

# DIFFERENT HIGH VARIABILITY PROCEDURES FOR TRAINING L2 VOWELS AND CONSONANTS

Angélica Carlet

Universitat Internacional de Catalunya (UIC)  
acarlet@uic.es

## ABSTRACT

This paper investigates the effectiveness of two high variability phonetic training methods aimed at improving the perception of a subset of English vowels (/i ɪ æ ʌ ɜ:/) and initial and final stops by Spanish/Catalan bilinguals. One-hundred L2 learners of English were divided into 4 experimental groups and a control group. Experimental groups differed in training method (forced-choice identification (ID), AX categorical discrimination (DIS) and trained segment (vowels, stops).

Participants were tested on their ability to perceive the target sounds presented in non-words and real words before training, after training and two months later. Results revealed that both training methods proved effective. However, while the ID trainees improved and generalized learning to a greater extent than the DIS trainees on the perception of L2 Vowels; both training methods were equally effective when training L2 stops.

These results suggest that modifying the perception of different types of segment might require different training procedures.

**Keywords:** High variability phonetic training, L2 speech perception, L2 vowels, L2 consonants.

## 1. INTRODUCTION

High variability phonetic training (HVPT) has been found to have a positive effect on the ability to correctly perceive L2 consonant sounds [7, 8, 10, 22] and L2 vowel sounds [9, 6]. Most studies found in the training literature focus on one segment only (vowels or consonants). The few HVPT studies that have trained both consonants and vowels in a controlled manner have reported different degrees of success with each segment [1, 3, 17]. For instance, Aliaga-Garcia and Mora [1] investigated the effect of six two-hour mixed-methods phonetic training sessions on the perception and production of the initial English stops (/p-b/ and /t-d/) and four English vowels (/æ-ʌ/ and /ɪ-i:/) by a group Catalan/Spanish native speakers. Whilst perception of all trained vowels was enhanced, only two consonant sounds improved as a result of the training regime.

In a different study, Cebrian and Carlet [3] assessed the effect of HVPT regime on the perception of four English consonant sounds (/v/-/b/ and /d/-/ð/) and two vowel pairs (/i/-/ɪ/ and /æ/-/ʌ/) by Catalan-Spanish advanced learners of English.

Results pointed to a significant positive effect of a phonetic training method for a subset of the target consonants and vowels, namely /v/, /d/, /i:/, /ʌ/, /b/. The investigators suggested that different factors, such as metalinguistic knowledge, word frequency and vowel duration might have affected the perception of the different sounds. These studies (among others) provide empirical evidence that despite some similarities between vowel and consonant learning [12], these two segments require different degrees of effort from L2 learners [18] and might require different training procedures [16].

Traditionally, most HVPT studies trained L2 learners' perception by means of discrimination (DIS) [15, 22] and/or identification (ID) training procedures [11, 13]. Despite the widely reported superiority of ID training [2, 1, 13], some studies that compared ID and Categorical DIS point to the fact that both training methods are effective when modifying learners' perception of L2 sounds [5, 17, 21]. For instance, Flege [5] directly compared these two procedures (AX categorical discrimination and forced-choice identification) in a study training Mandarin speakers in L2 English final stops. Results revealed that both types of training promoted gain, generalization and retention of learning to the same extent. The current paper reports the results of part of a study whose goal was to contrast the effect of two phonetic training regimes (ID and Categorical DIS) on the perception of L2 English vowels and L2 English stops by Spanish/Catalan native speakers. Moreover, the study aimed at assessing if learning acquired through training generalized and was retained over a period of two months.

## 2. EXPERIMENT

### 2.1 Participants

One hundred Spanish/Catalan learners majoring in English Studies in a public university in Barcelona participated in this study and were divided into four experimental groups and a control group (CG). The experimental groups were a vowel identification group (ID\_V), a vowel AX discrimination group (DIS\_V), a consonant identification group (ID\_C), and a consonant discrimination group (DIS\_C). All subjects reported normal hearing and received course credit at study completion. Importantly, 89 participants completed all training and testing tasks at T2 and only 63 at T3.

