The Production of English Aspirated Stops in Foreign Language Acquisition

Eleni Tsiartsioni
Aristotle University of Thessaloniki
eeltsiartsioni@hotmail.com

Abstract

This study examines the effect of teaching on the production of L2 English aspirated stops among Greek EFL (English as a Foreign Language) learners. Two groups of speakers were recorded, an experimental group who received the pronunciation teaching intervention and a control group who did not receive special pronunciation instruction. The data were analysed acoustically and a native-likeness criterion (Birdsong 2007, Flege et al. 1992) was used to investigate individual variability. Overall, this study reports a positive effect of pronunciation instruction on the production of aspirated stops, however the great individual variability suggests that systematic target-like production of stops constitutes a challenging goal. A tentative effect of learners’ gender is also obtained, as the female speakers produce greater VOT than the male ones after teaching.

1. Introduction

This paper is part of a larger research project on the acquisition of features of the English speech rhythm and stop voicing system among Greek learners of English. In particular, the present study examines the production of English word-initial aspirated stops by Greek students of ten to sixteen years old who learn English in a foreign language (FL) / formal setting context.

One of the differences between the Greek and English phonetic systems lies in the production of /p, t, k/ sounds. In English these sounds are produced as voiceless aspirated with long voice onset time (VOT, Lisker & Abramson 1964), whereas in Greek they are produced as voiceless unaspirated with relatively short VOT (Arvaniti 1987, 2001, Botinis et al 2000, Fourakis 1986, Nicolaidis 2002, Panagopoulos 1972). Given these differences, an interesting question is how these segments are realised by Greek learners of English and whether pronunciation instruction can have an effect on speakers’ production.
Previous research on the production of L2 aspirated stops by speakers whose L1 has unaspirated stops (for example, speakers of L1 Dutch, Italian, French and Portuguese) showed that the context of acquisition appears to be related to the degree of accuracy in L2 production. Studies which were conducted in a second language / naturalistic context of language acquisition (namely, in an environment where the learners live in a target-language community) showed that the L2 speakers generally produce longer VOT in L2 than in their L1, however, the L2 VOT is shorter than that of the target language (Caramazza et al. 1973, Flege & Eefting 1987, Major 1987, Williams 1979).

Studies which were conducted in foreign contexts of language acquisition (i.e. in environments where the learners are not exposed to the target language in their everyday communication) reported that the L2 learners had great difficulty producing target-like aspirated stops, as they tended to transfer L1 voiceless unaspirated stops into L2 (Riney & Takagi 1999, Poerdjianto 2003). However, exceptional cases of speakers who produced target-like VOT have been reported in foreign language acquisition studies. Riney and Takagi (1999), in particular, investigated the correlation between Global Foreign Accent (GFA, determined on the basis of native English listeners’ scores) and VOT for /p, t, k/ among L1 Japanese learners of English as a foreign language. The results showed that, generally, the L2 English VOT was very similar to that of L1 Japanese. A significant correlation was also found between GFA and VOT scores; in particular, more target-like GFA scores were associated with more target-like VOVT values. The individual speakers’ analysis showed that two L2 speakers produced VOT that was similar to L1 English. Riney and Takagi (1999) concluded that VOT is one aspect of pronunciation that some L2 speakers over the age of 12 can produce accurately. Approximation to target-like VOT in an FL context was reported by Lowie (2009) and Lowie and Reitsma (2009) regarding a group of Dutch children starting English at the age of 4 and being taught by a native English teacher. That study concluded that in FL contexts approximation to L2 VOT values can be achieved provided that early starting age and native speaker input are ensured.

