Identification vs. discrimination training: Learning effects for trained and untrained sounds

Angelica Carlet and Juli Cebrian
Universitat Autònoma de Barcelona
angelica.carlet@uab.es, juli.cebrian@uab.es

ABSTRACT

This paper presents the results of two high variability phonetic training regimes aimed at improving the perception of five English vowels, namely /æ, ʌ, ɪ, i, ɜː/ by Spanish/Catalan speakers learning English as a second language (L2). 60 learners of English as an L2 took part in a pre-test/training/post-test study and were assigned to one of two training methods (forced-choice identification training, AX discrimination training) and a control group. Vowel training was administered with CVC nonsense words. Pre- and post-tests involved consonant and vowel identification of nonsense words and generalization to real words.

Results indicate that both trained groups significantly outperformed the controls on the trained sounds, showing a positive effect of both training regimes. Identification trainees improved to a greater extent than discrimination trainees on the perception of trained segments. However, a tendency towards improvement with the untrained segments was observed for the learners who received discrimination training only.

Keywords: Phonetic training, L2 speech perception, trained vs. untrained segments.

1. INTRODUCTION

Research in adult second-language acquisition has shown that target language sounds that differ from native categories tend to be challenging, due to the learner´s attunement to L1 phonetic categories. Phonetic training procedures have been shown to be successful in helping learners improve L2 sound perception of unfamiliar speech sounds by focusing their attention on specific phonetic properties and by guiding their decisions through feedback [6]. This study adopted a High Variability Phonetic Training approach [7] (HVPT), which exposes trainees to several talkers in order to provide a natural phonetic variability within a phonetic category and to avoid reliance on the talkers' specific vocal characteristics. This approach increases the chances of generalizing the effects of training to novel talkers [2].

Phonetic training studies have mostly made use of discrimination [8, 9] and/or identification tasks [5, 7] to improve learners' perception and/or production of non-native sounds. Even though both training tasks might be useful in training different perceptual skills [6], the efficacy of identification training has been found to be superior to discrimination training [5]. One possible reason may be that identification tasks help listeners to attend to relevant between-category differences, whereas discrimination procedures focus on within-category variability [5]. Flege [4] directly compared these two types of methodologies (AX discrimination and forced-choice identification) in a study training Mandarin leaners of English to perceive English unreleased final /t/ and /d/. The results showed that learners benefited from both types of training equally, challenging previous results and views on each methodology.

This paper presents the results of part of a study whose goal was to contrast the effect of two phonetic training approaches on the perception of English vowels and consonants by Spanish/Catalan native speakers. Another goal of the study was to assess if training has an effect on trained as well as untrained sound contrasts. This was evaluated by means of a "cross-design", that is, by means of exposing trainees to both consonant and vowel contrasts throughout training, but having learners focus specifically on only one type of sound, in the current case on vowels. Thus, the specific research questions driving the present investigation were (a) whether there is a positive effect of training on the perception of the trained and/or untrained L2 sounds by Spanish/Catalan speakers, (b) which type of training (Identification or Discrimination) is more efficient in promoting improvement on trained and/or untrained segments and, (c) whether the effects of phonetic training carried out with nonsense words generalize to the perception of real words.

2. MATERIAL AND METHOD

2.1. Participants

Sixty Spanish/Catalan bilingual speakers who were second-year undergraduate students in English Studies at the *Universitat Autònoma de Barcelona* took part in this study. All subjects were enrolled in an introductory phonetics course and received course credit for their participation in the study. The 60 subjects were assigned to one of the three groups (N=20) evenly distributed following their initial scores at pre-test phase in order to assure three homogeneous groups. The three groups were an identification group (ID), an AX discrimination

group (DIS), and the control group. 20, 18 and 16 participants completed the training programme, respectively.

2.2 Design and Material

The two experimental groups and the control group were tested before and after a training period. The pre- and post-tests were the same and consisted of vowel identification in nonsense words (trained sounds), consonant identification in nonsense words (untrained sounds) and consonant and vowel identification in real words (generalization). Both experimental groups received training on vowel target sounds, whereas the control group was asked to transcribe texts phonemically on an online platform [3]. The training for the control group was intended to provide the same amount of L2 instruction as the other groups without specific training. Table 1 summarizes the general design of the study.

