Raising Perceptual Phonemic Awareness in the EFL Classroom

Esther Gómez Lacabex  
University of the Basque Country

Francisco Gallardo del Puerto  
University of Cantabria

Abstract

This study examined the impact of phonetic training on English lexical schwa vowel in primary school learners (aged 12) of English as a foreign language in Spain. Students’ perceptual awareness on the occurrence of schwa in unstressed position in English was tested in three groups: i) a group which underwent auditory discrimination and identification practice, ii) a group which underwent listen-and-repeat practice, iii) a control group with native exposure. The first two groups could consistently identify full vowels as incorrect in post-test. Few differences in perceptual awareness as a function of position of the unstressed vowel in the word (pre vs. post- tonic) were found. Word familiarity was found to affect pre-test performance but to be hindered after treatments. Results acknowledge the positive impact of controlled phonetic training on L2 sound perceptual awareness in classroom settings and contribute to the development of L2 pronunciation teaching and learning.

Recent research in L2 speech perception and production training research has been found to favour the development of new sound categories (Bradlow, 2008 for a review on /r/-/l/; Cenoz & Garcia Lecumberri, 1999; Hazan et al., 2006; Iverson et al., 2012; Pruitt et al. 2006; Wang, 2000; Wayland & Li, 2008). Much of this research has been motivated by a wish to investigate the nature of the acquisition process of the sounds of the
second/foreign language system in adulthood, given that late language learners experience difficulties in perceiving and producing L2 sounds (Flege & McKay, 2011). Many of these training studies have looked into variables such as robustness (Pruitt et al., 2006) and durability (Lively et al., 1994; Wang, 2000) of training effects as well as the relevance of feedback provision (Jamieson & Morosan, 1986; Strange & Dittman, 1984). Some have also explored the effects of different training procedures such as perception-based versus production-based training tasks (Leather, 1996), effects of perceptual discrimination versus perceptual identification tasks (Flege, 1995) or effects of training with different acoustic manipulations in natural and varied stimuli (Iverson et al., 2005). These variables have mainly been researched in rigid laboratory settings, where training tends to be fixed, short-termed, and with adult learners. Although these research lines may be of interest in pedagogically-oriented research, pronunciation teaching does not seem to have benefitted from them and is often referred to still hold a marginal position in applied linguistics these days (Barrera Pardo, 2004; Derwing & Munro, 2005; Jones, 1997; Setter & Jenkins, 2005). Those studies which have tested these training procedures in more instructional settings are scarce. Wang and Munro (2004) devised some training software based on perceptual identification tasks with synthetic and natural speech in various voices for some vowel contrasts. This protocol was administered to a group of university students who worked on the training blocks at their own pace over a period of two months and could replay them, if desired. The authors stated that such a procedure can be pedagogically applied as a self-study resource or extra computer-based practice as it addressed individual needs. In Spain, studies like Aliaga-García and Mora (2009) or Cenoz and García Lecumberri (1999) have also administered specific phonetic training based on controlled perceptual discrimination and identification practice as well as production practice within broader training courses which included lectures on phonetics. However, no studies have been carried out with younger foreign language learners without metaphonetic awareness. The availability of technological resources in primary and secondary education these days allows for the possibility of performing these training procedures in more pedagogically-oriented settings and exploring their impact in new and genuine learning environments.

The present study aims at contributing to research done in the field of L2 sound training outside laboratory settings by examining the impact of classroom phonetic training on English lexical schwa vowel by primary school learners of English as a foreign language. Students’ awareness on
the nature and occurrence of schwa in English unstressed syllables was tested in a perceptual identification task which presented both spliced correct schwas and spliced incorrect full vowels in unstressed syllables in English content words (ago, seven). Perceptual awareness was investigated as a function of position of the unstressed syllable within the word (pre-tonic and post-tonic) and as a function of word familiarity (Flege et al., 1996).

**METHOD**

**Participants**

Seventy-five Basque/Spanish 6th graders (aged 12) learning English as a foreign language and distributed in three intact groups of 25 students each took part in the experiment. Two of these groups were assigned two different phonetic training treatments: group A (10 boys and 15 girls) underwent a training regime based on perception tasks and group B (8 boys and 17 girls) underwent listen-and-repeat-practice. The third group (C) (8 boys and 17 girls) received no phonetic training but had a native teacher for English and Arts and Crafts. All the groups had started learning English at the age of 3 in the school, which was enrolled in a language programme that integrated three languages in the school curriculum averagely: Spanish, Basque (both official in the region) and English. The subjects which were taught in English were English, Arts and Crafts, P.E. and Science.

