THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF COMMUNICATION SCIENCES AND DISORDERS

DOES FINAL-CONSONANT DEVOICING AND HIGHLY CONFUSABLE AMERICAN ENGLISH VOWELS AFFECT THE PRODUCTION OF ENGLISH BY DUTCH-ENGLISH BILINGUALS

MEREL PADT
SPRING 2024

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Communication Sciences and Disorders
with honors in Communication Sciences and Disorders

Reviewed and approved* by the following:

Annie Olmstead
Assistant Professor
Thesis Supervisor

Carol Miller
Professor
Honors Adviser

ABSTRACT

When learning a second language, learners face several challenges: learning new syntactic structures and vocabulary and learning how to produce the phonology and phonetics of the language they are learning. Focusing on the production of English by native Dutch speakers, the production of English vowels by Dutch-English bilinguals and the transfer of the final-consonant devoicing present in Dutch (but not in English) to English were examined. The study specifically focused on their ability to distinguish the vowels /æ/ (as in *cat*) and /ɛ/ (as in *red*) and the consonants /t/ and /d/ in the word-final position. The vowel /æ/ is not present in the Dutch vowel inventory and is often produced as /ɛ/ by Dutch-English bilinguals. Dutch has final-consonant devoicing, meaning that regardless of a word's spelling, the final consonant is always produced voiceless. In our study, we examined how Dutch-English bilinguals produce English vowels and whether they transfer the final-consonant devoicing when speaking English, and whether this causes misperception.

Using a word-matching task, Dutch-English bilingual participant pairs interchangeably produced and perceived highly confusable English words (*bad, bat, bed, bet, bid, bit*). One participant said the word out loud, while the other chose the word they heard their partner say out of two potential answers. Acoustic analysis was conducted to measure the vowel, consonant closure, stop release durations, and formant values (F1 and F2) of the vowels. Out of all word pairs presented in the task, results showed the highest percentage of misperceptions in the *bid – bit* word pair due to final-consonant devoicing, as well as misperceptions of words containing the /æ/ – /ɛ/ contrast (e.g., *bet – bat* and *bad – bed*), indicating that these speech sounds were not always distinguished clearly in production. Analyses of consonant closure and stop release durations showed that words with /d/ as the final consonant were often produced more similarly to the duration of /t/, indicating that the Dutch-English bilinguals indeed devoiced the final consonant. On average, the vowel length of *bad* was significantly longer than the vowel duration in the other stimuli, and the difference between the vowel duration of *bat* and *bet* in the Dutch-English bilinguals' production was closer in duration than native English productions of these words observed in previous research. Additionally, the formant values of the Dutch-English bilinguals' productions of /æ/ had a great amount of spread and overlapped with the formant values of /ɛ/, indicating that these vowels were produced similarly spectrally. Although perception accuracy was generally high among all participants, the /æ/ – /ɛ/ contrast and /d/ – /t/ final-consonant contrast led to misperceptions of words containing these speech sounds. These findings show that Dutch-English bilinguals transfer features of their native Dutch phonetic inventory to English and provide support that this can lead to misperceptions of certain speech sounds.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First off, I would like to express my gratitude to the Center for Language Science for the invaluable opportunity to participate in the PIRE program (NSF OISE 1545900) to conduct research in the Netherlands and to the National Science Foundation for their generous grant, which provided essential support for this research project. The findings and conclusions reported in this thesis do not necessarily reflect the view of this funding agency.

I would like to thank everyone who has been a part of this research project and guidance in writing this thesis. I am incredibly thankful to my thesis advisor, Dr. Annie Olmstead. Thank you for your unwavering support and guidance. Your support, time, and dedication to this project has been tremendously uplifting and motivating. I also thank others who have played a significant role in my time as a scholar and during this project. Dr. Brouwer, thank you for your enthusiasm and warm welcome during my time in the Netherlands. Dr. Miller and Dr. Viswanathan, thank you for your continuous support and encouragement. Thank you to my fellow researchers, Catherine Pham and Jihee Lee, for the great time we had in the Netherlands conducting our research and for creating many precious memories!

I would also like to thank Dr. Jimin Lee, who generously provided me with research training since my first year. Dr. Lee, thank you for providing me with the knowledge, skills, encouragement, and insightful feedback throughout every stage of my research journey. To Dr. Lee and Dr. Olmstead, I am grateful for the opportunity to present our research at the 2022 ASHA conference because this experience had a transformative impact on my future career ambitions.

Thank you to the Schreyer Honors College and the Paterno Fellows Program for supporting my research and study abroad activities throughout my Penn State career. I also would like to thank the Center for Language Science for providing me with the Judith Kroll Award to present my research at the ASHA conference.

Last but certainly not least, I am deeply thankful to my family and friends for your support, patience, and encouragement throughout my undergraduate studies, and particularly when I was in “thesis-writing” mode — your encouragement kept me motivated and driven from start to finish.

Chapter 1 Introduction

1.1 The Phonological Systems of Languages

When learning a second language, learners face several challenges: learning new syntactic structures and vocabulary and learning how to produce the phonology and phonetics of the language they are learning. This thesis focuses on the production of non-native speech. When the native language phonetic inventory differs from the non-native inventory (e.g., a phoneme produced differently between two languages), producing the non-native speech sounds can be difficult. A well-known example is the /r/ – /l/ contrast, which is particularly difficult for Japanese speakers (Miyawaki et al., 1975). The challenge of accurately producing speech sounds unfamiliar in one's native language poses a significant obstacle to effective communication and can result in communication breakdowns and challenges.

The phonological systems of languages shape how we produce and perceive speech sounds. When speaking a newly learned language, individuals may rely on their native phonemic system when encountering unfamiliar sounds. For example, Bent et al. (2008) found that non-native speakers of English did not follow the same vowel-lengthening patterns as native speakers of English when producing the vowel preceding a consonant, and stuck to the temporal patterns in their native language. An individual carrying over their native speech sounds to another language often leads to them speaking with a foreign accent, which is common in the speech of non-native speakers. Producing words with subtle phonemic distinctions, like vowel and consonant pronunciation variations, may lead to misperceptions and communication breakdowns.

Over the years, English has emerged as a universal language, being widely spoken and taught from a young age in many parts of the world. The Netherlands is a country where children typically start learning English in school at around age ten, leading to Dutch speakers who are generally fluent in English. Dutch and English have phonological systems that are broadly similar, and speakers of Dutch do not typically have serious trouble when recognizing or pronouncing most English sounds (Tops et al., 2001). However, specific speech contrasts in English pose significant problems for Dutch-English bilinguals' perception and production of English. Such contrasts are specific vowel sounds and the voicing of the final consonant in a word, which both tend to be difficult for Dutch speakers to produce.

This thesis examines how Dutch-English bilinguals use their native phonetic system when producing specific English speech sounds. Specifically, I studied whether highly fluent Dutch-English bilinguals can distinguish the productions of specific vowel sounds (/æ/ and /ɛ/) and word-final consonants (/t/ and /d/). Dutch-English bilinguals' speech production of specific highly confusable English words was measured and analyzed to test this. The primary research questions revolve around how native Dutch speakers produce English vowels and whether they devoiced the consonant in the final position of a word. Additionally, I examined how Dutch speakers of English use vowel length to indicate final-consonant devoicing. This phenomenon can also be extended to examine how these phonetic features affect the perception of English sounds by Dutch-English bilinguals.

1.2 Dutch vs. English Vowel Space

To answer the question about the English production of vowels by Dutch-English bilinguals, it is crucial to be familiar with the differences between the Dutch and English vowel spaces. Vowel space refers to the range of possible vowel sounds a speaker produces within a language. The number of vowels and range of a language's vowel space varies across languages, resulting in unique phonetic characteristics. To map out the vowel space, the height (low, mid, high) refers to the vertical position and the frontness (front, central, back) refers to the horizontal position of the tongue in the mouth when producing a vowel. In the case of Dutch (Figure 1) and English (Figure 2), the number of vowels is similar, but some vowels have different locations in the vowel space — Dutch has a cluster of vowels in the mid-front area, whereas English does not.

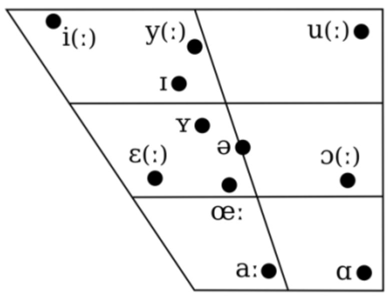


Figure 1 Dutch Vowel Chart

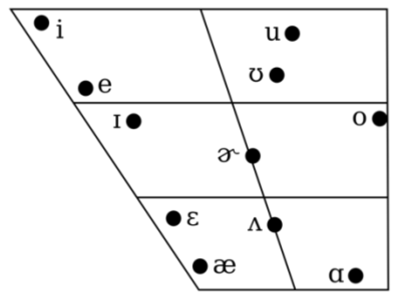


Figure 2 English Vowel Chart

This study focuses on Dutch-English bilinguals' English productions of the vowels /ɛ/, /ɪ/, and /æ/. These vowels are of particular interest due to their differences in pronunciation in Dutch and English. Firstly, the vowel /ɛ/ (as in *get, red, let*) is produced in the front mid position in Dutch and the front mid-low position in English, implying that it is slightly lower in the mouth compared to Dutch. Secondly, the vowel /ɪ/ (as in *kid, bin, sick*) has a more centralized position and is produced slightly higher in Dutch (high front-central) than in English (mid front). Finally, the front low vowel /æ/ (as in *sat, can, cat*) is not found in the Dutch vowel space, making this vowel challenging to accurately produce for a Dutch speaker.

Due to these differences in vowel space, some vowels are more difficult to produce than others in a native manner. For example, the vowel /ɑ/ (as in *father, not, barn*) has a highly similar location in both the Dutch and English vowel spaces, making the English production of this vowel typically easier to pronounce for a Dutch-English bilingual than vowels that are farther apart or non-existent in their native vowel space. The Dutch and English pronunciations of /ɛ/ and /ɪ/ are considered similar, as both languages produce these vowels in a similar position. Collins and Mees (1984) stated that the Dutch /ɪ/ can "pass straight into English without being modified" and should pose "no problem" for pronunciation. However, Dutch-English bilinguals struggle to distinguish the vowels /ɛ/ and /æ/ in production. These vowels are tricky for Dutch-English bilinguals because, as mentioned, Dutch has no vowel realized phonetically as /æ/, and /ɛ/ is the closest equivalent. Collins and Mees (1984) discussed that Dutch learners tend to substitute the Dutch /ɛ/ for the English /æ/ (e.g., the word *tan* may be pronounced as *ten*), but they did not quantify this observation. Because the /ɛ/ – /æ/ contrast is highly confusable for Dutch-English bilinguals, it is relevant to further examine the production of these vowels.