## 2.2 Design and Material

All groups were assessed before training, after training and two months later (T1, T2 and T3). The experimental groups received training either on vowels or consonants by means of non-word stimuli. Assessment involved identification of CVC non-words and real words, so that real words assessed a type of generalization. Moreover, all talkers from testing and training differed, so that generalization to novel talkers was assessed along with the main effect of training. The CG performed transcription exercises on *the Web transcription tool*, an open source transcription platform [4].

Recordings of six southern British speakers (3 F; 3 M) provided the training, testing and generalization stimuli. Recordings took place in a soundproof chamber in an English university and each word was recorded three times, with additional repetitions whenever necessary. Stimuli were embedded in the following carrier sentence *It rhymes with “real word”, “non-word”. I say “non-word” now; I say “non-word” again.*

### 2.2.1 Testing Stimuli

Testing stimuli consisted of unmodified CVC non-words and real words produced by two speakers unheard at the training phase. 30 non-words and 10 real words were used to test the perception of five southern British English vowels /æ, ʌ, ɪ, i, ɜ:/, and 24 CVC non-words and 24 real words to test the perception of stop consonants placed either on onset or coda position. In addition, 16 non-words and 8 real words involving the vowels /e/ and /ɑ:/ were included as testing fillers.

### 2.2.2 Training stimuli

Training stimuli consisted of 72 unmodified CVC non-words produced by four Southern British native talkers (2M, 2F). Every non-word contained one of the seven selected English vowels /æ ʌ ɪ i: ɜ: e ɑ:/ and one of the six English stop consonants /p t k b d g/ either initially or finally. The exact same words were used to train vowels and consonants; however, the consonant group was asked to attend to and identify the initial and final consonants that were part of the stimuli, whereas the vowel groups were asked to attend to and identify the vowel sounds present in the stimuli.

### 2.2.3 Procedure

The native Catalan subjects participated in five 30-minute training sessions delivered by the freeware software TP [19]. Participants were told that the purpose of the study was to increase their ability to perceive L2 sounds. The DIS groups were trained by means of AX discrimination tasks consisting of 288 trials (576 stimuli) and responded by clicking on “same” or “different”. The ID groups were trained

by means of a 7-alternative forced-choice identification task for the vowel group, and a 6-alternative forced-choice identification task for the consonant group. Training involved the same 576 stimuli, in order to ensure that all groups were exposed to the same set of stimuli throughout training. Immediate feedback was provided after each trial indicating if their perceptual answer was correct or incorrect. When incorrect, the correct answer was informed. Moreover, global feedback was provided at the end of each session indicating the total number of hits and errors.

## 3. RESULTS AND DISCUSSION

### 3.1 Main training effect and generalization to novel talkers

Since training made use of non-word stimuli, the main effect of training will be evaluated by looking at the amount of gain in perception for non-words. Since talkers from testing and training differed, this data also reports the effects of generalization to novel talkers. The data for vowel trainees will be shown first, followed by the data for the consonant trainees.

#### 3.1.1 Vowel training groups

Table 1 shows the percentage correct identification of non-words at T1, T2 and gain scores for the two groups trained on vowels (ID\_V, DIS\_V) and the CG. Since the groups did not differ statistically at pretest ( $F(2, 51) = .37, p = .68$ ), the effect of training was explored by comparing the amount of gain for each group by means of a generalized linear mixed model (GLMM). Analysis yielded a significant main effect of group ( $F(2, 51) = 53.29, p < .01$ ), and pairwise comparisons with a sequential Bonferroni correction revealed that both groups (DIS\_V and ID\_V) significantly outperformed the controls ( $p < .01$  in both cases) in their improved identification of vowels, and that the ID\_V significantly outperformed the DIS\_V ( $p < .01$ ). These results suggest that whilst both training methods are effective for training vowel perception, ID training may be superior to DIS training in directing the learners’ attention to specific L2 vowel sounds.

