Pit or phit? The production of English aspirated stops by speakers of Dutch
Broek, Merel van den

Citation

Leiden University

Pit or Pʰit? The production of English aspirated stops by speakers of Dutch

Thesis MA Linguistics: Applied Linguistics

Merel van den Broek, S3369684

1-7-2023

Under supervision of Dr. J. Caspers

Second reader Dr. E. D. Botma

Acknowledgements

First and foremost, I would like to thank the engineering company kind enough to give me a place to host my research and allowing their employees to take time out of their day to partake in this research.

Next, I would like to thank my parents for their unwavering belief in my abilities. My mom for showing that personal growth knows no bounds and it is never too late to become the person you need to be, despite everything and everyone. My dad for not only being the most hardworking and humble person I know, but also for showing that life may be tough, but that you can always be tougher (and for being such a cool dad that everyone is jealous of). I thank them both for raising me and making me who I am and how I am as a person today. I know I don't tell them enough so: mom and dad, I love you. You truly are the epitome of exceptional parents.

A special acknowledgement for one of my closest friends, Witteveen, arguably a rock amidst the wild currents of my life. Not only for the great friendship, nights in, and conversations, but also for the relentless (emotional and academic) support, calling me out on my bullshit and never ever losing faith in my capabilities. And for putting up with me of course.

Celeste and Willeke, my sincere gratitude for always being there when I needed to ask stupid questions and answering without judgement and for proofreading this thesis despite it being completely out of your academic fields. Your presence, both academically and socially, has profoundly shaped my growth as a person and an academic.

Thank you to Esther, who put up with my late nights, difficult questions, and the ups and downs of thesis writing. Your relentless support during the last 6.5 years have meant the world to me. I love you loads. To all my other friends, including Maarten, Nine, Marinda, Tessel, Moldir, Noor, Isa & Lars, I offer my heartfelt appreciation for making life be as fun and worthwhile as it can possibly be.

Lastly, I would like to thank myself. For both enduring and persisting. The past six years have not been an easy journey, but I am proud of the person I have become because of it all.

As a final note: I cannot overlook the insightful feedback from my girl scouts, whose candid appraisal described my research as "incredibly nit-picking and boring". You guys keep me humble.

To define is to limit.

Abstract

This thesis examines whether native speakers of Dutch can aspirate unvoiced word-initial plosives in English as a second language and if they are able to do so to the same degree as native speakers of English. Aspiration was measured in word-initial unvoiced plosives /p, t, k/ using Voice Onset Time (VOT). Influence of sex, age, age of onset of the second language, and self-reported general pronunciation ability in English were also examined. Native Dutch VOT generally lies between 0-20ms, while native English VOT is anywhere between 50-100 ms.

Using recordings from 19 participants who read both Dutch and English sentences out loud, it was found that participants increased their aspiration with an average of 10 ms when switching from Dutch to English. The analysis also shows that two participants out of 19 had an average VOT at a native-like level (>50 ms). Further investigation showed that sex, age, and age of onset did not influence VOT. Participants who graded their general pronunciation abilities with an 8 or higher did show a bigger increase in VOT, but there was no correlation to higher VOT in English when compared to other speakers. Other results showed that /p, t, k/ were not aspirated to the same extent, with /t/ being aspirated significantly more and /p/ significantly less. Lastly, the co-existence of words in both English and Dutch did not significantly alter the VOT, only when words started with /k/ did aspiration significantly improve when the exact word also existed in Dutch.

These results conclude that native speakers of Dutch increase aspiration and can aspirate to a native-like degree while speaking English. Any defining factors of the participants who were able to do so are not available due to the limited sample size.

Keywords: Aspiration, Dutch, English, Voice Onset Time (VOT), Cross-Linguistic Phonetic Interaction (CLI), plosives, Second Language Acquisition (SLA), pronunciation

Contents

List of Figures

List of Tables

List of abbreviations

AN   - Algemeen Nederlands, general Dutch pronunciation.

AoO   - Age of Onset

CLI   - Cross-Linguistic Phonemic Interaction

ENG   - English

F0   - Pitch contour

GA   - General American, pronunciation

IPA   - International Phonetic Alphabet

M   - Mean

ms   - Millisecond (0.001 second)

NED   - Dutch

RP   - Received Pronunciation, the general pronunciation of British English.

SE   - Standard Error

SLA   - Second Language Acquisition

VOT   - Voice Onset Time

1. Introduction

As of right now, the Dutch are among the world's most fluent non-native English speakers (EF English Proficiency Index, 2022). This is to be expected, considering there are English educational materials available for children from their first year of elementary school onwards, where most of the children are only four years old. English fluency rising is also due to 56.6% of current existing websites being in English (Usage Statistics and Market Share of Content Languages for Websites, 2023) and exposure to media significantly influencing English acquisition of non-native speakers (Muñoz, 2012; Pérez-Serrano et al., 2021). Frequent exposure to another language can cause learners to implement language traits into their own language learning process, often called Language Convergence or Incidental Learning (Al Zoubi, 2018). With the increasing growth of English fluency in the Netherlands, both through education and media exposure, a more specific light can be shed on proficiency on different levels. Most Dutch speakers' English communication goes through text, and very little through speech. Pronunciation is also a subject that is most often overlooked in acquiring a language, which makes it an engaging topic for further research.

Each language has its own setting, the general way in which the mouth is positioned during speech. The English setting is typically more relaxed, with relaxed tongue and throat muscles and loose lips, while Dutch has the opposite: tense tongue and throat muscles, tight lips, and throat constriction (Collins et al., 2018). The differences in the setting are one of the reasons why Dutch does not have aspiration while English does.[1] Aspiration is the short release of air after an unvoiced plosive (/p, t, k /) in syllable-initial position, which is also often referred to as an 'h-like sound'. An example of aspiration is in the English word 'pot', the aspiration makes the pronunciation be /pʰɒt/, where aspiration is indicated with a superscript 'h'. In Dutch, the same word is pronounced /pɒt/, without aspiration. Aspiration does not occur when /p, t, k/ is preceded by /s/ (such as in /spɒt/) or followed by /j, w, l, r/ (as in /plɒt/) (Gussenhoven & Broeders, 1997; Collins et al., 2018).

Up until this point, little to no research has been carried out to investigate the acquisition of aspiration in English from speakers of Dutch. Van den Doel (2006) argued that:

> […] the failure to aspirate […] may well give the impression of phoneme
> conflation to native speakers -aspiration being particularly important if RP is
> adopted as a model. […] Teachers of pronunciation should be aware that

---

[1] It is argued that in northern and eastern accents of Dutch (Twente, Groningen), aspiration does occur, but shorter and less intensely (Gussenhoven & Broeders, 1997).

approaches that do not go beyond the phoneme may in fact do learners a disservice […] (p. 225).[2]

Van den Doel highlights the importance of the accurate use of aspiration in English speech. More advanced pronunciation-related sources such as *English pronunciation for Student Teachers* (Gussenhoven & Broeders, 1997) and *Sounding Better, a Practical Guide to English Pronunciation for Speakers of Dutch* (Collins et al., 2018) highlight the importance of aspiration while acquiring a native-like pronunciation.

This thesis will dive into aspiration in English by native speakers of Dutch, analysing whether the average speaker of Dutch (with no specific English pronunciation education) aspirates in English and whether that happens close to a native-like level. Further sub-questions regarding age, age of onset (AoO), transfer, individual plosives and self-reported general pronunciation ability will also be discussed.

1.2 Overview
The second chapter contains a literature review and the research questions, chapter 3 follows up with the methodology, and chapter 4 the results and findings of the experiment. Chapter 5 will discuss these findings, and chapter 6 concludes this thesis.

---

[2] RP stands for Received Pronunciation and is the standard British English, also known as BBC English

2. Literature Review

Aspiration is the short puff of air which occurs in English after a word-initial unvoiced plosive is pronounced. Aspiration in the International Phonetic Alphabet (IPA) is portrayed by a superscript 'h' after the plosive (p^hit).