1.3 Final-consonant Devoicing

Another contrast between Dutch and English speech production is the voicing of the consonant at the end of a word. Voicing and devoicing refer to the vibration or absence of vibration of the vocal cords when producing a speech sound. During voicing, the vocal cords come together and vibrate, producing sounds such as vowels, nasals, and voiced consonants. Conversely, devoicing happens when the vocal cords do not vibrate, resulting in voiceless consonants and fricatives. Table 1 presents all English voiceless consonants and their voiced counterparts.

Table 1

English Voiceless and Voiced Consonants with Examples

Voiceless Consonants (no vibration)		Voiced Consonants (vibration)	
/p/	park	/b/	bark
/t/	town	/d/	down
/k/	coat	/g/	goat
/f/	fan	/v/	van
/ʃ/ ("sh")	sure	/ʒ/ ("zh")	treasure
/ʧ/ ("ch")	chain	/ʤ/ ("j")	Jane
/θ/ ("th")	thigh	/ð/ ("th")	thy

Dutch is one of several languages that exhibit final-consonant devoicing, meaning that in Dutch, the final consonant in a word is always voiceless, regardless of its spelling. For example, the Dutch word *zand* (*sand*, in English) is pronounced as zan[t] and the Dutch word *heb* (*have,* in English) is pronounced as he[p]. However, English preserves the final consonant's voicing and does not undergo final-consonant devoicing (e.g., the /d/ stays voiced in *rod*). Because of this, Dutch-English bilinguals may use their native final-consonant devoicing when saying English words. For example, in words like *glad* or *sad*, Dutch-English bilinguals may produce the /d/ as /t/, making these words sound like *glat* and *sat*. Simon (2010) examined whether Dutch-English bilinguals transfer the Dutch devoicing rules to English. The participants were native Dutch speakers who were all highly proficient in English. The participants were tested in pairs and talked to each other for 30-45 minutes about any topic they wanted, eliciting rapid and natural speech. The first conversation was in Dutch, and the second was in English, allowing the transfer of native phonological processes to their non-native language to be measured. Additionally, 10 out of the 16 participants read 40 Dutch words and an equal number of English words out loud, presented in isolation. Results show that devoicing of the final-consonant stops occurred in 48.67% of the tokens in spontaneous conversations and 8% in word-reading tasks. These results indicate that the participants have learned to suppress final-consonant devoicing to an extent in controlled conditions (here: word-reading task) but devoiced almost half of the tokens in a less restrictive environment (here: conversation).

The distinction between voiced and voiceless sounds is crucial for producing speech sounds and can change the meaning of a word in various languages, including English. For example, voicing the /f/ in *few* and *leaf* leads to pronouncing these words more as *view* and *leave.* Alternatively, devoicing the /g/ in *dog* makes the pronunciation sound more like *dock*. Because word-final consonants in Dutch are always voiceless, Dutch speakers are not accustomed to this distinction in voicing. To further examine whether Dutch-English bilinguals devoice word-final consonants in English, the current study focuses on the production discrepancy of /d/ (voiced) and /t/ (voiceless) stops at the word-final position.

1.4 The Effect of Final-consonant Devoicing on Preceding Vowel

In English, the length of the vowel preceding a voiced final consonant is longer than its voiceless consonant cognate. This was also found to be true in most other languages, including Dutch (Slis & Cohen, 1969). House (1961) examined the speech of three English monolingual speakers producing English vowels in voiced or voiceless consonant environments. This study included nonsense words with the CVC structure. The stimuli contained 12 common vowels of

English and 14 consonants that often appear in the initial and final positions in English syllables (that have voiceless cognates). Speech productions were analyzed, and all mean vowel durations were rounded to the nearest 10 ms. House found that lax closed vowels (like /ɪ/) had a vowel duration of 100 ms when the final consonant was voiceless and 190 ms when it was voiced. Additionally, lax open vowels (like /æ/ and /ɛ/) had a mean vowel duration of 120 ms when the final consonant was voiceless and 220 ms when it was voiced. When the lax closed vowels in voiceless contexts were compared in duration to the lax open vowels also in voiceless contexts, the former were shorter on average by 30 ms. These results show that the vowel length preceding a consonant can indicate the voicing of the final consonant. As mentioned above, devoicing the final consonant can change the production and meaning of the intended word. Take the example of *sad* and its devoiced final consonant counterpart, *sat.* The vowel duration in *sad* should be longer than *sat* because /d/ is a voiced consonant and devoicing the /d/ changes the word's meaning.

1.5 Acoustic Measurements

1.5.1 Vowel Quality

Vowel quality is measured through the acoustic signal's first formant (F1) and second formant (F2). According to the Acoustical Society of America (2004), a formant is "a range of frequencies in which there is absolute or relative maximum in the sound spectrum. The frequency at the maximum is the formant frequency." Formants are measured to determine a vowel's acoustic properties, and F1 and F2 give information about the speakers' use of vowel space when producing vowels. F1 corresponds to the articulatory dimension of vowel height, and F2 corresponds to the place of maximal constriction (i.e., front vs. back vowel). Due to differences in the length and size of vocal tracts and the shape of oral cavities between males and females, formant frequencies for females tend to be higher than those of males. Measuring these formant values gives crucial information for determining the acoustic properties of the vowel.

Vowel space directly correlates with a vowel's F1 and F2 values. When a vowel is produced lower in the mouth, it correlates to a higher F1 value and a lower F2 value than when a vowel is produced higher in the mouth. In English, the vowel /ɛ/ (mid-low) is produced slightly higher than /æ/ (low), and /ɪ/ (mid) is produced highest in the mouth among these vowels. Formant measurements of English productions of vowels were reported by Peterson and Barney (1952). In this study, 76 participants produced 20 words with the vowels in an /hVd/ context, and the formants of each production were measured. Results show that /æ/ (M = 660) has the highest F1 value (Hz), followed by /ɛ/ (M = 530), and then /ɪ/ (M = 390). As expected, F2 has the mirrored order where /ɪ/ (M = 1990) has the highest value, then /ɛ/ (M = 1840), and lastly /æ/ (M = 1720). Pols, Tromp, and Plomp (1973) reported the F1 and F2 values for Dutch productions of words. This study examined the productions of words in an /hVt/ context, and the speech of 75 speakers was measured. Their results show that the F1 of /ɛ/ (M = 583) is higher than /ɪ/ (M = 388), and the F2 of /ɪ/ (M = 2003) is higher than /ɛ/ (M = 1725). Both studies reported the F1 and F2 values of males in their results. The results of both studies show the same trend where F1 values are higher when a vowel is produced lower in the mouth, and F2 values are higher when the vowel is produced higher in the mouth.

Flege (1997) compared the results of the Dutch and English formant values reported in these studies. His ellipses show that the mean values for the Dutch /ɛ/ tokens fall into the upper portion of the English /æ/ space. The mean values for the Dutch /ɪ/ are reported to be very similar to the English production of /ɪ/.

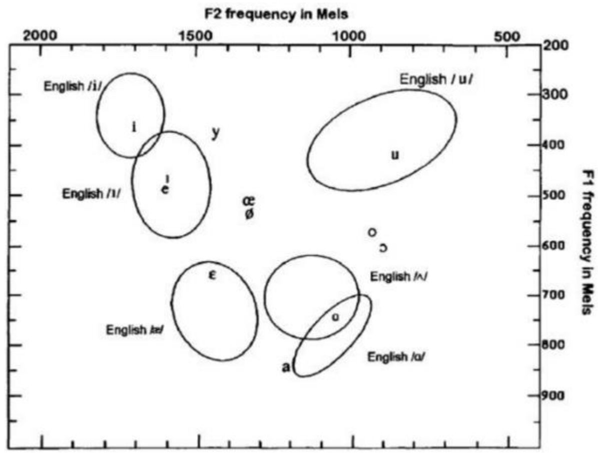


Figure 3 Acoustic relationship between American English /i/, /ɪ/, /u/, /ɑ/, /æ/, /ʌ/ and Dutch vowels. Adapted from "English vowel productions by Dutch talkers: more evidence for the 'similar' vs 'new' distinction," by J. E. Flege, 1997, p.19.

Although the /æ/ – /ɛ/ contrast is known to be highly confusable for Dutch-English bilinguals to produce, there has been some conflict in previous studies. Wang and Van Heuven (2006) recorded and measured the speech of 10 male and 10 female speakers for each of three languages: Mandarin, Dutch, and English. This study compared languages with a relatively small vowel inventory (Mandarin) with a language with a richer inventory and more similar to English (Dutch). Each speaker said vowels in an /hVd/ context, including /æ/, /ɛ/, and /ɪ/. Surprisingly, the results show a fair degree of separation in F1 and F2 values of the Dutch speakers' production of /æ/ and /ɛ/, but it is not as large as the native English speakers' separation. Results indicate that Dutch-accented English vowels were produced with a more English-like pronunciation than Chinese-accented English. It also contradicts the highly confusable /æ/ – /ɛ/ contrast and shows that the speakers spectrally distinguished these vowels.

1.5.2 Vowel duration

The duration of the vowel length indicates two main things: distinguishing between the production of different vowels and whether the vowel preceding a voiced or devoiced consonant differs in duration. Vowel durations of English were published by Peterson and Lehiste (1960), and on average, in connected speech across five speakers, /æ/ had a duration of 330 ms, /ɛ/ had a duration of 200 ms, and /ɪ/ had a duration of 180 ms. They also measured vowel durations of short vs. long vowels preceding different consonants. The duration of short vowels (like /ɛ/ and /ɪ/) preceding /t/ was 147 ms, and /d/ was 206 ms. The duration of long vowels (like /æ/) preceding /t/ was 210 ms, and /d/ was 318 ms. These data show that the durations of the vowels in English are significantly affected by the nature of the consonants following the vowel and that the vowel duration is shorter when followed by a voiceless consonant and longer when followed by a voiced consonant. As explained earlier and represented in this study, if the vowel duration preceding the voiced final consonant is shorter than the typical vowel duration preceding a voiced final consonant (e.g., the duration of /æ/ in *bad* is shorter than what is expected of a vowel preceding a voiced consonant), it is an indicator of devoicing of the final consonant.