**Table 1.** Percentage correct identification of non-words at T1, T2 and gain scores by vowel trainees and CG

	CG	DIS_V	ID_V
	% (SD)	% (SD)	% (SD)
<b>T1</b>	54.1 (9.9)	55.5 (6.5)	52.9 (9.5)
<b>T2</b>	57.8 (10.2)	65.3 (9.7)	79.1 (13.3)
<b>GAIN</b>	3.7	9.8	26.3

### 3.1.2 Consonant training groups

Correct identification scores obtained by the consonant trained groups and CG at T1 and T2 and the respective gain scores are shown in Table 2. A GLMM model with position (initial, final) and group as fixed effects was conducted on the gain scores. Results revealed a significant effect of position ( $F(1, 96) = 11.096, p < .01$ ), a significant effect of group ( $F(2, 96) = 9.662, p < .001$ ), and no group by position interaction ( $F(2, 96) = 2.494, p > .05$ ). Bonferroni pairwise comparisons revealed that both experimental groups significantly outperformed the controls ( $p < .01$ ) and that the trained groups did not differ significantly from each other ( $p > .05$ ). No effect emerged for the final consonants, despite the numerical advantage of the experimental groups (4.3-5.5% gain) over the controls (0.5%). This result indicates that both training methods were efficient in changing L2 learners' perception of initial stops to the same extent. However, no significant improvement was found with the final consonant stimuli in the present study. This may indicate that changing L2 learners' perception of initial and final consonants require different amounts of training time. Recall that the sessions that participants received were divided into two equal parts, so that both initial and final stops could be trained. Taken together, the different results for initial and final consonants suggest that such a short amount of training is sufficient to positively modify L2 VOT perception; however, it is not sufficient to modify the perception of final stops.

**Table 2.** Percentage correct identification of non-words at T1, T2 and gain scores by consonant trainees and CG

	CG	DIS_C	ID_C
Initial Stops	% (SD)	% (SD)	% (SD)
T1	78.1(12.4)	69.5 (9.6)	72.8 (9.5)
T2	79 (13.4)	82.3 (9.8)	88.7 (7.1)
GAIN	0.9	12.8	15.9
Final Stops	% (SD)	% (SD)	% (SD)
T1	69.1 (7.8)	70 (16.8)	69.9 (8.8)
T2	69.6 (6.2)	74.3 (7.3)	75.4 (7.0)
GAIN	0.5	4.3	5.5

## 3.2 Generalization to real words

### 3.2.1 Vowel training groups

Table 3 shows the percentage correct identification of L2 vowel sounds embedded in real words at T1, T2, and the amount of gain for each of the three groups. The results yielded a significant main effect of group, ( $F(2, 51) = 9.16, p < .001$ ). Sequential

Bonferroni pairwise comparisons confirmed that only the ID\_V group outperformed the CG,  $p < .05$ . Moreover, the ID\_V group outperformed the DIS\_V, indicating that generalization to L2 vowels embedded in real words only occurred after receiving identification training ( $p < .01$ ). This result points to two important facts. First of all, it provides further evidence of the robustness of the ID training method, and secondly, it indicates that an ID training method is superior to a categorical DIS method when promoting generalization to real words stimuli, in line with previous research [11, 22]. A possible explanation for this superiority might be connected to the presence of labels in the ID task, which provided learners with focus on phonetic form (i.e. phonetic symbols and/or orthography), which is said to impact speech perception [20].

**Table 3.** Percentage correct identification of real words at T1, T2 and gain scores by vowel trainees and CG.

	CG	DIS_V	ID_V
	% (SD)	% (SD)	% (SD)
T1	72.2 (11)	78.2 (9.7)	73.1 (11.2)
T2	79.5 (10.3)	79.7 (11.1)	88.5 (9.5)
GAIN	7.3	1.5	15.4

### 3.2.2 Consonant training groups

The percentage correct identification of L2 stops in real words at T1, T2, and the amount of gain for each of the three groups are shown in Table 4.