2.1 Measuring the Plosive

A plosive consists of three stages: closing, compression, and release. During the closing stage the articulators (the parts of the mouth responsible for the compression, e.g. the lips in case of /p/) are approaching each other, the compression stage indicates a state of compression, and the release of the burst indicates when the articulators are parting, which is ultimately what defines the sound of the plosive (Gussenhoven & Broeders, 1997). It is important to note that both the closing and compression stage are silent during unvoiced plosives.

2.1.1 VOT

What happens after the burst of the plosive indicates whether or not a plosive is aspirated. Voice Onset Time (VOT) is an acoustic measurement to analyse the time between release of the plosive (burst release) and the onset of voicing (vocal cord vibration). VOT was first established in 1964 by Lisker and Abramson who realised that only searching for an h-like sound on a spectrogram would provide inconsistent results. They found that VOT provided accurate and consistent measurements when measuring the stops in a stop + vowel arrangement. In their original research VOT also turned out to be accurate across different languages, including English and Dutch.

2.1.2 Praat

The VOT can be easily located with the use of audio analysis software such as Praat (Version 6.3.10, Boersma & Weenik, 2023). The release of the burst is where a plosive first starts to make a sound and the beginning of voicing. Voicing is characterized by formants (F0, F1, F2 etc.) which are concentrations of energy in specific frequency regions showing resonances in the vocal tract. The F0 value is also known as the pitch contour.

Conveniently, Praat has a built-in feature which can visualise the F0 with a blue line, thus helping to visualise the end of aspiration. In Figures 1 and 2 below, an example of a VOT measurement is shown. Figure 1 contains the English word "kind", while Figure 2 contains the Dutch word "kind" (child). Both words are spoken by the same native Dutch speaker. The F0 value is indicated by the blue line, thus creating the right-side border and indicating the end of

the VOT. The /k/ in the English noun has a VOT (indicated in red) of 110 ms, and the Dutch VOT is 20 ms.
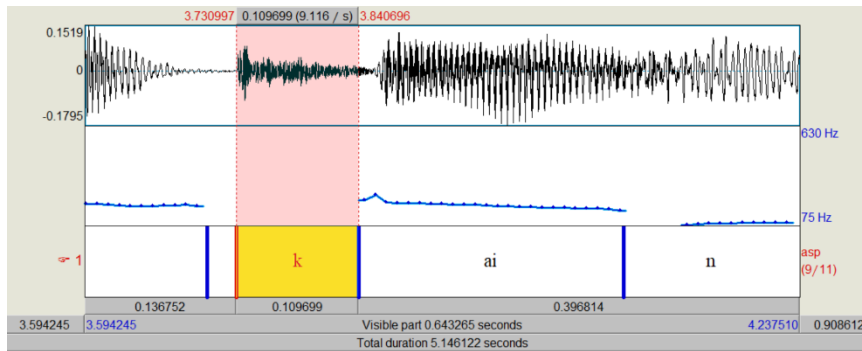
Figure 1: A screenshot of the English word "kind" in Praat. The highlighted red area shows the VOT of the /k/, which is 110 ms. The F0 is indicated by the horizontal blue line.
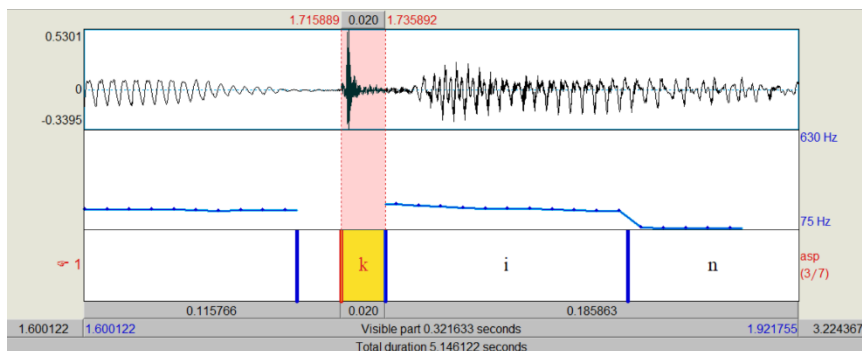
Figure 2: A screenshot of the Dutch word "kind" (child) in Praat. The highlighted red area shows the VOT of the /k/, which is 20 ms. The F0 is indicated the horizontal blue line.

2.1.3 Short and Long Lag

Measuring VOT has been established to be the most robust and efficient measurement of aspiration (Abramson & Whalen, 2017). The existence of VOT does not immediately guarantee the occurrence of aspiration. The measurement of VOT can be distinguished into three categories: prevoicing, short lag, and long lag. Prevoicing is the act of voicing before the release of the plosive, which causes a VOT to be negative. Prevoicing occurs both in voiced and unvoiced plosives in Dutch, whereas it does not occur at all in English. Lag is typically categorised as either short or long: a short lag VOT ranges from 0 to 25 ms, while long lag ranges from 60 to 100 ms and can also be referred to as aspiration (Ekelund, 2011; Stoehr et al., 2017). This categorisation and approach used by Ekelund and Stoehr et al. will be adopted during this research with the slight alteration of lowering the estimate to 50 ms based on the VOT averages in English (see p. 11).

The setting of Dutch and English causes these languages to employ different methods to differentiate between word-initial voiced and unvoiced stops. English differentiates these by aspiration (long lag), which postpones the voicing of the vowel in relation to the release of the burst (aspirating language), whereas Dutch employs prevoicing (voicing language) (Jansen, 2004).

The concept of lag has proven helpful to illustrate the behaviour of the plosives in Algemeen Nederlands (AN) (General Dutch) and English. While Dutch uses pre-voiced /b, d, g/ and short-lag /p, t, k/ stops, English differentiates short-lag /b, d, g/ and long-lag /p, t, k/ stops (Simon, 2010).

Figure 3 illustrates the three stages of a plosive and the corresponding VOT in Dutch and English. The black horizontal line signifies the complete plosive. The first area on the left indicates the closing stage and the first vertical line implies the moment of complete closure. The second area is the compression stage, and the third stage is the release of the plosive. The areas in blue are unvoiced and the yellow areas are voiced. It is important to note that in RP word-initial /p, t, k/ voicing only occurs after the plosive has ended. The VOT is measured from the burst of the plosive (right-side vertical line) and the beginning of voicing (colour switch from blue to yellow). The VOT in this figure is indicated by the red line, short lag in Dutch and long lag in English. Both /b, d, g/ in English and Dutch have no red line as they are prevoiced (negative VOT). [3]

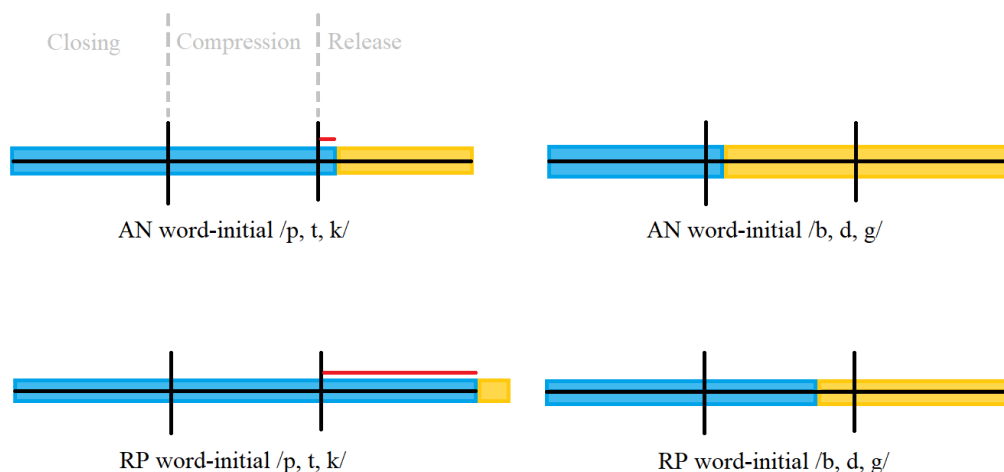

Figure 3: Visual representation of VOT. The horizontal line represents the full plosive, the blue areas are unvoiced while the yellow areas are voiced. The left vertical line represents the moment of complete closure, the right vertical line is the burst release. The red stripe is the VOT.