Vowel durations in Dutch were studied by Rietveld et al. (2004). Four native speakers of Dutch (3 males and 1female) aged between 23 and 35 years read 90 CVC words out loud, where the stimuli were combinations of 15 Dutch vowels (including diphthongs, long monophthongs, half-long, and short lax vowels) and the consonants /k/, /s/, and /m/. Each vowel was produced 24 times total, and they found that /ɛ/ had a slightly longer vowel duration (M = 107 ms) than /ɪ/ (M = 99 ms). Because final consonants are not voiced in Dutch, comparing the duration of vowels preceding voiced or voiceless final consonants was not measured.

The studies discussed above show that the productions of short vowels in Dutch were generally shorter than those of English, but it is important to note that the context in which the stimuli were produced was different. English vowel durations were measured in connected speech, and Dutch vowel durations were measured in CVC words. However, differences in vowel duration, along with other differences, are found to be important for distinguishing the vowels. The vowel duration differences between short and long vowels are significant (e.g., /æ/ should have a longer duration than /ɛ/). If there is no significant difference between the vowel durations of /æ/ in the words *bat* and *bad* produced by the same speaker, this may indicate that the final consonant had the same voicing in both productions. Short and long vowels were used in the present study, and vowel duration will indicate whether the final consonant was devoiced or stayed voiced.

1.5.3 Consonant Closure and Stop Release

Two factors affecting the final consonant are consonant closure duration and stop release duration. The articulators come together when producing stop consonants (like /t/ and /d/) and require a pressure buildup in the oral cavity before they are released. There is a moment of silence in the production of a word after the vowel and before a stop-consonant is released (if it is released), referred to as consonant closure. Many previous studies (e.g., Lisker, 1957) have shown that consonant closure durations are longer for voiceless than voiced stops. Stathopoulos and Weismer (1983) measured the closure duration of stop consonants in initial, medial, and final word positions. English speakers produced a series of nonsense words in the CVCVC structure (e.g., *debad* or *debat*) in the carrier phrase “Say ___ again.” Each speaker produced stimuli that included all possible combinations of stress, position, and place of articulation, and the duration of the consonant closure was measured. Results showed that in the word-final position, stressed voiced stop consonants (like /d/) had a mean of 40 ms, and unstressed voiced stop consonants had a mean of 58 ms. Voiceless stop consonants (like /t/) that were stressed had a mean of 58 ms, and unstressed ones had a mean of 76 ms. These data show that consonant closure of voiceless stop consonants has a longer consonant closure duration than voiced stop consonants in the word-final position. Therefore, a shorter consonant closure duration indicates that the final stop consonant was devoiced, and a voiced final stop consonant would have a longer consonant closure duration.

Another indicator of whether the consonant was voiced or devoiced is the duration of the release of the stop. The duration of the release of the stop is the period after the consonant closure when the final consonant is released. A voiced stop typically has a longer stop release duration, while a devoiced stop typically has a shorter one. In a study by Crystal and House (1988), stop consonants were measured in various word positions and contexts. Six speakers read two scripts out loud, approximately 600 words. In the word-final position, /t/ had a mean duration of 96 ms, and /d/ had a mean duration of 78 ms. Results showed that /t/ has a longer stop release duration than /d/, and the release duration was a distinguishing factor for word-final /t/ and /d/. The overall longer duration (both consonant closure and stop release durations) of voiced stops is due to the required vocal cord vibration, which voiceless stops lack. This vibration extends the duration of the stop's release phase, resulting in an overall longer duration. In English, however, the release burst is not always present (Byrd, 1993). If there is no release of the final consonant, neither the consonant closure duration nor stop release duration is measured.

1.6 Cues for Perception

The features of word production that have been discussed, vowel duration and final-consonant devoicing, are both critical indicators of a word, and therefore, an important cue for perception. Many previous studies have shown that vowel duration in English words with the CVC or CVCC structure plays a significant role in the perception of the voicing of the final consonant. Hillenbrand et al. (2000) examined how variations in vowel duration influence the recognition of vowel identity in English words. The perception of 300 /hVd/ words with four different vowel durations — the original duration, a neutral duration, a short duration, and a long duration — were measured. Listeners were asked to identify the words they heard, which were presented in a random order. The results show that altering vowel length in English vowels generally does not have a significant impact on perception of most vowel pairs (0%-2%), except for shifting among the vowels /ɔ/ (as in *caught*), /ɑ/ (as in *art*), /Λ/ (as in *but*) and /ɛ, æ/ (9%-43%). Roughly one-fifth of the tokens of /æ/ were perceived as /ɛ/ when the original duration was shortened, and /ɛ/ was perceived as /æ/ when the original duration was lengthened. The /æ/ – /ɛ/ pair shows a robust duration effect, while shortening or lengthening /ɪ/ had very little effect on vowel identity. Results show that certain vowels (like /æ/ – /ɛ/) have a greater reliance on duration for their separation, implying that if the vowel duration is altered, it could lead to misperception of the intended vowel or word.

Van der Feest and Swingley (2011) measured the perception of words with altered vowel durations in Dutch. This study examined native-speaker identification of English and Dutch words varying in their vowel duration. In each language, eight different vowels were presented in six real words and six non-words. Half the words were presented with natural vowel duration and half with shorter or longer duration, and participants listened to recordings of words and typed the word that they heard. Results show that overall, Dutch listeners were more strongly influenced by alterations in vowel duration (*M* = 29.4% of trials, *SD* = 7.8) than English listeners (*M* = 9.5%, *SD* = 5.9), and errors were more likely when vowels were shortened than when lengthened. The findings in this study show that alterations in vowel duration affect Dutch listeners' interpretation of words more than they affect English listeners'. This could be because vowel durations in Dutch do not change based on the voicing of the final consonant. In other words, the duration of the vowel stays consistent regardless of the consonant that follows it. Consequently, Dutch listeners use vowel duration as a primary cue for distinguishing and understanding words, so when the vowel duration is changed, they do not have other cues to rely on. In contrast, English listeners have other cues they can use to decipher words. While vowel duration is still a relevant cue in English, English listeners may also rely on other phonetic features, like voicing or devoicing of the final consonant, to distinguish words.

The phonetic contrasts between Dutch and English in vowel space and final-consonant devoicing may cause communication difficulty. For instance, Dutch-English bilinguals are likely to use final-consonant devoicing in English words and produce the vowel /æ/ as /ɛ/. As a consequence, their communication partner is likely to misunderstand the target word. Take the example of a Dutch-English bilingual saying the English word *bat* to a fluent English listener while pronouncing the /æ/ as /ɛ/. Instead of hearing the word associated with baseball or the animal, the listener might understand the word *bet.* Now considering final-consonant devoicing, if the Dutch-English bilingual says the word *bad* to the fluent English listener while pronouncing the /æ/ as /ɛ/ and /d/ as /t/, the listener might again understand *bet.* Given these phonological differences in production, perception is influenced, and communication partners may experience communication breakdowns or challenges.

To test the phenomenon of production features of Dutch-accented English, the participants in the present study produced target words in a word-matching task. This is a restrictive task where speakers produce highly confusable words to a partner, and their partner must select the word they think they just heard. The stimuli words included highly confusable English vowels (for Dutch speakers) and the voiced/devoiced stop-consonant pairs /d/ and /t/. Acoustic analysis using the program Praat (Boersma & Weenink, 1996) was conducted to measure the patterns of English speech produced by the Dutch-English bilinguals. Specifically, the vowel quality through formant frequencies, vowel duration, consonant closure duration, and stop release duration were measured.

To review, the differences in Dutch speakers' productions of English words compared to native English speakers' productions of English words are due to differences in vowel space, devoicing the final consonant, and the respective shortening in vowel duration. These differences affect the ability to correctly produce and perceive English words and may cause communication breakdowns between interlocutors. Dutch-English bilinguals' productions of highly confusable English words will be analyzed to examine whether these words undergo final-consonant devoicing and/or non-native vowel productions, which will potentially reveal which phonetic difference has the most significant impact on perception. These differences can be measured by examining vowel quality and the duration of consonant closure and release duration through acoustic analysis. This thesis aims to find a link between the production of highly confusable English vowels and final-consonant devoicing in Dutch-English bilingual speakers, thereby potentially influencing the perception of Dutch-English bilingual listeners.

Chapter 2 Methods

2.1 Participants

The participants in this study were 26 Dutch-English bilinguals (19 females and 7 males, *M* = 21.88 years, range: 19–28). The experiment was completed in pairs (13 pairs); six were familiar with one another, and seven were not. The participants were recruited from Radboud University's SONA subject pool, had no history of language disorders, and all reported normal hearing.

Consent for participation was obtained prior to the experiment. A Language History Questionnaire (LHQ; Li et el., 2014), Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012), and a Category Fluency task (Mitrushina et al., 2005) were used to measure proficiency scores for each participant. These tasks were completed asynchronously online, and a few participants completed the LexTALE after the experiment session.

2.2 Equipment

The computer software program Labvanced was used to administer the task, and participants communicated with one another through the online Zoom platform. Participants used Sony WH-1000XM5 noise-canceling headphones with a built-in microphone to record their speech, which was later used for acoustic analysis.

2.3 Procedure

The participants were asked to complete a word-matching task in pairs. These data were collected as part of a larger project, and all participants completed a Diapix task before the word-matching task. A list of words containing the vowels /ɛ/, /ɪ/, /æ/, the consonant /b/ in the word-initial position, and either /t/ or /d/ in the word-final position were used: bet, bed, bat, bad, bit, bid.

The participants were seated in separate sound isolation booths and could not see their partner's computer screen, and their cameras were turned off on Zoom. One participant was assigned a Matcher role, and the other was assigned a Director role. The roles randomly alternated throughout all trials. The Matcher and the Director's screens had two words, and the Director's screen had one word highlighted in red. The Director's task was to say the highlighted word on their screen to their partner, and the Matcher's task was to select the word they thought they heard the Director say. The two words on the screens were 12 potential combinations of the stimuli (e.g., *bed* was on the left, and *bet* was on the right of the screen). After the Matcher selected a word, both participants received feedback on whether they were correct or incorrect. Before starting the actual task, the participants did a practice trial to become accustomed to the task. After the practice trial, the participants started the actual task.

