Authored by **Mishko Bozhinoski**
Supervised by **Marijn van 't Veer**

The moet-moed split

An undocumented split in the province of Antwerp

2021

Preface

While we all stand on the shoulders of giants, rarely do we consciously acknowledge their presence. Today, I would like to do just that. The following people are the giants that I owe my accomplishments to, be it for this present study or my success in getting on the path to becoming someone I can refer to as a linguist:

Prof. Dr. Paul Boersma, whose illuminating insights directly shaped the trajectory of this project.

Dr. Marijn van 't Veer, who has not only ensured the smooth progress of this paper, but has also helped me maintain my curiosity and enthusiasm for phonology.

My family: my mother, father, brother, grandmother and uncle, to whom I owe my personal development, nurture and opportunities, and who have sacrificed their livelihoods to ensure that I can pursue what gives meaning to my life.

Asia Alfasi and Manel Vallet Buisan, whose kindness and generosity have directly made me mature as a person, and who ought to belong in the above.

and last but certainly not least,
My dearest Sarah Van Gelder, who inspired this undertaking, whose support, care and love I will be paying back for the rest of my life, and whom I now call my most important family.

Abstract

In contrast to other speakers of Dutch, Antwerpian speakers of the Flemish regional standard (tussentaal) show signs of a phonemic split, whereby the words *moet* and *moed* are no longer homophonous, but differ in the length of their vowel. This paper serves to ask and explore the following questions: (1) Is this phenomenon real? (2) If so, how can one account for it?

The pronunciation of 41 words containing /u/ of eighteen participants from the area of Antwerp was recorded and annotated using Praat. After analyzing the data using a Python implementation of Jenks natural breaks optimization, this study concludes that this distinction is real and is in fact phonemic. Further analysis reveals that the distribution of /u/ is in one part inherited from the Brabantian variety now spoken by a minority of Antwerp, and in one part the result of a shortening of /u/ before tautosyllabic /k/, /p/ and /j/.

1. Introduction

The Dutch-speaking area of provincial Antwerp seems to have undergone a process whereby what in the standard variety manifests as /u/ (spelled <oe>) appears in two forms: a short and a long variant (presumably [u] and [u:] respectively). This persists despite the traditional Brabantian variety getting displaced by the regional standard - also known as *tussentaal* (van der Sijs, 2019, p. 185). Despite the influence and popularity of the Antwerpian dialect (De Schutter, 1999), documentation as well as theoretical accounts of the feature prove to be scarce to non-existent[1]. This does not mean, however, that this feature has never been noticed. In 2009, User heerMat on the online forum dutchgrammar.com notes the following (translated from Dutch):

(1) Post by heerMat (2009) - translated from Dutch (original in Appendix)

> The theory says that <oe> is always a short vowel [u], and that the long [u:] appears before /r/.
> In practice, however, I do observe a distinction in my own pronunciation between a long and a short <oe>, elsewhere than before an /r/.
>
> Can anyone confirm this on the basis of their own experience, and potentially provide an explanation for it?
>
> In my case:
> - long: zoeken, boeten, zoet, roet, voet, roepen, doen, schoen, groen, loens, koen...
> - short: boeken, doeken, koeken, hoek, stoep, “tjoep”, groep…
>
> Actually, I find <oe> more often long than short. :)

Although it is difficult to guarantee phonetic awareness of individual speakers or judge it through text, it can be safely assumed that user heerMat is not talking about the Antwerpian variety of Brabantian, as the words “zoeken”, “zoet” and “groen” should then not contain [u:], but [y:] (Camerman, 2007), and the word “doen" should in fact be [duŋ], rather than [dun], as suggested by the spelling. He is rather referring

[1] One such scarce example being *De Schutter, 1999*, who mentions it casually and only in the context of the traditional Brabantian dialect

to the standard language, which has not undergone i-umlaut and has retained a back vowel [u] in the words listed above.

Perhaps one's first instinct would be to check for a direct etymological explanation of the two sounds. Both of these vowels, however, come from the same source in Proto-West-Germanic: the vowel */o:/. This includes both instances of short [u] ([buk] 'book' < OLF.[2] *buoc* < PWG */bo:k/) and long [u:] ([vu:t] 'foot' < OLF. *fuot* < PWG */fo:t/).

If this state of affairs is indeed a result of a split, there must be a conditioning environment that accounts for the distribution of the vowels. That environment, however, proves to not be easy to isolate[3]:

LONG			**SHORT**		
[zu:k]	'seek'	< PWG. **sōki*	[buk]	'book'	< PWG. **bōk*
[ru:p]	'call (out)'	< PWG. **hrōp*	[stup]	'pavement'	< Mnl.[4] *stoep*
[mu:t]	'courage'	< PWG. **mōd*	[mut]	'must'	< PWG. **mōt*
[plu:x]	'plow'	< PWG. **plōg*	[ɣənux]	'enough'	< PWG. **ganōg*

The vowels of each word appear in the same form in both Proto-Germanic, in which it is */ɔ:/, and Middle Dutch, in which it is orthographically represented as <oe>.

1.1 Hypotheses

If there is no directly evident etymological or environmental account of the distribution of these two sounds, it is best to test multiple plausible hypotheses motivated by similar processes in other languages, other productive processes in Dutch, and in general to analyse the phonological environment in which these two sounds occur. This study examines four such accounts.

1.1.1 Voicing

It is widely attested in many languages that voiced consonants tend to be preceded by longer vowels (Kluender et al., 1988). In some languages such as English, vowel length is sufficient to perceptually distinguish voiced and voiceless consonants (Denes, 1955; Raphael, 1972; Port & Dalby, 1982). On the face of the above, it is possible that (underlying) voicing has an effect on vowel length. The pair <moet>

[2] Old Low Franconian. All etymologies come from *etymologiebank.nl* (Retrieved 2021)
[3] The examples listed are based on the forum post mentioned hereinabove, confirmed and further enlarged by a female informant of age 28 who lives in the area. They are not to be taken as definitive.
[4] Middle Dutch (from *Middelnederlands*). All etymologies come from *etymologiebank.nl* (Retrieved 2021)

[mut] *'must'* and <moed> [mu:t] *'courage'* differs not just in vowel length but underlying voicing of the final consonant, which suggests the potential for voicing as a cue.

1.1.2 Syllable structure

As in all West Germanic languages (Prokosch, 2009, p.140), syllable structure is intimately related to vowel length in Dutch. More specifically, (historically) open syllables contain long vowels while (historically) closed syllables contain short vowels (Linke, 2020c). This connection is strong enough that there have been accounts positing that underlying syllabic structure is sufficient to account for the distribution of long vs. short vowels in Dutch (Botma & van Oostendorp, 2012). Due to the limited scope of this study, only words of structure (s)CCVC are taken as targets.
This is further discussed in section 5.

1.1.3 Place and manner

If this feature is the result of a sound change, it is important to examine other environmental components as potential causes, such as place and manner. As the distinction seems to have a primary length component, and moraic structure is not affected by syllable onsets (Auer, 1989), it is best to assume that syllable codas are more likely to have triggered a sound change than syllable onsets. As the scope of this study is limited, the effect of onsets will not be looked into, but could be grounds for future research.

1.1.4 Connection between etymology and neighboring cognates

Further insights can be gained by looking at the historical forms each word appears in in Proto-West-Germanic alongside the synchronic forms in Antwerp Brabantian, as well as surrounding variants such as Limburgish and North Brabantian. While the traditional Brabantian variety spoken in Antwerp is now spoken by a minority of speakers in Antwerp (further discussed in section 2), it is possible that this feature is accounted for by its influence on the regional standard variety.

1.2 Goal

The goal of the present work is two-fold. Firstly and most primarily, it is to document the existence of the phenomenon: is there indeed a salient length difference? Secondly - if such a length contrast exists, how could one account for its distribution?

2. Literature review

2.1 Dutch (Booij, 1999)

Dutch is a West Germanic language, spoken natively by around 24 million people, chiefly spread across the Netherlands (~17 million speakers) and the Flemish Region of Belgium (~6.5 million people). Historically, the areas in which Dutch is spoken were inhabited by speakers of a range of Low Franconian and Low Saxon varieties the native Dutch call "dialects". Presently, however, younger people speak a much more unified "Dutch" language, with each region speaking a version that has been "colored" by their historical dialects, despite the dialects themselves being spoken by an increasingly smaller minority of speakers.

Below is a broadest summary of the phonology of Dutch, as well as some more precise points which are relevant to the present study.

2.1.1 The consonants

Figure 1 - The consonants of Standard Dutch

	Bilabial	**Labiodental**	**Alveolar**	**Palatal**	**Dorsal**	**Glottal**
Plosive	p b		t d		k (g)	
Fricative		f v	s z		x/χ ɣ	h
Nasal	m		n		ŋ	
Liquid			l r			
Glide		ʋ		j		

/p/, /t/ and /k/ are voiceless and unaspirated

/b/, /d/ and [g] are fully voiced

Speakers with a uvular realization of the dorsal fricative do not have a voiced counterpart

Obstruents are devoiced word-finally (ex. [hɛbə(n)] 'to have' but [(ɪk) hɛp] '(I) have')

2.1.2 The vowels

Dutch vowels can be categorized as such[5]:

Figure 2 - The vowels of Standard Dutch

Short vowels	ɪ ɛ ɔ ʏ ɑ
Long vowels	i y u eː øː oː aː
Schwa	ə
Diphthongs	ɛi œy ɔu

While underlyingly belonging to the class of long vowels, the high vowels /i/, /y/ and /u/ are realized as short in most regiolects.

Flemish speakers have a tendency to realize /ɛi/, /œy/, and /ɔu/ as monophthongs, whereby they are realized as a lengthened version of the first element: [ɛː], [œː], and [ɔː].