---

[3] /b, d, g/ in English are sometimes prevoiced (as in Figure 3) and sometimes receive short lag.

2.1.4 VOT per Language

The VOT duration of aspiration differs per language. For example, Korean differentiates between unaspirated and aspirated /p, t, k/, unaspirated with a general VOT of 20, 25 and 50 ms respectively while aspirated /p, t, k/ generally have a VOT of 90, 95 and 125 (Lisker & Abramson, 1964). The VOT will be used as a measurement to determine whether the plosives are aspirated or not. The VOT of word-initial unvoiced stops in Dutch typically ranges from 0-20 ms, as reported in several studies (van Alphen & McQueen, 2006; van der Feest, 2007; Simon, 2009; Simon, 2010). On the other hand, the VOT of word-initial unvoiced stops in English generally ranges from 50-100 ms (Gussenhoven & Broeders, 1997; Simon 2010; Ekelund, 2011; Riney et al., 2007). Generally, the /p/ has the shortest aspiration while /k/ has the longest (Lisker & Abramson, 1964; Winitz & LaRieviere, 1975; Byrd, 1993; Yao, 2009; Chodroff et al., 2017; Chodroff & Wilson, 2017).

Since the general average VOT differs per language before a plosive is considered aspirated, the average VOT in RP will be used as a comparison. Aspiration is most significant in word-initial plosives, so only word-initial plosives will be measured in this research.

2.1.5 Possible Complications regarding VOT

The method of measuring VOT was created by Lisker and Abramson in 1964 in an attempt to measure the difference between voiced and voiceless stops. In their paper, they also discussed three possible complications with VOT: individual differences, different languages and the overlap between aspiration and voicing.

The first problem Lisker & Abramson discussed was the influence of individual differences in articulation. This will not be a problem in this research as the speakers' VOT is measured in Dutch too. This provides the possibility to examine the difference in English and Dutch aspiration per person, rather than analysing individual performances against a referenced average.

The second problem concerns different languages, but as previously mentioned and also mentioned again by Abramson & Whalen (2017), English and Dutch are two languages perfectly suitable for using a VOT measurement. The third and final complication is the possibility of overlap between aspiration and voicing. Lisker and Abramson argue that this is not a big problem on its own as long as the decision whether to measure to the end of aspiration or to the beginning of voicing is made clear and stays consistent during the entire research. In this research, VOT is measured to the initiation of voicing, which coincides with the appearance of the pitch contour (F0).

F0 value, duration of voicelessness and amplitude have also been considered while analysing VOT but have been found to add unnecessary complexity to VOT and have little to no influence on the original simple version in English and Dutch (Abramson & Whalen, 2017).

Abramson & Whalen (2017) concluded that now, 60 years later, measuring VOT using spectrograms in relation to voicing still lives up to modern-day standards. All possible limitations and further improvements in using VOT as a measurement for stops are considered, yet they conclude that the original concept of VOT, measuring the left boundary (silence) and the right boundary (voiceless to voiced), "has proven to be a robust measure of the acoustic realization of consonantal voicing distinctions in most languages" (p. 13). These thorough reviews and assessments justify that using VOT to measure the degree of aspiration is the right choice for this research.

2.2 Interlanguage and transfer

The native language of an L2 speaker also significantly influences acquisition and performance. Linguistic errors made by second language (L2) speakers are referred to as the *interlanguage* of the speaker (Selinker & Rutherford, 2013). This interlanguage is influenced by the native language of the speakers and the targeted language. Interlanguage in terms of phonetics is dependent on the amount of phonetic contrast between different sounds (Gass et al., 2013). Similarities between the L1 and L2 of a speaker can also cause *transfer*, a term used when knowledge of the L1 is applied to the L2 with the presumption of it being identical. For instance, words common in both English and Dutch might showcase the difficulty speakers of Dutch may have with aspiration, since word-initial /p, t, k/ also occurs in Dutch, except without aspiration. Whether or not speakers of Dutch aspirate less in words that exist both in Dutch and English will also be examined in this research.

Stockwell and Bowen (1968) established a hierarchy of phonological difficulty, which predicts that learners of a language will have the most difficulty with (obligatory) phonemic contrasts which do not occur in their L1. Though aspiration is allophonic and not phonemic, it is a new concept to Dutch speakers which could influence the difficulty in acquisition. This holds true for unaspirated word-initial plosives in Dutch and both aspirated and unaspirated plosives in English. The non-existence of aspiration in Dutch makes aspiration one of the most challenging aspects to learn, possibly more so as aspiration can go unnoticed by learners. The phonetic interaction between languages is known as *Cross-Linguistic Phonetic Interaction* (CLI) (Schwartz, 2022).

Systematic variation further complicates the issue of aspiration in terms of SLA (Gass, 2013). Notably, aspiration only occurs in specific contexts and does not occur when preceded by /s/ or followed by /j, r, w, l/, which induce devoicing. This means that a learner would not only have to learn of the existence of aspiration but also its rules and exceptions.

2.2.1 Equivalence classification

The ability to distinguish between two phonetically similar sounds is a critical factor in accurate pronunciation. Not being able to do this is also known as *equivalence classification* (Flege, 1995; Flege & Bohn, 2021). When two sounds are too similar, L2 learners might initially perceive them as one sound, seemingly considering them the same. For speakers of Dutch, this would mean that they need to be able to distinguish aspirated and unaspirated initial plosives. This distinction can be either active knowledge (awareness of aspiration) or subconscious knowledge (considering it to be part of the accent, yet not knowing why or how). Failure to recognise the difference may lead the VOT to be indistinguishable in English and Dutch plosives, which can lead to inaccurate pronunciation.

A similar idea was proposed by Skehan (1989) who stated that one of the components essential for SLA is phonemic coding ability. The phonemic coding ability enables learners to differentiate between foreign sounds and remember them in a way in which they can recall them when necessary. This is crucial for learning and acquiring a near-native accent in the target language. Skehan rightfully highlights the fact that learners need to be able to differentiate between foreign sounds (such as aspiration) before they can incorporate them into their own speech. Recognising smaller aspects of a language will ultimately lead to better pronunciation abilities. Therefore, this research will also examine the correlation between speakers' self-assessed general pronunciation abilities in English compared to the degree of aspiration.

2.3 Acquisition and Age

The acquisition of an L2 before the age of seven has been found to significantly reduce the likelihood of developing an accent in that language (Loewenthal, 1981). As age increases, the probability of accentless L2 acquisition decreases significantly (Flege et al.,1995; Abrahamsson & Hyltenstam, 2009). An experiment by Abrahamsson & Hyltenstam found that out of 41 speakers who passed as native speakers in their L2, only three speakers passed as native when their accents were thoroughly examined by putting them through tests. These three speakers had an AoO of 3, 7 and 8 years old. The age at which nativeness seemed to decrease was at an average AoO of 12. Two different studies by Flege et al. (1995; 1999) found no Italian speakers

with English as their L2 and an AoO above 16 sounding like a native speaker, and similar for Korean speakers with English as their L2 and an AoO above 10. Furthermore, Sebastián Galles and Díaz (2012) state that: "[…] learners who are exposed early and/or have had extended exposure to a new language generally attain a better command of the L2 than those who are exposed late and/or have received less exposure." (p. 132). Hence, exploring the relationship between aspiration use and AoO would be of significant interest.

A rather unexpected challenge of acquiring aspiration is motor skills. While adults are generally faster learners than children during the early stages of SLA, they may encounter difficulty in adapting the fine motor skills necessary to implement new phonological features and these motor skills seem to decline rapidly. Moyer (1999) argues that this occurs because: "late learners may face neurological or motor skill constraints, such as entrenched articulatory habits or restricted perceptual targets for phonetic categories, that render the possibility of native-like attainment highly unlikely or impossible" (p. 82). Speech requires many fine motor skills, as accurately pronouncing phonemes unfamiliar to the native tongue often proves to be a great challenge, let alone allophones such as aspiration. Therefore, it is hypothesised that speakers who have started English education at a younger age will have less difficulties with aspiration than those who learned English at a later age.