Another key issue in the present study is the effect of pronunciation instruction. The previous literature on pronunciation teaching shows that this is an area that has not been adequately explored in applied linguistics (Goodwin et al. 1994, Pardo 2004). It is, therefore, not surprising that teachers have reported being unsure about the importance and effectiveness of teaching pronunciation to their students (Pardo 2004). A summative research of twenty-five studies on the effect of pronunciation instruction, however, showed that there is a positive effect of well-organised pronunciation teaching, as long as carefully planned techniques are used (Pardo 2004). The positive effect of pronunciation teaching has been documented both in naturalistic settings (as in the vast majority of studies reviewed by Pardo 2004) and in foreign language contexts of language
acquisition, as well (Ekstrand 1982, Olson & Samuels 1982, Thogmartin 1982). With regard to the learning of L2 aspirated stops, Reis and Nobre-Oliveira (2008) and Aliaga-Garcia and Mora (2008) reported an increase in L2 VOT after teaching among Brazilian-Portuguese, Spanish and Catalan learners of English respectively, however, the L2 VOT did not reach the target values, even after teaching.

Another parameter that has been explored in L2 phonological acquisition research is that of speakers’ gender. With regard to this issue, no conclusive results have been obtained. Whereas certain studies found that female speakers tended to speak with a more target-like accent (Tahta et al. 1981, Thompson 1991), other studies reported no effect of speakers’ gender on L2 pronunciation (Flege & Fletcher 1992, Purcell & Suter 1980, Snow & Hoefnager-Höhle 1977, Suter 1976).

In view of the aforementioned literature, the aims of the present study are:

1. to investigate the production of L2 aspirated stops in a foreign / formal context of language acquisition and, in particular, to examine the effect of a specific pronunciation teaching intervention among Greek learners of English,
2. to investigate the potential effect of learners’ gender in the acquisition of English VOT.

2. Methodology

2.1. Subjects

In the present study two groups of subjects were recorded, an experimental group (n=36), which received pronunciation instruction, and a control group (n=36) which followed the regular English classes at school. An equal number of male and female speakers participated in each group, in order to investigate the potential effect of learners’ gender. Each group comprised students of three different ages: twelve ten-year old students, twelve thirteen-year old students and twelve sixteen-year old students. Given the aims of the present paper (section 1), the potential effect of speakers’ age is not addressed in the present paper. However, it is acknowledged that the effect of learners’ age on L2 phonological acquisition is an interesting issue, therefore this will be addressed in future publications. Data for L1 Greek were also obtained from thirty L1 Greek children. L1 English data were obtained from fifteen Greek-American bilingual speakers who were born in the USA and lived in Greece at the time of the experiment. The L1 speakers were of similar ages to the L2 speakers.
2.2. Speech Materials / Recording Procedure / Data Analysis

Recordings of students’ speech samples were made twice, once before and once after the teaching intervention (time 1 and time 2 respectively). For L1 and L2 English, the speakers were recorded on the words ‘Pat’, ‘tab’, ‘cap’ embedded in the carrier phrase ‘______ is the next word’. For L1 Greek, the words ‘πάσα’ / ‘pasa’ (‘pass’), ‘Τάσα’ / ‘tasa’ (‘Tasa’, a female name), ‘κάσα’ / ‘kasa’ (‘box’), were recorded in the carrier phrase ‘____       ___ λέω πάλι’, /____ ‘leo ‘pali’, (_____ I say again’). Six repetitions of each word were measured for each speaker. A total of 3402 items were measured for the present study (for L1 Greek: 30 speakers*3 words*6 repetitions=540 items, L1 English 15 speakers*3 words*6 repetitions=270 items; L2 English: 72 speakers*3 words*6 repetitions*2 recordings=2592 items).

The recordings were analysed acoustically with the use of waveforms and digital spectrograms generated by the speech analysis software PRAAT (Boersma & Weenink 2007). The VOT duration was measured from the beginning of the release burst to the first positive peak in the periodic portion of the waveforms. For the data of the experimental group at time 2, the ‘native-likeness criterion’ (Birdsong 2007, Flege et al. 1992) was used, based on the plus / minus one standard deviation from the average native English VOT. The percentage of stop tokens that belonged to this range was calculated.

