UNIVERSITY OF TARTU

DEPARTMENT OF ENGLISH STUDIES

THE ASPIRATION OF VOICELESS PLOSIVES BY ESTONIAN LEARNERS OF ENGLISH: A CASE STUDY

BA thesis

LIIS ANTONS
SUPERVISOR: *Lect.* REELI TORN-LEESIK (PhD)

TARTU
2023

ABSTRACT

Aspiration is a feature accompanying English plosives but is not a part of the Estonian sound system. Jenkins (2000, cited in Walker 2018) considers aspiration an essential aspect of English pronunciation. The aspiration of English voiceless plosives by Estonian speakers is a feature that has been studied by Hallaste (2008) and Ader & Miljan (2015) via auditory analysis. This thesis aims to investigate L1 Estonian students' pronunciation of English voiceless aspirated plosives via acoustic analysis. The aims of this thesis are summarized in a single research question: based on an acoustic analysis, do upper secondary school students aspirate initial voiceless plosives, indicating that they can separate voiceless plosives from voiced plosives? A pronunciation test was created and conducted on six upper secondary school students. The recordings from the students' pronunciation test were analyzed in terms of their VOT values and compared to native speaker VOT values, the presence of aperiodic aspiration noise on waveform and spectrogram was determined.

This thesis begins with an introduction containing the motivation and core concepts of this thesis. The literature review explores the definition of aspiration, its acoustic properties, and gives an overview of previous studies conducted on L1 production of L2 aspirated plosives. The second main section explores the methodology used for conducting the study and analyzing the students' recordings, the results of the acoustic analysis will be presented and discussed. The final part of this thesis is the conclusion which summarizes the findings of this thesis.

ACKNOWLEDGEMENTS

The author of this thesis would like to acknowledge the many people who assisted in the making of this thesis. Firstly, the author would like to thank Eva Liina Asu-Garcia for advising the pronunciation test of this thesis, and Pärtel Lippus for providing the author access to study materials regarding the use of PRAAT and advising the author on some of the acoustic analysis methods used in this thesis. Secondly, the author would like to thank Pire Teras for her in-person instruction in the program PRAAT and for assisting in the methodology section of this thesis. The author would also like to thank Anton Malmi for his instruction in the University of Tartu's phonetics lab and for supervising the first pronunciation test. Of course, the author is extremely grateful to the six students that participated in this study and to the schoolteacher who agreed to recruit the students for this thesis. And finally, the author could not have undertaken this journey without the help and support of the supervisor of this thesis, Reeli Torn-Leesik.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AmE	American English
BrE	British English
EFL	English as a Foreign Language
L1	first language, meaning the native language of a person
L2	second language of a person, which is not their native language and is learned
LFC	Lingua Franca Core
ms	milliseconds
NS	native speaker
NNS	non-native speaker
VOT	voice onset time

INTRODUCTION

English is most often utilized as *a lingua franca*, which is defined as a "contact language" between speakers who do not speak a common first language (Mauranen 2018: 7). Mauranen (2018: 7) writes that English is more often spoken by non-native speakers rather than native speakers of English. When it comes to the spread of English on a global scale, Estonia is no exception. Self-reported language proficiency and active language use (including comprehension, speaking, and writing) in English increased from 37% in 2017 to 47% in 2020 compared to those who report not speaking English at all, which was 26% in 2017 and dropped to 18% in 2020 (Kivistik 2020: 32–33). Foreign language learning is also a part of the national curriculum in Estonian basic education, where most students learn English as their first (A) foreign language (Ministry of Education and Research n. d.). Graduating upper secondary school students must also take a written and oral B-level examination in their A foreign language, (Education and Youth Board n.d.), which often is English.

The Estonian and English phonetic systems have many differences in the pronunciation of vowels, consonants, features of connected speech, sentence stress, and intonation. Many younger Estonian speakers have shown tendencies towards more native-like English pronunciation (showing less L1 interference) (Ader & Miljan 2015). Although more native-like English pronunciation might be seen as a success from an educational standpoint, according to Walker (2018), pronunciation teaching has shifted from a rigid native-level goal to one of intelligibility. According to Jenkins (2000, cited in Walker 2018), the aspiration of word-initial voiceless plosives is an important ability for an L2 English speaker to learn for mutual intelligibility in conversations with both native and non-native speakers of English. Despite the much-improved English skills amongst EFL students in Estonia, both Ader & Miljan (2015) and Hallaste (2008) found that the aspiration of voiceless plosives has remained a difficult area in English pronunciation. When it comes to teaching practices in Estonia, Teder (2023: 34,

40–41) found that most Estonian EFL teachers, even those who think that accent should be eliminated, agree that intelligibility should be the goal in teaching pronunciation. However, aspiration along with assimilation were the least taught pronunciation aspects (Teder 2023: 39).

Although L2 English aspiration and plosive production being a well-researched topic in other languages, L1 Estonian speakers' production of L2 English plosives is an area that has received little attention despite the fact that English is taught as a foreign language in schools and there are significant differences in voicing and aspiration categorization of plosives between the two languages. The aim of this small case study is to perform an acoustic analysis of the pronunciation of voiceless aspirated plosives by Estonian learners of English in terms of their VOT values and presence of aspiration using the free speech analysis software PRAAT. Comparing the findings of L1 Estonian pronunciation of L2 English to the plosive production of native English speakers can provide detailed empirical evidence on what specific features Estonian EFL learners might struggle with and whether they can phonetically differentiate between voiced and voiceless plosives.

The aims of this thesis were summarized into a research question:

- Based on an acoustic analysis, do upper secondary school students aspirate initial voiceless plosives, indicating that they can separate voiceless plosives from voiced plosives?

It is important to mention that the thesis at hand is a small case study composing of only six test subjects, so it is therefore not possible to draw any general conclusions on whether the pronunciation of plosives amongst Estonian EFL learners has improved on a larger scale. It can, however, add to the knowledge on Estonian production of English plosives in terms of acoustical measurements.

This thesis is divided into two main sections. The literature review that makes up the first section will first introduce the significance of aspiration in English and in EFL, the

production of aspiration and plosives, and the differences between Estonian and English plosives. Then it will introduce the findings of previous studies on L2 English plosives and provide measurements on the plosive production of native English speakers.

The second part of this thesis is dedicated to analyzing the pronunciation of English plosives by Estonian learners. The compilation of the pronunciation test and the background questionnaire as well as the parameters by which the acoustic analysis is conducted are explained. Secondly, the speech samples from the pronunciation test are analyzed and compared to the VOT and aspiration measurements of native English speakers to see what aspects of English plosive and aspiration production Estonian learners of English struggle with. The final part includes the discussion of the results.

1. LITERATURE REVIEW

The first part of the following section will give an overview of the significance of aspiration in English and in EFL, aspiration and plosive production, acoustic representations of aspiration on waveform and spectrogram, and a comparison between Estonian and English plosives. The second part of this section will summarize the findings of studies done on English plosives as well as studies conducted on the pronunciation of English plosives by Estonian and other EFL speakers.

1.1. Intelligibility and aspiration

Aspiration is a feature accompanying plosives in different contexts across many languages, including English. It refers to a puff of air that occurs after a plosive, resulting in a lag in voicing (Roach 2000: 34, Ogden 2009: 102), which in English separates voiceless /p t k/ from voiced /b d g/ (Lisker & Abramson 1964: 384).

Walker (2018: 64) claims that in the 1980s native-speaker accuracy was the norm for EFL pronunciation, with Received Pronunciation (BrE) and American English (AmE) being most favored. There was a shift in the perception on L2 English pronunciation in the 2000s when Jennifer Jenkins suggested a collection of pronunciation features called the *Lingua Franca Core* (LFC), which would ensure intelligibility between non-native speakers (NNS) (Walker 2018: 65). The LFC is an evidence-based syllabus composed of pronunciation features necessary for intelligibility for mainly NNS, since according to Jenkins (2005: 200) communication between NNS happens more frequently amongst EFL learners than conversing with native speakers (NS). When discussing the LFC, Wells (2005: 101) affirms that teaching aims depend on who the EFL speakers will be conversing with; however, Dauer (2005: 544) notes that since NNS interactions are the most common interactions EFL learners encounter, there is essentially no need for an EFL speaker to acquire a near-native accent. Therefore,

pronunciation mistakes are not seen as a deviation from NS pronunciation, but as a deviation from intelligibility described in Jenkins' LFC (Walker 2018: 65). Jenkins (2005: 200) concludes that the LFC was designed to ensure intelligibility in NNS interactions.

When it comes to aspiration, Roach (2000: 34) claims that aspiration is an important aspect from an NS standpoint since word-initial voiceless unaspirated plosives might be perceived as voiced plosives /b d g/ by native speakers. Furthermore, aspiration is deemed an important phonological feature even from an NNS intelligibility standpoint. Namely, the list of intelligibility features included in the LFC summarized in Walker (2018: 65) are as follows:

- all of the consonants except /θ/, /ð/ and /r/, where variation was acceptable;
- the aspiration of /p/, /t/ and /k/ in initial position in stressed syllables;
- the correct treatment of word-initial consonant clusters;
- vowel length differences, especially the shortening of vowels preceding voiceless consonants (e.g., cap–cab or wait–wade);
- nuclear stress placement (i.e., 'I love speaking English' versus 'I love speaking English');

The LFC states that aspiration is necessary for mutual intelligibility even between NNS speakers. The discussion around aspiration and both NS and NNS intelligibility can further be highlighted in Estonian EFL learners since Estonian does not make this distinction between plosive consonants: English has aspiration distinction between voiced and voiceless plosives (Lisker & Abramson 1964: 394), while Estonian recognizes only voiceless plosives (Asu et al. 2016: 65) and does not recognize voicing contrast in plosives (Ariste 1984: 49). As a result, Estonian EFL speakers may struggle with English plosives.

When it comes to pronunciation teaching tendencies amongst EFL teachers in Estonia, Estonian teachers' views on pronunciation teaching were described in Teder's thesis (2023). Teder (2023: 26–27) investigated the pronunciation practices of 20 native Estonian English teachers between the ages of 23 and 62. When it comes to accented English, the majority agreed

that achieving a native-like accent should not be a strict goal and the overwhelming majority agreed that intelligibility should be the primary goal of pronunciation (Teder 2023: 31–34). Despite this, aspiration was one of the most rarely taught aspects of English pronunciation, while half of the teachers taught voiced and voiceless consonants regularly, which was mentioned separately (Teder 2023: 38–41).

1.2. Aspiration and plosive production

Aspiration is the h-like sound produced by air passing through the glottis, which occurs between the release of a plosive and the voicing of a following vowel (Roach 2000: 34, Ogden 2009: 102). To better understand the features of aspiration, it is important to first look at various aspects of plosive production, such as plosive phases, voicing, the context in which aspiration takes place, and some acoustic characteristics of aspiration.