Gass (2013) further states that "there is a general consensus that most older individuals cannot reasonably hope to ever achieve a native accent in an L2. There is no such consensus about other areas of language" (p. 436). The notion that perfect pronunciation can rarely be achieved by non-natives further proves that pronunciation, though often overlooked, may well be the hardest part of second language acquisition.

This research will examine both the influence of AoO and general age, since younger speakers graduated more recently, creating a shorter window between their English education and the moment of research.

2.4 Previous research on speakers of Dutch

Previous research has shown significant variations in aspiration usage among Flemish (a Dutch dialect) L2 learners of English. Simon (2004) found that one speaker never used aspiration in free speech, while another speaker aspirated 93% of the time. These findings demonstrate the variability that exists among L2 speakers. Simon focused on a yes/no approach on aspiration usage and had Flemish native speakers as subjects. The degree or duration of aspiration was not mentioned and therefore lacks specific data in the degree of aspiration. The paper displays the

great variability among speakers, yet the lack of any specific or statistically significant results makes it of limited use to this research.

2.5 Research Question

Based on the aforementioned information, this thesis aims to examine and measure the degree of aspiration used by native speakers of Dutch when speaking English. Aspiration in word-initial unvoiced plosives is quite reputable in English, and it would be noticeable if a speaker would not use it (Van den Doel, 2006). This thesis will measure the degree of aspiration used by average native speakers by measuring the VOT and compare them to the speech habits of native English speakers. These results will provide insight into the speech patterns of Dutch speakers and their ability to adapt to different sound patterns in English. The main research question is:

> Do average native speakers of Dutch acquire aspiration in unvoiced word-initial plosives in English?

There are a few sub-questions this thesis aims to answer apart from the main research question:

1. Can native speakers of Dutch aspirate to a native-like level (50-100 ms)?
2. Do speakers of Dutch aspirate less in words that exist both in English and Dutch?
3. Are English /p, t, k/ aspirated to the same extent?
4. Is there a correlation between age of onset and degree of aspiration?
5. Is there a correlation between age and the degree of aspiration?
6. Is there a relationship between the degree of English aspiration and self-reported general English pronunciation abilities of speakers of Dutch?

2.6 Hypotheses

The hypothesis for the main research question is that there is a significant amount of variability regarding native Dutch speakers' aspiration in English. It is hypothesised that there will be speakers who aspirate to a native extend and speakers who will not aspirate at all. The hypotheses for the previous sub-question are as follows:

1. Despite the complexity of acquiring a native-like accent, it is possible to aspirate to native-like level, but it will not be common for speakers with no specific English pronunciation training.

2. Based on the occurrence of transfer in speakers of an L2 language it is hypothesised that speakers of Dutch aspirate less when words occur both in Dutch and English.

3. Based on the aforementioned literature, it is hypothesised that the plosives do not receive equal aspiration, but are in line with previous findings that /p/ is aspirated least and /k/ the most (Lisker & Abramson, 1964; Winitz & LaRieviere, 1975; Byrd, 1993; Yao, 2009; Chodroff et al., 2017; Chodroff & Wilson, 2017).

4. Based on the findings of Flege et al. (1995) and Abrahamsson & Hyltenstam (2009) it is hypothesised that earlier age of onset will have a positive influence on the degree of aspiration.

5. Younger speakers generally aspirate more when speaking English based on the idea that they have had explicit English education more recently.

6. Speakers with better self-reported general English pronunciation will aspirate better when speaking English as aspiration is a key part in English pronunciation.

3. Methodology

This research is a quantitative study. The study collects data from native speakers of Dutch, including VOT measurements and relevant background information and uses graphs to display the results.

3.1 Participants

19 native Dutch speakers participated in this research in May 2023, among which nine women and ten men. The ages varied between 20 and 65, with an average of 40. The majority of the participants are employees at a Dutch engineering company based in the middle region of the Netherlands. All participants live in the Netherlands, in various places spread over the country and have Dutch as their native language. None of the speakers had ever had any type of English pronunciation education.

The participants were carefully chosen to create an accurate 'slice of society'. Though it is important to note that society will contain more outliers, such as students of English performing much better or Dutch speakers who do not speak English at all. All participants read and signed a consent form, as made mandatory by the University of Leiden.

3.2 Materials

The research consists of two parts, the survey, and the recordings. The participants were asked to answer a few questions (see section 3.2.1) to gain background information and factors which could possibly influence the results. After the survey, they were asked to read 27 sentences out loud: 12 Dutch (see Appendix B), and 15 English (see Appendix C) (see section 3.2.2 and 3.2.3). To eliminate intrapersonal differences, each subject in this research will have both their average Dutch and average English VOT measured, and the Dutch VOT averages provided by literature will only be used as general indications of normal performance.

3.2.1 Background Questions

The survey consisted of a few questions regarding the linguistic background of the speaker and possible other factors that could influence the results. The list of questions can be seen in appendix A but will be discussed here. Question 1 and 2 inquire about the age and gender of the speaker. Question 3 and 4 inquire about the current place of residence and place of residence during childhood to ensure that speakers were not from specific regions in which Dutch dialects might carry aspiration. Question 5 is a control question to ensure that the native language of the speaker is Dutch, and question 6 inquires about specific dialects spoken by the speaker or people close to them for the same reason as questions 3 and 4. Question 7 inquires about other

17

languages with high fluency that might influence the results. Question 8 inquires about the speakers' highest level of education.

Question 9 asks about AoO, and question 10 notes down the amount of exposure to intentional English instruction in school. Questions 10 and 11 inquire about exposure to English before adulthood (10) and in recent years (11). Question 13 inquires about any possible influences on their English abilities, such as English friends, their job, or their family. Question 14 and 15 ask the speakers to rate themselves on a scale of 1 to 10 on general English ability (14) and general pronunciation abilities (15).

3.2.2 Recordings

The recordings were created and saved (.wav) using a Sennheiser PC 131 head-mounted microphone to ensure quality and consistency. The sentences were recorded in Praat (version 6.3.10), with mono sound and 44100Hz and 32-bit resolution, which is the standard quality for CDs and enough to measure VOT accurately (Abramson & Whalen, 2017). All recordings were made in a quiet office with no (notable) echo. The researcher and participant were the only people in the office at the time of recording. All sentences were recorded in one long recording.

3.2.3 Sentences

The participants were asked to read 27 sentences in total. These sentences can be found in appendix B and C. Twelve are in Dutch (see appendix B), number 1 and 7 are filler sentences. The first filler sentence is also the first sentence they are asked to read to get the speakers accustomed to the situation. The second filler sentence is after the halfway point. These filler sentences are used as a neutral point between the target sentences to prevent possible bias or anticipation by the speakers.

The fifteen other sentences are in English (see appendix C) with four filler sentences: sentences 1, 6, 12 and 15. The first filler sentence is important because it is the switch between the Dutch and English sentences. The first sentence could be influenced by the Dutch setting of the mouth, so the irrelevancy of the sentence makes for a proper leeway. The sixth and twelfth sentences are used as neutral sentences between the target sentences. The last filler sentence is to prevent any pronunciation influences of it being the last sentence in the recording, e.g. possible rushed reading or reading it with a sigh. The filler sentences are in light grey and none of the filler sentences contain word-initial /p, t, k/. The Dutch sentences contain ten instances of /p, t, k/ as word-initial (thirty total per speaker). The English sentences contain 15 instances of word-initial /p, t/ and 16 instances of word-initial /k/ (46 total per speaker), this is because

the data on the English plosives is more relevant for the research and more data allows for an improved generalisation.

English sentences 2 and 8 are taken from pronunciation instruction guide *English Pronunciation for Student Teachers* by Gussenhoven & Broeders (1997). Sentences 3 and 11 are loosely inspired by practice sentences from *Sounding Better: A Practical Guide to English Pronunciation for Speakers of Dutch* by Collins et al. (2018). These sentences all contained a significant number of plosives and the level was moderate enough for participants less skilful in English to pronounce.