Finally, a series of t-tests were conducted in order to investigate the statistical significance of the difference between: (a) L1 Greek and L1 English VOT, (b) L1 and L2 VOT, (c) L2 VOT between time 1 and time 2, (d) the VOT of the experimental and the control group and (e) the VOT of the male and the female speakers.

2.3. Pronunciation Teaching Intervention

The framework of pronunciation teaching which was used in the present study was adopted from Celce-Murcia et al. (1996: 36), who proposed five teaching stages for pronunciation teaching which range from controlled to free activities. These stages include the following: (1) description and analysis of features, (2) listening discrimination activities, (3) controlled practice, (4) guided practice and (5) communicative practice.

The students of the experimental group received fifty-one lessons of pronunciation instruction on English stops and speech rhythm. Twenty-two of these lessons were devoted to the teaching of word-initial phonologically voiceless and voiced stops. Each pronunciation lesson lasted 10-15 minutes and it was embedded in the English classes at school. The lessons were conducted by the author, who is a non-native speaker of English and was also the main English teacher of the experimental groups.

The demonstration of the phenomenon of aspiration (stage 1) started with the comparison of the pronunciation of Greek and English words. Segmentally
similar Greek and English words were chosen (e.g. ‘Pie’-Πάει ‘/paɪ/’), the former produced with aspirated word-initial stops, while the latter produced with unaspirated stops. A multimodal activity was then used with visual and tactile reinforcement, as the students were asked to observe the puff of air associated with the production of English aspirated stops, as they were holding sheets of paper in front of their mouths.

During the following stages (stages 2-5) the production of aspirated stops was practised in parallel with the production of devoiced word-initial /b, d, g/. The activities included, among others, listening discrimination of minimal pairs presented both in isolation (e.g. ‘coat’ - ‘goat’) and in context (e.g. ‘He had the peach to himself’ - ‘He had the beach to himself’), reading-aloud dialogues with emphasis on the particular sounds, information-gap activities, role-plays and creation of dialogues. Nursery rhymes (for the 10-year-old students) and limericks (for the 13- and 16-year old students) with a number of occurrences of the particular sounds were also used in class.

3. Results

This section presents the results on the production of word-initial voiceless stops in L1 and L2 production. Figure 1 presents the results for the VOT duration in L1 Greek and L1 English averaged across the speakers of the three ages. The figure indicates that, in accordance with the literature, Greek /p, t, k/ are voiceless unaspirated with short VOT values, whereas English /p, t, k/ are produced as voiceless aspirated with long VOT. The difference in the VOT of the two languages is statistically significant for all places of articulation (independent-sample t-tests, p<0.001 for all places of articulation).

![Figure 1. Mean VOT (ms) and standard deviation in L1 Greek and English.](image-url)
The results on L2 English production reveal that before teaching both the experimental and the control groups produced unaspirated voiceless stops with short VOT (figure 2). No significant difference is observed between the VOT of the experimental and the control groups at time 1 (independent-sample t-tests, $p>0.05$). It appears that the speakers generally produced Greek-like VOT at time 1, since the VOT of L2 English at time 1 is significantly different from L1 English (independent-sample t-tests, $p<0.001$ both for the experimental and the control group) and no different from L1 Greek (independent-sample t-tests, $p>0.05$ for the experimental and the control group, compare figures 1 and 2).

The results at time 2 indicate that there is a statistically significant increase in VOT duration for the experimental, but not for the control group (paired-sample t-tests comparing VOT at time 1 and 2, $p<0.001$ for the experimental group; $p>0.05$ for the control group). As a result, at time 2 the VOT of the experimental group is different from that of the control group to a significant extent (independent-sample t-tests, $p<0.001$). The comparison between L1 and L2 production reveals that after teaching the VOT of the experimental group is significantly different from L1 Greek ($p<0.05$) and no different from L1 English ($p>0.05$). The VOT values of the control groups at time 2 remain significantly different from L1 English ($p<0.05$) and no different from L1 Greek ($p>0.05$), suggesting no change in VOT for the particular group.