A plosive is a consonant in which a complete closure forms between two articulators to block the airflow, then after a short period of time the closure is suddenly released, causing an "explosion" of the air built up behind the articulators (Ogden 2009: 16; Roach 2000: 32). The English sound system recognizes six plosives /p t k b d g/, which are divided into voiced /b d g/ and voiceless /p t k/ plosives (Ogden 2009: 96). Sounds, especially consonants, are determined by their place of articulation that specifies which articulators come together to produce the sound. According to the place of articulation, English plosives are divided into bilabials /p/ and /b/, alveolars /t/ and /d/ and velars /k/ and /g/ (Ogden 2009: 96). Ogden (2009: 96) divides the production phase into three phases: closing, hold, and release (C, H, R). The closing phase consists of bringing two articulators together to form a closure. The vocal tract is completely sealed off during the hold phase as the lungs are still forcing air out of the vocal tract. During the release phase, built up air is released from behind the closure during the final release phase, resulting in a plosion.

Voicing is considered a distinguishable feature in English pronunciation, and that

includes plosives. Plosives can be fully voiced (throughout C, H, and R phases), partially voiced, and voiceless (Ogden 2009). English phonology divides plosives into voiced /b d g/ and voiceless /p t k/, where /p t k/ are always voiceless, but /b d g/ can be fully voiced, partly voiced and voiceless (Roach 2000: 33). When it comes to voicing consonants, it is measured with VOT (voice onset time), which is the time between the release of the plosive and the beginning of voicing (Ogden 2009). Lisker & Abramson (1964) also add that the force (lenis & fortis) aspect also exists and is regarded as a third distinction between voiced and voiceless plosives on top of aspiration and voicing, but similarly, as Roach (2000) states, there are no grounds for differentiating nor proving the aspect of "force" in the articulation of plosives.

English voiceless plosives /p t k/ are accompanied by aspiration before the voicing of the following vowel begins (Ogden 2009: 102). Aspiration takes place at the glottis (which is the small opening between the vocal folds) after the release of a plosive when the vocal folds are not yet vibrating, but air is passing through the open glottis (Karlsson 2002: 69–70). When it comes to the voicing of aspirated plosives, Lisker & Abramson (1964) note that the VOT is considerably longer for voiceless aspirated plosives. Roach (2000: 34) and Ogden (2009: 103) summarize the criterion in which a plosive is aspirated as follows: in word-initial stressed positions preceding a vowel. However, if the voiceless plosive is preceded by /s/, then it remains unaspirated (Roach 2000: 34). Therefore, aspiration and aspirated plosives are realized as allophones of voiceless plosives, not as separate phonemes since aspiration accompanies plosives only in specific contexts (McMahon 2002: 18).

Acoustically, aspiration is described by silence during the hold phase, a transient burst during the release phase, and then aspiration friction after the release (Ogden 2009: 98–99). Aspiration is produced by turbulent airflow passing through the glottis, resulting in a breathy voice (Johnson 2003: 138–139; Ogden 2009: 102). Turbulence is an aperiodic sound which appears in fricatives as well as aspiration. According to Ogden (2009: 34) both friction and

turbulence are represented by an aperiodic soundwave which has no repeating waveform pattern (whereas voicing is represented by a repeating waveform pattern). On the spectrogram, Lisker & Abramson (1964: 386) and Klatt (1975: 688) categorize aspiration as a "noise" at the second and third formant frequencies and missing at the first formant frequency.

To summarize, aspiration takes place after the release phase of a plosive and ends with the start of voicing of the following vowel. Aspirated plosives are realized as allophones of voiceless plosives and accompanies plosives only in stressed syllables preceding a vowel. On waveform and spectrogram, aspiration is indicated by an aperiodic sound missing a voicing bar that lies between VOT.

1.3. Estonian and English plosives compared

While Estonian orthography distinguishes between *p, t, k,* and *b, d,* g, Estonian phonology, however, only realizes four plosives /k p t t̪/, which are voiceless and unaspirated (Asu et al. 2016). Despite the fact that separate phonemes are not recognized based on voicing, research has shown that Estonian plosives have voiceless, partially voiced as well as fully voiced allophones in intervocalic positions of spontaneous speech (Ermus 2017). Unlike the English sound system, variation in voicing does not result in different phonemes in the Estonian sound system (Ariste 1984: 49). The Estonian sound system only distinguishes plosives based on fortis/lenis contrast: short plosives that appear in a voiced setting are categorized as lenis plosives; long plosives that appear in word-initial positions are categorized as fortis plosives (Asu et al. 2016: 65–66). The voicing and aspiration differences between English and Estonian are highlighted by Mutt (1978: 45–49), who has provided a comparison between Estonian and English consonants, taking voicing and aspiration into consideration. Mutt (1978: 49) instructs Estonian learners to put more effort into voicing voiced plosives and aspirating initial voiceless plosives.

When comparing voicing in Estonian and English, the approximate VOT in Estonian

word-initial voiceless plosives for /p/ is 18 ms, for /t/ is 31 ms, and for /k/ is 37 ms, but can vary depending on the following vowel (Eek & Meister 1996: 242–243, cited in Asu et al. 2016: 83). In the case of English, the average VOT for word-initial voiceless aspirated plosives in single isolated words is 58 ms, 70 ms, and 80 ms for /p t k/, respectively (Lisker & Abramson 1964: 394). Although it is not mentioned whether Eek & Meister (1996, cited in Asu et al. 2016: 83) collected the VOT values in connected speech or isolated words, the contrast between VOT values of English aspirated and Estonian unaspirated plosives by Eek & Meister (1996, cited in Asu et al. 2016: 83) and Lisker & Abramson (1964: 394, 410, 411) are illustrated in Table 1.

Table 1. VOT values of English voiceless aspirated plosives and Estonian voiceless unaspirated plosives (Eek & Meister 1996: 242–243, cited in Asu et al. 2016; Lisker & Abramson 1964: 394, 410, 411)

Voiceless phoneme	**English voiceless aspirated plosives in isolated words**	**English voiceless aspirated plosives in speech (in non-initial sentence positions)**	**English voiceless aspirated plosives in speech (in initial sentence positions)**	**Estonian voiceless unaspirated plosive**
/p/	58 ms	34 ms	28 ms	18 ms
/t/	70 ms	37 ms	39 ms	31 ms
/k/	80 ms	49 ms	43 ms	37 ms

From Table 1 it is apparent that word initial VOT values of English aspirated and Estonian unaspirated plosives are different, with English VOT values lasting longer in running speech than Estonian plosives due to aspiration. The minimum VOT for aspirated /p/ were recorded in initial sentence positions, which was 28 ms. The shortest duration for the English /t/ were recorded in non-initial sentence positions, which was 37 ms. The shortest duration for English /k/ was recorded in initial sentence position, which was 43 ms. Compared to Estonian voiceless plosives, the English /p/ is at least 10 ms longer than the Estonian equivalent, the English /t/

and /k/ are both at least 6 ms longer than the Estonian equivalent.

In summary, the Estonian and English sound systems have vastly different categorizations for plosives based on voicing, lenis/fortis contrasts and aspiration.

1.4. Previous studies on VOT and aspiration of English plosives

The production of VOT, voicing, and aspiration of L2 English plosives by EFL learners are features that have been studied across many languages (Shahidi & Rahim 2011; Kim 2012; Abdelaal 2017) while others have carried out cross-linguistic comparisons of VOT and aspiration distinctions between different languages, including English (Lisker & Abramson 1964, Deterding & Nolan 2007).

Some of the most valuable information for this thesis (especially regarding VOT) comes from the research of Lisker & Abramson (1964). Lisker & Abramson (1964: 388) carried out a cross-language study to see how well VOT distinguished between languages that have two, three, or four separate categories dedicated to plosives (17 speakers and 11 languages in total). The study examined plosives in word-initial positions before vowels in both isolated words as well as in sentences (Lisker & Abramson 1964: 388–389). The study included four speakers of American English who were instructed to read the word containing the targeted phonemes out in isolation and were then instructed to use the words in sentences in initial and non-initial positions (Lisker & Abramson 1964: 388, 391, 394, 407). VOT was deemed a sufficient measurement for differentiating between stop categories for English (Lisker & Abramson 1964: 420), but it was not sufficient in other languages, such as Hindi and Marathi, which differentiate between both aspirated and unaspirated categories of voiced and voiceless plosives, in which case VOT measurements were deemed insufficient (Lisker & Abramson 1964: 403, Abramson & Whalen 2017: 3).

Klatt (1975: 686) also set out to expand the knowledge on VOT and the burst frication noise in word-initial consonant clusters. His study analyzed 25 different word-initial consonant

clusters as well as words beginning with a single plosive consonant embedded in a carrier phrase to represent VOT in conversational speed (Klatt 1975: 687). The study investigated aspiration in word-initial consonant clusters by measuring the VOTs of plosives produced by three native speakers (Klatt 1975: 686–687). The results concluded that the VOT was longer before sonorants (such as plosives /r/) and high vowels (such as /i u/) and shorter before mid- and low vowels (such as /ɛ/) as well as detected aspiration in /s/-clusters (Klatt 1975: 686). The VOT measurements in voiceless aspirated plosives /p t k/ were found to be 47 ms, 65 ms, and 70 ms, respectively (Klatt 1975: 687, 689), which are longer than the VOT in conversational speed listed in Lisker & Abramson (1964).

When it comes to the duration of aspiration specifically, Deterding & Nolan (2007) compared the duration of aspiration and voicing of six Mandarin Chinese plosives to RP British English plosives. In terms of the acoustic measurements for aspiration, it was measured from the release of the plosive until the end of the release burst (Deterding & Nolan 2007: 386). Deterding & Nolan (2007: 386) also add that labeling the aspiration measurements as VOT would be inaccurate since voicing often started before the end of the burst measurement. The average duration of aspiration was found to be 71.4 ms, 80.4 ms, and 97.3 ms for /p t k/, respectively. These measurements are quite significantly longer than those presented in Table 1, but the authors note that this might be because of the British informants' work as lecturers (Deterding & Nolan 2007: 385). The most significant difference occurred in the voicing of plosives during closure in intervocalic positions, but there was little to no difference in the aspiration of plosives in both languages (Deterding & Nolan 2007: 385, 388).

Kim (2012) sought to determine if the place of articulation affects aspiration in voiceless English plosives produced by both English and Korean speakers. In this study aspiration was measured and determined in terms of their VOT values (Kim 2012: 2). The words under scrutiny included 36 words with plosives in word-initial position and plosives following /s/

(Kim 2012: 5). Although VOT values in English produced by Korean speakers were appropriate for English aspirated plosives, the Korean language does have two categories out of three categories dedicated to plosives (Kim 2012: 10, 13).

Abdelaal (2017) set out to examine the production of voiceless stops by native Arab speakers of English to assist Arab EFL Students. The main measurement used to analyze plosives in this study was VOT. According to the study, /t/ was produced with a shorter VOT and weaker aspiration than that of native speakers probably due to the interference of L1, but there was little difficulty in producing /d/ (Abdelaal 2017: 11–12). /k/ was mostly produced with an appropriate VOT value, while four out of the ten participants produced /k/ with a smaller VOT value (Abdelaal 2017: 12). The VOT value of /g/ was too high, probably due to it not being in the Arabic sound system (Abdelaal 2017: 12–13). Discrepancies in VOT values were said to be due to the phonetic differences between the two languages, for example, /g/ does not exist in Arabic or /t/ is a dental plosive in Arabic (Abdelaal 2007).