2.1.3 Phonotactics

Dutch syllable structure can be summarized by the following template: (C)(C)V(V/C)(C). Tautosyllabic clusters follow the Sonority Sequencing Principle, whereby members of a syllable must be progressively more sonorous as they approach the syllable peak. In other words, complex onsets must have rising sonority (e.g. /pl-/ and /tr-/) and complex codas must have falling sonority (e.g. /-mp/, /-nk/). Complex codas cannot co-occur with tautosyllabic long vowels, thus /dɑmp/ is a possible word, while /daːmp/ is not.[6]

Relevant to this study is the correlation between vowel length and the voicing of fricatives. With very few exceptions, long vowels and diphthongs may only be followed by (underlyingly) voiced fricatives, while short vowels may only be followed by voiceless fricatives (Linke, 2020a). This is exemplified hereunder:

(2) Example 1

Long vowels + voiced fricatives	*Short vowels + voiceless fricatives*
/ɦøːvəl/ ‘hill’	/knʏfəl/ ‘hug’
/veːzəl/ ‘fiber’	/tʏsə(n)/ ‘between’
/voːɣəl/ ‘bird’	/lɑxə(n)/ ‘laugh’

[5] “Long” and “short” have also been referred to as “tense” vs “lax” and “A-class” vs “B-class” (Linke, 2020c)
[6] Very rare exceptions exist, such as the word /tʋaːlf/ ‘twelve’.

One rare exception to the rule is the loanword /pʏzəl/ ‘puzzle’, which many speakers pronounce as [pyzəl]. This further demonstrates the preference for long vowel + voiced fricative sequences.

2.2 The language of Antwerp

2.2.1 ‘t Aentwaerps (De Schutter, 1999)

The local dialect of provincial Antwerp (‘Antwerp Brabantian’, henceforth also AB) is part of the Brabantian branch of Low Franconian (Triest, n.d.), alongside the regions *Vlaams Brabant* and *Noord-Brabant*, the former located to the south of Antwerp and the latter located to its north, where Belgium borders with the Netherlands.

The consonant inventory of AB is almost identical to that of Standard Dutch (differences bolded):

Figure 3 - The consonants of Antwerp Brabantian

	Bilabial	**Labiodental**	**Alveolar**	**Palatal**	**Velar**	**Glottal**
Plosive	p b		t d		k (g)	
Fricative		f v	s z		**x ɣ**	
Nasal	m		n		ŋ	
Liquid			l r			
Glide	**w**			j		

Unlike Standard Dutch, there is a tendency to realize the velar /k/ /x/ /ɣ/ and /ŋ/ as palatal or strongly palatalized.

The vocalic system of AB differs significantly from that of Standard Dutch:

Figure 4 - The consonants of Antwerp Brabantian

Short vowels		
	Front	Back
Close	i y	u
Mid	e ø	o
Open	a	

Long vowels		
	Front	Back
Close	i: y:	u:
Mid	ɛ: œ:	o:
Open	a:	ɒ:

Diphthongs		
	Front	Back
Close	iə	uə
Mid	ɛə	oə
Open	æə	aə

Notable for this dialect group is its participation in historical i-umlaut, a form of progressive vowel harmony whereby [i] and [j] caused the vowel of the previous syllable to front. Standard Dutch does not show these sound changes. Examples of this phenomenon are presented hereunder:

(3) Examples: i-umlaut

	AB	Standard Dutch	Proto-West-Germanic	English
_____	[ɣryːn]	[ɣrun]	*grōnī	green
	[vyːlə(n)]	[vulə(n)]	*fōlijan	feel
cp.	[ɣuː(t)]	[ɣut]	*gōd	good

Despite its historical presence in AB, it is synchronically no longer productive.

2.2.2 Current status of Antwerp Brabantian

The traditional dialect of Antwerp is commonly taken as the dominant dialect of Flanders, widely understood by everyone in Flanders and commonly represented in the media (De Schutter, 1999). This is humorously referenced in the Antwerpian expression *Aentwaerps is een wêreldtaal* "Antwerpian is a world language" (Camerman, 2009).

Despite its status, increasingly fewer inhabitants of provincial Antwerp speak the Brabantic dialect in actuality. In her work *15 eeuwen Nederlandse taal*, Nicoline van der Sijs (2019, p. 187) points this decline out. In 1979, 91% of the population of Antwerp claimed to be able to speak the local Brabantian dialect well. This is juxtaposed to the data of 2014, where this number was only 39%.

The variety that is displacing Antwerp Brabantian is what the local Flemings commonly refer to as *tussentaal*, literally "inter-language". This is a version of Standard Dutch that has been colored by the local varieties historically spoken by the population of each Flemish region (van der Sijs, 2019, p. 185).

3. Method

3.1 Pilot

In order to probe for potential difficulties and inconsistencies, a pilot study was conducted on 5 participants. The pilot study was identical to the main experiment, which is described below.

3.2 Recruiting

The recruiting of participants was carried out online. More precisely, the study was promoted on the online platform Reddit - specifically the subreddits r/Belgium, r/Vlaanderen and r/Antwerp (URLs in references) and multiple Flemish speaking Discord communities. The promotional post on both communities was identical (Appendix A). The post asked participants only to apply if they come from the province of Antwerp.

Each participant received an informational brochure (Appendix B) laying out the experiment. Once each participant felt informed, their experiment was scheduled on one of the days following recruitment and was given a consent form (Appendix B), which they were instructed to read carefully and sign before the date of the experiment.

To anonymize the data, each participant was given a code consisting of two random letters followed by two random numbers (e.g. AF53). The participants' answered questionnaire and recorded data were thereonforth referred to only by this code.

3.3 Questionnaire

Before the experiment, each participant was presented a questionnaire (Appendix E). The questionnaire asked for the following:

- Age
- Gender
- Their native language or languages
- Whether the participant grew up in the province of Antwerp
- Whether the participant still lives in Antwerp
- Whether the participant has lived outside of Antwerp for more than 5 of the last 10 years
- Whether the participant suffers from a language disorder, or a disorder that can have direct influence on language acquisition, such as deafness

3.4 Experiment

3.4.1 Stimuli

The stimuli consist of 3 monosyllabic words per possible coda consonant, all of which contain /u/ (spelled as <oe>). As there are too many possible etymological components to account for, etymology was not taken into account during word-selection.

The underlying form of a phoneme was determined by its form when the infinitival or plural morpheme /-ən/ is present. Thus, the words <hoed>, <vroeg>, <groef> and <hoes> were said to end in /d/, /ɣ/, /v/ and /z/ despite their surface forms ending in [t], [x], [f] and [s] respectively (Linke, 2020a). The phoneme /ɦ/ cannot appear in the syllable coda (Sebregts, 2020), and no words were found for underlying /s/, /b/, /ŋ/ or /x/. That leaves the following possible codas, with their respective words in Figure 5.

Figure 5 - Target words by final consonant

∅	/t/	/d/	/k/	/p/	/m/	/n/
moe	moet	moed	boek	groep	bloem	doen
koe	groet	hoed	zoek	stoep	roem	groen
hoe	stoet	voed	doek	snoep	doem	zoen
/z/	/v/	/f/	/ɣ/	/r/	/l/	/j/
hoes	hoef	sloef	vroeg	boer	koel	gloei
bloes	boef	stoef	boeg	hoer	boel	boei
moes	groef	toef	kroeg	stoer	boel	roei

The fillers were the following monosyllabic words: haas, kaas, pel, boot, gooi, baan, ben, dij, zij, pijl, kop, lach, hut, peuk, geur, zeug, dek, maan.

The target words (42) and fillers (18) together add up to 60 words, and the stimuli set contains each 3 times, which adds up to 180 stimuli.

Each target word was presented using the carrier sentence *Schrijf het woord _ eens op*[7].

[7] English “Write the word _ down”

3.4.2 Procedure

Most participants were recorded on the online communication platform Zoom (Zoom Video Communications, Inc.). One participant preferred to use Skype (Skype Technologies S.A.R.L), and another preferred to use Discord (Discord Inc.). The recording itself was done using the audio editing software Audacity (The Audacity Team), and each participant used their own microphone to record themselves.

Each participant was presented with the stimuli in randomized order. The participants were instructed to read each individual sentence out loud, despite only word one changing between sentences. After 60 and 120 stimuli, the participant was given the option to take a break.

The presentation of the stimuli was carried out using a Praat script, which is presented in Appendix D. The script randomizes the order of inputted stimuli and shows the word "PAUZE" (English *break*) after the 60th and 120th stimulus in order to signal when the participant may take a break. The Praat script additionally outputs a text file whose name is the unique code of the participant (mentioned in section 3.2). The text file contains all stimuli in the order they were presented to the participant.

3.4.3 Post-experiment interview

After the experiment, each participant was asked a few questions based on the preliminary impressions of the data. An example of such questions is: "You pronounced 'zoek' one way, and then another. Do both sound equally natural? Do they both mean the same?". The answer to those questions was noted down.

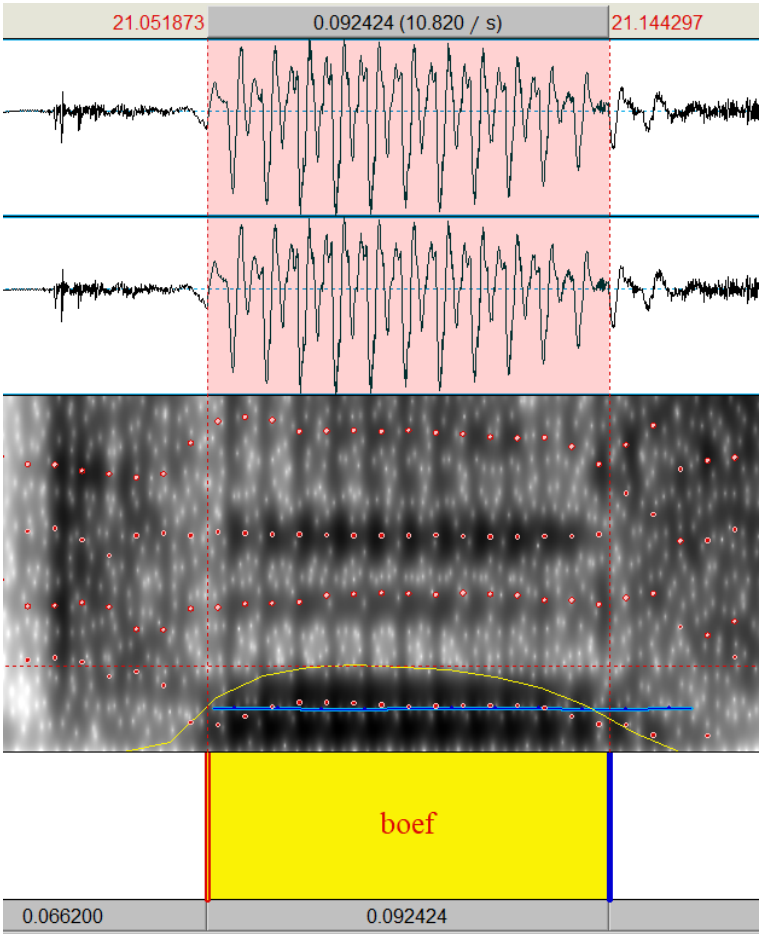
It is important to note that this interview was quite informal and is based on the participants' own impressions. It is therefore not to be taken as part of the main dataset and only serves as documentation upon which further ideas and research may be developed.