In the actual task, there were 96 trials and a break halfway through. Each trial had a fixation cross for 50 ms, and both stimuli were presented onscreen to the Director and Matcher for 10,000 ms total. After the first 2,000 ms, the word on the Director's screen turned red, which they said aloud. The Matcher had 8,000 ms to make their selection, and then feedback (correct or incorrect) was given for 2000 ms. After completion of both the Diapix task and word-matching task, all participants received either SONA credit or VVV vouchers for their participation.

Chapter 3 Results

Proficiency data were collected for each participant using a Language History Questionnaire (LHQ; Li et el., 2014), Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012), and a Category Fluency task (Mitrushina et al., 2005). An LHQ was obtained to assess the frequency and onset of English language use and learning (Table 2). According to the self-reported LHQ, participants started acquiring Dutch in early childhood (age two was the oldest). Participants started acquiring English at a young age (*M* = 9.23 years), the earliest at age six and the latest at age twelve. Overall, participants considered their Dutch proficiency higher than their English proficiency. On a scale of 1–10, they ranked their level of Dutch proficiency including speaking, reading, understanding, and writing (*M* = 9.50) and English (*M* = 8.16). They also indicated their mean age of becoming fluent in English (*M* = 16.13), and Dutch (*M* = 6.63). With increased globalization, the use of other languages (especially English) is becoming more frequent, and 88.46% of participants reported that they code-switch between Dutch and English. Those who reported code-switching do it mostly with friends and some with family and coworkers/professors. On average, they used Dutch most of the day (61%) and English for the rest (39%).

The LexTALE score represents the average percentage of correct responses from a vocabulary test consisting of both real words and nonwords. The LexTALE score was relatively high for all participants (*M* =82.74%, range: 58.75%–100%). A higher percentage indicates a higher level of proficiency. Lastly, the Category Fluency task measured how many items a participant named in specific categories (animals and fruits/vegetables) within a certain time (*M* = 17.40 for categories combined). A greater number of responses is indicative of greater English proficiency. LexTALE and Category Fluency scores were broken down by each participant

(Table 3). Results from all proficiency tasks show that participants are generally highly fluent in English.

Table 2

Results From the Language History Questionnaire

Category	English				Dutch			
	Mean	*SD*	*Min*	*Max*	*Mean*	*SD*	*Min*	*Max*
Overall proficiency	8.16	0.72	6.75	9.5	9.5	0.57	7.5	10
Reading	8.73	0.87	7	10	9.69	0.55	8	10
Speaking	7.35	0.98	6	9	9.58	0.64	8	10
Understanding	9.00	0.89	7	10	9.88	0.43	8	10
Age of Acquisition	9.23	2.30	5	12	1.50	0.71	1	2
% of Daily Use	39.23	18.31	10	80	60.77	18.31	20	90

Table 3

Individual Participant and Mean LexTALE and Category Fluency Scores (Correlation = .39, $p < .05$)

Participant	LexTALE	Category Fluency
1	68.75	18
2	87.50	20
3	60.00	13.5
4	96.25	19
5	82.50	19
6	100.00	24

7	97.50	22
8	86.25	15.5
9	87.50	16.5
10	80.00	12.5
11	63.75	17.5
12	75.00	8.5
13	81.25	10.5
14	93.75	20
15	90.00	18.5
16	58.75	14.5
17	73.75	20
18	95.00	23
19	82.50	12.5
20	81.25	18
21	95.00	11.5
22	86.25	19
23	78.75	19.5
24	87.50	23.5
25	85.00	14.5
26	77.50	21.5
Mean	82.74	17.4

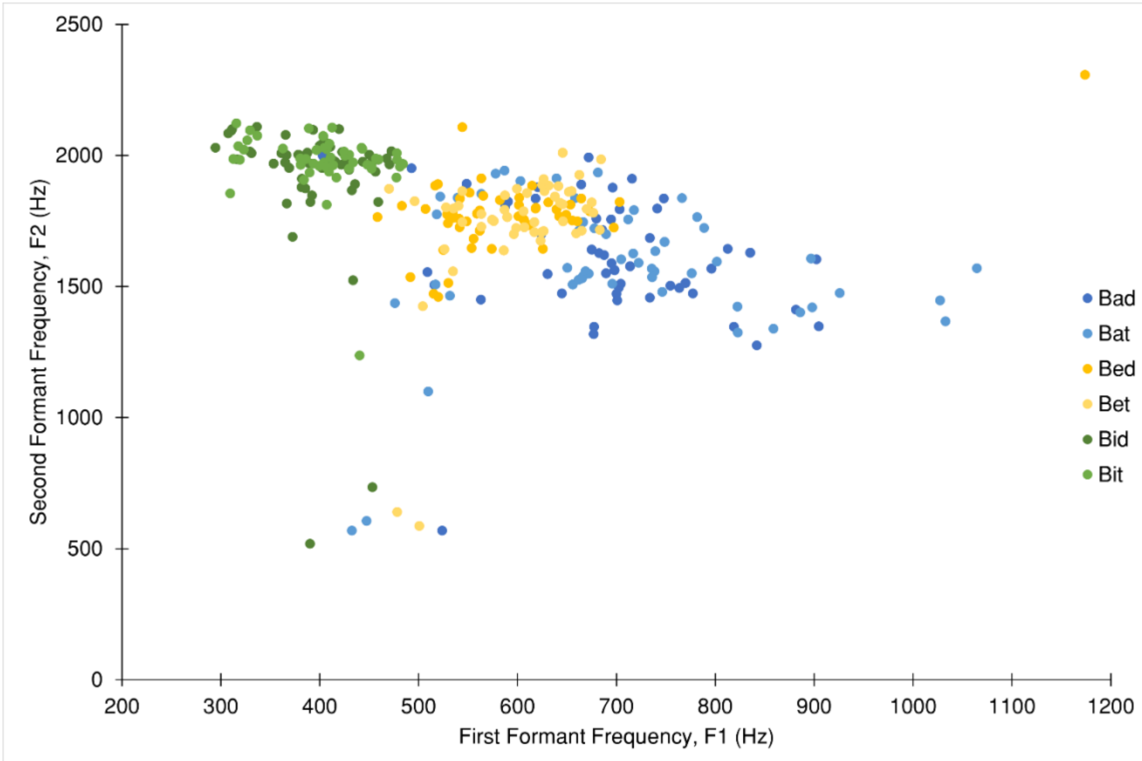
Using the Praat speech processing software (Boersma & Weenink, 1996), the onset and offset of the vowel, consonant closure, and stop release were determined manually using oscillographic and spectrographic displays, as well as listening to the production. An oscillogram is a visual display of sound data that shows the relationship between time and amplitude. A spectrogram provides a display of the amplitude and frequency over time. Some productions were omitted from acoustical analysis for several reasons: audio difficulties with participants not being able to hear each other, participants asking clarifying questions after the task started, poor audio quality of the recording, or instructions were not followed by participants (e.g., spelling out the word or using the word in a sentence).

3.1 Vowel Quality

Formant values (F1 and F2) of the vowel were calculated for each production by the Burg LPC algorithm implemented in Praat, and the F1 and F2 values were extracted at the temporal midpoint of the target vowel. It is important to note that the values calculated by the Burg LPC algorithm in the program Praat were not manually corrected. This built-in algorithm is generally reliable but may have some errors. Additionally, it is not guaranteed that all productions were produced with the correct vowel, as participants may have misspoken or misread the word. These potential inconsistencies could be the explanation for why some extraneous points on the graphs do not fit the general trend. Mean formant values were calculated for males and females separately. For males (Figure 4), the vowel /æ/ had the highest F1 value (*M* = 705 Hz, *SD* = 124), then /ɛ/ (*M* = 590 Hz, *SD* = 74), and lastly /ɪ/ (*M* = 397 Hz, *SD* = 43). The vowel /ɪ/ (*M* = 1954 Hz, *SD* = 193) had the highest F2 value, then /ɛ/ (*M* = 1759 Hz, *SD* = 180), and lastly /æ/ (*M* = 1597 Hz, *SD* = 253).

Figure 4

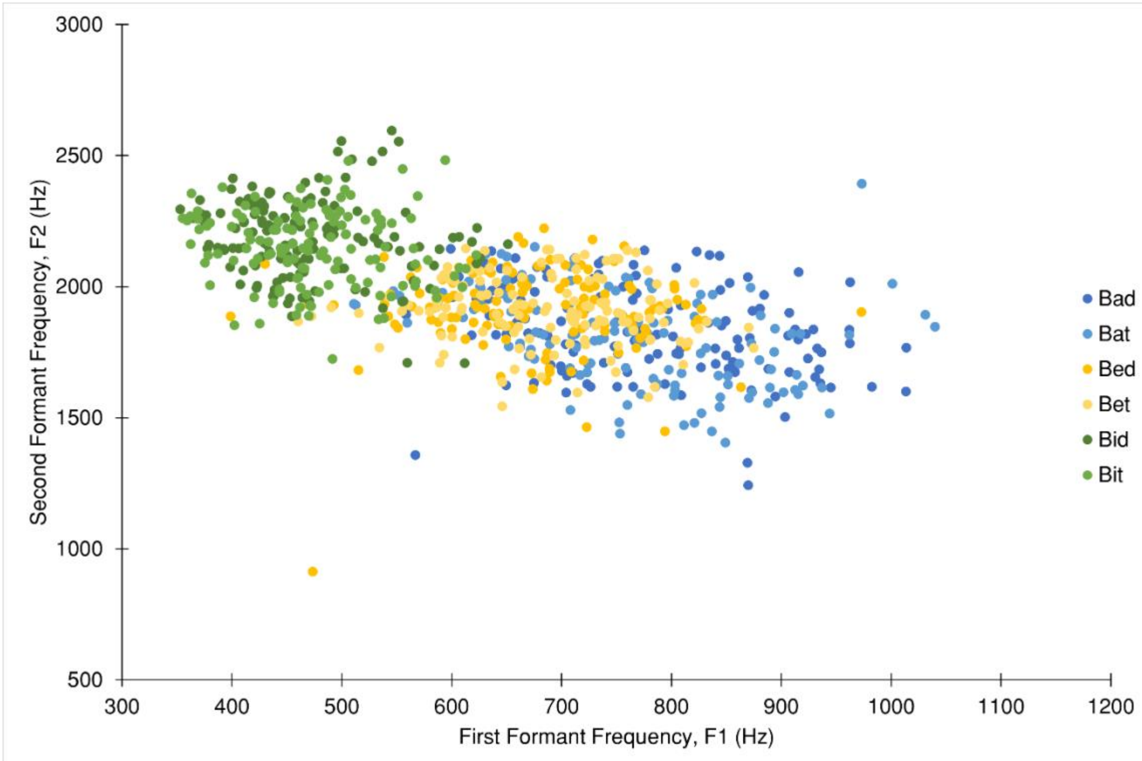
First and Second Formant Frequencies for Males



For females (Figure 5), the highest to lowest order is the same as males for F1, and vice versa for F2. The vowel /æ/ had the highest F1 value (M = 771 Hz, SD = 168), then /ɛ/ (M = 675 Hz, SD = 89), and lastly /ɪ/ (M = 471 Hz, SD = 63). The vowel /ɪ/ (M = 2166 Hz, SD = 147) had the highest F2 value, then /ɛ/ (M = 1910 Hz, SD = 144), and lastly /æ/ (M = 1823 Hz, SD = 223).