Shahidi & Rahim (2011) addressed the pronunciation differences between Malay and English plosives and the performance of English plosives by Malay EFL speakers. Although the comparison of the sound systems of Malay and English reveals that Malay plosive stops are unaspirated, VOT was the only measurement taken for this study and aspiration was not included as a research goal (Shahidi & Rahim 2011). The results concluded that there are voicing differences in the sound systems of both languages and Malay accented English highlighted by L1 Malay VOT patterns (Shahidi & Rahim 2011).

From the studies by Abdelaal (2007), Deterding & Nolan (2007), Shahidi & Rahim (2011), and Kim (2012), we can see VOT measurements being used as an acoustic measurement for detecting aspiration and voicing contrasts across many studies, even in those languages that do not distinguish between plosive categories based on aspiration. Furthermore, Suarez (2008: 148) claims that VOT is the measurement with which the degree of aspiration for voiceless

stops is defined, although other sources describe more specific measurements (e.g., Lisker & Abramson 1964, Klatt 1975; Deterding & Nolan 2007). There is also evidence to support the fact that aspiration can be difficult for EFL students whose native language does not aspirate plosive consonants.

The most in-depth study on L1 Estonian pronunciation of L2 English plosive consonants was carried out by Hallaste (2008), where she conducted a contrastive analysis of high-school students' production of English plosives in terms of aspiration, voicing, and place of articulation. The author used mainly auditory analysis through listening and interpretation of the subjects' pronunciation (Hallaste 2008: 47). The study found that only 33% of voiceless plosives were produced as intended and 27% were weakly aspirated (Hallaste 2008: 66). Most students displayed a weak aspiration and the students who produced the studied phonemes without aspiration, produced them with voicing (Hallaste 2008: 66–67). In terms of voicing, the percentage of properly voiced consonants was 44%, although the voicing was inconsistent (Hallaste 2008: 67). Hallaste (2008: 73) summarizes that voicing was the least difficult aspect of pronunciation, although less than 44% of the phonemes were produced with correct voicing. Only a third of the aspirated phonemes were produced correctly (Hallaste 2008: 72). Hallaste (2008: 72–73) concludes that 12th grade students studying English as their A-language in 2008 could not accurately produce the different English phonemes in terms of aspiration and voicing.

The most recent study on L2 English pronunciation by L1 Estonian learners was conducted by Ader & Miljan (2015). In addition to analyzing the pronunciation mistakes of two age groups, the study also analyzed the effects of L1 phonological features on L2 pronunciation. The informants' pronunciation was analyzed with the help of a native speaker of British English hinting at auditory analysis being implemented. In the case of aspirated voiceless plosives, Ader & Miljan (2015: 28) found that the subjects left them unaspirated, and they were produced as L1 Estonian sounds. For example, from the younger group of participants with an average age

of 20.5 years, 67% pronounced /k/, and 63% pronounced /p/ without aspiration, but only 13% produced /t/ without aspiration. The study also found that some voiced plosives /g/, /b/ and /d/ were produced as aspirated, which is not a characteristic of Estonian nor English (Ader & Miljan 2015: 25–28).

Although many instruction manuals for L1 Estonian learners of English (e.g., Mutt 1978, Kostabi 2004) have emphasized the importance of aspirating (initial) voiceless plosives when covering the topic of plosives, only two studies (Ader & Miljan 2015, Hallaste 2008) have tackled the issue of aspirated plosives and have concluded that English plosives produced by Estonian speakers has remained a difficult aspect. Despite this, Estonian speakers of English have shown improvement towards more native-like pronunciation. For example, Asu (1999) reported a strong accent amongst first-year students of English language and literature in higher education in 1997. In contrast, Ader & Miljan's (2015) findings reported the younger group of subjects exhibiting less L1 interference than the group of middle-aged subjects. More native-like pronunciation is explained by an earlier onset of learning and exposure to natural communication in the target language via English speaking media (Ader & Miljan 2015). Kuppens (2010) adds that better general language skills amongst young EFL learners might be explained by positive attitudes towards English due to their association with leisure activities. In terms of variety, Ader & Miljan (2015) concluded that the American variety was especially present in younger subjects because of American TV shows and the US film industry. Despite a trend of more native-like pronunciation amongst Estonian and other EFL learners, Stölten (et al. 2015) found that native-like L2 plosive production (in terms of VOT) was difficult to obtain even when being perceived as NS by other native speakers. It is important to add that the study was conducted on L1 Spanish (unaspirated plosives) production of L2 Swedish (aspirated) plosives (Stölten et al. 2015: 71).

To summarize this literature review, there are significant differences between Estonian

and English plosives in terms of voicing/voiceless and lenis/fortis contrasts, the most significant being the aspiration and voiced distinction not being realized in the Estonian sound system. Many previous acoustic cross-language studies have used mainly VOT as a measurement to provide detailed information about voicing and aspiration. While Estonian English pronunciation has improved (Ader & Miljan 2015), little attention has been given to Estonian pronunciation of voiceless aspirated plosives (Hallaste 2008, Ader & Miljan 2015) and it has received no acoustic attention.

2. ANALYSIS OF UPPER SECONDARY SCHOOL STUDENTS' ASPIRATION OF VOICELESS PLOSIVES

The aim of the second main section of this thesis is to first explain the methodology used for this study, the selection of participants, the principles of compiling the language background questionnaire, the compilation of the speech test, the tools for the acoustic analysis, and the parameters by which the presence of aspiration was determined. The section will then focus on the results of the analysis of the students' recordings and present the findings.

2.1. Selection of participants

The participants in the study consisted of six adult students studying in upper secondary school. The criteria for the subjects that participated in this study included being at least 18 years of age, being a native speaker of Estonian, being a student in upper secondary school and, studying English as a B1 or a B2 language. According to the Council of Europe (2001: 74), a person that has reached B2 level in a foreign language should be able to "interact with a degree of fluency and spontaneity that makes regular interaction, and sustained relationships with native speakers quite possible without imposing strain on either party. /…/", meaning the student should have acquired a level of pronunciation proficiency to be able to be intelligible to both NNS and NS.

The participants were six upper secondary school students. Five of the students were in their last year of upper secondary school (12^{th} grade) and one student was in the second year of upper secondary school (11^{th} grade). The mean age of the participants was 18.33 years ranging from 18 to 19. The participants were acquired through both formal and informal channels. Four participants were recruited through a representative in a school who agreed to recruit students for this study. Two participants were acquired through convenience sampling that took part in the study based on their proximity to the author. Since the number of participants in this study is small, acquiring participants through formal and informal channels can give a broadened

overview of Estonian students' pronunciation. The author suspects that the students who agreed to participate through the formal channel agreed to do so due to being confident in their English-speaking ability. The students who participated due to being affiliated with the author might not be as confident in their pronunciation but still agreed to participate. This gave an opportunity to test the pronunciation of students that might not be as confident in their pronunciation.

To ensure the anonymity of the participants in this study, the students were given a code, which consisted of S (short for student) and a number based on the order of their participation in the pronunciation test. Therefore, the participants in this study will be referred to as S#1, S#2, S#3, S#4, S#5, and S#6 based on the order of participation.

2.2. Language background questionnaire

The questionnaire was designed to determine the students' language learning background and their attitudes towards English pronunciation. The students were asked to state their age, gender, and the grade they were currently in. Questions 1-2 focused on the students' language background in English, asking when they started formally learning English and how good their self-reported proficiency in English was. Questions 3-5 aimed to determine how often and where the students report using English the most and to assess their pronunciation proficiency. Question 6 asked what variety of English the students report using the most (American, British, or other). Questions 7-8 focused on whether the students had received pronunciation instructions in school. Questions 9-10 asked the students whether they strive for a native-like pronunciation while speaking English and whether they think other English speakers should strive for a native-like English pronunciation. Question 11 asked the students to describe the most important aspect to them when speaking English.

The background questionnaire revealed that five subjects out of six subjects were in their last year of upper-secondary school (i.e., 12th grade) and one (S#2) was a second-year upper-secondary (i.e., 11th grade) student. All the participants in this study were female.

As for the students' English learning background (Q1), three students started formal English lessons in the first education level (grades 1–3)[1] , S#1 started formal English education in the second level of primary education (grades 4–6), and S#5 started formal English education in kindergarten. Regarding pronunciation teaching in school (Q7), two students (S#1 and S#3) were taught in detail how to pronounce different phonemes. Two students (S#2 and S#6) report being taught sounds that do not appear in the Estonian sound system, and two students (S#4 and S#5) report being corrected when pronunciation was unintelligible.

Self-reported proficiency (Q2) revealed that four participants assessed their proficiency as "very good". One of the participants (S#6) reported their English proficiency as being "good", and one (S#2) reported their proficiency as "satisfying". When assessing pronunciation skills (Q5), only S#1 reported her pronunciation skills as being "very good". Four students reported their pronunciation as being "good". Only S#2 reported their pronunciation as being "satisfying".

When it comes to variety (Q6), half of the students reported using BrE, two students (S#2 and S#6) reported opting for AmE variety, and one student (S#4) reported not focusing on any variety.

When asked about personal pronunciation goals (Q9), three (S#1, S#4, and S#5) students said that correct pronunciation was important to them. The other three students (S#2, S#3, and S#6) said that correct pronunciation was not that important to them. When inquiring about the students' general view on L2 spoken English and accent (Q10), the majority (S#1, S#2, S#3, S#5, S#6) agreed that strictly sticking to the pronunciation customs was not as important, but correct pronunciation was something an L2 English speaker should strive

[1] The education levels in the Estonian education system are classified as follows:
I education level (grades 1–3)
II education level (grades 4–6)
III education level (grades 7–9)
Upper secondary education (grades 10–12)

towards. Only one student (S#4) thought that correct pronunciation was something an L2 speaker of English should use.

The last question (Q11) was an open-ended question asking the students to state their most important aspects while speaking English. Four students mentioned (mutual) intelligibility in some capacity. S#1 added that she wished to be thought of as a native speaker based on her pronunciation. S#4 mentioned using a wide range of vocabulary as her focus when speaking English, and S#6 highlighted correct grammar and sentence structure as important to her while speaking English.

2.3. Pronunciation test

The pronunciation test included testing the three voiceless aspirated plosives /p t k/ as well as the three voiced plosives /b d g/ in English. Three words were selected for each of the three voiced and voiceless English plosives. There were 18 main words altogether. The test also included 18 filler words to distract the subjects. The test included three warm-up words which were presented before the actual test to allow for the subjects to get accustomed to the test format. In summation, the test included 18 main words, 18 filler words, and three warmup words, 39 words altogether. The four types of words and their purpose can be summarized as:

i. three warm-up words at the beginning of the test so the subjects could get acquainted with the format of the test;
ii. words intended to test the pronunciation of voiceless plosives and aspiration;
iii. words including the pronunciation of voiced plosives;
iv. filler words to distract the subjects from figuring out the target words;

The words with the targeted phoneme were repeated at least three times to get an accurate representation of the subjects' pronunciation of the plosive. The 18 filler words were also repeated 3 times each.