4. Analysis

4.1 Annotation

After the data collection period, each audio recording was annotated using Praat into a TextGrid file. The annotations served to mark the start and end of the vowel in each target word. The label of each annotation was the target word, as exemplified by Figure 6:

Figure 6 - Waveform, spectrogram and TextGrid annotation, ordered from top to bottom.



After annotation, each recording was run through a Praat script (Appendix J), in which the annotations were extracted and outputted into one file per recording. The format of the output file was as follows:

<label of annotation 1> <length of annotation 1 in seconds>

<label of annotation 2> <length of annotation 2 in seconds>

<label of annotation 3> <length of annotation 3 in seconds>

...

A real example from the data:

(4) Example: data as extracted from Praat

kroeg 0.12843682181022587

zoek 0.05643932314367106

stoet 0.09154571467168893

…

This data was fed into a Python script (Appendix F), which is explained further in detail in the section below.

4.2 Statistical analysis

Further statistical analysis was performed using the scripting language Python, more precisely Python 3.9 (Van Rossum & Drake, 2009). The script used to perform said analysis is included in Appendix F.
As each target word appeared three times in the stimuli, each occurrence was first averaged out in order to get one value per target word. Each target word was averaged out using an arithmetic mean.

4.2.1 Problems and solutions

The way the statistical analysis was to be approached was complicated by three difficulties, namely:

1. The problem of uneven representation: Long vowels were overrepresented in the dataset. In other words, the data set contained more long vowels than short ones. This means that an arithmetic mean would be biased towards long vowels.

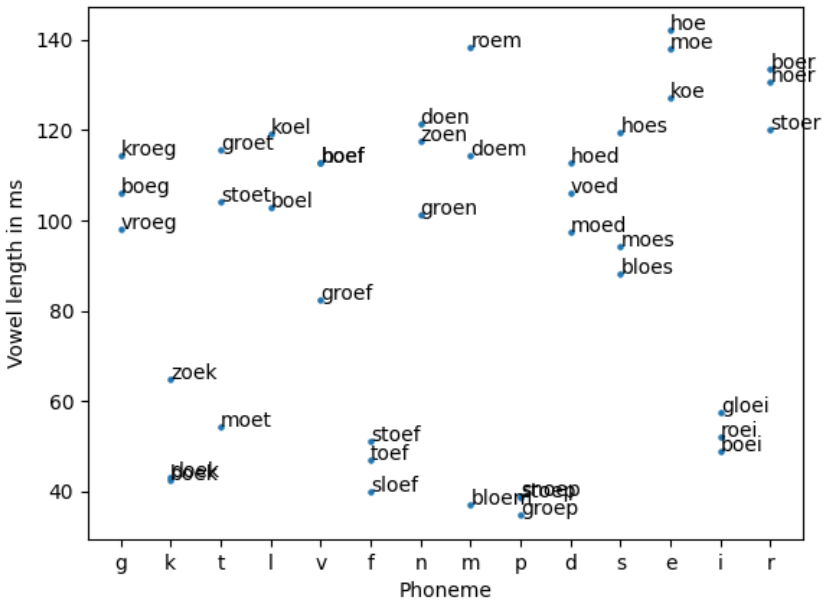
2. The problem of class inconsistency: Some words occurred both with long and short realizations. This varied both between participants and within participants. Most especially, but not exclusively, the words *zoek* and *snoep* occurred both with a long and short vowel, not only between participants, but sometimes within the same participant.

3. The problem of uneven deviation: Long vowels showed much greater length variation than short vowels. In other words, the difference between the shortest long vowel and the longest long vowel was much larger than the difference between the shortest short vowel and longest short vowel. This means that even if there was even representation of long vs short vowels, the arithmetic mean would be biased towards long vowels.

The combination of these three factors meant that a simple arithmetic mean could not be used to distinguish short from long vowels; i.e. one cannot take the arithmetic mean of the dataset and define short vowels as those shorter than the mean, and long vowels as those longer than the mean. The solution to all of these problems is discussed in sections 4.2.2 and 4.2.3.

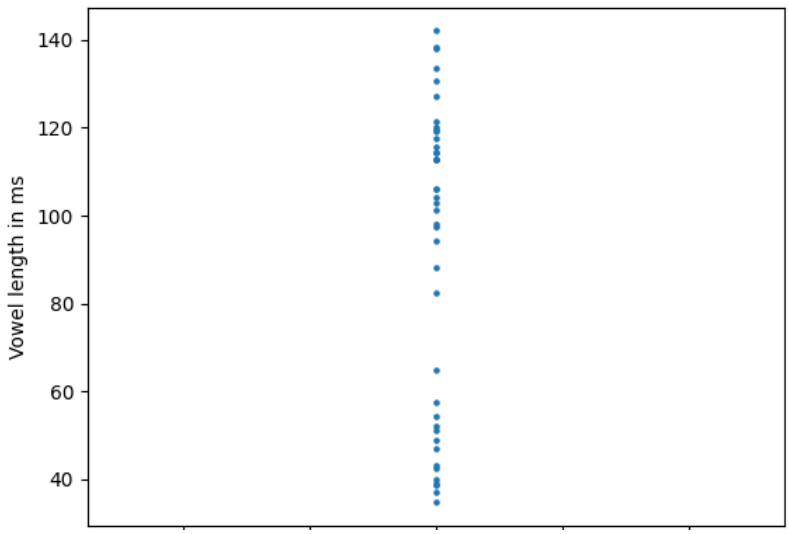
If we briefly disregard the strict formalization of the distinction between long and vowels, it is possible to look at the data impressionistically and see what stands out. Below is the a graph of all token words of one 27-year-old female participant in the study:

Figure 7 - Y axis is vowel length in ms, X axis is the coda phoneme



While this representation of the data may not show a clear pattern, collapsing the data such that all data points are represented on one vertical line reveals the following:

Figure 8 - Collapsed graph of Figure 7



Impressionistically, the graph shows a bimodal distribution in density: there are two center points on the Vowel-length axis around which vowels cluster and get progressively denser towards. Additionally, there is an observable gap between around 65ms and 85ms.

4.2.2 Formalization and the solution to the problems of representation and deviation

While the impressionistic view suggests a bimodal distribution with a gap between long and short vowels, it is important to formalize such distributions as consistently as possible. To solve the previously introduced problems 1 and 3, a consistent method of discriminating between short and long vowels must be used.

A potential rigorous solution is to use a data clustering method called *Jenks natural breaks optimization* (Jenks, 1967), whereby one-dimensional data can be grouped into a given number of clusters by looking for so-called *natural breaks* in the data. The algorithm, in the simplest of terms, works as such:

1. Divide the data into an N amount of arbitrary groups (called *classes*).
2. Calculate the sum of squared deviations[8] between classes (henceforth σ^c).
3. Calculate the sum of squared deviations from the mean of the dataset (henceforth σ^a)
4. Calculate σ^c-σ^a; this is the sum of the squared deviations from the class means (henceforth $\Delta\sigma$).
5. Expand or shrink groups (by transferring items from one group to another) until the lowest possible value of $\Delta\sigma$ is achieved.

A Python implementation of this algorithm can be found in Appendix G (MacWright, 2016).

Applying the above algorithm requires specification of the amount of classes the dataset is to be divided in. As the algorithm will always divide the data into a given amount of classes, it is necessary to justify why the data is to be divided into two. In other words, the assumption that the words can be divided into ones having a “short” versus a “long” vowel must be justified. One such way could be the following:

1. For “short” and “long” to exist as categories, one must assume that vowel length is a distinctive property.
2. If there exist two words that differ in nothing but vowel length, then it follows that length is a distinctive property.

If two such words were to exist in our dataset, this would be enough to justify dividing the set into classes. Since three-way vowel length distinctions are exceedingly rare, it is safe to assume that the

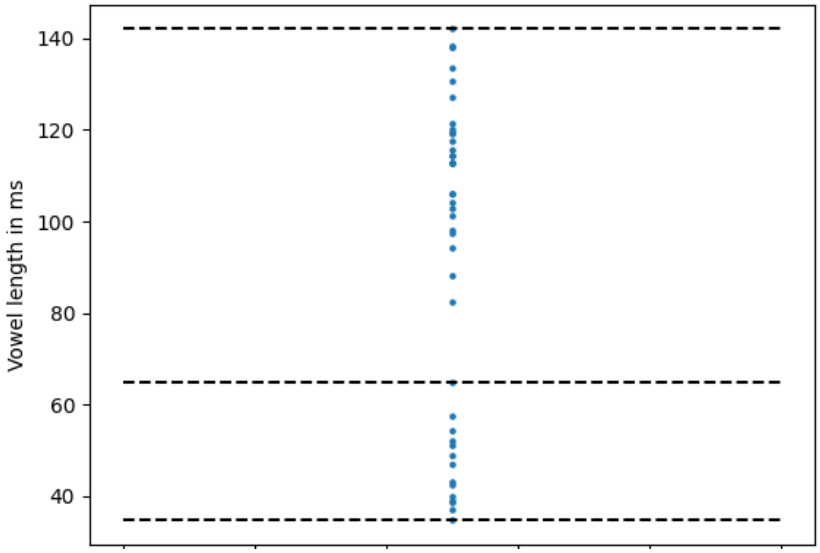
[8] The sum of squared deviations is the sum of the squares of the difference between each data point and the average of the dataset. Formalized as $\Sigma(x_i-\bar{x})^2$

presence of a vowel length distinction implies a bilateral short-long distinction. The impressionistic examination supports this as well.

The data does indeed show such a pair: <moet> [mut] vs <moed> [mu:t]. Across all participants, the average length of <moed> is 110ms, while the average length of <moet> is 53ms. Therefore, length must be a distinctive category, and the data can therefore be divided into "long" and "short" groups.