**Stimuli and Testing Procedure**

Fifty real two-syllable words which included a lexical schwa in pre-tonic (alarm) and post-tonic (salad) unstressed position were selected in a 1 to 2 distribution respecting the frequency of occurrence of each pattern in English: 34 items were chosen for post-tonic lexical schwa and 16 items for pre-tonic lexical schwa. The items were recorded in a sound-proof booth by three native speakers of English with knowledge of Spanish (two females and one male). These speakers were first asked to elicit the English words twice, which were presented to them orthographically, and then they were asked to elicit the same words mimicking Spanish pronunciation by trying to colour the schwas towards Spanish vowel qualities. A native Spanish speaker highly proficient in English supervised these recordings and supported the informants, when necessary. Schwa-
full vowel minimal pairs were created by splicing the mimicked Spanish 
full vowel and the schwa of the second elicitation and pasting them to the 
same baseline for each informant’s word using PRAAT. The audio input 
chosen for the testing sessions was one of the female voices. The audio 
files recorded by the other two speakers were used in the training 
sessions. All audio sound files were normalised to the same Root Mean 
Squared (RMS) level.

The perceptual awareness test (47 items¹) presented words with a 
spliced correct schwa (eg: /əgrə:/) and words with a spliced incorrect full 
vowel (e.g: /əgrɪs/) in a random manner (along with a set of 8 distractors) 
which students had to judge as a correctly pronounced word (click on a 
happy smiley) or incorrectly pronounced word (click on a sad smile). The 
elicited word was presented orthographically for each listening trial and 
the vowel in the unstressed syllable was underlined (woman) as 
instructions explained that the potential incorrectness would always be in 
the underlined sound, a vowel, never in the rest of the word. Instructions 
were delivered using both English and Spanish to ensure understanding 
of the task. Prior to the pre-test, word familiarity was controlled by means 
of a written test in which the students ticked whether they knew the word 
or not. This information was used to distribute the items averagely (same 
amount of known and unknown words by means of calculating median 
value) in each training session. The items were also averagely and 
proportionally distributed into training sessions and tasks according to 
speaker voice (male/female) and stress pattern (pre/post-tonic).

Training

Four training periods were programmed to be delivered in 6 sessions so as 
to cater for absentees or slower learners. An introductory session 
reviewed the notions of syllable, stressed syllable, unstressed syllable and 
unstressed vowel and provided instruction for the practice sessions. In 
this session, the group undergoing listen-and-repeat practice was 
additionally given two pronunciation tips: i) prepare your lips for the first 
consonant in words which start with a vowel, as in amend and ii) imagine 
the underlying score moves up and crosses the vowel in between 
consonants, as in raisin, so the consonants are pronounced subsequently.

¹ Three items had to be eliminated because cross-splicing produced too much artifact in 
the voice of the speaker used in the testing. This did not happen in the voices of the two 
speakers used in training; it was decided not to eliminate them from training.
Sessions were always supervised by an instructor and a technician, both of whom ensured that tasks were completed sensitively and that audio and microphone equipments were working.

The training regime was based on perceptual practice and included discrimination and identification tasks of the type: i) same/different discrimination task (12 items), ii) oddity discrimination task (13 items), iii) correct/incorrect identification task (25 items), iv) correct/incorrect identification task with confidence rating (12 items). Feedback was delivered by flashing a happy smile and providing a cumulative star every time a trial was correct in the first attempt. If the trial was incorrect, the sequence would play again and the student had to provide a new answer. No stars were given in subsequent attempts.

The training regime for group B was based on imitation practice. Students heard the word, which they also had on the screen orthographically signaled so as to reinforce vowel focus, and repeated it in a MATLAB interface, which recorded their productions. The programme forced them to listen to their own recording once and they could repeat the elicitation two more times if they were not satisfied with their productions. The supervising instructor monitored the productions throughout the sessions and provided positive feedback to those students who attempted at producing a reduced vowel successfully and reminded those students who did not about the pronunciation tips presented in the introductory session.

RESULTS

Figure 1 presents an overall analysis of results which included all variables. It displays percentage of correct perceptual identifications of schwa words (identified as correct) and full vowel words (identified as incorrect) as indicated by the t-test analyses carried out. There were not any significant differences in the identification of a spliced schwa vowel as correct between testing stages in any of the tested groups [A (pre-test: 81.8%; post-test: 80.6%), B (pre-test: 74.6%; post-test: 79.3%) and C (pre-test: 74.4; post-test: 74.6%)] or when analysed together [ABC (pre-test: 76.9%; post-test: 78.1%)].
Figure 1. Percent correct identification of all items for schwa words (top) and full vowel words (bottom) for groups A, B and C.