Figure 5

First and Second Formant Frequencies for Females



Formant values for the participant pair with high proficiency (participants 5 & 6) and the highest accuracy in the identification of words and the participant pair (participants 9 & 10) with lower proficiency and lower accuracy in the identification of words were calculated separately and shown in Figures 6 and 7. The graphs display formant values of both participants who had the same gender.

Figure 6

First and Second Formant Frequencies of two Participants with High Accuracy in Word Identification

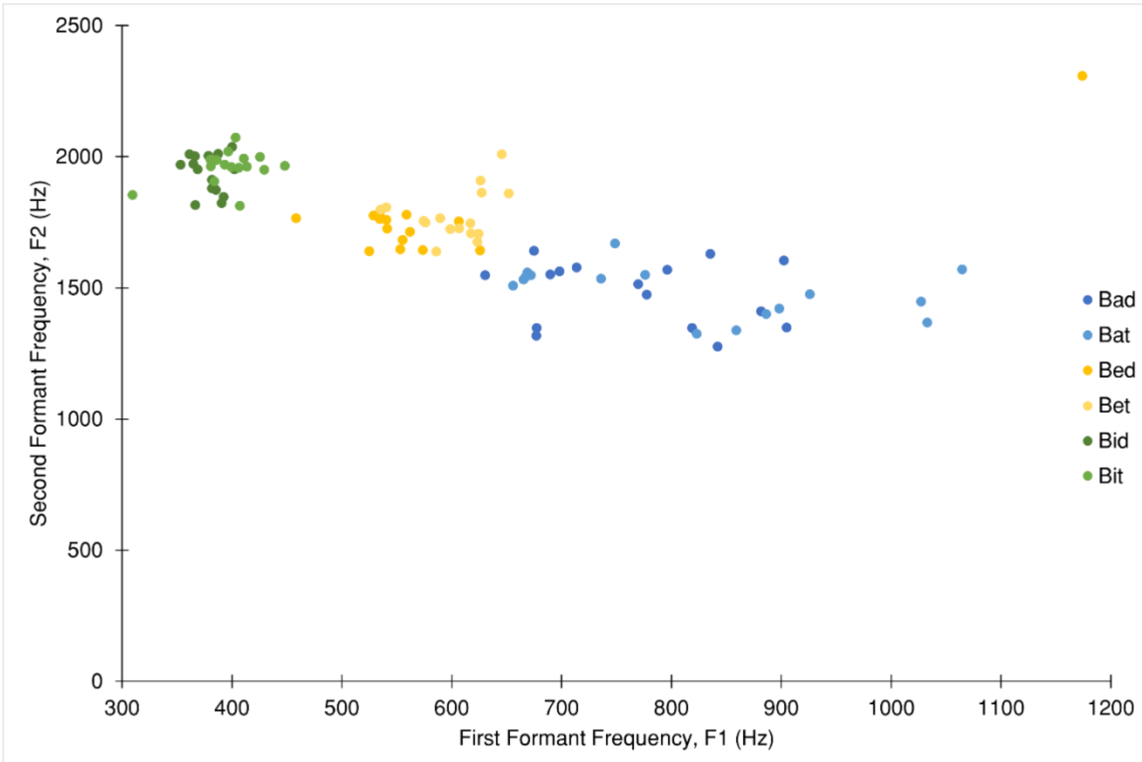
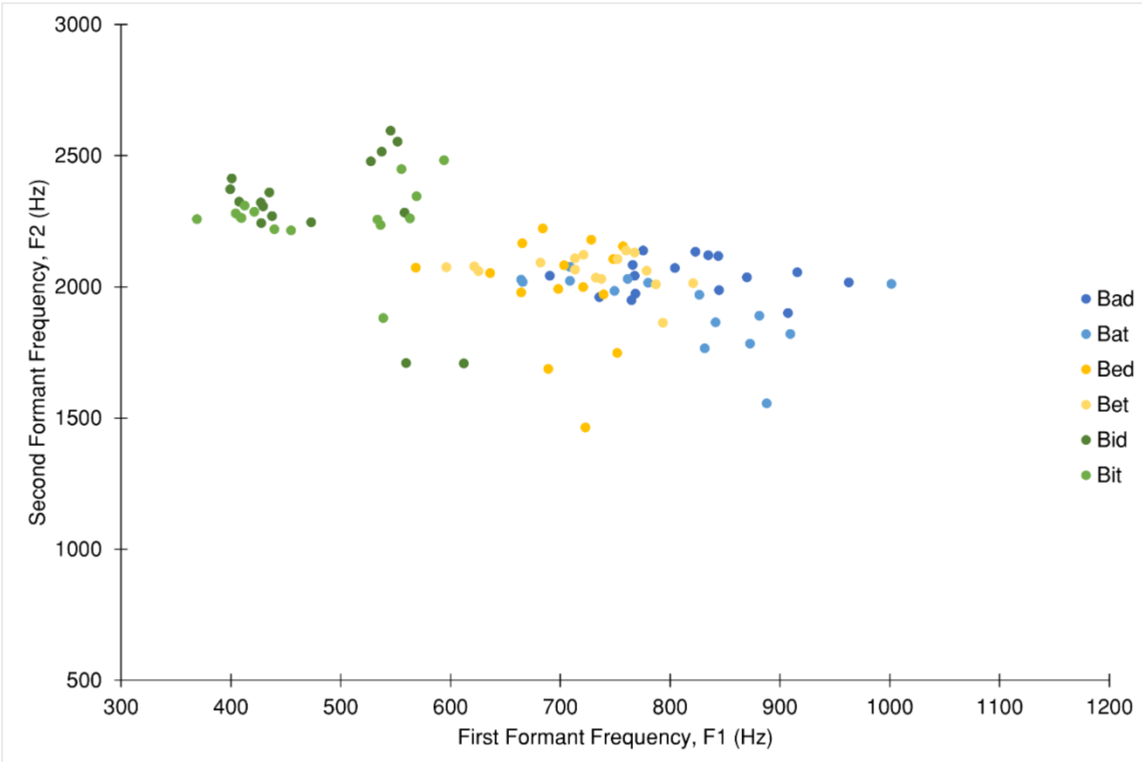


Figure 7

First and Second Formant Frequencies of two Participants with Lower Accuracy in Word Identification

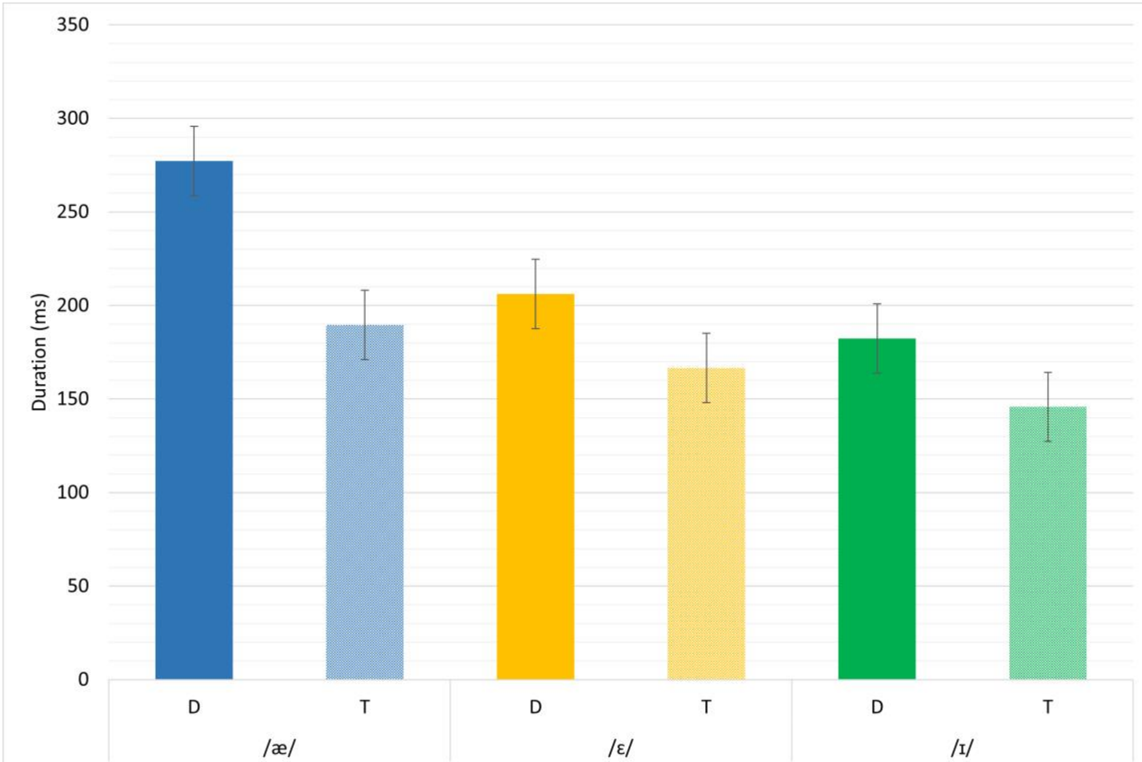


3.2 Vowel Duration

The average duration of the vowels produced by all speakers was calculated and differentiated by the final consonant it was preceding, as shown in Figure 8. The *bad – bat* contrast in vowel duration (ms) was the greatest (difference of 87.516 ms), where /æ/ preceding /d/ (M = 277.14) was longer than /æ/ preceding /t/ (M = 189.63 ms). The difference between the *bed – bet* (difference of 39.586 ms) and *bid – bit* (difference of 36.593 ms) contrasts was not large. The /ɛ/ preceding /d/ (M = 206.09 ms) was longer than /ɛ/ preceding /t/ (M = 166.50 ms). The duration of /ɪ/ preceding /d/ (M = 182.26 ms) was longer, and the duration of /ɪ/ preceding /t/ (M = 145.76 ms).