Finally, applying the Jenks algorithm to the dataset presented in Figure 8 yields the graph of Figure 9:

Figure 9 - Figure 8 with cluster analysis (striped lines represent natural breaks)



4.2.3 Solution to the problem of class inconsistency

As of now, the analysis does not account for the categorical length disparity individual words show: while the majority of words were consistent in whether or not they were long or short, an impressionistic look at the data already suggests that some words appeared both as long and short. To account for this problem, the following measures were taken:

1. A point is defined between the highest value of the "short" cluster and the lowest value of the "long" cluster. This is effectively the midpoint of the "gap" between the two clusters. This leaves us with a partially arbitrary but serviceable cutoff point.
2. Each token is defined as "long" or "short" depending on whether or not it lies above or below the cutoff point.

3. The distance of each token from the cutoff point is recorded. This is to prevent certain clusters of tokens to be categorized as differing in length when in fact they are all situated around the cutoff line. These words will be marked as being ambiguous.

The absolute deviation of each target word's tokens is not useful, as length variation increases rapidly with average vowel length.

4.2.4 Application of analysis

Figure 10 shows an example of the analysis applied to the same dataset as in Figures 7-9. +X represents Xms above the cutoff line, -X represents Xms below the cutoff line. To make it easier to read at a glance, values above the cutoff line were bolded.

Figure 10 - Example of processed length data

Participant code: IG27				Cutoff length: 73.70ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	**+39 +49 +104**	moet	-30 -17 -11	moed	**+12 +28 +31**	boek	-41 -31 -22
koe	**+28 +63 +69**	groet	**+31 +42 +54**	hoed	**+23 +37 +57**	zoek	-17 -16 **+8**
hoe	**+64 +69 +72**	stoet	**+18 +21 +52**	voed	**+24 +35 +40**	doek	-33 -30 -28
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-41 -39 -36	bloem	-42 -36 -32	doen	**+25 +49 +69**	hoes	**+26 +55 +57**
stoep	-39 -36 -30	roem	**+58 +60 +76**	groen	**+21 +27 +35**	bloes	-21 **+29 +35**
snoep	-45 -41 -19	doem	**+30 +41 +51**	zoen	**+27 +44 +61**	moes	**+8 +27 +28**
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	**+26 +35 +56**	sloef	-38 -36 -26	vroeg	**+22 +23 +27**	boer	**+33 +72 +74**
boef	**+25 +46 +46**	stoef	-26 -25 -16	boeg	**+30 +32 +34**	hoer	**+46 +61 +64**
groef	**+1 +9 +16**	toef	-36 -23 -22	kroeg	**+33 +35 +55**	stoer	**+43 +45 +51**
/l/				/j/			
Word	Distance			Word	Distance		
koel	**+37 +39 +60**			gloei	-25 -24 -1		
boel	-17 **+30 +34**			boei	-31 -30 -14		
				roei	-33 -23 -9		

Next, the data of each participant is divided into three categories:

- Words whose tokens were always above the cutoff line (labelled "consistently long")
- Words whose tokens were always below the cutoff line (labelled "consistently short")
- Words of which some of the tokens below, and some above (labelled "inconsistent")

In Figure 11 is a table with the data presented in the above manner. The rest of the participants' data is in Appendix H.

Figure 11 - Words divided into classes based on length

Participant code: IG27		
Consistently long	**Consistently short**	**Inconsistent**
moe; koe; hoe groet; stoet moed; hoed; voed roem; doem doen; groen; zoen hoes; moes hoef; boef; grief vroeg; boeg; kroeg boer; hoer; stoer koel	moet boek; doek groep; stoep; snoep bloem sloef; stoef; toef gloei; boei; roei	boel bloes zoek

4.2.5 Overview

Finally, once the above steps are applied to all participants, what is left is to get a full overview of all words. This is how this was done:

1. Select a word from the target word list.
2. Iterate through each participants' processed data (i.e. words divided into "consistently short", "consistently long" and "inconsistent").
3. Tally the amount of times each word appeared as consistently long, consistently short or inconsistent.

The final data is represented in a table as the one in Figure 12:

Figure 12 - Example of final data

∅				/t/				/d/				/k/			
Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I
moe	18	0	0	moet	0	18	0	moed	17	0	1	boek	0	18	0
koe	18	0	0	groet	18	0	0	hoed	16	0	2	zoek	4	4	10
hoe	18	0	0	stoet	11	1	6	voed	18	0	0	doek	0	18	0

Where "Wrd" represents the target word, L represents "amount of consistently long occurrences", S represents "consistently short occurrences" and I represents "inconsistent occurrences". For example, if a word has L=5, S=10 and I=3, that means that 5 participants pronounced it consistently long, 10 pronounced it consistently short and 3 pronounced it inconsistently (i.e. sometimes long, sometimes short). The data for all target words of the study is given in the next section.

5. Results

5.1 Participants

The total number of people participating in the study was 21. Out of those, the data of three was unusable: two of them grew up outside of Antwerp and one knew what the topic of the study was. That leaves 18 participants whose data is usable.

Out of those 18 participants, 4 were women and 14 were men. The age range of all the participants is 18 to 57, with a mean of 28.61 and median of 27.5.

5.2 Word analysis

Below is the overview of all target words represented as per section 4.2.5. Individual speakers' distributions are presented in Appendix H.

Figure 13 - overview of all words

∅				/t/				/d/				/k/			
Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I
moe	**18**	0	0	moet	0	**18**	0	moed	**17**	0	1	boek	0	**18**	0
koe	**18**	0	0	groet	**18**	0	0	hoed	**16**	0	2	zoek	**4**	**4**	**10**
hoe	**18**	0	0	stoet	**11**	**1**	**6**	voed	**18**	0	0	doek	0	**18**	0
/p/				/m/				/n/				/z/			
Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I
groep	0	**18**	0	bloem	0	**18**	0	doen	**18**	0	0	hoes	**18**	0	0
stoep	0	**18**	0	roem	**18**	0	0	groen	**18**	0	0	bloes	**13**	0	**5**
snoep	**5**	**9**	**4**	doem	**18**	0	0	zoen	**18**	0	0	moes	**12**	**2**	**4**
/v/				/f/				/ɣ/				/r/			
Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I	Wrd	L	S	I
hoef	**17**	0	1	sloef	0	**17**	1	vroeg	**18**	0	0	boer	**18**	0	0
boef	**17**	0	1	stoef	0	**18**	0	boeg	**17**	0	1	hoer	**18**	0	0
groef	**16**	0	2	toef	0	**17**	1	kroeg	**16**	0	2	stoer	**17**	0	1
/l/								/j/							
Wrd	L	S	I					Wrd	L	S	I				
koel	**16**	0	2					gloei	0	**17**	1				
boel	**15**	0	3					boei	0	**18**	0				
								roei	0	**18**	0				

6. Discussion

6.1 Interpretation of results

Looking at the data above at face value, one can say the following:

6.1.1 Long vowels

The following words were consistently long for all participants:

moe, koe, hoe, groet, voed, roem, doem, doen, groen, zoen, hoes, vroeg, boer, hoer

The following words were consistently long for the vast majority of speakers (at least 15), and pronounced inconsistently by at most 2 of them

moed, hoed, hoef, boef, groef, boeg, kroeg, stoer, koel, boel

6.1.2 Short vowels

The following words were consistently short for all participants:

moet, boek, doek, groep, stoep, bloem, stoef, boei, roei

The following words were consistently short for 17 of the participants, and inconsistent for 1:

sloef, toef, gloei

6.1.3 Other

The following words showed significant variation, both across- and within-speakers:

zoek was long for 4 participants, short for 4, inconsistent for 10

snoep was long for 5 participants, short for 9, inconsistent for 4

The following words lean strongly towards long or short, but had non-neglectable variation:

stoet was long for 11 participants, short for 1, inconsistent for 6

moes was long for 14 participants, short for 2, inconsistent for 4

bloes was long for 13 participants, short for 0, inconsistent for 5

6.2 Hypotheses

In section 1, we introduce possible accounts of the distribution of the long and short /u/. They can be summarized as follows:

- The (underlying) voicing of the following consonant
- The syllabic structure of the word (eg. open and closed syllables, complex codas, etc.)
- Phonological environment - most importantly the nature of the following consonant
- Etymology - look into corresponding cognates in modern Antwerp Brabantian and other surrounding dialects, as well as the corresponding words in older forms, such as Proto-West-Germanic.

Below are more detailed examinations of the above mentioned hypotheses.

6.2.1 Voicing

As discussed in section 2.1, Dutch exhibits final obstruent devoicing. As the experiment of the present study only covered words with the structure XV(C), this means that the full scope of the effects of voicing on vowel length remains unexamined. It is possible, however, to make some statements about underlying voicing and vowel length.

As briefly discussed in section 3.4, this study maintains that word-final obstruents that have been devoiced are underlyingly voiced.

While phonetically voiced, sonorants are not underlyingly specified for voicing (Visser, 2020). Therefore, any effects of underlying voicing should only be apparent in obstruents.

Below is a table dividing obstruent-final words with respect to vowel length as present in the experimental data. The inconsistent words *zoek* and *snoep* are listed in both groups, but underlined.

Figure 14 - overview of words with final obstruents

Plosives			
Long		**Short**	
Final C voiced	**Final C voiceless**	**Final C voiced**	**Final C voiceless**
voed, moed, hoed	groet, stoet*, zoek, snoep		moet, boek, doek, zoek, groep, stoep, snoep
Fricatives			
Long		**Short**	
Final C voiced	**Final C voiceless**	**Final C voiced**	**Final C voiceless**
hoes, vroeg, boef, groef, boeg, kroeg, moes*, bloes*			stoef, sloef, toef

*Long for most speakers, but with at least 5 deviant speakers

In plosives, the distribution of long vowels is not accounted for by the distribution of final voiced obstruents. Bar the words *zoek* and *snoep* (which are further discussed in sections 5.2.2 and 5.3), it is important to note that long [u:] before underlyingly voiceless plosives is limited to before /t/, whose atypical structural distribution is to be acknowledged, such as its word-final extrasyllabicity (Booij, 1999). Neither short [u] nor long [u:] was found before /b/, and /g/ is at most a marginal phoneme in Dutch.

Before fricatives, the experimental data showed long [u:] appearing exclusively before voiced fricatives[9] and short [u] appearing exclusively before voiceless fricatives. One should however be careful drawing causal conclusions, as there is a strong tendency for all long vowels to precede underlyingly voiced fricatives, and all short vowels to precede underlyingly voiceless fricatives (see section 2.1.3). It is unclear if long vowels somehow affect the voicing of those fricatives, if fricative voicing somehow affects the length of the preceding vowel, or if any causality is to be inferred in the first place.