**Table 4.** Percentage correct identification of real words at T1, T2 and gain scores by consonant trainees and CG

	CG	DIS_C	ID_C
Initial Stops	% (SD)	% (SD)	% (SD)
T1	80.9 (11)	77.5 (10.3)	75.2 (20.8)
T2	83.2 (10)	81.9 (11.4)	88.3 (5.9)
GAIN	2.3	4.4	13.1
Final Stops	% (SD)	% (SD)	% (SD)
T1	65.5 (8.2)	69.3 (11.4)	66.7 (8.6)
T2	68.3 (8.6)	72.3 (7.8)	71.7 (8.0)
GAIN	2.9	3.1	5.1

A GLMM model with position (initial, final) and group as fixed effects was conducted on the gain scores and yielded no significant effect of position ( $F(1, 96) = 2.27, p > .05$ ), no significant effect of group ( $F(2, 96) = 2.76, p > .05$ ), and no group by position interaction ( $F(2, 96) = 1.79, p > .05$ ). These results report no evidence of generalization to real words for any of the consonant trained groups. This

might be due to the reduced training time allocated for each consonant segment.

### 3.3 Retention effects at T3

This phase of the study assessed whether the improvements observed as a result of training were maintained after the training regime was over. To that effect, participants performed a delayed post-test two months after the completion of training (T3). Importantly, fewer participants completed this last test and the total number of participants at this phase was less homogeneous among groups. There were 9 controls, 17 ID\_V trainees, 12 DIS\_V trainees, 12 ID\_C participants and 13 DIS\_C participants at T3. The results for the subset of participants that completed all training stages are shown next.

#### 3.3.1 Vowel training groups

Correct identification percentages by the vowel trained groups and the controls at T1, T2 and T3 were calculated and can be seen in Table 5. The three groups performed either numerically higher or similarly to T2 in T3, including the CG.

**Table 5.** Percentage correct identification of non-words at T1, T2 and T3 by vowel trainees and CG.

	CG	DIS_V	ID_V
	%(SD)	%(SD)	%(SD)
<b>T1</b>	56.7 (11.3)	53.0 (4.2)	51.8 (9.7)
<b>T2</b>	61.9 (11.1)	62.8 (9.4)	79.7 (9.3)
<b>T3</b>	63.3 (14)	60.4 (8.2)	80.1 (8.3)

Analyses with time as a fixed effect (T1, T2, T3) for each group showed no significant effect of time for the CG ( $F(2, 48) = 1.84, p > .05$ ), confirming that this group performed similarly across all three testing times. Regarding the trained groups, results in each case yielded a significant effect of time (ID:  $F(2, 48) = 51.35, p < .01$ ; DIS:  $F(2, 33) = 7.62, p < .01$ ). Importantly, Bonferroni adjusted pairwise comparisons confirmed that the performance at T1 significantly differed from the performance at T2 and T3 ( $p < .001$  in both cases). Moreover, the results at T3 did not differ from T2 results, revealing that L2 vowel learning acquired through training was maintained over a period of two months.

#### 3.3.2 Consonant training groups

Correct identification scores for initial stops by the subset of participants at the three testing times are shown in Table 6. All groups performed numerically similarly or better at T3 than at T2, showing that no group experienced a large decline on the identification scores after two months.

**Table 6.** Percentage correct identification of non-words at T1, T2 and T3 by vowel trainees and CG.

	CG	DIS_C	ID_C
Initial Stops	%(SD)	%(SD)	%(SD)
<b>T1</b>	79.2 (15)	71.8 (9.4)	71.0 (10.9)
<b>T2</b>	82.6 (14.7)	84.6 (9.3)	88.4 (6.8)
<b>T3</b>	80.2 (13.9)	86.4 (9.7)	84.3 (14.4)

Analyses showed a significant effect of time for both trained groups (ID: ( $F(2, 33) = 11.46, p < .001$ ) and DIS: ( $F(2, 36) = 9.369, p < .01$ )). Bonferroni adjusted pairwise comparisons confirmed that the performance at T1 significantly differed from the performance at T2 and T3 ( $p < .001$  in both cases). Moreover, the results at T3 did not differ from T2 results. These results revealed that both groups were able to retain initial consonant learning two months after the training ended, which indicates that robust learning has taken place [13].