Figure 8

Average Vowel Duration by Final Consonant (in ms)



3.3 Consonant Closure and Stop Release Durations

Consonant closure duration was manually determined by the last glottal pulse for the preceding vowel and the initial burst of the following stop. On the spectrogram, the portion containing the consonant closure portion appeared mostly blank due to the temporary silence. On both displays, the stop release appeared as a burst of sound. If a final consonant was not released, neither the consonant closure nor the stop release durations were measured. Each condition's boundary was set using the TextGrid, and the data was extracted for statistical analysis. The average consonant closure duration (ms) of /d/ and /t/ among all words is shown in Figure 9, and the average stop release duration is shown in Figure 10. The consonant closure duration among all words with /d/ as the final consonant (M = 49.61 ms) was slightly shorter than /t/ as the final consonant (*M* = 53.19 ms). The average stop release duration among all words with /d/ as the final consonant (*M* = 108.73 ms) was shorter than /t/ as the final consonant (*M* = 157.86 ms).

Figure 9

Average Consonant Closure Duration of /d/ and /t/ (in ms)

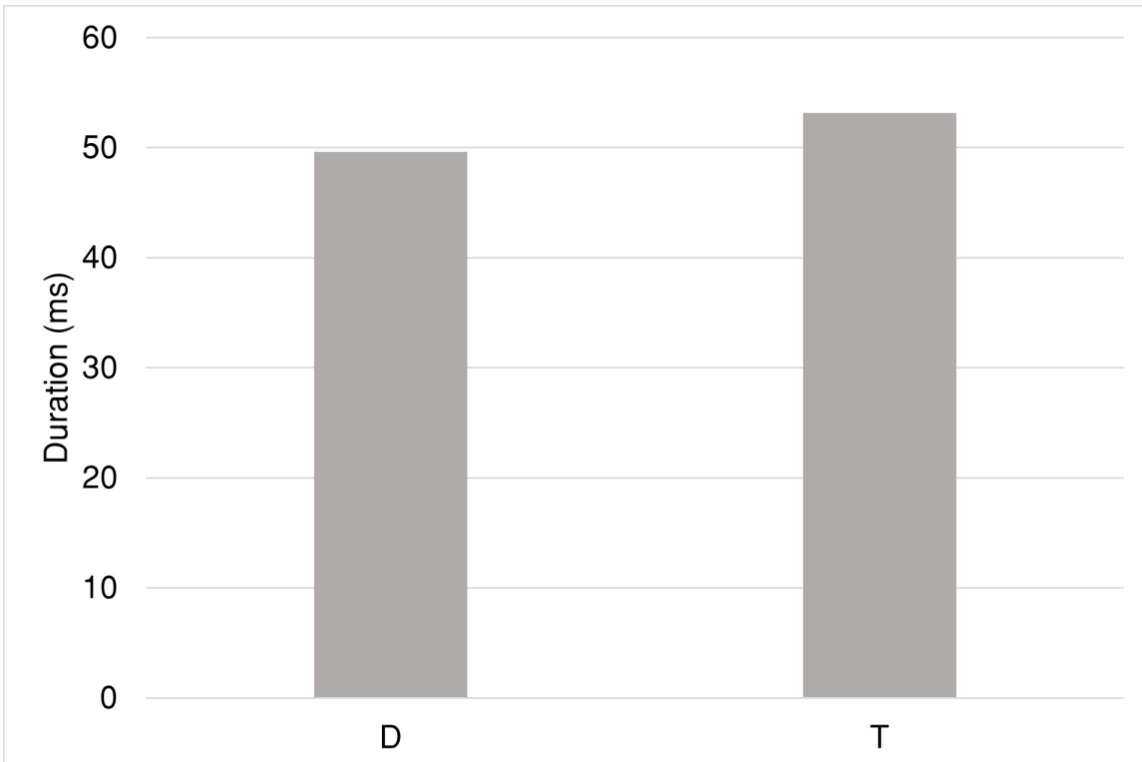
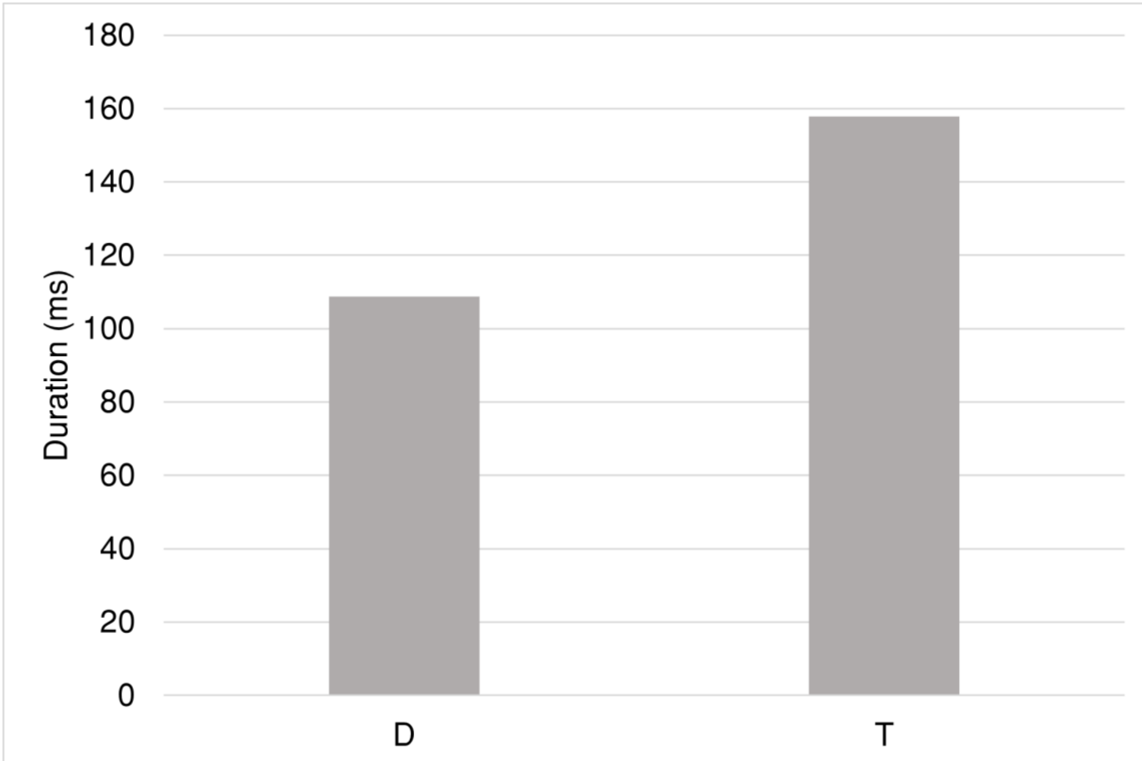


Figure 10

Average Stop Release Duration of /d/ and /t/ (in ms)

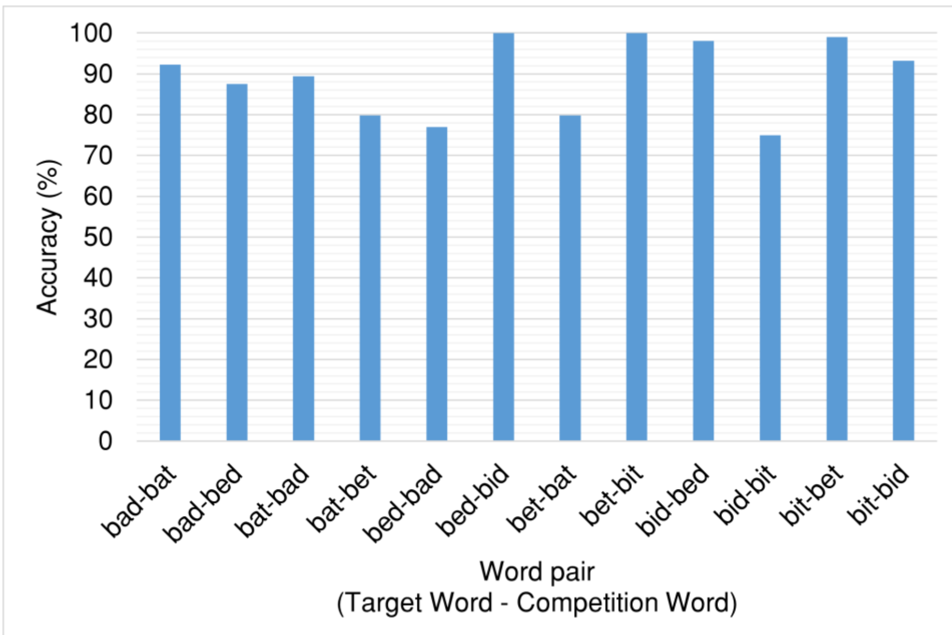


3.4 Perception

In the word-matching task, accuracy in the identification of words was high across all participant pairs (*M* = 89.26%). The participant pair with the lowest accuracy in identification was 78.13%, and the pair with the highest was 98.96%. Figure 11 shows the accuracy of identification broken down for each word pair, where percentages are shown for the number of word pairs correctly identified. The higher the bar, the better participants distinguished the word pair. The word pairs with the highest percentage of correct identifications were *bed – bid* (100%), *bet – bit* (100%), *bit – bet* (99.04%), and *bid – bed* (98.08%). Word pairs with the lowest percentage of correct identifications were *bid – bit* (75.00%), *bed – bad* (76.92%), *bat – bet*, and *bet – bat* (79.81% for both). Word pairs where the final consonant in the target word was voiceless and the competition word was voiced did not cause many misperceptions (*bat – bad* (89.42%) and *bit – bid* (93.27%)). Word pairs where the target word was *bad* also had a relatively high percentage of correct identifications (*bad – bat* (92.31%) and *bad – bed* (87.50%)).