Table 1: Study design and tasks per group

Group	Pre_test	Training	Post_ test
	V + C	V only	V + C
ID	✓	7 FC_ID	✓
DIS	✓	AX_DIS	✓
Control	✓	Transcription	✓

Note. V=vowels and C= consonants.

2.2.1 Speakers and Elicitation procedure

A total of 6 southern British speakers (3 females and 3 males) produced the training and testing stimuli. Nonsense words were elicited making use of rhyming real words embedded in the following carrier sentence: It rhymes with X, Y. I say Y now, I say Y again. All instances were monitored by the researcher in order to guarantee the desired pronunciation of each target segment.

2.2.2 Testing material

Testing material consisted of non-modified CVC real and nonsense words produced by 2 southern British English native speakers (a male and a female). Stimuli from these speakers were not used in the training corpus, so that the pre- and posttests measured generalization to new speakers. The pre-test and the post-test were identical. Stimuli involved 15 minimal pairs of CVC nonsense words to test the perception of five southern British English vowels /æ, ʌ, ɪ, i, ɜː/ (trained segments) and 12 minimal pairs of CVC nonsense words to test the perception of stop consonants (untrained segments). 17 minimal pairs of real words (5 testing vowel perception, and 12 testing consonant perception) were also part of the testing stimuli and were selected in order to test generalization to new real words. Some examples of real words used in testing are: *back, buck, bag, bug*, which were used to test the vowels /æ, ʌ/ and the final stop consonants in different trials. In addition, 32 words (16 real and 16 nonsense words) involving the vowels /e/ and /ɑː/ were included as testing fillers.

2.2.3 Training material

Training material consisted of 576 stimuli in total (72 nonsense words x 4 speakers x 2 repetitions). Non-modified CVC nonsense words produced by 4 Southern British native speakers (two males and two females) were recorded in order to provide stimuli variability, which is characteristic of HVPT. All stimuli contained one of the 7 selected English vowels /æ, ʌ, ɪ, i, ɜː, e, ɑː/ and one of the 6 English stop consonants /p, t, k, b, d, g/ either initially or finally, in a balanced design. Trainees were exposed to both contrasts involving the trained segments and contrasts involving untrained segments in the course of the training task. For example, the nonsense words *deedge, teedge, didge, tidge* were used to train the vowels /ɪ, i/, while providing a balanced exposure to the initial stop consonant contrast /d, t/. The decision of using nonsense words for training was driven by the need to eliminate any word knowledge bias, as nonsense words require the listener to rely solely on their auditory skills, as well as accomplishing a balanced corpus of stimuli with same number of word pairs for each vowel and each consonant contrast.

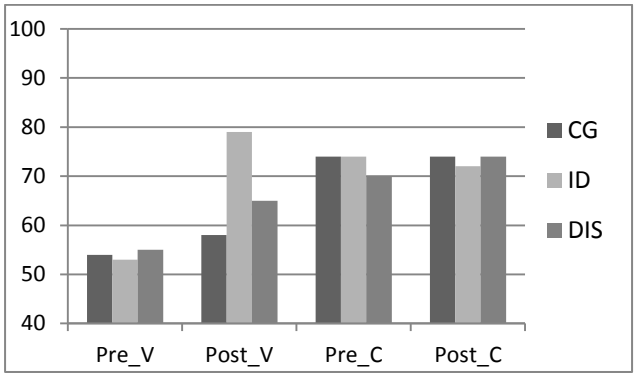
2.2.4 Procedure

Training consisted of five 30-minute sessions over a 7 week-period and it was administered using TP software [10] from weeks 2 to 5. The DIS group was trained by means of AX discrimination tasks consisting of 288 trials (576 stimuli). Participants responded by clicking on "same" or "different". The ID group was trained by means of a 7-alternative forced-choice identification task involving the same 576 stimuli in order to ensure that both groups were exposed to the same set of stimuli throughout training. All subjects were informed that they were being trained on vowel sounds and that their general ability to perceive English sounds would be re-assessed at the end of the training regime. The trainees received immediate feedback after each trial indicating if their vowel perception answer was correct or incorrect and global feedback at the end of each session indicating the total number of hits and errors. Feedback has been shown to be important in training studies as it enables subjects to determine whether what they are doing is appropriate or not, and decide to continue responding in the same way as before or change the way they are responding [6].