For the full vowel words corpus, intra-group analyses revealed that all three groups underwent significant improvement (A: $t = -8.45; p<.001$; B: $t = -5.17; p<.001$; C: $t = -3.04; p<.01$) in their identification of a spliced full vowel as incorrect in post-test. Pre-test means (A: 27.9%; B: 36.5%; C: 33.1%; ABC: 32%) indicated that performance was at low chance levels, possibly revealing a bias towards ‘correct’ identification. Post-test means (A: 73.5%; B: 60.4%; C: 43.3%; ABC: 58%) indicated that groups A and B could more consistently identify a full vowel as incorrect after training whereas group C remained in chance performance levels.

Figure 2 shows results in terms of position of the vowels within the word: pre-tonic and post-tonic. In the case of spliced schwa words, results did not reveal intra-group differences for any of the groups in either pre-tonic or post-tonic position.
Figure 2. Percent correct identification of schwa words (top) and full vowel words (bottom) for groups A, B and C in post-tonic and pre-tonic positions.

As for spliced full vowel words, significantly higher post-test scores were found in the three groups [A (t = -7.43; p<.001), B (t = -5.05; p<.001) and C (t = -3.25; p<.01)] in post-tonic position (pre-test means: A: 27.9%; B: 36.5%; C: 33.1%; post-test means: A: 73.5%; B: 60.4%; C: 43.3%) and for A and B only [(A (t = -9.41; p<.001), B (t = -3.16; p<.01)] in pre-tonic position (pre-test means: A: 21.5%; B: 35.5%; C: 31.3%; post-test means: A: 71.5%; B: 53.3%; C: 40.7%).

Inter-context (pre-tonic vs. post-tonic) comparisons revealed few significant differences. Groups A and B exhibited significantly lower scores in correct schwa vowel identification in pre-tonic position than in post-tonic position in post-test (A: t = 2.7; p<.05; B: t = 2.32; p<.05).

As for word familiarity, as can be seen in Figure 3, no consistent intra-group differences were observed in the schwa vowel words corpus in the
groups for known words (pre-test means: A: 89.2%; B: 81.5%; C: 84.3%; post-test means: A: 80%; B: 78.6%; C: 79%) or unknown words (pre-test: A: 72.2%; B: 65.6%; C: 67.6%; post-test: A: 77.1%; B: 73.8%; C: 69.1%), with the exception of group A’s significant decrease in post-test for the known corpus (t = 2.25; p<.05).

![schwa vowel words](image1)

![full vowel words](image2)

**Figure 3.** Percent correct identification of *schwa* words (top) and full vowel words (bottom) for groups A, B and C in known and unknown words.

As for identification of an incorrect full vowel in known (pre-test means: A: 18.9%; B: 34.2%; C: 26.4%; post-test means: A: 72.4%; B: 57.2%; C: 40.1%) and unknown (pre-test: A: 32.6%; B: 38%; C: 39.6%; post-test: A: 73.4%; B: 59.1%; C: 44.7%) corpora, intra-group differences were significant both in the known (A: t = -8.4; p<.001; B: t = -4.06; p<.001; C: t = -3.2; p<.01) and the unknown (A: t = -6.79; p<.001; B: t = -4.13; p<.001) corpora for all groups except group C’s scores in the unknown corpus.

Inter-variable analyses (known vs. unknown) showed that the subjects tended to judge known words as ‘correct’ more often than unknown words in pre-test both for *schwa* words (A: t = 4.36; p<.001; B: t = 3.8;
p<.001; C: t = 3.72; p<.001) and for full vowel words (A: t = -3.37; p<.001; C: 
t = -3.19; p<.01), which is revealed by the very low scores reported. In post-
test, this tendency was only significant in the schwa vowel corpus for 
group C only (C: t = 2.28; p<.05).