English sentences 2 and 3 are direct translations of the Dutch sentences for three reasons. First, these parallel sentences can be used as a controlling factor as these sentences have the same sentence structure. It helps to isolate the aspiration rather than being influenced by differences in sentence structures or content. The second reason is that it can provide a more direct comparison between the two languages, making space for a cross-linguistic analysis. The third reason is that these sentences had a translation which worked in both languages while keeping the general difficulty level low and still including the required phonemes. This was not possible to do for all sentences as it would create too much repetition.

The use of /p, t, k/ is equally spread throughout the sentences, and special attention was paid to avoid abundant alliteration, something which could influence the outcome. The stop-initial words occur in every position in the sentence in an attempt to mimic natural speech and naturally occurring stress patterns in both Dutch and English.

All sentences were written with a speaker with moderate proficiency in English in mind, making them accessible to everyone who participated in this research. The participants were not aware they had to read English sentences out loud before they sat down to participate. This was done on purpose to ensure participants would not refrain from participating based on the idea that their English was not good enough. Out of all speakers, only one speaker had mispronounced a word (peas).

To address sub-question three, the words "post", "pen", "test", "taxi", "cool", "kind"[4] (meaning child in Dutch), and "cake" were added in the English sentences. This means that per plosive, 2 words (3 in case of /k/) could be compared to the other 13 (see Appendix E).

---

[4] The word "kind" was considered after the sentences were written, as the realisation set in that this word (despite having a very different pronunciation) could still cause recognition for speakers of Dutch, hence the reason there are three words starting with /k/.

3.3 Data Analysis

The data analysis of this research consists of two parts: measuring the VOT and the statistical analysis of those measurements.

3.3.1 Measuring the VOT

As previously stated, this research will consider the beginning of voicing (F0 appearance) the end of aspiration. The software Praat will be used to measure the VOT, as it is one of the most widely used audio analysis software to date.

VOT measurement could be done manually or through a pre-made script. Abramson and Whalen (2017) analysed over seven different pre-made scripts. The system created by Keshet, Sonderegger and Knowles (2014) is the most commonly used and is reported to have an accuracy of around 80% agreement in automated vs manual measurements in within 5 ms and 90% at 10 ms. Their system was used in research regarding Scottish English (Stuart-Smith, et al., 2015) which resulted in 62.6% accuracy. 15.8% had to be corrected by hand and 21.6% of the measurements were not usable (p. 518). Because of the success rate and the limited amount of input, the choice was made to measure the VOT manually.

Manual measuring allows for a more detailed analysis as the researcher can inspect every aspect of the waveform and spectrogram and create consistency, something an automatic script might miss. Measuring by hand also creates the opportunity for the researcher to manually adjust if the circumstances are suboptimal or if background noise influences the original measurement. Manual measuring also adds a level of quality control. Speech is fluent and manually selecting the boundaries creates consistency and validity across all data points because each VOT is checked manually. Out of 1450 data points, 16 were found to be unfit to include in the results, reasons include a complete lack of voice, creaky voice (which has an unstable F0) and possible background noise created by the participant such as moving or pushing the papers around on the table.

3.3.2 Statistical Analysis

The VOT durations of each speaker will be split up into the /p, t, k/ categories for both English and Dutch to be able to analyse their average VOT per plosive in either language. The English averages will be compared to the standard range of aspiration (50-100 ms) in English to get an image of how Dutch speakers perform. The difference in VOT when switching from Dutch to English will also be calculated and used for statistical analysis. Correlation tests will be used to compare the average Dutch VOT to the average English VOT, and the English VOT to the participants' proficiency levels. Further correlation tests between the average difference in VOT

averages between Dutch and English and secondary data such as age, sex, self-reported pronunciation abilities, and time of first exposure to English education were used to answer the sub-questions (see Chapter 4).

These statistical tests will use the data to generate results to answer the research questions and provide insight into native Dutch speakers' aspiration acquisition.

4. Results

Table 1 and Figure 4 are presented in milliseconds (ms), the standard unit of VOT measurement. Representing the data in milliseconds instead of seconds allows for a more straightforward interpretation of results. The main research question was answered using all data points, while the sub-questions are answered using each speaker's average (see Appendix D). The overall difference in VOT (average English VOT – average Dutch VOT) when comparing Dutch to English is also used in the statistical analysis, these numbers can also be found in Appendix D. Using the average value provides a representative value which represents the central tendency of the data and minimizes the influence of outliers or extreme values.

4.1 Overall Results

Table 1 depicts the mean, standard deviation, and standard error of the VOT in ms, divided by language and by plosive and the average (AVG) of the Dutch and English plosives.

		Mean	**Standard Deviation**	**Standard Error**	**N**
Dutch	**P**	22.41	8.587	1.970	187
	T	29.37	6.481	1.487	190
	K	28.36	5.392	1.237	189
	AVG	26.72	5.093	1.168	566
English	**P**	26.56	11.190	2.730	280
	T	44.46	11.262	2.584	281
	K	39.92	9.925	2.277	299
	AVG	37.06	9.588	2.200	860

Table 1: Mean, SD, SE, and N value of the VOT in ms, separated by language and plosive.

Figure 4 shows a boxplot comparing the values of the Dutch and English plosives in comparison to each other. The /p/ is shown in blue, /k/ in red and /t/ in green. The darker shade indicates the English VOT, while the lighter shade indicates the Dutch VOT.

To answer the main research question, the findings indicate that, on average, participants aspirated 10 ms longer in English (M=37.06, SE=2.200) than they did in Dutch (M=26.72, SE=1.168). A paired-samples t-test revealed that this difference is significant, $t(18)=4.679$, $p < .0005$. This shows that speakers generally have a longer VOT in English. Out of 19 participants, two participants achieved a VOT average at a native-like level (>50 ms) (see Appendix D), answering sub-question 1.
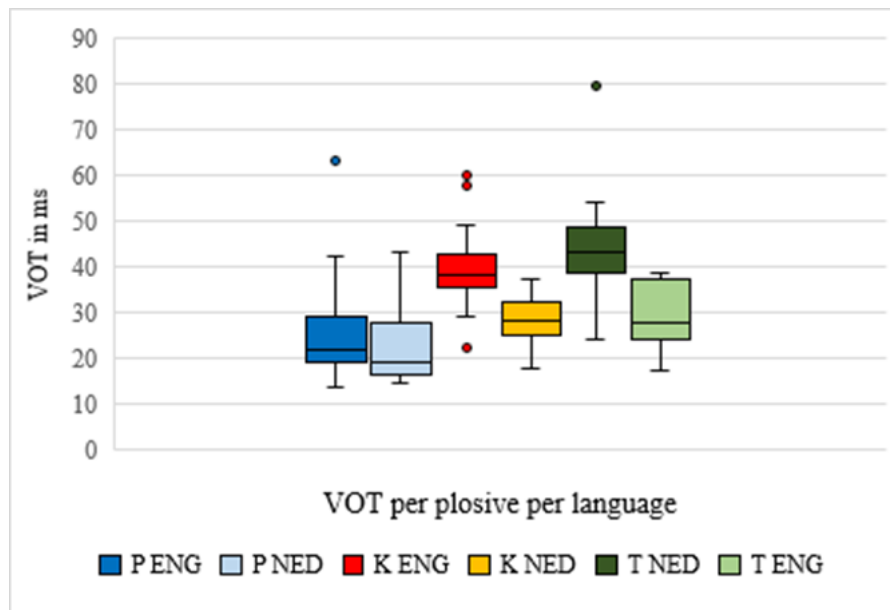
Figure 4: Boxplot of the VOT in ms values by plosive and language.

To answer the main research question, the findings indicate that, on average, participants aspirated more in English, which can also be seen in the boxplot in Figure 4. While there were two speakers with a native-level VOT, there were also two speakers (see Appendix D) whose VOT actually decreased while switching from Dutch to English. This confirms the hypothesis that there is a great variability among speakers in terms of aspiration and VOT duration.

4.2 Individual Plosives

The second sub-question, whether words that coexist in Dutch are aspirated less than words that only exist in English, was answered using paired samples t-tests. To answer this question, the words "post", "pen", "test", "taxi", "cool", "kind" (meaning child in Dutch), and "cake" were considered in the English sentences. The averages of these words can be found in Appendix E and were compared against the general average of the corresponding plosives in a paired samples t-test.