## 4. CONCLUSIONS

This study assessed the effects of two perceptual methods and it demonstrated positive changes in L2 learners' perceptual abilities as a result of high variability phonetic training (HVPT). More specifically, the present investigation provided evidence that identification (ID) training is more effective than discrimination (DIS) training at improving perception of L2 vowel sounds, in line with some previous studies. This was true both with the perception of non-words as well as when testing generalization to real words. Nevertheless, categorical DIS training was also effective in improving vowel perception, even if to a lesser extent than ID training. Besides this, both methods were similarly found to retain the learning acquired through training. As regards to training consonant sounds, performance with final stops was not successfully enhanced by any of the two short training regimes under investigation. On the other hand, both ID and categorical DIS methods were found to promote gain and retention effects to a similar extent when training initial stops differing in VOT. This suggests that modifying the perception of different types of segment might require different training procedures and amounts of training time.

This research was supported by research grants FFI2017-88016-P and FFI2013-46354-P from the Spanish Ministry of Economy and Competitiveness, and by the research group Grant 2017SGR34 from the Catalan Government.

## 5. REFERENCES

- [1] Aliaga-García, C., Mora, J. C. 2009. Assessing the effects of phonetic training on L2 sound perception and production. *Recent research in second language phonetics/phonology: Perception and production*, 2-31.
- [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] Cebrian, J., Carlet, A. 2014. Second-language learners' identification of target-language phonemes: A short-term phonetic training study. *Canadian Modern Language Review*, 70(4), 474-499.
- [4] Cooke, M., García Lecumberri, M.L., Maidment, J. Ericsson, A. 2005. The web transcription tool. <http://wtt.org.uk/>
- [5] Flege, J.M. 1995. Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics* 16, 425-442.
- [6] Gómez-Lacabex, E., Lecumberri, M. L. G., Cooke, M. 2008. Identification of the contrast full vowel-schwa: training effects and generalization to a new perceptual context. *Ilha do Desterro, A Journal of English Language, Literatures in English and Cultural Studies*, 55, 173-196.
- [7] Hardison, D. M. 2003. Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24(4), 495-522.
- [8] Hazan, V., Sennema, A., Iba, M., Faulkner, A. 2005. Effect of audio-visual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, 47, 360-378.
- [9] Iverson, P., Evans, B. 2007. Auditory training of English vowels for first-language speakers of Spanish and German. In Trouvain, J., Barry, W. (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences (pp.1625-1628)*. Saarbrücken: Universität des Saarlandes.
- [10] Iverson, P., Evans, B. G. 2009. Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, 126(2), 866-877. doi:10.1121/1.3148196
- [11] 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.
- [12] Lerdpaisalwong, S. 2015. *Perception training of Thai learners: American English consonants and vowels*. (Unpublished doctoral dissertation). The University Of Wisconsin-Milwaukee, Wisconsin, USA.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] Nishi, K., Kewley-Port, D. 2007. Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech Language and Hearing Research*, 50(6), 1496-1509.
- [17] Nozawa, T. 2015. Effects of training methods and attention on the identification and discrimination of American English coda nasals by native Japanese listeners. In the Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow. ISBN 978-0-85261-941-4
- [18] Pereira, Y. I. 2014. *Perception and production of English vowels by Chilean learners of English: Effect of auditory and visual modalities on phonetic training* (Unpublished doctoral dissertation). University College London, London, UK.
- [19] Rauber, A.S., Rato, A., Kluge, D., Santos, G. R. 2011. TP software <http://www.worken.com.br/sistemas/tp/>
- [20] Saito, K. 2015. Variables affecting the effects of recasts on L2 pronunciation development. *Language Teaching Research*, 19(3), 276-300.
- [21] Shinohara, Y., Iverson, P. 2018. High variability identification and discrimination training for Japanese speakers learning English /r/-/l/. *Journal of Phonetics* 66, 242-251.
- [22] Strange, W., Dittmann, S. 1984. Effects of discrimination training on the perception of /r/ and /l/ by Japanese adults learning English. *Perception and Psychophysics* 36, 131-145.