The words for the pronunciation test were selected based on the criteria for aspirated

plosives described in Ogden (2009) and Roach (2000): word-initial position preceding a vowel in a stressed syllable. The words employed in the pronunciation test are presented in Table 2.

Table 2. Table 2. Words used in the pronunciation test according to phoneme.

/k'/	**1) C**at /**k**æt/ **2) K**ite /**k**aɪt/ **3) C**up /**k**ʌp/
/p'/	**4) P**ig /**p**ɪg/ **5) P**ear /**p**eər/;/**p**er/ **6) P**an /**p**æn/
/t'/	**7) T**iger /ˈ**t**aɪgə(r)/ **8) T**eabag /ˈ**t**iːbæg/ **9) T**oy /**t**ɔɪ/ - mid, long
/g/	**10) G**oat /**g**əʊt/; /**g**oʊt/ **11) Gh**ost /**g**əʊst/; /**g**oʊst/ **12) G**alaxy /ˈ**g**æləksi/
/b/	13) **B**all /**b**ɔːl/; /**b**ɑːl/ 14) **B**anana /**b**əˈnɑːnə/; /**b**əˈnænə/ **15) B**ook /**b**ʊk/
/d/	**16) D**og /**d**ɑːg/; /**d**ɒg/ **17) D**eer /**d**ɪər/; /**d**ɪr/ **18) D**iamond /ˈ**d**aɪəmənd/

Each word was represented with three separate images in the pronunciation test (111 images altogether, including the warmup words). These images were mostly acquired from two free photography websites called Unsplash and Pixabay, though some personal photographs were also included. All the images had the targeted word spelled out on the images to avoid confusion. The pronunciation test was first compiled via PowerPoint presentation, since it had the appropriate editing software to display the spelling of the word on the image. The pronunciation test was then compiled by taking screenshot images of the PowerPoint slides to then use in the free recording software program Speech Recorder BAS developed by the Bavarian Archive for Speech Signals at Ludwig Maximilian University (Ludwig Maximillian University n.d.), which is the standard software used in the phonetics lab at the University of Tartu. Speech Recorder BAS allows for a smooth and sufficient speech recording process and later acoustic analysis. The targeted words were embedded in a carrier phrase *I can see a* ___ so that the targeted word could be preceded by a vowel sound. Having the aspirated plosives

following a vowel allows for a more sufficient acoustic analysis, since the repeating wave pattern of the vowel is easily distinguishable from the silence of the closing and hold phases of the following plosive. The vowel creates a clear barrier between the previous phoneme and the burst noise of the plosive.

The final pronunciation test was piloted on one person. The pilot test subject was asked to comment on the clarity of the instructions and images. After piloting the test some changes were made regarding the clarity of the instructions and some images considered ambiguous or confusing were changed. Piloting the pronunciation test also provided information about the approximate duration the test might take, which was 6 minutes and 18 seconds for the final version.

The pronunciation test was conducted in the phonetics lab run by the Institute of Estonian and General Linguistics at Lossi 3, Tartu. The participants were first asked to fill out the language background questionnaire and were then briefed on the nature of the pronunciation test. The students were then led to the recording booth of the phonetics lab and given instructions by the author. The students were instructed to say in English what they see on the image by using the sentence "I can see a ___" with each image and were given an example sentence. The students were also instructed not to rush and to try and keep a level tone throughout the carrier phrase. Then the signals of the program were explained, and the students were instructed not to start speaking until the signal turned green, since only then did the program start recording. The images were changed by the author in the control room during the test and the students only saw the speaker window on their computer (see Appendix 2). The students were first shown three images as a warmup task and asked if they had any questions about the test during the break. The pronunciation test commenced after the break. The first pronunciation test was conducted under the supervision of a research fellow in the phonetics lab.

2.4. Methods of acoustic analysis of the speech samples

A combination of auditory and acoustic analysis was implemented, i.e., a combination of listening, waveform, and spectrogram analysis. The main focus was on acoustic analysis. The speech was analyzed using the program called PRAAT. PRAAT is a free software application created by Paul Boersma and David Weenink, designed for speech analysis through spectrogram and waveform (Boersma & Weenink n.d.). Version 6.3.08 of PRAAT was used to conduct the analysis.

The presence of aspiration was determined by the presence of two criteria: an appropriate VOT measurement and the presence of aspiration noise on waveform and spectrogram. The VOT measurements on waveform and spectrogram were taken according to the acoustic parameters described in Lisker & Abramson (1964), Klatt (1975), and Ogden (2009). VOT measurement borders were mainly determined by analyzing the waveform. Lisker & Abramson (1964: 389) marked VOT as the time interval between the release of the stop and the onset of voicing, which is represented by a periodic sound, i.e., a repeating wave pattern (Ogden 2009: 34). Therefore, on the waveform, the borders of VOT measurements will be determined by a sudden burst of sound on the waveform, and the beginning of voicing were determined by the appearance of a repeating wave pattern. On the spectrogram, the burst release is indicated by the appearance of a dark vertical strip and the beginning of voicing is indicated by the appearance of the voice bar, which is a dark horizontal strip appearing at the bottom of a spectrogram at low frequencies indicative of voicing (Fulop & Disner 2011: 2). The shortest VOT measurements of aspirated plosives in English described by Lisker & Abramson (1964: 410–411) depicted in Table 1 were used as an indicator of whether there was enough time for the speaker to display appropriate aspiration. These measurements were 28 ms, 37 ms, and 43 ms for /p t k/, respectively (Lisker & Abramson 1964). If the VOT values lie in the preferred time-range, the presence of aspiration was determined by the presence of an aperiodic wave on

the waveform described in Ogden (2009) and Johnson (2003). Aspiration on the spectrogram was determined by the absence of the voicing bar described in Lisker & Abramson (1964) and Klatt (1975). The final decision on aspiration was concluded via listening. The PRAT program has the option of playing a selected part of the recording. After VOT borders were marked, the VOT section was played and listened to detect if there was a puff of audible air after the release of the plosive.

Image 1 below is a magnified sample of the aperiodic aspiration noise on the waveform. The amplitude of aspiration may vary depending on the subject and the place of articulation of the plosive.

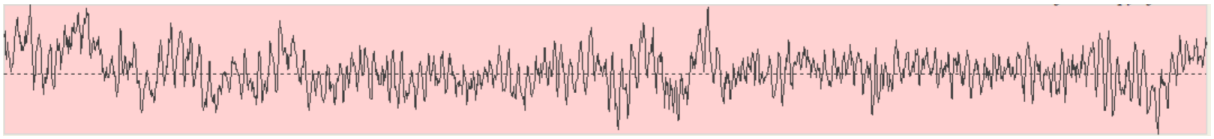


Image 1. Magnified sample of the aperiodic aspiration noise on the waveform.

Image 2 and Image 3 show S#5's pronunciation of the words *teabag* and *pig*. The vertical strip across the waveform and spectrogram highlighted with a red hue indicates the VOT measurement, including the period of aspiration. Note that the amplitude in the aspiration noise in Image 3 is not as strong as in Image 2 but is still considered as aspirated.

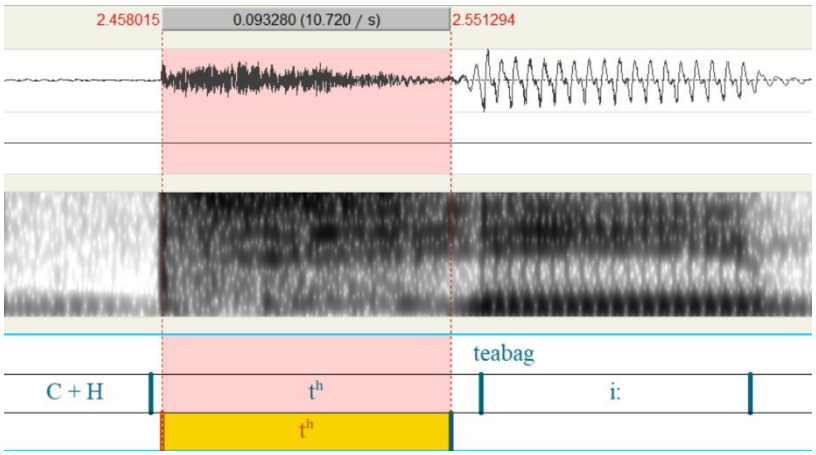


Image 2. Sample of S#5's pronunciation of the word *teabag*.

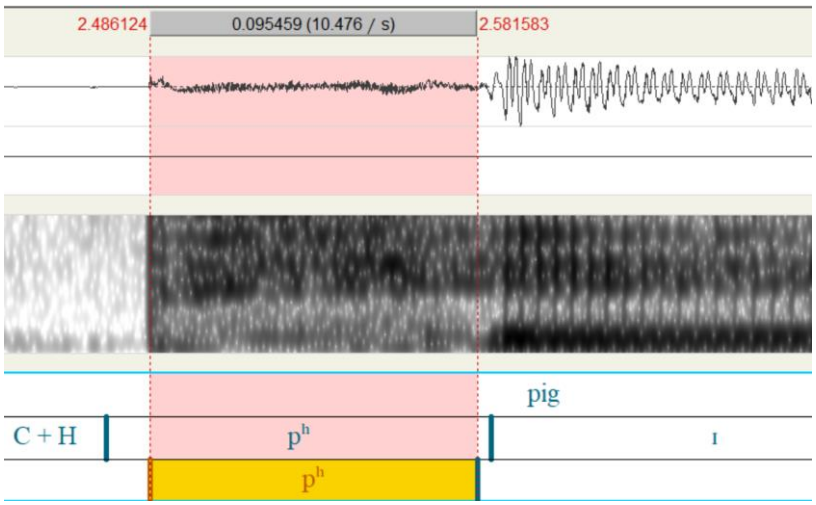


Image 3. Sample of S#5's pronunciation of the word *pig*.

2.5. Results of the acoustic analysis of the students' speech samples

The following section focuses on analyzing the VOT and aspiration of the subjects' speech samples by comparing those findings to the VOT values of native speakers presented in Lisker & Abramson (1964) and detecting the presence of aspiration noise on waveform and spectrogram. The instances of voiceless plosives (27 instances per subject, 162 instances across all six participants) were divided into three categories based on VOT and aspiration: (1) instances where VOT values were sufficient, and aspiration was detected; (2) instances where VOT values were shorter than the shortest values presented by Lisker & Abramson (1964), but could be considered as aspirated; (3) and instances where plosives were produced as lenis plosives according to VOT (and the absence of aspiration noise). The second (2) category of shorter VOT measurement results will be referred to as having "shorter" aspiration by the standards set for this thesis. It must be noted that this thesis does not consider speaking rate and compares VOT of aspirated plosives to previous studies. It would be ideal if recordings of native speakers undergoing the same pronunciation test in the exact same conditions as the L2 Estonian students were available since the length of this pronunciation test might produce even shorter results than that of Lisker & Abramson (1964). Therefore, marking the results as having

"shorter" VOT might not be entirely accurate, especially considering that the results in Lisker & Abramson (1964) were the mean values of four speakers. Despite this, it is trusted that the measurements provided by Lisker & Abramson (1964) accurately represent VOT and aspiration in the English language in running speech.