[9] Note that the words ‘hoes’, ‘bloes’, ‘moes’ and ‘groef’ end in underlyingly voiced consonants.

6.2.2 Syllable structure

Not much can be said about the relationship between syllable structure and the length of /u/ on the basis of the experimental data alone. The post-experimental interviews do give us some preliminary insight into potential directions the nature of that relationship may go in:

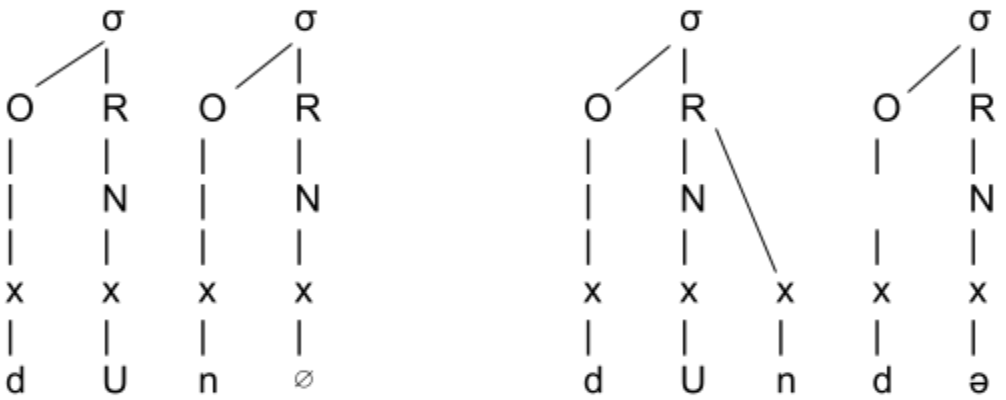
- Out of the participants that did not pronounce the words *zoek* and *snoep* consistently long, all but one reported that they would always pronounce the words *zoeken* 'to seek' and *snoepen* 'to snack (on candy)' with a long vowel. One participant claimed to pronounce *snoepen* with a short vowel, but that pronouncing it with a long vowel "sounds fine too".
- Almost all participants whose pronunciation of *zoek* and *snoep* was either short or inconsistent expressed the nouns *zoek* 'searching' and *snoep* 'candy' were short, while the verbs *(ik) zoek* '(I) search' and *(ik) snoep* '(I) eat candy' had a long vowel. One participant also added that the form *(hij) zoekt* '(he) searches' was pronounced short, while the form '(ik) zoek' was pronounced long.
- One participant mentioned that, while the words '(ik) hoef' *(I) need to*, 'groen' *green* and 'doen' *to do* had a long vowel, the derived words 'behoefte' *need*, 'groente' *vegetable* and 'voldoende' *sufficient* had short vowels.

It is important to acknowledge the anecdotal and impressionistic nature of the above statements. Drawing conclusions based on any of the above claims should be done with caution.

Having said that, the interviews suggest that syllabic structure is relevant to the distribution of /u/. Firstly, if the anecdotal account is accurate, whether or not the final consonant is tautosyllabic affects the length of /u/. This is exemplified in *snoep*, which the experimental data showed more often short than long, versus *snoepen*, which all participants claimed to pronounce long.

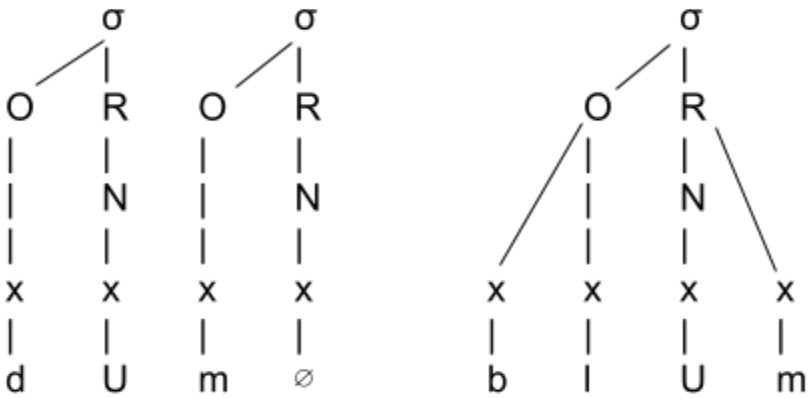
Secondly, while words of the structure UC# may have a short or a long /u/, words with the structure UCC or UC.C may only allow a short vowel. One way of accounting for this is through the analysis of Botma & van Oostendorp (2012), whereby the final consonant of words of the structure CV:C is not a coda at all, but the onset of a degenerate syllable. In other words, the /u/ in "doen" is in an open syllable, while the /u/ in "(vol)doende" is in a closed syllable. Visually illustrated, this would look as such:

Figure 15 - syllable structure of *doen* vs *(vol)doende*



This would additionally imply that the words “bloem” and “doem” differ in underlying syllabic structure, whereby “bloem” has a coda-m, while “doem” does not:

Figure 16 - syllable structure of *doem* vs *bloem*



6.3.3 Place and manner

As discussed in the introduction, it is taken as uncontroversial that if a vowel’s length is affected by an adjacent consonant, it is unlikely that it is the consonant that precedes the vowel. This is the reason why the following consonant was taken into account in the selection of target words, and the preceding consonant was not.

The experimental data shows that long [u:] was the more common realization among all final consonants, with the exception of /k/, /p/ and /j/, where short [u] was strongly preferred. No word was unambiguously long before /k/, /p/ and especially /j/.

6.3.4 Etymology and cognate comparisons

The instances of /u/ contained in this dataset mostly come from Proto-West-Germanic /o:/. With the exception of two words that come from PWGerm /u:/, through borrowing from an eastern dialect, where PWG /u:/ had not fronted to /y:/. Lastly, five of the words are French loans, whereof the resulting /u/ comes from French /u/.

Synchronically, the words were compared to their cognates in Antwerp Brabantian, Limburgish and North Brabantian (full table of comparison as well as sources for all lexical entries in Appendix G). Neither Limburgish nor North Brabantian are a single unified dialect. In this case, the representative dialects for Limburgish are those of Venlo and Thorn, and for North Brabantian those of Tilburg and Den Bosch. The selection of these dialects was based on the accessibility of dictionaries and descriptions.
Below is a summary of the correspondences.
The following sounds all correspond to the Standard /u/:

1. Antwerp Brabantian /y:/, /u/, and /u:/
2. Thorn & Venlo /ø:/, /u:/, /o/, and /o:/
3. Den Bosch & Tilburg /y:/, /u:/, and /u/

Most striking is the correspondence between short AB /u/ and short tussentaal /u/. If the Standard Dutch word is identical to the word in AB, then a word with short /u/ in AB is always pronounced short in the tussentaal:

(5) Length correspondence of /u/ between AB and Standard Dutch

AB [u] vs tussentaal [u]

AB [mut]	vs	tussentaal [mut]	‘must’
AB [blum]	vs	tussentaal [blum]	‘flower, flour’

AB [u:] vs tussentaal [u:]

AB [mu:t]	vs	tussentaal [mu:t]	‘courage’
AB [hu:t]	vs	tussentaal [hu:t]	‘hat’

Compare (5) to instances in which the corresponding words between AB and Standard Dutch were not identical:

(6) Length correspondence in words differing between AB and Standard Dutch

AB [sny:p]	vs	tussentaal [snup][10]	‘candy’
AB [duŋ]	vs	tussentaal [du:n]	‘do’.INF

Short [u] in AB corresponds to non-i-umlauted long *ō in Proto-West-Germanic before /k/, /p/ and /j/. The word [mut] ‘must’ is most likely the result of an unrelated process of early and widespread shortening of PWG long *ō. This change seems to have happened in only some highly frequent words and morphemes (Paul Boersma, p.c.). Another example of such shortening is in the morpheme *-*dōm* with its English reflex -dom (ex. *free<u>dom</u>*).

6.3 Summative account

Below is a summative account of the distribution of /u/ based on the examination of the abovementioned hypotheses.

1. When Antwerp Brabantian has short [u], so does the regional standard.
2. In words without short [u] in Antwerp Brabantian, instances of /u/ before /k/, /p/ and especially /j/ are short, while the rest are long.
3. The words *zoek* and *snoep* are notable exceptions, where some people pronounce them short (as expected from the above rule) and some pronounce them long. The long pronunciations could be explained by paradigm levelling: the words *zoek* and *snoep* appear as the verb forms *zoeken* and *snoepen*, where /k/ and /p/ are not tautosyllabic, resulting regularly in long [u:]. In the name of paradigm uniformity, the expected short [u] in *zoek* and *snoep* gets replaced by long [u:].

6.4 Problems

Given the nature of the present study, it is important to acknowledge all of its problems and shortcomings in order to understand in which way it is to be interpreted and utilized.

Problem 1 - Length as a cue

Because of the COVID-19 pandemic, the most viable way the study could be conducted is through online platforms. This led to the recording of data being carried out through the individual microphones of each participant. In other words, there is no way to guarantee that each participant provide high quality audio.

[10] Note that long [snu:p] also occurred in the dataset, but with lesser frequency

Thus, the only viable aspect in which each vowel can be analyzed is phonetic length. Phonetic length, however, is not the only cue that distinguishes so-called “long” vs “short” vowels neither in the traditional Antwerpian dialect (De Schutter, 1999) or the standard language (Booij, 1999). This quality difference is usually even sufficient to distinguish short vs long vowels, especially for Flemish speakers (Kloots et al., 2010).

Problem 2 - Limited word structure

Due to the time and resource limitations of the present study, it was not feasible to explore the effects of syllabic structure on the distribution of /u/. While this was partially touched upon with the help of the post-experimental interviews, very little of the experiment was designed to provide useful insights into syllable structure.

Problem 3 - Lack of documentation

Perhaps paradoxically, the lack of documentation of this feature makes it more difficult to study. For example, if one were to want to dedicate an experiment to examining the historical relationship between the length of /u/ and i-umlaut, one would need to know that i-umlaut could even have anything to do with the process in the first place.

In other words, the lack of documentation of this phenomenon hinders the development of a standard hypothesis-driven study.