Figure 11

Accuracy in Identification of Word Pairs (in Percentages)



Chapter 4 Discussion

This thesis examined whether Dutch-English bilinguals transfer features from their native language Dutch to English. Specifically, the productions of the vowels /ɛ/, /ɪ/, /æ/ and final-consonant devoicing were explored. Dutch-English bilingual participants worked in pairs in a word-matching task. They interchangeably produced highly confusable words to their partner, while the other matched the word they heard their partner say. Their productions were acoustically analyzed to measure vowel and consonant durations as well as vowel quality. The results are used to decipher whether Dutch-English bilinguals transfer features of their native language Dutch to English and how this can lead to the misperception of certain words.

4.1 Vowel Quality

Vowel quality obtained through acoustic analysis shows the spectral differences between the vowel productions. Formant values relate to the vowel space, and indicate where the vowel was produced in the mouth. As discussed previously, formant values for males and females differ due to differences in the length and size of vocal tracts and the shape of oral cavities, which causes females and males to have different formant values. Due to this, formant values for males and females in this study are compared separately, but general trends in the formants remain the same. For both males and females, results show that /æ/ has the highest mean F1 value, followed by /ɛ/, then /ɪ/, and F2 has the mirrored order of /ɪ/, /ɛ/, then /æ/, respectively. The mean F1 and F2 values for the male participants can be compared to those of male English monolinguals published by Peterson and Barney (1952). Mean F1 and F2 values for each vowel examined in the current study were produced by the male Dutch-English bilinguals and English monolinguals from Peterson and Barney (1952), and the differences between the two groups are presented in Table 4.

Table 4

Mean Formant Values (in Hz) of Dutch-English Bilinguals and English Monolinguals

Vowel	Dutch		English		Difference	
	F1	F2	F1	F2	F1	F2
/æ/	705	1597	660	1720	45	123
/ɛ/	590	1759	530	1840	60	81
/ɪ/	397	1953	390	1990	7	37

Note. Using results from Peterson and Barney (1952), mean F1 and F2 values for vowels /æ/, /ɛ/, and /ɪ/ are compared between Dutch-English bilinguals (Dutch) and English monolinguals (English). The difference between the two types of speakers is in the final column.

The mean formant values for /ɪ/ produced by the native English speakers and the Dutch-English bilinguals are highly similar, which was unsurprising given that Dutch speakers tend to produce this vowel similarly to English. The mean amount of separation between the Dutch-English bilinguals' and the English monolinguals' production of /æ/ and /ɛ/ were surprisingly similar. The difference in F1 values of /ɛ/ between Dutch-English bilinguals' and English monolinguals' productions was higher than that of /æ/, but the difference in F2 values was greater for /æ/ than /ɛ/. Based on previous research, Dutch-English bilinguals tend to produce /ɛ/ similarly to the English production but struggle with the production of /æ/, so one would expect the difference between formant values of the production of /æ/ between the two speaker groups to have the biggest difference. As discussed in the Introduction (Section 1.5.1), Pols, Tromp, and Plomp (1973) reported the F1 and F2 values for Dutch speakers' productions of the vowels /ɛ/ and /ɪ/. The mean value of /ɛ/ they reported is very close to the mean value of /ɛ/ produced by the Dutch-English bilingual males in the current study. The reported mean of /ɪ/ is closer to the native English speakers' production but remains close to the mean male Dutch-English bilingual speakers' production.

Because this study examines the /æ/ – /ɛ/ contrast, the difference in formant values of these vowel productions in both the Dutch-English bilinguals as well as the native English speakers is important to examine. Because both /æ/ and /ɛ/ are present in the English vowel inventory, English speakers can clearly distinguish these vowels in production. Since /æ/ is not in the Dutch vowel inventory, it is expected that Dutch-English bilinguals may not have as much distinction and produce /æ/ more as /ɛ/. When comparing the difference between the Dutch-English bilinguals' production of /æ/ and /ɛ/, the F1 values differ by 115 Hz, and the F2 values differ by 62 Hz. When comparing these values in the native English speakers' production, F1 values differ by 130 Hz and F2 values by 120 Hz. The differences among the two male groups in F1 and F2 values are not large, meaning that the male Dutch-English bilinguals had a similar amount of separation in production between /æ/ and /ɛ/ compared to the native English speakers. These comparisons are based on the formant values of the males. Therefore, specific formant values cannot be compared between the English monolingual study (all males) and the female participants in the present study. However, the magnitude of the difference in formant values between /æ/ and /ɛ/ can be compared. Females had the smallest difference between mean F1 and F2 values between /æ/ and /ɛ/, where F1 had a difference between 97 Hz, and F2 had a difference of 87 Hz. Although the mean differences in the present study's male and female participants did not differ substantially from the differences in the native English productions, there was a great amount of visible spread in the formant value graphs for both males and females (see Figures 4 and 5).

In the F1-F2 graphs for the male (Figure 4) and female (Figure 5) speakers, productions of the vowel /ɪ/ are the most clustered (in both groups of speakers), indicating that this vowel was produced most similarly among participants and distinguished clearly from /æ/ and /ɛ/. Visually, however, there is considerable overlap in formant values of /æ/ and /ɛ/. This overlap shows that the Dutch-English bilinguals did not all systematically distinguish the productions of /æ/ and /ɛ/ in production, and /æ/ often had a vowel quality closer to /ɛ/. The vowel /ɛ/ is also relatively clustered, but the vowel /æ/ is much more dispersed. This indicates that the production of /æ/varied among the speakers, where some distinguished the production from /ɛ/, whereas others produced it more similarly. Participants who were more fluent in English had more distinction in the formant values of each vowel than less fluent participants (see Table 3, $r = .39$, $p < .05$). There are visible differences between formant value trends between participants with high accuracy (Figure 6) in perception and those with lower accuracy (Figure 7). Both participants (male) in the pair with the highest accuracy in perception (99%) showed clear distinctions between each vowel. However, participants with one of the lowest accuracies (80%) had separation for /ɪ/ productions but some overlap between /æ/ and /ɛ/ productions. Based on these mean results of vowel quality, Dutch-English bilinguals show a surprising amount of distinction between the productions of /æ/ and /ɛ/. However, the visual data shows considerable overlap between all productions of /æ/ and /ɛ/. The productions of /æ/ were much more varied than the other vowels, meaning that some participants did not have spectral differences between the production of /æ/ and /ɛ/, whereas some did.

4.2 Changes in Vowel Duration Based on Voicing of the Final Consonant

In addition to vowel quality, the difference in the production of vowel duration based on the voicing of the final consonant can be used to distinguish the production of different vowels. Previous studies have shown (House, 1961) that the vowel duration preceding a voiced final consonant is longer than preceding a voiceless consonant. This was also the case in the current study. Among all vowels (/ɛ, ɪ, æ/), the duration of the vowels preceding the voiced final consonant (/d/), on average, was longer than the voiceless final consonant (/t/). A noticeable feature of the vowel duration graph (Figure 8) is that the vowel duration in *bad* was significantly longer than the rest of the stimuli, and there was a substantial difference between the vowel duration in *bad* compared to *bat*. However, there was no considerable change in the vowel duration of /ɛ/ and /ɪ/ based on the voicing of the final consonant. These results can be compared to those of House (1961), who examined how vowel duration changes based on voicing the final consonant in English monolinguals. The study did not report durations for individual vowels, but rather all vowels in the vowel category based (tense vs. lax, open vs. closed). Because there is a significant difference in production between the lax open vowels in this study (/æ/ and /ɛ/), they are compared separately. Mean values in the House (1961) study were rounded to the nearest 10 ms, so mean values from the current study are also rounded. Table 5 displays the mean English vowel durations from the House (1961) study and the mean Dutch vowel durations from the current study.

Table 5

Mean English and Dutch Vowel Durations

	English		Dutch		
	Lax open (/ɛ/, /æ/)	Lax closed (/ɪ/)	Lax open (/æ/)	Lax open (/ɛ/)	Lax closed (/ɪ/)
Voiceless (/t/)	120 ms	100 ms	190 ms	170 ms	150 ms
Voiced (/d/)	220 ms	190 ms	280 ms	210 ms	180 ms
Difference	100 ms	90 ms	90 ms	40 ms	30 ms

Note. Using results from House (1961), mean vowel durations from English monolinguals (English) and Dutch-English bilinguals from this study (Dutch) are presented in ms. The difference between voiceless and voiced consonants for each language is presented in the final row.

The difference in vowel duration preceding a voiceless and voiced final consonant between the English production of lax open vowels and the Dutch production of /æ/ was similar. However, the difference in duration between /ɛ/ and /ɪ/ in the Dutch productions was substantially smaller than the difference in duration between the same vowel categories in English productions. These comparisons show that the Dutch-English bilinguals followed the native English duration when producing the lax open vowel /æ/, but did not when producing the lax open vowel /ɛ/ nor the lax closed vowel /ɪ/. Given that the voicing of the final consonant impacts the duration of the preceding vowel, and the vowel duration of /ɛ/ and /ɪ/ did not change much when the final consonant was voiced or voiceless, it indicates that the Dutch-English bilinguals did not change the vowel duration much based on the voicing of the final consonant.

Due to the lack of change in vowel duration, the final consonant /d/ was often devoiced. The vowel /æ/ is not found in the Dutch vowel inventory, so participants may have especially elongated the production of this vowel in *bad* to distinguish it from *bed* (which has a shorter vowel duration). Additionally, they may have elongated /æ/ to distinguish the production of /d/ from /t/.

4.3 Consonant Closure and Stop Release Durations

Because Dutch has final-consonant devoicing, it was hypothesized that Dutch-English bilinguals may devoice the final consonant at the end of a word in English. Before releasing a stop consonant at the end of a word, there is a buildup of oral pressure characterized by a moment of silence, called consonant closure. Determining whether the stop consonant at the end of the word was devoiced can be indicated by the durations of the consonant closure and the stop release. Voiceless consonants have a more extended moment of silence before the stop is released than voiced stops, meaning that /t/ has a longer consonant closure duration than /d/. Based on previous research by Stathopoulos and Weismer (1983), who measured consonant closure duration in English words ending in either voiced or voiceless consonants, voiced stops have a shorter consonant closure duration than voiceless stops. Their study showed a difference between consonant closure durations of 18 ms. However, the consonant closure durations of /t/ and /d/ in the current study are very similar, with a difference of only 4 ms. This similarity indicates that /t/ and /d/ may have often been produced with similar voicing. Both durations in the current study are close to the duration of a voiceless consonant, meaning that the consonant closure before the release of /d/ was similar to its voiceless counterpart /t/.

