and vowel set on learning non-native vowel categories. *Journal of the Acoustic Society of America, 100, 4*, 2724-2750.

Kluender, K. R., Lotto, A. J., Holt, L. L., Bloedel, Z. L., (1998). Role of experience for language specific functional mappings of vowel sounds , *Journal of the Acoustic Society of America, 104, 6*, 3568-3582.

Maye, J. (2000). *Learning speech sound categories on the basis of distributional information.* Unpublished doctoral dissertation. University of Arizona, Tuscon.

Maye, J. and Gerken, L., (2001). Learning Phonemes: How far can the input take us? In A. H.-J. Do et al. (eds.), *BUCLD 25 Proceedings*, 480-490. Somerville, MA: Cascadilla Press.
Maye, J., Werker, J. F. and Gerken, L., (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101-B111.
McClaskey, C. L., Pisoni, D. B. and Carrell, T.D., (1983). Transfer of training of a new linguistic contrast in voicing. *Perception & Psychophysics, 34, 4*, 323-330.
Mochizuki, M. (1981). The identification of /r/ and /l/ in natural and synthesised speech. *Journal of Phonetics, 9, 3*, 283-303.
Mullennix, J. W., Pisoni, D. B., Martin, C. S., (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustic Society of America, 85,* 365-378.
Pallier, C., Bosch, L., Sebastián-Gallés, N., (1997). A limit on bihavioral plasticity in speech perception. *Cognition, 64*, 9-17.
Pisoni, D. B., (1990). Effects of talker variability on speech perception. Implications for current research and theory. Proceedings of the 1990 *International Conference on Spoken Language Processing*, Kobe, Japan, 1399-1407.
Polka, L., Werker, J. F., (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.
Redington, F. M., Chater, N., Finch, S., (1998). Connectionist and statistical approaches to language acquisition. A distributional perspective. *Language and Cognitive Processes, 13*, 129-191.
Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
Saffran, J. R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by adults and infants. *Cognition, 70*, 27-52
Tilkov, D. (1970). *Le vocalisme bulgare: les mouvements articulatoires et leur effet acoustique dans la formation des voyelle bulgares*. Paris: Klincksieck
Wang, Y. (1994). Training American listeners to perceive Mandarin tones, *Journal of the Acoustic Society of America, 106, 6*, 3649-3658.
Werker, J., Tees, R., (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustic Society of America, 75, 6,* 1866-1878.

Appendix 1: Table with the number of assimilationes obtained in the 20-itemed identification test, where group 1 is bimodal+instruction; group 2 is bimodal-instruction; group 3, monomodal+instruction and group 4, monomodal–instruction.

subjectN	group	aas_pre	aas_post	iis_pre	iis_post
1	1	11	11	19	20
2	1	11	14	13	15
3	1	15	12	18	15
4	1	11	11	20	20
5	1	10	17	14	19
6	1	14	17	10	6
7	1	11	11	18	17
8	1	14	11	18	18
9	1	11	11	20	20
10	1	13	11	12	14
11	2	14	13	12	13
12	2	12	10	14	14
13	2	16	15	19	18
14	2	15	11	20	18
15	2	12	11	17	14
16	2	12	10	16	17
17	2	13	10	16	16
18	2	12	11	18	18
19	2	11	5	7	13
20	2	13	11	16	17
21	3	14	17	18	18
22	3	20	20	20	18
23	3	14	11	20	17
24	3	15	14	16	20
25	3	11	10	17	16
26	3	11	12	20	20
27	3	10	11	14	14
28	3	17	17	17	16
29	3	9	11	15	16
30	3	11	11	15	15
31	4	9	9	15	14
32	4	11	11	17	17
33	4	13	12	11	9
34	4	8	11	19	19
35	4	12	11	19	20
36	4	13	12	16	12
37	4	17	12	20	20
38	4	11	10	15	16
39	4	12	10	15	16
40	4	20	20	20	20

Appendix 2: Table with the correct responses obtained in the 20-itemed identification test.