7. Conclusion

What one may take away from this study can be summarized under three findings. Firstly, the current research demonstrates that there is indeed a phonemic length distinction in the close back vowel /u/ in the regional standard speech of speakers from provincial Antwerp. Secondly, the distribution of the short and long /u/ of the regional standard is intimately related to the distribution of the short and long /u/, as well as its i-umlauted form /y:/ in the now minority dialect of Antwerp Brabantian. Finally, instances of short /u/ that correspond to a long vowel in Antwerp Brabantian seem to be a result of a shortening of /u/ which occurs before tautosyllabic /k/, /p/ and /j/. Instances of short or long /u/ that are not explained by their Brabantian parallel or following /k/, /p/ or /j/—such as *zoek* and *snoep*—may be the result of paradigm leveling.

References

Auer, P. (1989). Some ways to count morae: Prokosch's Law, Streitberg's Law, Pfalz's Law, and other rhythmic regularities. *Linguistics 27 (1989), S. [1071] - 1102*, *27*. https://doi.org/10.1515/ling.1989.27.6.1071

Bokken, P., Tonnaer, M., & Werkgroep Thoears Woeardebook. (2011). *Thoears woeardebook*. Werkgroep Thoears Woeardebook.

Booij, G. (1999). *The Phonology of Dutch*. Clarendon Press. https://books.google.be/books?id=LT6E6YdAh-MC

Botma, B., & van Oostendorp, M. (2012). A propos of the Dutch vowel system 21 years on, 22 years on. In B. Botma & R. Noske (Eds.), *Phonological Explorations*. DE GRUYTER. https://doi.org/10.1515/9783110295177.135

Camerman, F. (2007). *Antwerps schrijven: Spelling en grammatica van het 21e-eeuwse Antwerps*. De Vries-Brouwers.

Camerman, F. (2009). *Antwerps.be - alles over de Antwerpse taal.* http://www.antwerps.be/uitdrukkingen.

De Schutter, G. (1999). Het Antwerps. Een schets van zijn evolutie tussen 1898 en 1998. In J. Kruijsen & N. Van der Sijs (Eds.), *Honderd jaar stadstaal* (pp. 301–315). Contact.

de Vries, J. W., 1937-2008 (viaf)41885808, Hogenhout-Mulder, M., & de Tollenaere, F. J. M. L., 1912-2009 (role)aut (viaf)46795555. (1971). *Nederlands etymologisch woordenboek*. Leiden : Brill. http://lib.ugent.be/catalog/rug01:002018492

Denes, P. (1955). Effect of Duration on the Perception of Voicing. *The Journal of the Acoustical Society of America*, *27*(4), 761–764. https://doi.org/10.1121/1.1908020

Jenks, G. (1967). *The Data Model Concept in Statistical Mapping*.

Kloots, H., Verhoeven, J., Coussé, E., & Gillis, S. (2010). The role of duration in vowel categorization. Flemish vs. Dutch listeners: A case study. *Papers of the Linguistic Society of Belgium*, *5*, 1–14.

Kluender, K. R., Diehl, R. L., & Wright, B. A. (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics*, *16*(2), 153–169. https://doi.org/10.1016/S0095-4470(19)30480-2

Kruijsen, J., & Sijs, N. van der (Eds.). (1999). *Honderd jaar stadstaal*. Contact.

Lahiri, A., & Dresher, B. E. (1999). Open Syllable Lengthening in West Germanic. *Language*, *75*(4), 678–719. JSTOR. https://doi.org/10.2307/417730

Linke, K. (2020a). *Co-occurrence restrictions in rhymes*. https://taalportaal.org/taalportaal/topic/pid/topic-14008432337635282

Linke, K. (2020b). *Coda*. https://taalportaal.org/taalportaal/topic/pid/topic-14008431983824103

Linke, K. (2020c). *Nucleus*. https://taalportaal.org/taalportaal/topic/pid/topic-13998813314542255

MacWright, T. (2013). Literate Jenks Natural Breaks and How The Idea Of Code is Lost. Tom MacWright. https://macwright.com/2013/02/18/literate-jenks.html

Port, R. F., & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. Perception & Psychophysics, 32(2), 141–152. https://doi.org/10.3758/BF03204273

Prokosch, E. (2009). *A comparative Germanic grammar*. Tiger Xenophon.

r/belgium. reddit. (n.d.). https://www.reddit.com/r/belgium/. (Retrieved May 2021)

r/antwerp. reddit. (n.d.). https://www.reddit.com/r/antwerp/. (Retrieved May 2021)

r/vlaanderen. reddit. (n.d.). https://www.reddit.com/r/vlaanderen/. (Retrieved May 2021)

Raphael, L. J. (1972). Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word‐Final Consonants in American English. *The Journal of the Acoustical Society of America*, *51*(4B), 1296–1303. https://doi.org/10.1121/1.1912974

Sebregts, M. van, Koen; Oostendorp. (2020). *The glottal fricative /h/*. https://taalportaal.org/taalportaal/topic/pid/topic_20150916115116266

Sijs, N. van der. (2019). *15 eeuwen Nederlandse taal*. Sterck & De Vreese.

Swanenberg, J. (2011). *Brabants-Nederlands: Nederlands-Brabants : handwoordenboek*. Europese Bibliotheek.

User: heerMat. (2009, May 9). *Lange en korte <oe>?* (Retrieved May 2021) https://www.dutchgrammar.com/forum/viewtopic.php?f=9&t=2624&sid=413ab96beb060a0083100e0cf63db9c1

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Visser, W. (2020). *Sonorants and obstruents and the features ±voice and ±cont*. https://taalportaal.org/taalportaal/topic/pid/topic-14020545877732644

Appendix A - Original forum post by user heerMat

De theorie zegt dat <oe> altijd een korte klank is , maar lang [u:] wordt uitgesproken vóór /r/.

In de praktijk neem ik bij mijn eigen uitspraak echter wél een onderscheid tussen een lange en een korte <oe> waar, elders dan vóór /r/.

Kan iemand dit beamen op basis van eigen ervaring en hier eventueel een verklaring voor geven?

Bij mij zijn:

- lang: zoeken, boeten, zoet, roet, voet, roepen, doen, schoen, groen, loens, koen...

- kort: boeken, doeken, koeken, hoek, stoep, "tsjoep", groep...

Eigenlijk vind ik <oe> vaker lang dan kort. :)

Appendix B - Information and consent form (in Dutch)

Informatiebrochure voor

The moet-moed split

Geachte participant / Beste deelnemer,

U gaat deelnemen aan het onderzoek *The moet-moed split* van de Universiteit van Amsterdam - Faculteit der Geesteswetenschappen uitgevoerd door Mishko Bozhinoski onder begeleiding van Marijn van 't Veer. Voordat het onderzoek begint, is het belangrijk dat u kennis neemt van de procedures die in dit onderzoek worden gevolgd. Leest u deze brochure daarom zorgvuldig door.

Doel van het onderzoek

Het Vlaams dat in de Antwerpse regio gesproken wordt, lijkt momenteel een proces te ondergaan dat zelden voorkomt in de talen van de wereld. Dit verschijnsel lijkt ook nog moeilijk te verklaren vanuit zowel historisch als synchroon aspect.

Het meest verrassende is misschien wel dat deze verandering zelfs door de academische wereld grotendeels onopgemerkt lijkt te zijn gebleven.

Het doel van dit onderzoek is dit proces en de eigenschappen ervan te documenteren.

Voor het afnemen van het experiment kunnen we nog geen nadere informatie geven over de factoren waarin we geïnteresseerd zijn. Na afloop van het experiment ontvangt u hierover meer informatie.

Wie kan er aan dit onderzoek meedoen

Voor dit onderzoek worden volwassen sprekers van het Nederlands uitgenodigd, die in de provincie Antwerpen geboren zijn en hier het grootste deel van hun leven hebben doorgebracht. Wij vragen u voorafgaand aan het onderzoek naar uw gehoor en gezichtsvermogen. Gezien de aard van het onderzoek is het van belang dat u goed kunt horen en zien. Het dragen van een bril is wel toegestaan. Verder stellen wij u een aantal vragen over uw taalachtergrond. U kunt deelnemen aan dit onderzoek als Nederlands uw moedertaal is. Verder is het van belang dat u, voor zo ver bij u bekend, geen taalprobleem heeft, zoals dyslexie of een specifieke taalstoornis.

Instructie en procedure

Het onderzoek zal in twee fasen worden uitgevoerd: een proefexperiment en het eigenlijke experiment. Zowel het proefexperiment als het eigenlijke experiment worden op exact dezelfde manier uitgevoerd; het eigenlijke experiment zal echter meer stimuli bevatten.

Het experiment vindt plaats in een besloten bijeenkomst via het online communicatieplatform Zoom. Het bestaat uit twee delen: een korte enquête en het experiment zelf. Tijdens de enquête worden u enkele vragen gesteld over uw persoonlijke en talige achtergrond: uw leeftijd, geslacht, of u een taalstoornis heeft, of u bent opgegroeid in de provincie Antwerpen en of u gedurende een lange periode elders hebt gewoond. Dit duurt ongeveer 5 minuten.

Het experiment zelf bestaat uit een reeks zinnen, die door de deelnemer moeten worden ingesproken. Dit duurt max. 15 minuten.

Vrijwilligheid

U doet vrijwillig mee aan dit onderzoek. U kunt dan ook op elk moment gedurende het onderzoek uw deelname stopzetten. Dit zal geen gevolgen voor u hebben en u bent in geen geval verplicht de hierboven beschreven procedures af te ronden. Tevens kunt op eender welk moment na het onderzoek uw deelname alsnog intrekken. Als u uw deelname staakt of toestemming intrekt, worden alle tot dan toe verzamelde gegevens definitief verwijderd.

Ongemak, Risico's en Verzekering

De risico's van deelname aan dit onderzoek zijn niet groter dan die in dagelijkse situaties thuis. Uit ervaring met voorgaande, vergelijkbare onderzoeken is gebleken dat er van enig ongemak voor de deelnemers niet of nauwelijks sprake is. Bij elk onderzoek van de Universiteit van Amsterdam geldt een standaard aansprakelijkheidsverzekering.

Vertrouwelijkheid van de onderzoeksgegevens

De gegevens die in dit onderzoek worden verzameld, zullen door de onderzoekers alleen worden gebruikt voor nadere analyse en voor publicatie in wetenschappelijke tijdschriften. Hierbij wordt geen gebruik gemaakt van uw persoonlijke gegevens en uw anonimiteit blijft onder alle omstandigheden gewaarborgd.