The quantitative data regarding the students' VOT values are summarized in Table 3, which represents whether the instances of aspirated plosives surpass the native VOT threshold per student. Note that there were 27 instances per student.

Table 3. Sufficient and insufficient instances of VOT values per student.

Student	**Instances with insufficient VOT values**	**Instances with appropriate VOT values**	**Percentage of insufficient VOT values**
S#1	7	20	25.9%
S#2	3	24	11.1 %
S#3	0	27	0%
S#4	0	27	0%
S#5	0	27	0%
S#5	0	27	0%

Due to the small number of participants in this thesis, the results should be discussed individually. Regarding VOT measurements, out of 162 instances of aspirated plosives across six subjects, there were 10 instances where VOT measurements were considered too short compared to those of native speakers. All these instances appeared in the pronunciation of two subjects: S#1 and S#2. The rest of the four subjects had sufficient VOT values and present aspiration noise.

In the case of S#1, 7 out of the 27 instances of voiceless aspirated plosives were produced with a shorter VOT. The insufficient VOT values appeared in /p/ and /k/ plosives, where VOT values were shorter than native VOT values, yet aspiration noise was present on the waveform. The second instance of /p/ in the word *pig* and the second instance in the word *pan* was produced with a shorter VOT, which did not affect the mean value of the three

instances in either word. In the case of /k/, the first instance of the word *cat* and the third instance of the word *kite* had shorter VOT values; however, the mean values remained sufficient. In the instances mentioned above, the VOT values were a maximum of 4 ms shorter than those mentioned in Table 1 set as the VOT threshold. Therefore, the shorter VOT might not be as significant. To summarize, aspiration noise was present in the instances with shorter VOT values, and the plosives were produced as aspirated voiceless plosives. A more significant difference in VOT appeared in the case of S#1's pronunciation of the word *cup*, where all three instances had a shorter VOT value than the standard native VOT mentioned in Table 1. The mean VOT value in the word *cup* was 34 ms, which is 9 ms shorter than the average native VOT value of 43 ms. Despite this, aspiration noise was present. To summarize, all 7 instances of plosives with shorter VOT in S#1's pronunciation had aspiration present, meaning they belong in category (2). The following table (Table 4) shows the VOT values of S#1. The mean value for S#1's plosive /k/ came out to be 42.2 ms, which is 0.8 ms shorter than the native 43 ms native average. Compared to the results of S#1's background questionnaire, S#1 assessed her English speaking and pronunciation (Q2 and Q5) as "very good". Although 25.9% of S#1's instances of aspirated plosives belong in category (2), all the instances of aspirated plosives were still produced with aspiration, although with shorter aspiration. S#1's VOT values are presented in Table 4. The green hue indicates that VOT values surpass native VOT values, and the orange hue indicates shorter VOT values than those by native speakers. The last column of the table indicates the deviation from the mean VOT values of native speakers.

Table 4. VOT values and mean values of voiceless plosives of S#1.

Plosive p'	Instance 1	Instance 2	Instance 3	Mean	Deviation
Pig	35 ms	25 ms X	28 ms	29.3 ms	1.3 ms
Pear	51 ms	48 ms	34 ms	44.3 ms	16.3 ms
Pan	46 ms	24 ms X	30 ms	33.3 ms	5.3 ms
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	Deviation
Tiger	60 ms	62 ms	43 ms	55 ms	18 ms
Teabag	78 ms	84 ms	76 ms	79.3 ms	51.3 ms
Toy	77 ms	38 ms	64 ms	59.7 ms	22.7 ms
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	Deviation
Cat	40 ms X	53 ms	46 ms	46.3 ms	3.3 ms
Kite	44 ms	54 ms	41 ms X	46.3 ms	3.3 ms
Cup	36 ms X	34 ms X	32 ms X	34 ms X	-9 ms

In the case of S#2, 3 instances out of 27 were produced with a shorter VOT value. The first instance of the word *pear* was produced with a VOT of 9 ms, which was too short for aspiration to occur as it could not be detected. The VOT for English /b/ is 4 ms in non-initial sentence positions and 7 ms in initial sentence positions (Lisker & Abramson 1964: 410, 411). Despite being 2 ms longer than native English speaker VOT of voiced /b/, auditory analysis concluded that this instance is considered an unaspirated voiced plosive, since 9 ms is too short for aspiration. S#2 also had another peculiar instance of a short VOT in the second instance of /t/ in the word *tiger*, which was produced with a VOT of 16 ms. The VOT for English /d/ is 9 ms in initial and 7 ms in non-initial positions (Lisker & Abramson 1964: 410, 411); therefore, S#2 did not produce it necessarily as a voiced plosive. Despite this, the instance of /t/ was produced with very weak aspiration (the VOT was too short to determine aspiration), and auditory analysis would classify it as an unaspirated plosive, perhaps even a voiced plosive /d/. These two instances of /p/ and /t/ are placed in category (3). Furthermore, the second instance of /k/ in the word *kite* was produced with a smaller VOT, but aspiration noise was still present and did not affect the mean value, being therefore placed in category (2). The following table (Table 5) shows VOT values of S#2's pronunciation. The green hue in Table 5 indicates

appropriate VOT, the orange hue indicates a shorter VOT than the native average, and the dark orange hue indicates plosives whose VOT was considered too short to be aspirated. This information of S#2's pronunciation can be compared to the information acquired from S#2's background questionnaire in which she assesses her level of English pronunciation and English-speaking skills (Q2 and Q5) as "satisfying". Despite this, only 11.1% of S#2's VOT values were insufficient, meaning 88.9% of the plosives were produced with sufficient VOT values.

Table 5. VOT values and mean values of voiceless plosives of S#2.

Plosive p'	Instance 1	Instance 2	Instance 3	Mean	Deviation
Pig	45 ms	57 ms	53 ms	51.7 ms	23.7 ms
Pear	9 ms X	43 ms	34 ms	28.7 ms	0.7 ms
Pan	77 ms	57 ms	64 ms	66 ms	38 ms
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	Deviation
Tiger	52 ms	16 ms X	64 ms	44 ms	7 ms
Teabag	80 ms	58 ms	64 ms	67.3 ms	30.3 ms
Toy	79 ms	72 ms	73 ms	74.7 ms	37.7 ms
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	Deviation
Cat	47 ms	63 ms	70 ms	60 ms	17 ms
Kite	65 ms	39 ms X	61 ms	55 ms	12 ms
Cup	57 ms	59 ms	43 ms	53 ms	10 ms

Despite some shortcomings in terms of VOT values in S#1 and S#2's VOT values, the differences between the two students' mean VOT values and native average VOT values suggest that they are capable of appropriately producing English aspirated plosives. In many cases S#1 and S#2's VOT are close to the measurements observed in Klatt (1975) where the use of the carrier phrase was similar to that used in the current thesis. These measurements were 47 ms, 65 ms, and 70 ms for /p t k/, respectively.

Therefore, 152 instances out of 162 instances of aspirated plosives were produced with the presence of aspiration and sufficient VOT values, meaning that 93.8 % of plosives belong in category (1) of appropriately aspirated plosives; 8 instances (4.9 %) were had a shorter VOT value, but aspiration noise was present, belonging in category (2); and finally, two instances

(1.2%) had a very short VOT and without aspiration, where they could be considered as a voiced, and were placed in category (3). Despite 10 instances being placed in categories (2) and (3), the majority of the mean VOT values per word (except for /k/ for S#1) as well as per phoneme surpassed the appropriate VOT value threshold across all six students.

The mean values of all the students' VOT per phoneme and native VOT values (in isolated words and running speech) are summarized in Figure 1.

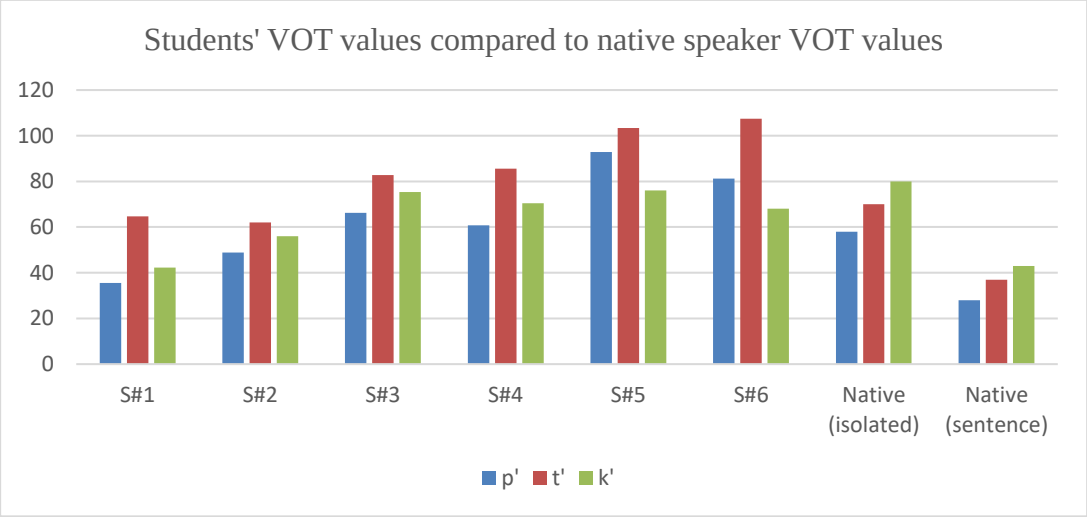


Figure 1. Students' mean VOT values per plosive compared to the VOT values of native speakers.

Figure 1 suggests that the native VOT values are usually the shortest for /p/ and the longest for /k/, with /t/ lying in the middle. In the case of Estonian students' VOT values, all the students display the longest VOT values for /t/. In the case of S#5 and S#6, the VOT values for both /p/ and /t/ are longer than the VOT values for /k/. To better analyze the VOT values of the students, Table 6 will summarize the students' VOT values along with native VOT values.