On average, words that start with /p/ which also coexist in Dutch (M=26.8, SE=2.5) are more aspirated than words that start with /p/ that only exist in English (M=26.6, SE=2.7). This difference is not significant, t(18)=0.209, p=0.836.

On average, words that start with /t/ which also coexist in Dutch (M=42.6, SE=3.3) are less aspirated than words that start with /t/ that only exist in English (M=44.5, SE=2.58). This difference is not significant, t(18)=-0.679, p=0.506.

On average, words that start with /k/ which also coexist in Dutch (M=43.9, SE=2.5) are more aspirated than words that start with /k/ that only exist in English (M=39.9, SE=2.3). This is significant t(18)=2.604, p=0.018. These results reject the hypothesis that words that co-exist in English and Dutch are less aspirated. Only words that also coexist in Dutch starting with /k/ are pronounced significantly different than the words which only exist in English starting with /k/ and they are more aspirated instead of less, which was hypothesised. The differences between the /p/ words and /t/ words are not statistically significant, also rejecting the hypothesis.

A paired samples t-test was carried out to answer the third sub-question. The means of the English VOTs of /p, t, k/ were compared to each other. All values had a significant difference between them. Table 2 depicts the results of the paired samples t-test. The results indicate that all the VOT of all plosives are significantly different. /p/ and /k/ had the smallest difference between them but are still significantly different. This confirms the hypothesis that /p, t, k/ are not all aspirated to the same extent.

Plosive	Compared to	p-value
p (M=26.6, SE=2.73)	t	.000
t (M=44.5, SE=2.58)	k	.047
k (M=39.9, SE=2.28)	p	.000

Table 2: Results of a paired samples t-test comparison of the VOT of /p, t, k/, showing the M, SE, and p-value per compared pair.

4.3 Individual Speakers

The fourth sub-question is answered by running a Pearson correlation test to see whether AoO has an influence on general VOT duration and VOT improvement. There was no significant relationship between VOT difference and AoO (r=-0.400, N=19, p=0.090). There was also no correlation between average English VOT and AoO (r=-0.419, N=19, p=0.074). This rejects the hypothesis that a younger AoO would positively influence either general VOT duration in English or the difference between Dutch and English VOT. Though it is important to note that there is a trend (r=-419), so these results might be statistically significant if there were more speakers and more variability among speakers.

The fifth sub-question can be answered by running a Pearson correlation to determine the relationship between age and difference in VOT in English and Dutch. There was no

significant correlation between age and difference in VOT (r=-0.255, N=19, p=0.292). There was also no correlation between age and average VOT in English (r=-0.237, N=19, p=0.328), which rejects the hypothesis that young people would aspirate more or have a larger difference in English and Dutch VOT when speaking English.

A Spearman's rank correlation was used to assess the relationship between the difference in VOT and self-reported pronunciation abilities and between average VOT in English and self-reported pronunciation abilities for sub-question six. The average VOT difference of the four highest-rated and four lowest-rated participants was used (see Appendix F). The four lowest-rated participants all assessed their pronunciation abilities with a 5 or lower, while the four highest-rated participants all assessed their abilities with an 8 or higher.

There was a positive correlation between the difference in VOT and self-reported pronunciation ability, r(19)=0.592, p=0.008. There was no correlation between average VOT in English and self-reported pronunciation ability, r(19)=0.222, p=0.361.

These results partially confirm the hypothesis, while participants with a higher self-reported pronunciation ability do not have a longer VOT in general, they do have a bigger difference in their VOT when comparing their Dutch VOT to their English VOT.

4.4 Additional Results

The additional information from the background questions were also analyzed to determine whether other factors influenced the results.

An independent t-test was carried out to see if sex had any effect on VOT difference. There was no significant effect for sex t(19)=1.785, p=0.092, despite women (M=14.18, SE=3.9) having a difference almost twice as big in their VOT between Dutch and English than men (M=6.59, SE=2.0).

Overall, there were no speakers with birthplaces, current residence or contact with dialects that could influence the findings. Since their educational background was also roughly similar, this was not analysed in the statistic section as it would not yield any relevant or generalising results.

The question regarding exposure to English, both before ending puberty and current exposure, was answered with such variability that it was unattainable to create a scale or any other method of grouping in order to be able to generalise it for statistical analysis. While it can still explain results when referring to a speaker in specific, this factor combined with other remarks regarding English exposure were unfit to be generalised or used for statistical tests across all speakers.

5. Discussion

The results of this research have provided valuable insight into the English VOT of speakers of Dutch. The sections below will provide a comprehensive analysis of the research findings. Additionally, the limitations in the study design and data collection will be addressed, as well as highlighting areas that could be further refined or explored in future research.

5.1 Overall Discussion

The results indicate that speakers of Dutch aspirate more when speaking English (M=37.06, SE=2.200) than when speaking Dutch (M=26.72, SE=1.168). Out of 19 participants, two participants reached the native-like aspiration threshold of >50 ms. One of these participants had worked in a company where English was the main language for 15 years, and the other participant is currently studying in an international environment in which English is the main language. Though it is important to mention that other participants with similar exposure to English did not achieve higher VOT. Two speakers actually decreased their VOT when switching to English, which is in line with the main hypothesis that there is a significant amount of variability between speakers. These results also stand in line with the results of Simon (2004).

These results are in line with the hypothesis of sub-question one, which stated that speakers of Dutch can achieve native-like aspiration, but the notion that only two out of the 19 participants were able to do so indicates that aspiration still poses a challenge for non-native speakers, even when exposed to English more extensively.

Another significant finding is the average VOT in Dutch being 26.72 ms, contradicting literature that the average VOT in Dutch is between 0 and 20 ms (van Alphen & McQueen, 2006; van der Feest, 2007; Simon, 2009; Simon, 2010). A possible explanation could be that the recordings created for this research were not free speech, and pronunciation was perhaps more emphasised than it would be in free speech.

Including the analysis of the Dutch VOT has proven to be the right choice with significant benefits, considering the results would be significantly different if it was just assumed that the average VOT in Dutch would be in between 0 and 20 ms.

Despite these findings answering the main research question, it is important to acknowledge the limitations. The small sample size limits the generalizability, especially for the two participants who reached native-like levels with their VOTs. Because there are only two of them, it is nearly impossible to create a general explanation as to why these participants reached a native-like level and the others with similar backgrounds did not. Still, the main

research question was successfully addressed and despite the scope of its limitations, the results have provided a meaningful answer to the research question.

5.2 Individual Plosives

The results for sub-question 2 demonstrate that words which co-exist both in Dutch and English do not receive less aspiration than words that only exist in English. For both /p/ and /t/ the difference in aspiration was not statistically significant, and for /k/ it even had the opposite effect, making it significantly more aspirated than the other English words. These results reject the hypothesis that co-existing words would be less aspirated. The recognition of Dutch words did not trigger the participants into aspirating less, demonstrating that equivalence aspiration and cross-linguistic phonetic interaction are not as prevalent regarding aspiration.

The results from sub-question 3 indicate that /p, t, k/ are not aspirated to the same extent. The plosives were all pronounced significantly different from each other. These results are not in line with previous research, in which native speakers were analysed and the order (shortest VOT to longest) was /p, t, k/, while in these findings it is /p, k, t/ (Lisker & Abramson, 1964; Winitz & LaRieviere, 1975; Byrd, 1993; Yao, 2009; Chodroff et al., 2015; Chodroff & Wilson, 2017). It might indicate that aspiration for /t/ (M=44.5 ms, SE=2.58) comes more naturally to speakers of Dutch compared to /k/ (M=39.9, SE=2.28) and that /p/ (M=26.6, SE=2.73) feels the most unfamiliar to aspirate.