De verzamelde onderzoeksgegevens zullen gecodeerd opgeslagen worden, apart van uw persoonlijke gegevens. Alleen medewerkers aan het onderzoek hebben toegang tot deze gegevens en de codering.

De verzamelde eye-tracking-, video- en audio-opnamen zullen nooit zonder uw schriftelijke toestemming publiekelijk getoond worden. U ontvangt een apart formulier waarop u desgewenst deze toestemming kunt verlenen..

Nadere inlichtingen

Als u nog verdere informatie wilt over dit onderzoek, dan kunt u zich wenden tot
Mishko Bozhinoski (telefoon: +32 476 71 01 31; e-mail: miskobozinoski@gmail.com; Vrouwvlietstraat 37, 2800 Mechelen, België).

Met eventuele klachten over dit onderzoek kunt u zich wenden tot de secretaris van de Commissie Ethiek van de Faculteit Geesteswetenschappen van de Universiteit van Amsterdam; email: commissie-ethiek-fgw@uva.nl ; telefoon: +31 20 - 525 3054; Kloveniersburgwal 48, 1012 CX Amsterdam.

Toestemmingsformulier

'Ik verklaar hierbij op voor mij duidelijke wijze te zijn ingelicht over het onderzoek *A new phonological split in Antwerp?* van de Universiteit van Amsterdam, Faculteit der Geesteswetenschappen uitgevoerd door Mishko Bozhinoski onder begeleiding van Marijn van 't Veer, zoals uiteengezet in de informatiebrochure. Mijn vragen zijn naar tevredenheid beantwoord.

Ik stem geheel vrijwillig in met deelname aan dit onderzoek. Ik behoud daarbij het recht deze instemming weer in te trekken zonder dat ik daarvoor een reden hoef op te geven. Ik besef dat ik op elk moment mag stoppen met het onderzoek en op eender welk moment na het onderzoek mijn deelname alsnog kan intrekken. In het geval dat ik mijn deelname staak of toestemming intrek zullen alle tot dan toe verzamelde gegevens definitief worden verwijderd.

Indien mijn onderzoeksresultaten gebruikt zullen worden in wetenschappelijke publicaties, dan wel op een andere manier openbaar worden gemaakt, zal dit volledig geanonimiseerd gebeuren. Mijn persoonsgegevens zullen niet door derden worden ingezien zonder mijn uitdrukkelijke toestemming.

Als ik nog verdere informatie over het onderzoek zou willen krijgen, nu of in de toekomst, kan ik me wenden tot Mishko Bozhinoski (telefoon: +32 476 71 01 31; e-mail: miskobozinoski@gmail.com; Vrouwvlietstraat 37, 2800 Mechelen, België).

Met eventuele klachten over dit onderzoek kan ik me wenden tot de secretaris van de Commissie Ethiek van de Faculteit Geesteswetenschappen van de Universiteit van Amsterdam; email: commissie-ethiek-fgw@uva.nl; telefoon: +31 20 - 525 3054; Kloveniersburgwal 48, 1012 CX Amsterdam.'

Aldus in tweevoud getekend:

.....	.....	.....
Naam proefpersoon	Datum	Handtekening

'Ik heb toelichting verstrekt op het onderzoek. Ik verklaar mij bereid nog opkomende vragen over het onderzoek naar vermogen te beantwoorden.'

.....	.....	.....
Naam onderzoeker	Datum	Handtekening

Appendix C - Promotional post (in Dutch)

Hoi! Ik hoop dat ik dit hier mag plaatsen.

Ik doe onderzoek naar een zeldzaam soort klankverschuiving die in de provincie Antwerpen lijkt te zijn begonnen.

Voor mijn onderzoek heb ik deelnemers uit de provincie Antwerpen nodig (niet alleen de stad). Het onderzoek bestaat uit een heel korte vragenlijst (~30s invullen) en een experiment (5-10min) waarin de deelnemer een reeks zinnen moet inspreken. Om de resultaten van het experiment niet te beïnvloeden kan ik jullie helaas niet meer vertellen over het onderwerp van het onderzoek, maar als iemand er wel interesse in heeft nadat het experiment uitgevoerd is, kan ik er wel dieper op ingaan :).

In totaal offer je dus een klein kwartiertje van je tijd op aan de taalwetenschap! Als je interesse hebt en uit de provincie Antwerpen komt, stuur mij een DM!

Appendix D - Praat experiment script

```
carrier$= "Schrijf het woord X eens op"
targetPos= 4
file$= "list.txt"
resultFolder$= "Result"

maxFilledIn= 1
subject$= ""
repeat
  filledIn= 0
  beginPause ("Input")
     sentence ("Subject", "1")
  clicked = endPause("Submit", 1)
  if length(subject$) > 0
     filledIn+= 1
  endif
until filledIn = maxFilledIn

screenBackColor$= "Teal"
screenForeColor$= "Blue"
screenFontHeight= 24
screenFont$= "Times"

demoWindowTitle ("Experiment ACLC")


buttonMax= 1

button1Width= 10
button1Height= 5
button1ID= 1
button1BackColor$= "Yellow"
button1ForeColor$= "Black"
button1Item$= "Next"
button1FontHeight= 18
button1L= 50 - button1Width/2
button1R= button1L + button1Width
button1T= 20
button1B= button1T - button1Height


r= Read Table from tab-separated file: file$
r_nor= Get number of rows
Randomize rows
Save as tab-separated file: "'resultFolder$'/'subject$'.txt"

@splitSentence(carrier$)

@ShowPage("Start of experiment", 50, 80)

for i to r_nor
    target$= object$[r, i, "Target"]
    sentence$= ""
    for j to splitSentence.wordCount
      if j<> targetPos
        sentence$+= splitSentence.word'j'$
      else
         sentence$+= unicode$(34) + target$ + unicode$(34)
      endif
      if j <> splitSentence.wordCount
        sentence$+= " "
      endif
    endfor

    @ClearScreen()
    demo Text: 50, "centre", 70, "half", sentence$
     if i mod 60 = 0
              demo Text: 50, "centre", 40, "half", "PAUZE"
     endif
```

```
    #Two options to call
    #@CreateButton(button1ID, "", 0, "", "")
    @CreateButton(button1ID, button1Item$, button1FontHeight, button1BackColor$, button1ForeColor$)


    demo Select inner viewport: 0, 100, 0, 100
    demo Axes: 0, 100, 0, 100
    while demoWaitForInput( )
      if demoClicked ( )
        if demoClickedIn(button1L, button1R, button1B, button1T)
          goto NEXT_WIN
        endif
      else
        if demoKeyPressed()
          if demoKey$() = " "
            goto NEXT_WIN
          endif
        endif
      endif
    endwhile

    label NEXT_WIN
endfor

@ShowPage("End of experiment", 50, 80)

demo Erase all

exit

procedure splitSentence(.text$)
   .wordCount= 0
   .start= 1
   .text$= replace_regex$(.text$, "^\s*", "", 0)
   .text$= replace_regex$(.text$, "\s*$", "", 0)

   repeat
      .pos= index_regex(.text$, "\s+")
      if .pos > 0
        .wordCount+= 1
        .word'.wordCount'$= mid$(.text$, .start, .pos)
        .text$= right$(.text$, length(.text$)-.pos)
      endif
   until .pos= 0
   if length(.text$) > 0
     .wordCount+= 1
     .word'.wordCount'$= .text$
   endif
endproc

procedure ClearScreen
  demo Erase all
  demo Select inner viewport: 0, 100, 0, 100
  demo Axes: 0, 100, 0, 100
  demo Paint rectangle: screenBackColor$, 0, 100, 0, 100
  demo 'screenFont$'
  demo Font size: screenFontHeight
endproc

procedure ShowPage(.text$, .x, .y)
  @ClearScreen()
  demo Text: .x, "centre", .y, "half", .text$
  while demoWaitForInput( )
    if demoKeyPressed()
      goto NEXT_WIN_ShowPage
    endif
 endwhile

 label NEXT_WIN_ShowPage
endproc

procedure CreateButton(.id, .text$, .fontHeight, .backColor$, .foreColor$)
```

```
    demo Select inner viewport: button'.id'L, button'.id'R, button'.id'B, button'.id'T
    demo Axes: 0, 100, 0, 100

    if length(.backColor$) = 0
      .backColor$= button'.id'BackColor$
    endif
    demo Paint rectangle: .backColor$, 0, 100, 0, 100

    if length(.foreColor$)= 0
      .foreColor$= button'.id'ForeColor$
    endif
    demo Colour: .foreColor$

    if .fontHeight= 0
      .fontHeight= button'.id'FontHeight
    endif
    demo Font size: .fontHeight

    if length(.text$)= 0
      .text$= button'.id'Item$
    endif
    demo Text: 50, "centre", 50, "half", .text$

    demo Font size: screenFontHeight
endproc
```

Appendix E - Questionnaire (In Dutch)

Leeftijd:

_____

Geslacht:

Mannelijk Vrouwelijk Anders

Is Nederlands uw moedertaal?

JA NEE

Spreekt u nog een andere taal als moedertaal?

JA NEE

Welke?

_____ Zeg ik liever niet.

Bent u in de provincie Antwerpen opgegroeid?

JA NEE

Woont u in de provincie Antwerpen?

JA NEE

Heeft u meer dan 5 van de voorbije 10 jaar buiten de provincie Antwerpen doorgebracht?

JA NEE

Heeft u een taalstoornis, of lijdt u aan iets wat directe gevolgen heeft voor de taalverwerving (bv. doofheid)?