Figure 2. Mean VOT (ms) and standard deviation in L2 English in the experimental and control groups before and after teaching.

Figure 2 presents the results for the male and the female speakers before and after teaching, averaged across the three places of articulation. The results on speakers’ gender indicate that before teaching no significant difference
is observed between the VOT of the male and female speakers, either in the experimental or in the control group (independent-sample t-tests p>0.05 for the experimental and the control groups). However, at time 2 the VOT of the female speakers is significantly greater than that of the male ones for the experimental group (p<0.001). No difference is found for the control group with regard to speakers’ gender.

**Figure 3.** Mean VOT (ms) in L2 English for the male and female speakers before and after teaching.

Figure 2 indicates that considerable standard deviation is prevalent in L2 data, especially with regard to the VOT of the second recording of the experimental group. This finding suggests that inter-speaker variability is evident in this group of speakers, since not all speakers performed in a uniform way.

In order to explore individual variability in detail, the criterion of ‘native-likeness’ (Birdsong 2007, Flege et al. 1992) was adopted. This criterion was based on the plus / minus one standard deviation from the native English average VOT scores and, according to this, the percentage of stop tokens that belonged to the native-like range was estimated. Table 1 presents the native-like VOT range obtained as a result of the application of the native-likeness criterion.

**Table 1.**
The VOT range for voiceless stops based on the plus / minus one standard deviation from the average L1 English VOT values (in ms).

<table>
<thead>
<tr>
<th>Native-likeness criterion</th>
<th>/p/</th>
<th>/t/</th>
<th>/k/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>47.8-77.2</td>
<td>59.2-88.7</td>
<td>60.5-87.5</td>
</tr>
</tbody>
</table>
In the experimental and the control group at time 1 and in the control group at time 2 there are no speakers whose average VOT falls within the native-like range, as they all produce short VOT values, similar to L1 Greek, except for one ten-year-old speaker who produces target-like VOT for \(/k\)/ at time 1.

After the teaching intervention, the implementation of this criterion led to the identification of four types of voiceless stop production in the experimental group. These four types of stop production are the following: (a) native-like production (according to the native-likeness criterion), (b) longer VOT values than those determined by the native-likeness criterion, (c) intermediate VOT between L1 Greek and L1 English (estimated as the VOT values that fell between L1 Greek average VOT plus 1 standard deviation and L1 English average VOT minus 1 standard deviation) and (d) Greek-like VOT. From these categories the first three can be interpreted to denote an improvement in VOT production, whereas the last one shows no change from L1 Greek VOT. In needs to be noted that type (b) is interpreted to denote improvement, since similarly high VOT scores have been reported in the literature for adult monolingual English speakers (Allen et al. 2002, Deterding & Nolan 2007). Figure 4 presents the percentage of stop tokens that belong to each type of stop production in the experimental group after teaching.

![Figure 4](image)

**Figure 4.** The percentage of stop tokens that belong to each type of voiceless stop production in the experimental group at time 2. The percentages are averaged across all places of articulation and all age groups.

Figure 4 indicates that a greater percentage of stop tokens is intermediate between the VOT of L1 Greek and L1 English than the tokens that belong to each of the other three types of production. Also a considerable percentage of tokens are indicative of longer VOT than that of the L1 English speakers of the
present study, whereas approximately one fourth of stop tokens belong to the native-like range. The least frequent type of stop production is that of Greek-like unaspirated stops. The greatest percentage of stop tokens, i.e. 83.2%, belongs to the types (a), (b) and (c), which are considered to denote an improvement in production.