3. RESULTS AND DISCUSSION

Correct identification scores were calculated for each group at pre-test and post-test on trained sounds (vowels) and untrained sounds (consonants). The same procedure was followed for the generalization tests (real words). Figure 1 below displays the percentage correct identification of vowels and consonants in nonsense words for each group at pretest and posttest. (Figure 3 below shows the results for real words (generalization)).

Figure 1: Percentage correct identification of nonsense words per group at pre-test and posttest.



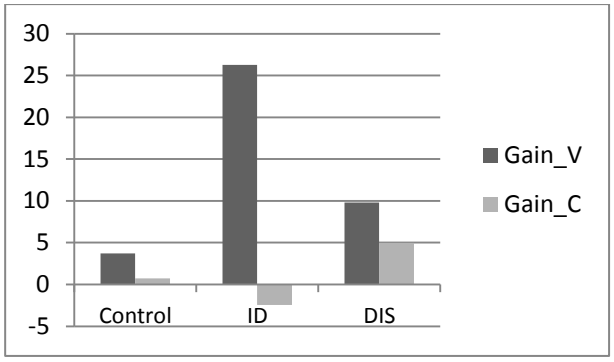
Note. V= vowels and C= consonants.

At pre-test, learners appeared to identify consonants more successfully than vowels, and real words better than non-sense words. Comparing pre-test to post-test, it appears that training had a greater effect on vowel identification than on consonant identification, particularly for the identification training group. As mentioned above, the groups did not differ statistically at pretest, $F(2,51) = .37$, $p = .68$. Therefore, the effect of training was explored by comparing the amount of gain (i.e., the difference between post-test and pre-test scores) for each group. The analysis for the non-sense words is presented first, followed by the analysis of real words (generalization).

3.1 Effects of training on trained and untrained segments: non-sense words

Figure 2 shows the percentage of gain for each of the three groups for consonant sounds and vowel sounds in non-sense words. A mixed design 3x2 ANOVA exploring the effect of group as a between subjects factor (ID, DIS, Control) and sound type (consonants, vowels) as a within subjects factor was conducted on the gain scores for non-sense words. The analysis yielded a significant effect of group, $F(2,51) = 25.02$, $p < .005$, as well as a significant effect of sound type, $F(1,51) = 84.19$, $p < .005$ and a group x soundtype interaction, $F(2,51) = 41.01$, $p < .005$.

Figure 2: Percentage of gain per group for nonsense words



Note. V= vowels and C= consonants.

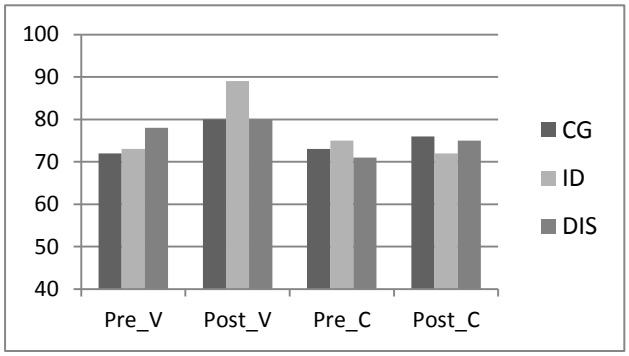
Tests of simple effects showed that the identification group's gain scores were significantly higher than those of the other two groups, and the discrimination group also outperformed the control group ($p < .001$ in all cases). The effect of sound type can be explained by the fact that the trained segments (vowels) obtained higher gains than consonants overall. The identification group's much greater difference in gain between vowels and consonants possibly accounts for the interaction.

A series of follow-up one-way ANOVAs on group (Id, Dis, and Control) with Bonferroni post hoc tests for trained and untrained sounds separately confirmed that the ID group and the DIS group outperformed the controls in their improved identification of vowels ($p < .001$ and $p < .05$, respectively) and the ID group outperformed the DIS group ($p < .05$). These results confirm previous findings that identification training is superior to discrimination training in directing the learner's attention to specific target sounds. Regarding gain scores for consonants (untrained sounds), the analyses showed that while the identification group did not differ from the controls, the discrimination group outperformed the other two groups, especially for final consonants ($p < 0.5$ in both cases). This shows an advantage of discrimination training in the perception of untrained sounds. This may indicate that while identification training may be an overall more effective method, discrimination training promotes generalization to untrained structures to a greater extent.