5.3 Individual Speakers

The results from sub-question 4 demonstrate that there is no correlation between AoO and VOT duration and difference in VOT. Similar to the data from sub-question 5 which demonstrate that there is also no significant correlation between age and average VOT in English and average difference in VOT. The results from this research disprove the hypothesis that younger people would aspirate more than older people and that AoO has an influence on aspiration. Yet, as previously mentioned, there is a visible trend and the data is nearing significance, so these results might be influenced by the limited number of participants.

These results slightly contradict Flege et al. (1995) and Loewenthal (1981), who stated that the probability of being able to speak an L2 without an accent decreases after an AoO above 7. The only participant who was taught English before the age of 7 (6) had an average VOT of 48 ms in English, coming close to a native-like level. Both Gass (2013) and Moyer (1999) argued that the odds of late learners ever achieving a "native" accent are generally deemed impossible. General age (younger speakers having a closer proximity to English education) did

in the end not have a significant influence on performance, rejecting both the hypotheses for sub-questions 4 and 5. These results might also be influenced by the limited variability and number of participants.

The results from the sixth sub-question demonstrate that higher self-reported pronunciation ability does significantly predict a difference in average VOT but does not necessarily predict the average English VOT on its own. Participants with a higher self-reported pronunciation ability were probably more aware of their accent in English and aspects they could change to improve it, creating a bigger difference in their VOT, yet they had no significant VOT duration longer than the other participants. It is noteworthy that neither of the two participants with native-like aspiration were in the group of participants with a high self-reported pronunciation ability.

5.4 Further Research

Further research in the field should aim to also include participants from Friesland and the eastern regions of the Netherlands. As previously mentioned, northern and eastern dialects do have a shorter version of aspiration and speakers surrounded by Frisian may be influenced as well. None of the participants in this research were from that region so that influence could not be analysed. It can also further generalise the findings and investigate potential regional variations.

A larger sample size would also provide an opportunity to examine the influences of education, AoO, and the characteristics of those who do achieve native-level aspiration. The sample size used for this research is not large enough to make claims about these factors.

6. Conclusion

This research aimed to answer the question of whether speakers of Dutch could aspirate unvoiced word-initial plosives. The results of this research and its data show that native speakers of Dutch can do so while speaking English. The usage of aspiration was recorded in both Dutch and English sentences for word-initial unvoiced plosives /p, t, k/ and the Voice Onset Time (VOT) was analysed for each language. The averages of this data showed that some participants were even able to have a VOT average above 50 ms, which is the threshold for English native-like aspiration. In general, participants had their VOT duration increased with 10 ms after switching from Dutch to English.

Further research questions showed that the co-existence of words in both Dutch and English did not influence aspiration usage except in words beginning with /k/, in which the coexistence improved aspiration. While age and AoO also did not have a significant effect on aspiration in this research, it is likely that the limited variability and number of speakers influenced this outcome, since both factors were nearing statistical significance and showed a favourable trend towards younger age and younger AoO.

Plosives /p, t, k/ are not all aspirated to the same extent, /t/ is significantly more aspirated than /p/ and /k/, and /p/ is statistically less aspirated than /t/ and /k/, something that is contradictory to native speakers of English, as they aspirate the /t/ more than the /k/ (Lisker & Abramson, 1964; Winitz & LaRieviere, 1975; Byrd, 1993; Yao, 2009; Chodroff et al., 2017; Chodroff & Wilson, 2017). Lastly, participants with better self-reported general pronunciation abilities did show a significantly larger difference in their VOT between English and Dutch but did not have a significantly longer VOT than the other participants.

Despite the sample size and corresponding limitations, these findings still successfully addressed the main research question. As northern and eastern dialects of the Netherlands can carry aspiration, further research would be necessary to address these potential variations.

This research is one of the first to explore English aspiration by native speakers of Dutch. These findings fill an academic gap in the literature regarding aspiration, while also examining surrounding factors such as age, age of onset, English pronunciation abilities in general and the possible influence of transfer. Moving forward, future research can build upon these findings addressing the remaining questions and improving the generalisability of these results.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language learning*, *59*(2), 249-306.

Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of phonetics*, *63*, 75-86.

Al Zoubi, S. M. (2018). The impact of exposure to English language on language acquisition. *Journal of Applied Linguistics and Language Research*, 5(4), 151-162.

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.10, retrieved 3 May 2023 from http://www.praat.org/

Byrd, D. (1993). 54,000 American stops. *UCLA working papers in phonetics*, *83*, 97-116.

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics*, *61*, 30-47.

Chodroff, E., Godfrey, J., Khudanpur, S., & Wilson, C. (2015). Structured variability in acoustic realization: a corpus study of voice onset time in American English stops. In *ICPhS*.

Collins, B. S., Hollander, S. D., Mees, I. M., & Rodd, J. (2018). *Sounding better: A practical guide to English pronunciation for speakers of Dutch*. Walvaboek, Holten.

EF English Proficiency Index. (2022). In *https://www.ef.com/wwen/epi/*. Education First. Retrieved May 15, 2023, from https://www.ef.com/wwen/

Ekelund, M. (2011). Aspiration in Japanese Speakers' English: A study of the acquisition of new phonetic categories in a second language [BA Dissertation]. Stockholms Universitet Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-58417.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E. (1999). Age of learning and second language speech. In *Second language acquisition and the critical period hypothesis* (pp. 111-142). Routledge.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). *Second language speech learning: Theoretical and empirical progress*, *10*(9781108886901.002).

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*(5), 3125–3134. https://doi.org/10.1121/1.413041

Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second Language Acquisition: An Introductory Course*. Routledge, New York.

Gussenhoven, C., & Broeders, A. (1997). *English Pronunciation for Student Teachers*. Noordhoff Uitgevers Groningen.

Jansen, W. (2004). Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch. Doctoral dissertation, Groningen Dissertations in Linguistics.

Keshet, J., Sonderegger, M., Knowles, T. AutoVOT, v 0.93. 2014. https://github.com/mlml/autovot/

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384-422.

Loewenthal, S. T. (1981). Foreign accents: factors relating to transfer of accent from the first to a second language. *Language and speech* (24 part 3), 256-272.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in second language acquisition*, 21(1), 81-108.

Muñoz, C. (2012). The Significance of Intensive Exposure as a Turning Point in Learners' Histories. In M. Matter (Red.), *Intensive exposure experiences in second language learning,* 65.

Pérez-Serrano M, Nogueroles-López M, Duñabeitia J.A. (2021). Incidental Vocabulary Learning with Subtitles in a New Language: Orthographic Markedness and Number of Exposures. *Plos One* 16(2):0246933.

Riney, T. J., Takagi, N., Ota, K., & Uchida, Y. (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics*, *35*(3), 439-443.

Schwartz, G. (2022). Asymmetrical cross-language phonetic interaction: Phonological implications. *Linguistic Approaches to Bilingualism*, *12*(2), 103-132.

Sebastián-Gallés, N., & Díaz, B. (2012). First and second language speech perception: Graded learning. *Language Learning*, *62*, 131-147.

Selinker, L., & Rutherford, W. E. (2013). *Rediscovering interlanguage*. Routledge, New York.

Simon, E. (2004). De verwerving van het Engelse stemcontrast door moedertaalsprekers van het Nederlands. *Handelingen-Koninklijke Zuid-Nederlandse maatschappij voor taal-en letterkunde en geschiedenis*, *58,* 87-104.

Simon, E. (2009). Acquiring a new second language contrast: An analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research*, *25*(3), 377-408.

Simon, E. (2010). Child L2 development: A longitudinal case study on Voice Onset Times in word-initial stops. *Journal of child language*, *37*(1), 159-173.

Skehan, P. (1989) *Individual differences in second-language learning*. London: Edward Arnold.

Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1968). The sounds of English and Spanish, University of Chicago Press, London & Chicago.

Stoehr, A., Benders, T., van Hell, J. G., & Fikkert, P. (2017). Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Language Research*, *33*(4), 483–518. https://www.jstor.org/stable/26375898

Stuart-Smith J, Sonderegger M, Rathcke T, Macdonald R. The private life of stops: VOT in a real-time corpus of spontaneous Glaswegian. *Laboratory Phonology*. 2015; 6:505–549.