4. Discussion

The present study examined the VOT for word-initial voiceless stops produced by Greek learners of English before and after a specific pronunciation instruction. The first aim of the study was to investigate L2 production and, in particular, the effect of teaching in a foreign language context. The analysis of the results shows that, before the teaching intervention, interference from L1 Greek was prevalent in L2 production, as both groups tended to transfer the production of voiceless unaspirated stops into L2 English. In particular, the L2 VOT was not significantly different from that of L1 Greek at time 1. Our finding that Greek learners of English tend to transfer the L1 VOT in L2 production does not agree with previous studies that reported intermediate VOT values between L1 and L2 in the speech of the L2 speakers (Caramazza et al. 1973, Flege & Eefting 1987, Major 1987, Williams 1979). However, one major difference between our study and the earlier ones is that the latter examined speakers who acquired the L2 in a naturalistic setting, as they were exposed to the target language in their every day communication. The extensive exposure to the target language and to native-speaker input can explain the approximation to the target-like VOT values for the speakers examined previously; however, the foreign language context of acquisition may not provide learners with adequate native-like input and exposure to L2 in order to help them approximate the L2 VOT values, as a result speakers tend to resort to L1 transfer.

As for the effect of teaching, at time 2 a significant increase was observed for the VOT of the experimental, but not of the control groups. Also the VOT of the experimental groups was not significantly different from English at time 2, whereas the VOT of the control groups was not significantly different from Greek, similarly to time 1. These results suggest that well-organized and planned pronunciation teaching can improve students production, even if it occurs in a foreign language environment with all the limitations that characterise it, for example lack of native English teacher, limited exposure and practice outside classroom, limited interaction in the target language in everyday life and limited amount of time devoted to English lessons at schools.

It needs to be noted, however, that even though the statistical analyses revealed English-like VOT for the experimental groups at time 2, the standard deviation, as well as the implementation of the native-likeness criterion revealed
great variability in VOT production. In particular, the use of the native-likeness criterion proved a useful tool for the investigation of individual variability, as it helped identify four interlanguage patterns, one of which was closer to L1 Greek, two closer to L2 English and one was intermediate between the two. This finding suggests that, even though approximately one fourth of stop tokens were realised as target-like after teaching, the speakers’ production was still not systematic. This lack of systematic production can be related to the teaching intervention adopted, together with the foreign context of language acquisition. Perhaps the particular number of lessons might not have been adequate enough for aspiration to be produced systematically in a target-like way and that longer and more intense practice, as well as greater exposure to the language are required before students’ VOT production shows systematic improvement.

Regarding the role of speakers’ gender, our results provide preliminary evidence of female students’ advantage over the male ones in the production of aspirated stops after the teaching intervention. In particular, although no difference between the male and female speakers appeared before teaching, after the teaching intervention the female speakers produced significantly longer VOT values than the male speakers. However, it needs to be noted that both the male and the female speakers produced longer aspiration at time 2 than at time 1, which suggests that speakers of both genders benefited from pronunciation instruction.

Providing an explanation of the tentative advantage of the female students in the production of the particular feature appears quite challenging, as the examination of students’ background questionnaire fails to provide a clear answer. However, a tentative explanation comes from the teachers’ observation that, as a general trend, the female students, and in particular those who participated in the present research, tended to be more diligent and more attentive during the lessons than the male ones. Also according to the teacher’s observations, the female students generally tended to show greater enthusiasm and interest in the pronunciation lessons than the male students. This could potentially explain the advantage of the female over the male speakers, therefore it might be useful for future studies to investigate students’ gender in relation to attitudinal and motivational factors before firmer conclusions are obtained on the issue.

5. Conclusion

Overall, this study has established the need for teaching the phenomenon of aspiration to Greek learners of English, since without teaching Greek learners of English tend to transfer the Greek pattern of unaspirated stop production. The results indicate that the students can benefit from pronunciation instruction, as they generally produced longer VOT values after the teaching intervention than
before teaching. Also a considerable percentage of stop tokens were produced in a native-like fashion, which indicates that native-like production of aspirated stops is an achievable goal even for students in a foreign context of language acquisition. Nevertheless, native-like performance was not systematic, possibly due to constraints related to the particular teaching intervention and the foreign context of language acquisition. Finally, an advantage of the female over the male students was reported and this was tentatively interpreted in relation to attitudinal and personality factors.
References