Table 6. Students' mean VOT values per plosive compared to the VOT values of native speakers

Student/plosive	/p'/	/t'/	/k'/
S#1	35.6 ms	64.7 ms	42.2 ms
S#2	48.8 ms	62 ms	56 ms
S#3	66.2 ms	82.8 ms	75.4 ms
S#4	60.8 ms	85.6 ms	70.4 ms
S#5	92.8 ms	103.4 ms	76.1 ms
S#6	81.2 ms	107.6 ms	68 ms
Native VOT (sentence)	28 ms	37 ms	43 ms
Native VOT (isolated)	58 ms	70 ms	80 ms

In Table 6, the yellow hue indicates the mean VOT values per phoneme, which surpassed the minimal VOT value threshold in sentences but did not surpass the threshold of native VOT values in isolated words. The blue hue indicates that the mean VOT value per phoneme has surpassed the VOT value in isolated words from Lisker & Abramson (1964). S#1 and S#2 did not surpass the isolated VOT threshold, and none of the students' VOT values for /k/ surpassed the isolated VOT threshold. However, S#3, S#4, S#5, and S#6 did pass the isolated VOT threshold regarding /p/ and /t/. Especially high mean VOT values for /p/ and /t/ were recorded in S#5 and S#6's pronunciation. The mean value for /p/ for S#5 was 92.8 ms, which was 34.8 ms longer than the native average in isolated words, and the mean value for S#6 was 81.2 ms, which was 23.2 ms higher than the native average. The VOT values were especially high for /t/, which for S#5 was 103.4 ms, which was 33.4 ms higher than native average, and for S#6 was 107.5 ms, which was 37.5 ms higher than the native average. In summation, out of 18 mean VOT values per phoneme across six students, 8 mean VOT values, meaning 44.4% of the mean VOT values per phoneme surpassed the native average in isolated words.

2.6. Discussion of the findings

The results of the analysis show a high success rate of students' aspiration of voiceless plosives in terms of VOT values and presence of aperiodic aspiration noise, indicating that they

can produce English aspiration. Hallaste's (2008) thesis studying the pronunciation of 32 students' aspiration of voiceless plosives found that only 33% of plosives were appropriately aspirated, 27% were weakly aspirated, and 40% were left unaspirated. Ader & Miljan (2015: 27–28) found that the younger group of participants produced in 63% instances of /p/, 67% instances of /k/, and 13% instances of /t/ without aspiration. Although the small number of subjects in this thesis does not allow for drawing general conclusions, 93.8% of the plosive instances being produced with appropriate aspiration could indicate that students performed better in this study than the previous ones. Only 1.2% (2 instances produced by S#2) can be considered unaspirated compared to Hallaste's (2008) 40%. Despite the limitations of this study, the students performed with 93.8% instances of plosives being appropriately aspirated compared to the results of Hallaste's (2008) 33% appropriately aspirated and 27% weakly aspirated results.

These differences between this study and Hallaste's (2008) and Ader & Miljan's (2014) study can be explained by a smaller number of participants, as well as differences in research methods. Hallaste (2008) implemented mainly auditory analysis, dividing plosives into aspirated, weakly aspirated, and unaspirated categories. Ader & Miljan (2014) implemented auditory analysis with the help of a native speaker. The current thesis separates plosives into aspirated, aspirated with a shorter VOT, and unaspirated plosive categories relying almost solely on acoustic details in detecting aspiration in terms of VOT and the presence of aperiodic noise. During the analysis of this thesis, it was noted that auditory analysis of the whole carrier phrase, including the targeted word might categorize many instances of aspirated plosives as "weakly aspirated". Despite aspiration not being as audible in some cases, VOT and the aperiodic noise did suggest that aspiration was present in all instances. The amplitude (or loudness) of aspiration can vary from plosive to plosive depending on the place of articulation and the speaker. However, this thesis does not account for the amplitude or loudness of the

aspiration noise and did not make this distinction between "weaker" aspiration (in terms of amplitude). Solely focusing on acoustic measurements can lead to such a contrast between this study and previous studies. In turn, if Ader & Miljan's (2015) and Hallaste's (2008) studies had implemented acoustic analysis, the results could be different depending on VOT values and the presence of aspiration.

The mean VOT values presented in Table 6 show mostly appropriate VOT values for English aspirated plosives (except for S#1's value for /k/). Furthermore, two of the students' mean VOT values (in the case of /p/ and /t/) surpassed those of native speakers in even isolated words and the VOT values in Deterding & Nolan (2007: 387), which were 71.4 ms, and 80.4 ms for /p/ and /t/. High VOT measurements were recorded for /t/, which were higher than the VOT measurements for /k/ in all the subjects. This can be somewhat correlated with the findings in Ader & Miljan's (2015: 27) study, where only 13% of /t/ instances were left unaspirated, meaning students have displayed a smaller number of errors in the pronunciation of /t/ before. In the case of two subjects (S#5 and S#6) the VOT measurement for /p/ surpassed even that of /k/.

These differences can be explained by the nature of the targeted words and the following vowel. Klatt (1975) found that aspiration was longer before high vowels and shorter before mid- and low vowels. Most of the targeted words with aspirated plosives were followed by mid and low vowels, except in words *pig* and *teabag*. Furthermore, in the case of /t/, which had the highest VOT values across all six students, the targeted words for /t/ were the only words consisting of more than one syllable (*teabag, tiger*) one of which had a high long vowel following the plosive (*teabag*) and the other had a diphthong following a plosive (*tiger*). Although vowel length has not been mentioned as having an impact on the length of aspiration (Klatt 1975: 694), the two syllables could mean that the students paid more attention to the pronunciation of those two words, especially considering that the pronunciation of /t/ in *teabag*

had the longest VOT measurement In S#1's pronunciation. The other five subjects recorded the highest VOT values for the word *toy*, which includes a long vowel and is the only monosyllabic word ending with a vowel. Additionally, vowel height can also explain S#1's short mean VOT value for the word *cup*, where the plosive /k/ is followed by a short vowel and ends with a voiceless plosive. Vowels are shortened preceding voiceless plosives (Roach 2000: 35). This feature can be observed with other instances of S#1's pronunciation of category (2) plosives, such as *cat*, and *kite* (although it does not explain the short VOT value for *pig* and *pan*). Perhaps vowel length can affect the speaking rate of the speaker, although there is no evidence to support that claim. However, there seems to be a small correlation between the unaspirated plosives by S#2 between shorter aspiration and the following mid- and low vowels in the words *pear* and *tiger*.

The unusually high VOT values can be explained by hypercorrection, which was also mentioned in the production of voiced plosives by Ader & Miljan (2015). According to Crystal (2008: 232), hypercorrection in second language pronunciation is a term used to describe a phenomenon that takes place when a speaker "over corrects" in pursuit of correct pronunciation and goes beyond the norm set by native speakers. These "overproduced" aspects might not even appear in the native language of the speaker, in which case perfectly suits the aspiration difference between Estonian and English. The findings of Eckman (et al. 2013) suggest that hypercorrection is indicative of a more advanced stage of second language acquisition. Hypercorrection could be explained by the students' English learning background: three of the students that surpassed the native isolated VOT average started formal English education in the first education level, and one started in kindergarten. However, all of those students reported their English pronunciation skills as being "good", but the meaning of "good" pronunciation can vary from person to person.

In conclusion, the 93% success rate suggests that the students did well in the

pronunciation test. Nevertheless, comparison to previous research in the pronunciation of English plosives by Estonian learners (Hallaste 2008, Ader & Miljan 2015) suggests that there is a gap between research methods. Implementing acoustic and auditory analysis to the full extent could better detect if Estonian learners of English can appropriately aspirate English plosives. Therefore, the issue with L1 Estonian production of L2 English plosives is not the length of the plosive, but rather the volume of the aspiration. Some of the shortcomings of this thesis include the small number of participants and the absence of native speakers participating in this study to provide material for comparison. If the methodology includes measuring VOT values, the speaking rate could be neutralized by implementing the method suggested by Stölten (et al. 2015) of comparing VOT percentages from the duration of the whole word instead of just comparing VOT values. Therefore, the author of this thesis has some suggestions regarding future research into the topic of English plosive by Estonian EFL students. Firstly, acoustic and auditory analysis should be used to the full extent to figure out which specific features of English aspirated plosives Estonian students struggle with. Including native speakers in a pronunciation test would ensure appropriate material for comparison. Furthermore, looking deeper into the aspect of hypercorrection could explain abnormally high VOT values in future research. Finally, it would also be beneficial to test aspirated plosives by Estonian speakers in both sentences and isolated words.

CONCLUSION

The aim of this thesis was to investigate L1 Estonian students' pronunciation of English voiceless aspirated plosives /p t k/ via acoustic analysis. The motivation for conducting the study is results from aspiration being one of the features Jenkins (2000) includes in her list of pronunciation features necessary for mutual intelligibility. Moreover, aspiration is not present in the Estonian sound system (Asu et al 2016) and does not realize the voiced plosives /b d g/.

The pronunciation test was conducted on six Estonian students in upper secondary school (five students in 12th grade and one in 11th grade). The pronunciation test includes testing the three voiceless plosives in aspirated contexts, including three words per plosive which were repeated three times each. Therefore, there were 27 instances of aspirated plosives per student, 162 instances across six students. The acoustic analysis included comparing the VOT values of English plosives provided by the Estonian students to the VOT values of native English speakers provided by Lisker & Abramson (1964) and detecting the presence of aspiration. The results were divided into three categories: (1) instances of plosives with appropriate VOT values and aspiration; (2) instances of plosives with insufficient VOT values that still had aspiration noise present; and (3) instances of plosives with insufficient VOT values which were too short for aspiration.

Previous studies on plosives in English and other languages have shown VOT measurements used in many different contexts, such as comparing voicing and aspiration in different languages (Deterding & Nolan 2007, Shahidi & Rahim 2011, Abdelaal 2017), as well as using VOT as a measurement of aspiration (Lisker & Abramson 1964, Kim 2012, Suarez 2013) in English and other languages. English VOT and aspiration has also been studied extensively (Lisker & Abramson 1964, Klatt 1975). The production of English voiceless aspirated plosives produced by Estonian EFL learners has only been studied via auditory

analysis by Hallaste (2008) and Ader & Miljan (2015). Hallaste (2008) found that only 33% of plosives were aspirated, 27% were weakly aspirated, and 40% were unaspirated. Ader & Miljan (2014) found that the younger group of participants produced /p/ in 63% of instances, /k/ in 67% of instances, and /t/ in 13% of instances without aspiration.

Compared to previous studies, the current study found that 93.8% of plosives were aspirated (category (1)), 4.9% were produced with shorter aspiration (category (2)), and only 1.2% were produced without aspiration and as voiced plosives (category (3)). The unaspirated and voiced plosives appeared in the pronunciation instances of a single student, whose background questionnaire revealed that she was not that confident in her English skills to begin with. The acoustic analysis concludes that VOT values as well as the aspiration noise suggest that Estonian EFL learners can pronounce English voiceless aspirated plosives and are aware of the distinction between voiced and voiceless plosives. Furthermore, many of the mean values per plosive surpassed the mean values in isolated words produced by native English speakers. Especially high VOT values were recorded for /t/, which surpassed the VOT values of /k/ in all six students. This can be explained by the nature of the targeted words as well as by students displaying hypercorrection in their pronunciation. The results conclude that the students can produce English voiceless plosives in terms of VOT values, but the author suggests that further research is needed.

The suggestions for further research include testing native speakers alongside L2 speakers, a larger number of participants, implementing both acoustic and auditory analysis, and testing plosives in both isolated words and connected speech.