Usage Statistics and Market Share of Content Languages for Websites. (2023). W3Techs. Retrieved May 15, 2023, from https://w3techs.com/technologies/overview/content_language

Van Alphen, P. M., & McQueen, J. M. (2006). The effect of voice onset time differences on lexical access in dutch. *Journal of experimental psychology: Human perception and performance*, *32*(1), 178.

Van den Doel, W. Z. (2006). *How Friendly are the Natives? An Evaluation of Native-speaker Judgements of Foreign-accented British and American English* [PhD Dissertation]. Universiteit Utrecht.

Van der Feest, S. V. (2007). *Building a phonological lexicon: The acquisition of the Dutch voicing contrast in perception and production*. [PhD Dissertation]. Radboud University, Nijmegen.

Winitz, H., LaRiviere, C., & Herriman, E. (1975). Variations in VOT for English initial stops. *Journal of Phonetics*, *3*(1), 41-52.

Yao, Y. (2009). Understanding VOT variation in spontaneous speech. *UC Berkeley PhonLab Annual Report*, *5*(5).

Appendix

A: Background Questions

1. Age.

2. Gender.

3. Current place of residence.

4. Place(s) of residence during childhood (<18).

5. Mother tongue.

6. Possible dialects spoken by either the participant or in the participants direct surroundings, now and in the past.

7. Other languages they speak either better/as well or more often than English (apart from Dutch).

8. Highest level of education.

9. The age at which they received their first English education.

10. Exposure to English in school under the age of 18 in hours.

11. Exposure to English after and outside of school (none, a little, moderate, a lot, very much).

12. Current exposure to English (none, a little, moderate, a lot, very much).

13. Any possible further influences regarding English (such as English friends, working for an English company, English tertiary education or other significant exposures to English)

14. Self-reported English skill on a scale of 1-10

15. Self-reported English pronunciation abilities on a scale of 1-10

B: Dutch Sentences

1.   Gisteren heb ik mijn werk niet meer afgemaakt.

2.   De prijs van de **k**offie en **t**hee is gestegen.

3.   **K**an je een **p**aar **p**akketjes naar de **p**ostbode brengen?

4.   Het **t**oppunt van de vakantie waren de **t**ulpen.

5.   Ik neem een **p**aarse **t**axi naar **K**aatsheuvel.

6.   **P**eter **k**opte de bal recht langs de **k**eeper.

7.   De mooiste bloem uit de achtertuin is een madeliefje.

8.   De Griekse **t**empel lag midden in de hoofdstad.

9.   Er is een **k**ans dat **K**arin de **p**etitie **t**egelijk had gemaakt.

10.   De **p**ater **k**auwde met veel **t**egenzin op de **k**oude **p**izza.

11.   De **t**egenspraak kwam al **t**ijdens de speech van de **p**aus.

12.   De **p**iraat knalde met zijn boot **t**egen de **k**ade.

C: English Sentences

1. The next generation of children does not like eating healthy foods.

2. They've **p**ut up the price of **c**offee and **t**ea.

3. **C**an you **c**arry a **c**ouple of **p**ackages to the **p**ost office?

4. He forgot **t**o bring a **p**en **t**o the **t**est.

5. I will **t**ake a **t**axi with my **c**ool friend **P**ete.

6. I walked near the river yesterday.

7. **K**evin said it was **t**ime for **c**ookies.

8. How **k**ind of **P**aul to **c**ome to **C**ambridge next week.

9. I **p**aid five **p**ounds for the book.

10. It is **p**ossible that **K**ennedy wrote the **t**errible **p**etition.

11. A **p**acket of **p**eas is **k**ind of expensive **t**oday.

12. Jessica's brother baked a nice lunch for all of us.

13. **P**olly **t**alked **t**oo much about **c**arrot **c**ake.

14. The **p**olice **c**ar **p**arked next **t**o the **c**urb.

15. It feels like the years are going by much faster.

D: Participant Averages

This table shows the VOT average per participant in ms. The first three columns (skipping the speaker ID) show average VOT per plosive in English, same for the following three but in Dutch. Followed by their average VOT in English and Dutch, and the difference (VOT increase) between them.

VOT averages that are of native level (>0.050) are marked in light blue. Speakers of whom the VOT actually decreased when speaking English are marked in green.

ID	P ENG	T ENG	K ENG	P NED	T NED	K NED	AVG ENG	AVG NED	VOT DIF.
401	24.9	24.2	38.4	41.4	23.8	23.5	29.3	29.6	-030
402	63.3	53.9	58	32.7	37.4	25	58.4	31.7	26.7
403	21.9	30.7	29.6	16.4	24.4	28	27.5	22.9	04.6
404	34.5	39.3	40.1	43.2	37.4	34.4	38.1	38.3	-00.2
405	41.6	79.5	57.6	17.3	37.7	26.2	59.9	27.1	32.8
406	29.3	43.1	38.6	28.9	38.8	37.5	37	35.3	01.7
407	20.7	42.1	42.6	27.8	34.3	25.6	35.3	29.2	06.1
408	21.1	41.6	37	16.6	28.6	28.4	33.2	24.5	08.7
409	13.7	35.2	32.3	14.8	22.4	35.8	27.2	24.3	02.9
410	21.1	38.5	37	16.5	33.1	32.2	32.3	27.3	05.0
411	19.2	49.8	35.3	22.3	37.2	31.9	34.8	30.5	04.3
412	42.5	52.9	49	19.2	25.8	27.8	48	24.3	23.7
413	22.5	43.9	29.4	19.4	22	21.6	31.9	21	10.9
414	29.4	48.7	59.9	15.6	25.7	30.9	46.6	24.1	22.5
415	26.2	46.4	39.1	17.9	24.9	17.7	37.3	20.2	17.1
416	18.8	41.8	22.4	14.7	17.3	19.5	27.2	17.2	10.0
417	16.7	47.2	38.1	17.6	32.5	31.8	34.4	27.3	07.1
418	16.9	48.5	37.8	21.6	26.8	32.5	34.5	27	07.5
419	20.3	37.6	36.4	22	28	28.6	31.3	26.2	05.1

E: Averages of Words that co-exist in Dutch

VOT averages in ms of words that co-exist in Dutch in ms. Cells with a '-' indicate that the

VOT could not be measured and is therefore not taken into account with statistical analysis.

Post	Pen	Test	Taxi	Cool	Kind	Cake
15	27	29	31	41	21	51
42	61	79	43	59	81	4
22	13	84	22	26	37	28
46	34	38	62	6	29	41
23	6	84	49	8	4	8
2	17	35	46	28	48	45
24	18	23	32	66	05	33
23	2	34	33	88	3	-
07	23	2	34	32	49	12
16	21	27	49	44	29	46
24	22	53	41	79	25	41
47	44	57	61	46	5	55
23	2	32	22	42	32	46
34	16	47	36	62	59	46
34	4	37	56	46	55	21
21	23	29	13	19	15	36
21	18	27	39	67	32	35
2	21	66	74	53	4	42
26	34	37	37	42	46	-

F: Average English VOT, VOT difference, and self-reported Pronunciation of Participants
The average English VOT in seconds, the VOT difference in seconds and their self-reported pronunciation capabilities per speaker on a scale of 1 (worst) to 10 (best). The speakers who gave themselves insufficient scores (5 or below) are marked in blue. The speakers who gave themselves a good grade (8 or higher) are marked in green. These eight participants were used for the statistical analysis.

ID	AVG VOT ENG	VOT DIFFERENCE	Self-reported pron. abilities (1-10)
401	0.0293	-0.0030	6
402	0.0584	0.0267	7
403	0.0275	0.0046	4
404	0.0381	-0.0002	6
405	0.0599	0.0328	7
406	0.037	0.0017	4
407	0.0353	0.0061	6
408	0.0332	0.0087	6
409	0.0272	0.0029	6
410	0.0323	0.0050	2
411	0.0348	0.0043	5
412	0.048	0.0237	8.5
413	0.0319	0.0109	5.5
414	0.0466	0.0225	8
415	0.0373	0.0171	8
416	0.0272	0.0100	8.5
417	0.0344	0.0071	7
418	0.0345	0.0075	6
419	0.0313	0.0051	7