**DISCUSSION AND CONCLUSION**

Our study aimed at exploring the effect of both classroom phonetic 
training (perception vs. production based) and exposure to native accent 
on the perception of English lexical schwa by Basque-Spanish bilingual 
schoolchildren. Post-test results attested treatment effects on learners’ 
perceptual awareness of occurrence of schwa in unstressed syllables in 
English. The learners significantly improved their ability to identify 
incorrect spliced full vowels as such in post-test. In addition, the groups 
having undergone explicit phonetic intervention exhibited it in a more 
significant manner than the group with native exposure, which remained 
in chance performance levels in the contexts in which the other groups 
exhibited improvement, namely the full vowel corpora. Results seem to 
indicate that explicit computer-aided phonetic training programs can 
contribute to boosting phonemic awareness in classroom learning 
environments with younger learners. In addition, the present study 
provides evidence for the fact that two distinct phonetic training regimes 
(one based on perceptual skills and one based or oral/production skills) 
could raise learners’ perceptual awareness on L2 sound occurrence. These 
findings can actively contribute to research on the complex and versatile 
relationship between speech perception and production within the 
training domain. An exploration of inter-group behavior remains to be 
carried out in near future work on the part of the authors.

Results indicated that the schwa corpus did not undergo significant 
 improvement after training. As can be interpreted from Figure 1, students 
tended to identify both correct schwas and incorrect full vowels as correct 
(hence, high scores for schwa vowels and low scores for full vowels) in 
pre-test. Given that previous work (Gómez Lacabex, 2009) has attested 
that learners of a similar age can perceptually discriminate the contrast 
schwa-full vowel when presented in natural stimuli and not manipulated 
for durational or acoustic cues, we interpreted results as learners 
exhibiting a lack of awareness on the occurrence of these vowels in 
English unstressed syllables. However, no gain scores were observed for 
this corpus. It may be that participants are performing at a ceiling level or 
that the schwa/full vowel contrast may have acoustically approximated
(either in F1 and F2 values or in durational values) so as to cause trouble to the learners in some of the stimuli (Flege & Bohn, 1989; Gallardo del Puerto, 2005). Indeed, the nature of training types in this study did not control for vowel quality or durational features, which may have helped learners overcome this possible perceptual threshold level.

Few differences were observed as a function of position of the vowels within the word: pre-tonic (amend) and post-tonic (seven). Learners’ identification scores were rather consistent in pre- and post-tonic positions in the two testing stages and the two corpora. However, it was found that groups A and B’s identification of correct schwa vowels in pre-tonic position (allow) was significantly lower than in post-tonic position after treatment. While, on the one hand, this could be owing to the different number of items included in each corpus, on the other hand, it could be suggesting that the learners may be noticing some finer acoustic cues in pre-tonic position after training which they cannot successfully identify as correct. This study avoided including final schwa, traditionally described as articulated with an open-mid tongue position and closer to English vowels /ʌ/ and /ɜ/ (Cruttenden, 1994) and used items traditionally described as more centralized schwas (between open-mid and close-mid positions) such as in open, pre-tonic syllable, as in afraid, or closed syllable, as in pilot (also Cruttenden, 1994). However, schwas have also been reported to adhere acoustic qualities of neighboring sounds (Browman & Goldstein, 1992; Flemming, 2009). The results, thus, seem to call for a further acoustic exploration of the schwa vowels used as prompts in this study and for a re-analysis of results in terms of phonetic context rather than stress pattern only. This may elucidate whether this inter-contextual difference in post-test in the two groups which received phonetic training may be owing to the students being able to notice variations in schwa acoustic values in pre-tonic position but not being able to successfully identify them as correct schwas in post-test.

Finally, results did reveal some word familiarity effect as known words tended to be judged as ‘correct’ by the students more often than unknown words in pre-test. Interestingly, this was done so regardless of the words having a correct schwa or an incorrect full vowel. In post-test, this effect was only present in group C, which did not work with the items during the treatment period; the groups which underwent phonetic treatment (A and B), which worked with the items at a phonetic/phonological level but not at a lexical level, did not report this tendency. Results seem to suggest
that phonetic training can inhibit a bias towards identification of familiar words as correct.

To conclude, our results seem to attest that classroom learning contexts can benefit from L2 sound training protocols such as discrimination, identification or imitation, as these were able to boost young learners’ perceptual awareness on the occurrence of schwa in English unstressed syllables in our study. While data suggested that specific phonetic treatment may be able to fade a tendency to label known words as led us to suggest that *schwa* may be better analysed according to multi-contextual phonetic effects (eg: open/close syllable, adjacent consonant, preceding/succeeding vowel) rather than in terms of stress pattern only.

ACKNOWLEDGEMENTS

The research was supported by the following research grants: UFI11/06 from the University of the Basque Country (UPV/EHU); IT-311-10 from the Basque Government and FFI2009-10264 and FFI2012-31597-DIACEX by the Spanish Ministry of Economy and Competitiveness (MINECO).

REFERENCES