3.2 Effects of training on the generalization to real words

As can be observed by comparing Figures 1 and 3, all three groups obtained considerably higher scores when perceiving real words than nonsense words already at pretest, especially for vowel segments.

Figure 3: Percentage correct identification of real words (generalization) per group at pretest and postest.



Note. R= real words, V= vowels and C= consonants

A 3x2 mixed design ANOVA exploring the effect of group as a between subjects factor (ID, DIS, Control) and sound type (consonants, vowels) as a within subjects factor was conducted on the gain scores for real words. The results yielded a significant effect of soundtype ($F(1,51) = 19.95$, $p < .001$), confirming that vowels were better identified than consonant sounds. No significant effect of group emerged, but there was a group x sound type interaction ($F(2,51) = 16.98$, $p < .001$), meaning that different groups performed differently with different segments.

The interaction was further explored by means of a series of one-way ANOVAs as was done for the non-sense words. The results for vowel sounds yielded a significant main effect of group, $F(2,51) = 9.16$, $p < .001$. Bonferroni post hoc tests confirmed that only the identification group outperformed the control group, $p < .05$ indicating that generalization to real words for the trained sounds only occurred after receiving identification training. The results for consonant sounds also yielded a main effect of group, $F(2,51) = 7.72$, $p < .01$. The discrimination trainees' gain scores for consonants (real words) were significantly higher than the identification trainees' gain scores, $p < .001$, but it did not differ from the controls', $p > .05$. This shows a tendency for discrimination training to have a greater effect on untrained sounds than identification training, in line with the results obtained with the non-sense words.

4. CONCLUSIONS

The results of this high variability phonetic training study provide further evidence that identification training is more effective than discrimination training at improving perception of L2 sounds, in line with some previous studies. This was particularly true in the case of the trained sounds in this study (vowels), both with nonsense words as well as when testing generalization to real words. However, discrimination training was also effective in improving vowel perception, even if to a lesser extent than Identification training. Moreover, discrimination training was found to positively affect the perception of untrained sounds.

Possibly, by specifically directing learner's attention to trained segments, identification training might have abstracted learner's attention away from any other characteristics present in the stimuli. Considering that most of the training studies make use of CVC words, this may be an important finding, indicating that discrimination training might have a greater influence on the "cross trained" segments, as it yielded better results than identification training with consonant sounds overall.

This research was supported by a research grant from the Spanish Ministry of Economy and Competitiveness (FFI2013-46354-P) and by a grant from the Catalan Government (2014SGR61).

5. REFERENCES

[1] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5, 341–345.

[2] Bradlow, A.R. 2008. Training non-native language sound patterns: Lessons from training Japanese adults on the English /r/-/l/ contrast. In: Hansen Edwards, J.G., Zampini, M.L. (eds), *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins Publishing Company, 287-308.

[3] Cooke, M., García Lecumberri, M.L., Maidment, J. Ericsson, A. 2005. The web transcription tool. http: //wtt.org.uk/

[4] Flege, J.M. 1995. Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics* 16, 425–442.

[5] Jamieson, D., Morosan, D. 1986. Training non-native speech contrasts in adults: Acquisition of the English /h/ and /ð/ contrast by francophones. *Perception and Psychophysics* 40, 205–215.

[6] Logan, J.S., Pruitt, J. S. 1995. Methodological issues in training listeners to perceive non-native phonemes. In: W. Strange. (ed), *Speech perception and linguistic experience: Issues in cross language research*. Timonium: York Press, 351-377.

[7]Logan, J.S., Lively S.E., Pisoni, D.B. 1991.Training Japanese listeners to identify English /r/ and /l/: A first report. *J. Acoust. Soc. Am.* 89, 874–886.

[8] Miranda, S., Underbakke, M., Strange, W., Micceri, T. 1989. Training methods for the facilitation of Japanese students' perception of American English/r/and/l. *The J. Acoust. Soc. Am. 86*, S102–S102.

[9] Strange, W., Dittmann, S. 1984. Effects of discrimination training on the perception of /r/ and /l/ by Japanese adults learning English. *Perception & Psychophysics* 36, 131–145.

[10] Rauber, A.S., Rato, A., Kluge, D., Santos, G. R. 2011.TP software http://www.worken.com.br/sistemas/tp/