REFERENCES

Abdelaal, Noureldin Mohamed. 2017. Instrumental Analysis of the English Stops Produced by Arabic Speakers of English. *International Journal of Education & Literacy Studies*, 5: 3, 8–15.

Abramson, Arthur S. and Douglas H. Whalen. 2017. Voice Onset Time (VOT) at 50: Theoretical and Practical Issues in Measuring Voicing Distinctions. *Journal of Phonetics*, 63: July, 75–86.

Ader, Kristiina and Merilin Miljan. 2015. External Factors and The Interference of L1 Estonian on L2 English Pronunciation: An Apparent-time Study. *Estonian Papers in Applied Linguistics*, 11, 31–32.

Asu, Eva Liina. 1999. British or American English: A study of Estonian students' attitudes to pronunciation. In Pilvi Rajamäe (ed.). *New Britain: The Heritage of the Past and the Challenge of the Future: Proceedings of the 2nd International Tartu Conference on British Studies Held at the University of Tartu, August 24–25, 1998*, 22–30. Tartu: Tartu University Press.

Asu, Eva Liina, Pärtel Lippus, Karl Pajusalu & Pire Teras. 2016. *Eesti Keele Hääldus* [Pronunciation of the Estonian Language]. Tartu: University of Tartu Press.

Ariste, Paul. 1984. *Eesti Keele Foneetika II* [Estonian Phonetics II]. Tartu: Tartu Riiklik Ülikool.

Boersma, Paul and David Weenink. N.d. Praat: doing phonetics by computer. Available at https://www.fon.hum.uva.nl/praat/, accessed May 20, 2023.

Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Strasbourg: Cambridge University Press.

Crystal, David. 2008. *A Dictionary of Linguistics and Phonetics*. 6th ed. Malden: Blackwell Publishing.

Deterding, David & Francis Nolan. 2007. Aspiration and Voicing of Chinese and English Plosives. In *Proceedings of the 16th International Congress of Phonetic Sciences, Universität des Saarlandes Saarbrücken, Germany*, 385–388.

Dauer, Rebecca M.. 2005. The Lingua Franca Core: A New Model for Pronunciation Instruction? *Tesol Quarterly*, 39: 3, 543–550.

Eckman, Fred R., Gregory K. Iverson and Jae Yung Song. 2013. The Role of Hypercorrection in the Acquisition of L2 Phonemic Contrasts. *Second Language Research*, 29: 3, 257–283.

Education and Youth Board. N. d. State Examinations. Available at https://harno.ee/en/state-examinations, accessed November 2022.

Ermus, Liis. 2017. Eesti keele lühikeste klusiilide häälduse variatsioon ja seda mõjutavad tegurid [Variation in the pronunciation of Estonian short plosives and its affecting factors]. *Mäetagused*, 68, 27–52.

Fulop, Sean A. and Sandra F. Disner. 2011. Examining the Voice Bar. *Journal of the Acoustical Society of America: Proceedings of Meetings on Acoustics (162nd Meeting)*, 14: 3, 060002.

Hallaste, Anneli. 2008. *Oh, Gaptain, My Gaptain! Aspiration, Voicing, and Place of Articulation in the Production of the English Plosive Stop Consonants by Estonian Upper Secondary School Leavers with Special Reference to the CEFR B2 Level of Language Proficiency.* MA thesis. University of Tartu.

Jenkins, Jennifer. 2005. Misinterpretation, Bias, and Resistance to Change: The Case of the Lingua Franca Core. In n Kataryna Dziubalska-Kołaczyk & Joanna Przedlacka (eds). *English pronunciation models: A changing Scene*, 199–210. Bern: Peter Lang.

Johnson, Keith. 2003. *Acoustic and Auditory Phonetics*. 2nd ed. Malden: Blackwell Publishing.

Karlsson, Fred. 2002. *Üldkeeleteadus* [General Linguistics]. Tallinn: Eesti Keele Sihtasutus.

Klatt, Dennis H.. 1975. Voice Onset Time, Frication, and Aspiration in Word-initial Consonant Clusters. *Journal of Speech and Hearing Research*, 18: 4, 686–706.

Kostabi, Leili. 2004. *Pronunciation Practice for Advanced Students*. Tallinn: AS Koolibri.

Kim, Ji Yea. 2012. Production of Aspiration in English Voiceless Stops in Terms of Place of Articulation. *SNU Working Papers in English Language and Linguistics,* 10: 12, 1–15.

Kivistik, Kats. 2020. *Eesti Elanike Keeleoskus ja -kasutus, Inimestevahelised Kontaktid, Sallivus ning Eesti Keele Väärtustamine* [Language Skills and Usage if Estonian Residents, Interpersonal Contacts, Tolerance and Appreciation of the Estonian Language]. In The Estonian Integration Monitoring Report 2020: 28–43.

Kuppens, An H 2010. Incidental Foreign Language Acquisition from Media exposure. *Learning, Media, and Technology* 35:1, 65–85.

Lisker, Leigh & Arthur S. Abramson. 1964. A Cross-language Study of Voicing in Initial Stops: Acoustical Measurements. *Word* 20: 3, 384–422.

Ludwig Maximillian University. N.d. Speech Recorder. Available at https://www.bas.uni-muenchen.de/Bas/software/speechrecorder/, accessed May 2023.

Mauranen, Anna. 2018. Conceptualizing ELF. In Jennifer Jenkins, Will Baker and Martin Dewey (eds). *The Routledge Handbook of English as a Lingua Franca*, 7–24. Abingdon and New York; Routledge.

McMahon, April. 2002. *An Introduction to English Phonology.* Edinburgh: Edinburgh University Press.

Ministry of Education and Research. N. d. Foreign Language Learning in Estonia. Available at https://www.hm.ee/en/estonian-and-foreign-languages/foreign-languages/foreign-language-learning-estonia, accessed November 2022.

Mutt, Oleg. 1978. *Inglise Keele Foneetika* [English Phonetics]. Tallinn: Valgus.

Ogden, Richard. 2009. *An Introduction to English Phonetics*. Edinburgh: Edinburgh University

Press.

Roach, Peter. 2000. *English Phonetics and Phonology: A practical course.* 3rd ed. Cambridge: Cambridge University Press.

Shahidi, A.H. & Aman Rahim. 2011. An Acoustical Study of English Plosives in Word Initial Position Produced by Malays. *3L, Language, Linguistics, Literature: The Southeast Asian Journal of English Language Studies* 17: 2, 23–33.

Stölten, Katrin, Niclas Abrahamsson & Kenneth Hyltentstam. 2015. Effects of Age and Speaking Tate on Voice Onset Time. The Production of Voiceless Stops by Near-Native L2 Speakers. *Studies in Second Language Acquisition* 37: 1, 71–100.

Suarez, Philip. 2013. The Pronunciation of the Aspirated Consonants *P, T,* and *K* in English by Native Speakers of Spanish and French. In *College of Education Research Conference (COERC) 2008: Proceedings of The Seventh Annual College of Education Research Conference*: 144–150.

Teder, Johanna. 2023. *Teaching Pronunciation: Teachers' Perception and Aspects Taught in Estonian EFL Classes*. MA thesis. University of Tartu.

Walker, Robin. 2018. From Being Intelligible to Being Themselves: Pronunciation for Today. *Modern English Teacher*, 28: 3, 64–67.

Wells, John C.. 2005. Goals in Teaching English Pronunciation. In Kataryna Dziubalska-Kołaczyk & Joanna Przedlacka (eds). *English pronunciation models: A changing scene*, 101–110. Bern: Peter Lang.

APPENDICES

Appendix 1 – Questionnaire

Küsimustik õpilastele enne kõnekatset

Olen Liis Antons ning kirjutan bakalaureuse tööd inglise keele hääldusе teemal. Antud küsimustik on mõeldud kõnekatses osalejatele enne katset täitmiseks. Küsimustiku eesmärk on teada saada, milline on Teie inglise keele õpingute ja kõnelemise taust ning millised hoiakud on Teil inglise keeles kõnelemise suhtes. Antud küsimustik on anonüümne ning küsimustiku andmeid kasutatakse koos kõnekatse tulemustega minu bakalaureuse töös.

Küsimustikus „Muu…“ valikuvastuse juures võib julgelt enda kohta käiva vastuse lisada!

Klass: _____

Vanus: _____

Sugu: _____

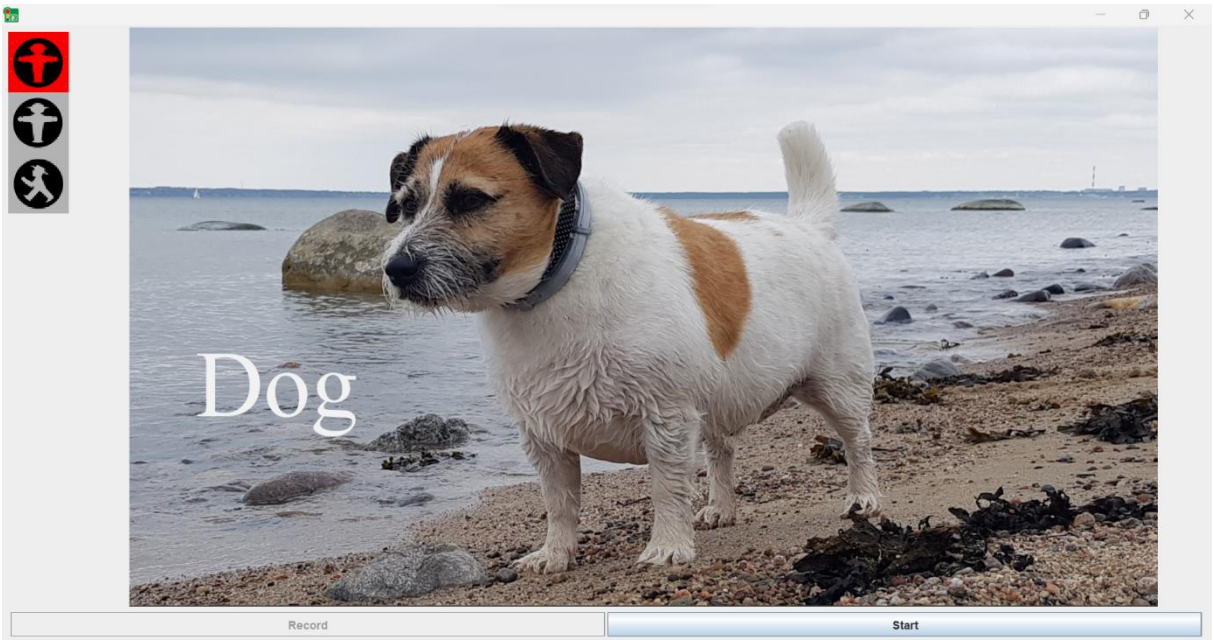
1. Millises kooliastmes Te inglise keelt õppima hakkasite?
    - Lasteaed
    - I kooliaste (1-3. klass)
    - II kooliaste (4-6. klass)
    - III kooliaste (7-9. klass)
    - Gümnaasium
2. Teie enda hinnangul, kui hea on Teie inglise keele tase?
    - Väga hea
    - Hea
    - Rahuldav
    - Mitterahuldav
3. Umbes kui tihti Te inglise keeles kõnelete?
    - Iga päev
    - Üle päeva
    - Kaks korda nädalas
    - Kord nädalas
    - Harvem
4. Kus Te kõige rohkem inglise keelt kõnelete?
    - Koolis inglise keele tundides
    - Kodus perega
    - Sõpradega suheldes
    - Muu: ……..
5. Teie enda hinnangul, kui hea on Teie inglise keele häälduso skus?
    - Väga hea
    - Hea
    - Rahuldav
    - Mitterahuldav
6. Millist inglise keele hääldusе varianti Te kõige rohkem kasutate (Ameerika hääldus, Briti hääldus, jne.)?
    - Ameerika häälдust
    - Briti häälдust
    - Otseselt ei keskendu ühelegi hääldusvariandile