subjectN	group	correct_A_pre	correct_A_post	correct_I_pre	correct_I_post
1	1	19	19	9	10
2	1	19	16	9	9
3	1	13	16	12	11
4	1	17	19	10	10
5	1	18	11	16	11
6	1	12	13	12	14
7	1	17	19	12	13
8	1	12	11	10	8
9	1	19	19	10	10
10	1	9	9	12	6
11	2	10	17	12	9
12	2	12	16	10	16
13	2	10	13	9	12
14	2	13	17	10	12
15	2	18	19	9	6
16	2	14	18	14	13
17	2	17	18	8	12
18	2	18	19	12	12
19	2	9	11	13	17
20	2	15	17	14	13
21	3	12	11	12	10
22	3	10	10	10	12
23	3	16	19	10	11
24	3	15	16	12	10
25	3	17	20	11	14
26	3	19	18	10	10
27	3	18	19	10	10
28	3	13	13	11	10
29	3	19	17	11	10
30	3	19	19	9	11
31	4	17	17	11	14
32	4	19	19	13	11
33	4	17	18	3	5
34	4	16	15	9	9
35	4	18	19	11	10
36	4	17	16	14	8
37	4	13	18	10	10
38	4	19	18	13	14
39	4	16	18	13	14
40	4	10	10	10	10

Appendix 3: Table of the correct responses obtained in the 35-itemed discrimination test (20 same pairs plus 15 different pairs), separately for the two contrasts.

subjectN	group	corrSame_A_pre	corrSame_A_post	corrDiff_A_pre	corrDiff_A_post
1	1	20	19	13	15
2	1	14	16	13	9
3	1	10	14	8	7
4	1	18	20	15	15
5	1	9	17	15	5
6	1	12	14	10	9
7	1	16	18	14	13
8	1	13	15	10	7
9	1	20	19	12	15
10	1	9	6	7	6
11	2	11	12	4	11
12	2	15	16	10	8
13	2	15	13	8	7
14	2	16	18	10	13
15	2	18	18	10	15
16	2	5	17	12	15
17	2	15	16	10	8
18	2	19	19	13	13
19	2	9	7	11	10
20	2	16	14	7	11
21	3	11	14	9	9
22	3	12	18	6	2
23	3	17	18	11	14
24	3	18	15	10	11
25	3	16	18	15	15
26	3	18	20	15	15
27	3	15	19	15	13
28	3	11	11	8	11
29	3	16	16	13	13
30	3	19	20	13	15
31	4	15	16	14	13
32	4	16	19	13	13
33	4	18	18	13	15
34	4	16	13	10	12
35	4	16	17	15	14
36	4	8	11	13	13
37	4	17	16	6	12
38	4	18	17	14	13
39	4	13	17	9	15
40	4	16	19	6	1

subjectN	group	corrSame_I_pre	corrSame_I_post	corrDiff_I_pre	corrDiff_I_post
1	1	20	18	1	7
2	1	15	15	5	3
3	1	13	14	6	4
4	1	20	20	0	0
5	1	18	18	4	3
6	1	15	13	7	9
7	1	19	20	2	4
8	1	18	14	2	4
9	1	20	20	0	0
10	1	8	11	13	5
11	2	8	10	7	8
12	2	16	16	3	6
13	2	19	16	3	7
14	2	20	17	0	6
15	2	14	15	4	11
16	2	10	10	11	10
17	2	16	16	3	6
18	2	19	19	2	2
19	2	11	11	10	11
20	2	14	17	5	9
21	3	10	15	6	4
22	3	20	20	3	1
23	3	17	19	3	2
24	3	19	14	1	5
25	3	17	15	3	9
26	3	20	16	1	6
27	3	12	12	7	5
28	3	14	14	4	4
29	3	10	15	10	8
30	3	11	12	5	7
31	4	15	12	5	10
32	4	18	15	2	3
33	4	14	15	7	10
34	4	18	16	8	7
35	4	20	20	0	0
36	4	12	7	5	7
37	4	20	20	0	0
38	4	15	16	5	8
39	4	17	14	6	5
40	4	20	20	0	0