After the buildup during the consonant closure, the pressure is released to produce the consonant. Voiceless stops typically have a longer stop release duration than voiced stops, so in the current study, /t/ should have a more prolonged release duration than /d/. Crystal and House (1988) measured the release duration of /t/ and /d/ in the word-final position in connected speech in English native speakers, and found that /t/ had an 18 ms longer duration than /d/. In the Dutch speakers of English tested in current study the release duration of /t/ was 49 ms longer than /d/. The difference in duration in the current results is substantially longer, likely because in English (especially in connected speech) voiceless stops are generally not released when they are in a word-final position. However, in Dutch the voiceless final consonants are released. The Dutch-English bilinguals are unfamiliar with producing a /d/ in the word-final position in Dutch, so they may elongate the duration of /t/ to differentiate it from /d/ in production. Participants in the current study had similar consonant closure durations of /d/ and /t/ but followed the same stop release duration pattern that native English speakers do. If there was no notable difference between the release duration of /t/ and /d/, and /d/ had a release duration closer to that of /t/, it would be evident that the participants devoiced the final consonant. However, the durations contradict each other.

4.4 Perception

To find which phonological aspect in the production of Dutch-accented English led to the highest levels of misperception, the accuracy of each word pair presented in the study was broken down and summarized for all participants (see Figure 11). Words that had the highest percentage of correct perceptions were *bed – bid, bid – bed, bet – bit,* and *bit – bet*. Given that these words have the same final consonant, final-consonant devoicing does not play a role in these word pairs, and perception was based on the vowel. These pairs all have the vowels /ɛ/ and /ɪ/, indicating that the participants can distinguish these vowels well. As Collins and Mees (1984) stated, the Dutch and English productions of /ɛ/ and /ɪ/ are highly similar and typically not confused, which the current results also present. However, the /æ/ – /ɛ/ contrast caused more confusion. Misperception results indicate that the productions of the /æ/ and /ɛ/ were commonly confused with one another, and word pairs containing this contrast that was the most misperceived were *bed – bad, bat – bet*, and *bet – bat*. Target words containing both the vowels /æ/ and /ɛ/ were misperceived, indicating that there was not always a clear distinction in production among these vowels.

Devoicing of the final consonant led to the highest misperception among all word pairs. When a word pair has the same vowel, but the final consonant in the target word is voiced, and the competition word is not, the listener perceiving the vowel relies on the consonant closure and release durations. The word pair with the highest misperception was *bid – bit*, indicating that the /d/ in *bid* was often devoiced and perceived as a /t/. However, the word pair *bad – bat* also has the same vowel and a target word ending in /d/ but had a high accuracy percentage in identification. Interestingly, the /d/ in *bid* was perceived as devoiced, but the /d/ in *bad* was not. This mismatch in perception can be due to multiple reasons, one being that not all words ending in /d/ were devoiced. Furthermore, in addition to consonant closure and release durations, the produced vowel length of /æ/ in *bad* could have been used to distinguish /d/ from /t/. Word pairs with *bad* as the target word had high perception accuracies, so elongating the vowel distinguishes /æ/ from /ɛ/ and /d/ from /t/ was an effective strategy.

The Dutch-English bilinguals had an overall good ability to distinguish the target word from the competition word and did not have a word pair with an extremely low percentage of misperceptions. Some words used in the study may have elicited the Dutch-English bilinguals to use their native language to produce them. For example, *bad* is also a word meaning *bath* in Dutch, and one participant consistently pronounced this word with Dutch pronunciation. Additionally, the words *bit* and *bet* are both very "Dutch-like," as these sounds are found in Dutch, making them easy for Dutch-English bilinguals to say. This may have caused participants to produce them in a Dutch way instead of an English way, but because their partner is also a Dutch-English bilingual, they recognize the Dutch pronunciation of the word. Future research can examine whether native English speakers can distinguish the production of the Dutch accented English productions of all the words. For native English speakers, the Dutch-accented production of *bad* might sound much more like *bet* than for Dutch-English bilinguals.

The globalization of English may have also impacted the ability of Dutch-English bilinguals to produce English sounds in a native-like manner. The participants are exposed to English for a large portion of their day, as the classes at the university they attend are taught in English (although not necessarily taught by English native speakers), and the overall growing globalization of English has made English more prevalent in daily life. It is common for Dutch people to consume media in English, such as TV shows and movies, music, or podcasts. Some English words and phrases have been carried over to the Dutch language over the years that people in the current generation are accustomed to saying and hearing. Examples of this may be words like *iPad, online,* and *Instagram*, which contain speech sounds not found in the Dutch inventory. Because of this, young adult Dutch-English bilinguals may have learned to better distinguish certain sounds (like the /æ/ – /ɛ/ contrast) than earlier generations. By using context cues in conversation (i.e., using other words or phrases within the sentence that provide information about the meaning of the word), devoicing the final consonant of a word or pronouncing /æ/ like /ɛ/ will likely not cause major communication breakdowns between speakers and listeners. Still, the results of this study do show that Dutch-English bilinguals transfer features of their native Dutch phonetic inventory to English.

References

American Standards Association. (1960). Acoustical terminology SI. *1-1960, American Standards Association*.

Bent, T., Bradlow, A. R., & Smith, B. L. (2008). Production and perception of temporal patterns in native and non-native speech. *Phonetica*, *65*(3), 131–147. https://doi.org/10.1159/000144077

Boersma, P., & Weenink, D. (1996). Praat, a system for doing phonetics by computer, version 3.4. *Institute of Phonetic sciences of the University of Amsterdam, Report*, *132*, 182.

Byrd, D. (1993). 54,000 American stops. *UCLA working papers in phonetics*, *83*(1), 97-116.

Collins, B., & Mees, I. M. (1984). *The sounds of English and Dutch*. Brill Archive.

Crystal, T. H., & House, A. S. (1988). The duration of American-English vowels: An overview. *Journal of Phonetics*, *16*(3), 263–284. https://doi.org/10.1016/s0095-4470(19)30500-5

Flege, J. E. (1997). English vowel production by Dutch talkers: More evidence for the “similar” vs “new” distinction. *Second-Language Speech*, 11–52. https://doi.org/10.1515/9783110882933.11

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America*, *108*(6), 3013–3022. https://doi.org/10.1121/1.1323463

House, A. S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America*, *33*(9), 1174–1178. https://doi.org/10.1121/1.1908941

Lemhöfer, K., & Broersma, M. (2011). Introducing lextale: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Li, P., Zhang, F., Tsai, E., & Puls, B. (2013). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/s1366728913000606

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, *33*(1), 42. https://doi.org/10.2307/410949

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment*. Oxford University Press.

Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [R] and [l] by native speakers of Japanese and English. *Perception and Psychophysics*, *18*(5), 331–340. https://doi.org/10.3758/bf03211209

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. https://doi.org/10.1121/1.1906875

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America*, *32*(6), 693–703. https://doi.org/10.1121/1.1908183

Pols, L. C., Tromp, H. R., & Plomp, R. (1973). Frequency analysis of Dutch vowels from 50 male speakers. *The Journal of the Acoustical Society of America*, *53*(4), 1093–1101. https://doi.org/10.1121/1.1913429

Rietveld, T., Kerkhoff, J., & Gussenhoven, C. (2004). Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics*, *32*(3), 349–371. https://doi.org/10.1016/j.wocn.2003.08.002 Simon, E. (2010). Phonological transfer of voicing and devoicing rules: Evidence from L1 Dutch and L2 English conversational speech. *Language Sciences*, *32*(1), 63–86. https://doi.org/10.1016/j.langsci.2008.10.001

Slis, I. H., & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction I. *Language and Speech*, *12*(2), 80–102. https://doi.org/10.1177/002383096901200202

Stathopoulos, E. T., & Weismer, G. (1983). Closure duration of stop consonants. *Journal of Phonetics*, *11*(4), 395–400. https://doi.org/10.1016/s0095-4470(19)30838-1

Tops, G. A., Dekeyser, X., Devriendt, B., & Geukens, S. (2001). Dutch speakers. *Learner English: A teacher’s guide to interference and other problems*, 1-20.

van der Feest, S. V., & Swingley, D. (2011). Dutch and English listeners’ interpretation of vowel duration. *The Journal of the Acoustical Society of America*, *129*(3). https://doi.org/10.1121/1.3532050

Wang, H., & Van Heuven, V. J. (2006). Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers. *Linguistics in the Netherlands*, *23*(1), 237-248.

Academic Vita

Merel Padt

Education
The Pennsylvania State University
B.S. Communication Sciences & Disorders | B.A. German
Minor Human Development and Family Studies
Expected date of Graduation: Spring 2024

Honors
Paterno Fellow (Spring 2022 – present)
Schreyer Honors Student (Spring 2022 – present)

Research Experience
Speech Production Laboratory | Penn State University | PI: Dr. Jimin Lee
Undergraduate research assistant (Spring 2021 – present)
Work under the supervision of Dr. Jimin Lee on examining and developing strategies and procedures for the clinical management of dysarthria. Responsibilities include analyzing speech data using TF32 and Praat software for acoustic analysis.
National Science Foundation (PIRE) Fellowship | Penn State University
Undergraduate research assistant (Summer 2023 – present)
Work under the supervision of Dr. Annie Olmstead on conducting research and analyzing data collected at Radboud University in Nijmegen, the Netherlands.

Leadership & Involvement
Language and Linguistics Club | Penn State University
Programming Director (Fall 2021 – Spring 2023)
On the executive board for Language and Linguistics Club. Responsibilities include outreach and recruitment, planning and organizing club events, and managing all social media platforms.
President (Fall 2023 – present)
On the executive board for Language and Linguistics Club. Responsibilities include presiding at all meetings, facilitating executive board meetings, and being the spokesperson for the organization.

Scholarships and Awards
Paterno Student Enhancement Fund (Spring 2022)

Student Enrichment Award from the Department of Germanic and Slavic Languages and Literatures (Spring 2022)

Student Enrichment Award from the Department of Germanic and Slavic Languages and Literatures (Summer 2022)

College of the Liberal Arts enrichment award (Summer 2022)

CLS Judith Kroll Award (Fall 2022)

National Science Foundation (PIRE) Fellowship to conduct research project at Radboud University, Nijmegen, the Netherlands (Summer 2023)

Major Student Marshal for the German Major (Spring 2024)

Languages
English (near-native)
Dutch (native)
German (proficient)