   - Muu: ……
7. Kas Teile on koolis õpetatud inglise keele häädlust?
   - Jah, meile õpetati detailselt kuidas erinevaid häälikuid hääldada
   - Meile õpetati eelkõige häälikuid, mida eesti keele häälduses ei esine
   - Pigem mitte, õpetaja parandas kui häädlus oli arusaamatu, kuid häädlusele pandi vähe rõhku
   - Ei, mulle pole koolis inglise keele häädlust õpetatud
   - Muu: ……
8. Kui Teile on koolis inglise keele häädlust õpetatud, siis millises kooliastmes (valige üks või mitu)?
   - Lasteaed
   - I kooliaste (1-3. klass)
   - II kooliaste (4-6. klass)
   - III kooliaste (7-9. klass)
   - Gümnaasium
   - Ei õpetatud koolis häädlust
   - Muu: ….....
9. Kui Te inglise keeles kõnelete, kas Te pöörate tähelepanu inglise keele korrektsele häädlusele (st. inglise keele häädlusnormidest kinni pidamist, aktsendi vältimine, jne.)?
   - Jah, õige häädlus on minu jaoks oluline
   - Enam-vähem, inglise keele õige häädlus ei ole minu jaoks kuigi tähtis
   - Ei, õige häädlus ei ole inglise keelt rääkides minu jaoks üldse oluline
10. Kas Teie arvates peaksid inglise keelt õppivad ja kõnelevad inimesed häädlusnormidest kinni pidama?
    - Jah, inglise keele kasutajad peaksid kasutama võimalikult korrektset häädlust
    - Otseselt mitte, kuid korrektse häälduse poole võiks pürgida
    - Ei, oluline on arusaadavus
    - Muu: ……
11. Mis on Teie jaoks inglise keeles kõnelemisel oluline?

……

Tänan vastuste eest!

Appendix 2 – Sample screenshot of the speaker window in Speech Recorder BAS



Appendix 3 – Students' VOT values summarized in tables

S#1					
Plosive p'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean**
Pig	0.035 s	0.025 s X	0.028 s	29.3 ms	35.6 ms
Pear	0.051 s	0.048 s	0.034 s	44.3 ms	
Pan	0.046 s	0.024 s X	0.030 s	33.3 ms	
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Tiger	0.060 s	0.062 s	0.043 s	55 ms	64.7 ms
Teabag	0.078 s	0.084 s	0.076 s	79.3 ms	
Toy	0.077 s	0.038 s	0.064 s	59.7 ms	
Plosive k'	I#1	I#2	I#3	Mean	**Plosive mean:**
Cat	0.040 s X	0.053 s	0.046 s	46.3 ms	42.2 ms
Kite	0.044 s	0.054 s	0.041 s X	46.3 ms	
Cup	0.036 s X	0.034 s X	0.032 s X	34 ms X	

S#2					
Plosive p'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Pig	0.045 s	0.057 s	0.053 s	51.7 ms	48.8 ms
Pear	0.009 s X	0.043 s	0.034 s	28.7 ms	
Pan	0.077 s	0.057 s	0.064 s	66 ms	
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Tiger	0.052 s	0.016 s X	0.064 s	44 ms	62 ms
Teabag	0.080 s	0.058 s	0.064 s	67.3 ms	
Toy	0.079 s	0.072 s	0.073 s	74.7 ms	
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Cat	0.047 s	0.063 s	0.070 s	60 ms	56 ms
Kite	0.065 s	0.039 s X	0.061 s	55 ms	
Cup	0.057 s	0.059 s	0.043 s	53 ms	

S#3					

Plosive p'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Pig	0.039 s	0.075 s	0.052 s	55.3 ms	66.2 ms
Pear	0.060 s	0.076 s	0.078 s	71.3 ms	
Pan	0.083 s	0.052 s	0.081 s	72 ms	
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Tiger	0.084 s	0.081 s	0.071 s	78.7 ms	82.8 ms
Teabag	0.064 s	0.081 s	0.085 s	76.7 ms	
Toy	0.103 s	0.087 s	0.089 s	93 ms	
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Cat	0.050 s	0.062 s	0.063 s	58.3 ms	75.4 ms
Kite	0.084 s	0.088 s	0.083 s	85 ms	
Cup	0.071 s	0.081 s	0.097 s	83 ms	

S#4					
Plosive p'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Pig	0.042 s	0.049 s	0.064 s	51.7 ms	60.8 ms
Pear	0.054 s	0.052 s	0.058 s	54.7 ms	
Pan	0.068 s	0.066 s	0.094 s	76 ms	
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Tiger	0.082 s	0.082 s	0.084 s	82.7 ms	85.6 ms
Teabag	0.080 s	0.076 s	0.071 s	75.7 ms	
Toy	0.108 s	0.094 s	0.093 s	98.3 ms	
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Cat	0.077 s	0.072 s	0.070 s	73 ms	70.4 ms
Kite	0.079 s	0.077 s	0.083 s	79.7 ms	
Cup	0.067 s	0.060 s	0.049 s	58.7 ms	

S#5					
Plosive p'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Pig	0.095 s	0.086 s	0.097 s	92.7 ms	92.8 ms
Pear	0.104 s	0.094 s	0.099 s	99 ms	
Pan	0.080 s	0.089 s	0.091 s	86.7 ms	
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Tiger	0.101 s	0.102 s	0.097 s	100 ms	103.4 ms
Teabag	0.093 s	0.097 s	0.095 s	95 ms	
Toy	0.115 s	0.123 s	0.108 s	115.3 ms	
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Cat	0.073 s	0.073 s	0.064 s	70 ms	76.1 ms
Kite	0.086 s	0.085 s	0.086 s	85.7 ms	
Cup	0.069 s	0.083 s	0.066 s	72.7 ms	

S#6					
Plosive p'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Pig	0.052 s	0.071 s	0.080 s	67.7 ms	81.2 ms
Pear	0.064 s	0.096 s	0.100 s	86.7 ms	
Pan	0.090 s	0.070 s	0.108 s	89.3 ms	
Plosive t'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**

Tiger	0.111 s	0.082 s	0.099 s	97.3 ms	107.6 ms
Teabag	0.107 s	0.106 s	0.106 s	106.3 ms	
Toy	0.144 s	0.101 s	0.112 s	119 ms	
Plosive k'	Instance 1	Instance 2	Instance 3	Mean	**Plosive mean:**
Cat	0.064 s	0.050 s	0.094 s	69.3 ms	68 ms
Kite	0.091 s	0.058 s	0.065 s	71.3 ms	
Cup	0.062 s	0.063 s	0.065 s	63.3 ms	

RESÜMEE

TARTU ÜLIKOOL
ANGLISTIKA OSAKOND

Liis Antons

The Aspiration of Voiceless Plosives by Estonian Learners of English: A Case Study
Eesti inglise keele õppijate helitute klusiilide aspiratsiooni hääldus: juhtumiuuring

Bakalaureusetöö
2023
Lehekülgede arv: 52

Annotatsioon:

Käesolev bakalaureusetöö uurib Eesti inglise keele õppijate helitute klusiilide aspiratsiooni. Töö eesmärk oli uurida kas ja kuidas Eesti inglise keele õpilased aspireerivad inglise helituid klusiile. Aspiratsiooni testimiseks koostas töö autor autorkõnekatse ja viis läbi akustilise analüüsi. Kuna inglise keele helituid klusiile on kuulmisanalüüsi põhjal Hallaste (2008) ja Aderi ja Miljani (2015) juba uurinud , sisaldas antud töö metoodika spektrograafi ja häälelaine analüüsi, mis mõõtis Eesti õpilaste helilisusviivet (VOT) ja aspiratsiooni olemasolu.

Töö sissejuhatus sätestab töö põhiprobleemi ning tutvustab põhimõisteid. Kirjanduse ülevaade selgitab lahti aspiratsiooni mõiste ja selle akustilisi parameetreid. Lisaks tutvustab kirjanduse ülevaade varasemaid uuringuid, mis käsitlevad klusiilide häälдust ja aspiratsiooni nii muudes keeltes kui ka Eesti keeles.

Analüütilise osa metoodika tutvustab kõnekatse ja taustaküsimustiku koostamist ning kõnekatses osalenud õpilaste värbamist ja värbamiskriteeriumeid. Metoodika osa seletab lahti inglise aspiratsiooni uurimiseks kasutatud parameetreid, milleks oli sobiv helilisusviibe mõõtmistulemus ning aperioodilise aspiratsiooni olemasolu. Analüüsi osa tutvustab kuue keskkooli õpilase kõnekatse tulemusi ning diskussiooni osa võrdleb neid tulemusi varasemate tulemustega.

Iga õpilase kohta oli 27 aspireeritud helitute klusiilide juhtumit ning kokku analüüsiti 162 aspiratsioonijuhtumit. 162st aspiratsiooni juhtumist häälдati 152 (93.8%) klusiili juhtumitest aspiratsiooniga. 8 (4.9%) juhtumit häälдati lühema aspiratsiooniga ning vaid 2 juhtumit (1.2%) jäid aspireerimata ning häälдati heliliste klusiilidena.

Märksõnad: inglise keele hääldus, võõrkeele hääldus, foneetika, akustiline analüüs, klusiilid, aspiratsioon, EFL kõnelejad

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Liis Antons,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

The Aspiration of Voiceless Plosives by Estonian Learners of English: A Case Study,

mille juhendaja on Reeli Torn-Leesik,

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace´i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Liis Antons

Tartus, 23.05.2023

Autorsuse kinnitus

Kinnitan, et olen koostanud käesoleva bakalaureusetöö ise ning toonud korrektselt välja teiste autorite panuse. Töö on koostatud lähtudes Tartu Ülikooli maailma keelte ja kultuuride kolledži anglistika osakonna bakalaureusetöö nõuetest ning on kooskõlas heade akadeemiliste tavadega.

Liis Antons

Tartus, 23.05.2023

Lõputöö on lubatud kaitsmisele.

Reeli Torn-Leesik

Tartus, 23.05.2023