JA NEE

Appendix F - TextGrid processing Python script

```
import pandas as pd
import matplotlib.pyplot as plt
import statistics as stats
import glob, os
import numpy as np
from random import uniform
from dip_test import dip

def get_jenks_breaks(data_list, number_class): ...

def check(arr, thing):
	for x in range(len(arr)):
		if arr[x][0] == thing:
			return True, x;
	return False, -1;

def deviant(arr, thresh):
	perc = []
	for x in arr:
		perc.append(x-thresh)
	return perc

def midPoint(arr, brk):
	for x in range(len(arr)):
		if arr[x][1] == brk:
			return (arr[x][1]+arr[x+1][1])/2
			break

Lines = []

###############
###READ FILE###
###############

#for file in glob.glob("*.data"):
#	dataFile = open(file, 'r')
#	lns = dataFile.readlines()
#	Lines = Lines + lns
#	dataFile.close()

dataFile = open('user.data', 'r')
lns = dataFile.readlines()
Lines = Lines + lns
dataFile.close()

lst = []
names = []
vls = []
std = []
zrs = []

rawData = []
rawNames = []
rawZrs = []
```

```
processed=[]
preproc=[]

for x in Lines:
	if len(x.split()) > 0:
		lst.append([x.split()[0], x.split()[1]])

for x in lst:
	isPresent, pos = check(processed, x[0])
	if isPresent:
		processed[pos].append(x[1])
	else:
		processed.append([x[0], x[1]])

	rawData.append(float(x[1]))
	rawNames.append(x[0])
	rawZrs.append(x[0][-1])

for x in range(len(processed)):
	processed[x] = [processed[x][0], processed[x][1:]]

for x in range(len(processed)):
	processed[x][1] = [float(i)*1000 for i in processed[x][1]]

for x in processed:
	preproc.append(x)

for x in range(len(processed)):
	processed[x] = [processed[x][0], stats.mean(processed[x][1])]

processed = sorted(processed, key=lambda x: x[1])

for x in processed:
	names.append(x[0])
	vls.append(x[1])
	#std.append(x[2])
	zrs.append(0.5)
	#if x[0] == "hoef" or x[0] == "groef" or x[0] == "boef":
	#	zrs.append('v')
	#else:
	#	zrs.append(x[0][-1])

for x in processed:
	print(x)
#print(std)
#print("\n\n\n\n\n\n"+str(len(processed)))
#print(check(lst, "stoet"))

fig, ax = plt.subplots();
#ax.scatter(rawZrs, rawData, s=5)
ax.scatter(zrs, vls, s=5)
#print(dip(vls))
#ax.set_yscale('log')

#print(stats.harmonic_mean(vls))

for i, txt in enumerate(names):
```

```
            ax.annotate(txt, (zrs[i], vls[i]))
            #print('skip')

#for i, txt in enumerate(rawNames):
#           ax.annotate(txt, (rawZrs[i], rawData[i]))

preproc = sorted(preproc, key=lambda x: x[0][-1])

breaks = get_jenks_breaks(vls, 2)

midPoint = midPoint(processed,breaks[1])
print("\n\n\n\n\n\n\n\n\n\n\nCutoff: "+str(midPoint)+"\n\n\n\n\n")

consLong=[]
consShort=[]
ambigu=[]

for x in preproc:
            lStr=str(x[0])
            x[1] = sorted(x[1])
            i=deviant(x[1], midPoint)

            ap=0
            for k in i:
                        ap+=np.sign(k)
            if ap>=len(i):
                        consLong.append(lStr)
            elif ap<=-len(i):
                        consShort.append(lStr)
            else:
                        ambigu.append(lStr)

            for y in i:
                        if y>0:
                                    lStr=lStr+" +"+str(round((y)))
                        else:
                                    lStr=lStr+" "+str(round((y)))
            print(lStr)

strLong="ConsLong: "
for x in consLong:
            strLong=strLong+x+", "
strShort="ConsShort: "
for x in consShort:
            strShort=strShort+x+", "
strAmb="Ambigu: "
for x in ambigu:
            strAmb=strAmb+x+", "
print(strLong)
print(strShort)
print(strAmb)

for line in breaks:
   plt.plot([line for _ in range(len(x))], 'k--')

plt.xlabel('Phoneme')
plt.ylabel('Vowel length in ms')
#plt.show()
```

Appendix G - Python implementation of Jenks

```
def get_jenks_breaks(data_list, number_class):
    data_list.sort()
    mat1 = []
    for i in range(len(data_list) + 1):
        temp = []
        for j in range(number_class + 1):
            temp.append(0)
        mat1.append(temp)
    mat2 = []
    for i in range(len(data_list) + 1):
        temp = []
        for j in range(number_class + 1):
            temp.append(0)
        mat2.append(temp)
    for i in range(1, number_class + 1):
        mat1[1][i] = 1
        mat2[1][i] = 0
        for j in range(2, len(data_list) + 1):
            mat2[j][i] = float('inf')
    v = 0.0
    for l in range(2, len(data_list) + 1):
        s1 = 0.0
        s2 = 0.0
        w = 0.0
        for m in range(1, l + 1):
            i3 = l - m + 1
            val = float(data_list[i3 - 1])
            s2 += val * val
            s1 += val
            w += 1
            v = s2 - (s1 * s1) / w
            i4 = i3 - 1
            if i4 != 0:
                for j in range(2, number_class + 1):
                    if mat2[l][j] >= (v + mat2[i4][j - 1]):
                        mat1[l][j] = i3
                        mat2[l][j] = v + mat2[i4][j - 1]
        mat1[l][1] = 1
        mat2[l][1] = v
    k = len(data_list)
    kclass = []
    for i in range(number_class + 1):
        kclass.append(min(data_list))
    kclass[number_class] = float(data_list[len(data_list) - 1])
    count_num = number_class
    while count_num >= 2:  # print "rank = " + str(mat1[k][count_num])
        idx = int((mat1[k][count_num]) - 2)
        # print "val = " + str(data_list[idx])
        kclass[count_num - 1] = data_list[idx]
        k = int((mat1[k][count_num] - 1))
        count_num -= 1
    return kclass
```

Appendix H - Participant data

Participant code: IG27				Cutoff length: 73.70ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+39 +49 +104	moet	-30 -17 -11	moed	+12 +28 +31	boek	-41 -31 -22
koe	+28 +63 +69	groet	+31 +42 +54	hoed	+23 +37 +57	zoek	-17 -16 +8
hoe	+64 +69 +72	stoet	+18 +21 +52	voed	+24 +35 +40	doek	-33 -30 -28
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-41 -39 -36	bloem	-42 -36 -32	doen	+25 +49 +69	hoes	+26 +55 +57
stoep	-39 -36 -30	roem	+58 +60 +76	groen	+21 +27 +35	bloes	-21 +29 +35
snoep	-45 -41 -19	doem	+30 +41 +51	zoen	+27 +44 +61	moes	+8 +27 +28
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+26 +35 +56	sloef	-38 -36 -26	vroeg	+22 +23 +27	boer	+33 +72 +74
boef	+25 +46 +46	stoef	-26 -25 -16	boeg	+30 +32 +34	hoer	+46 +61 +64
groef	+1 +9 +16	toef	-36 -23 -22	kroeg	+33 +35 +55	stoer	+43 +45 +51
/l/				/j/			
Word	Distance			Word	Distance		
koel	+37 +39 +60			gloei	-25 -24 -1		
boel	-17 +30 +34			boei	-31 -30 -14		
				roei	-33 -23 -9		

Participant code: IG27		
Consistently long	**Consistently short**	**Inconsistent**
moe, koe, hoe, groet, stoet, moed, hoed, voed, roem, doem, doen, groen, zoen, hoes, moes, hoef, boef, grief, vroeg, boeg, kroeg, boer, hoer, stoer, koel	moet, boek, doek, groep, stoep, snoep, bloem, sloef, stoef, toef, gloei, boei, roei	boel, bloes, zoek

Participant code: MD90				Cutoff length: 101.98ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+55 +66 +79	moet	-54 -40 -39	moed	+19 +51 +52	boek	-56 -50 -44
koe	+67 +128 +136	groet	+23 +42 +42	hoed	+31 +60 +65	zoek	-56 -43 +48
hoe	+44 +91 +110	stoet	+30 +32 +46	voed	+30 +49 +56	doek	-69 -52 -44
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-64 -52 -52	bloem	-65 -60 -59	doen	+31 +37 +83	hoes	+26 +69 +72
stoep	-60 -58 -50	roem	+31 +41 +43	groen	+35 +48 +57	bloes	+30 +34 +46
snoep	-19 +19 +48	doem	+5 +34 +52	zoen	+22 +47 +60	moes	+22 +29 +53
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+41 +42 +48	sloef	-57 -47 -46	vroeg	+55 +62 +64	boer	+47 +48 +584
boef	-25 +37 +39	stoef	-59 -41 -40	boeg	+29 +60 +61	hoer	+86 +87 +115
groef	+20 +30 +31	toef	-41 -23 -17	kroeg	+42 +60 +79	stoer	+46 +52 +88
/l/				/j/			
Word	Distance			Word	Distance		
koel	+28 +41 +84			gloei	-50 -30 -21		
boel	+24 +30 +35			boei	-34 -15 -11		
				roei	-57 -41 -32		

Participant code: MD90		
Consistently long	**Consistently short**	**Inconsistent**
voed, hoed, moed, hoe, koe, moe, groef, hoef, vroeg, boeg, kroeg, koel, boel, roem, doem, doen, groen, zoen, stoer, hoer, boer, moes, hoes, bloes, stoet, groet	toef, stoef, sloef, gloei, roei, boei, doek, boek, bloem, stoep, groep, moet	boef, zoek, snoep

Participant code: AG56				Cutoff length: 117.36ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+17 +43 +89	moet	-61 -50 -30	moed	+1 +23 +25	boek	-62 -43 -32
koe	+43 +67 +71	groet	+15 +21 +23	hoed	+26 +29 +32	zoek	-57 +15 +33
hoe	+76 +89 +89	stoet	+4 +13 +16	voed	+32 +42 +44	doek	-60 -60 -49
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-55 -42	bloem	-57 -48 -36	doen	+3 +30 +38	hoes	+23 +26 +69
stoep	-61 -48 -29	roem	+31 +39 +58	groen	+20 +33 +113	bloes	+21 +22 +33
snoep	-7 +13 +14	doem	+8 +33 +45	zoen	+2 +26 +33	moes	-6 +51 +79
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+3 +51 +89	sloef	-45 -35 -18	vroeg	+17 +19 +31	boer	+57 +89 +102
boef	+44 +62 +84	stoef	-56 -55 -49	boeg	+21 +47 +571	hoer	+13 +45 +84
groef	+32 +48 +48	toef	-38 +20 +27	kroeg	+8 +32 +51	stoer	+45 +65 +76
/l/				/j/			
Word	Distance			Word	Distance		
koel	+23 +45 +46			gloei	-20 -18 -14		
boel	+11 +13 +19			boei	-24 -12 -6		
				roei	-17 -12 -13		

Participant code: AG56		
Consistently long	**Consistently short**	**Inconsistent**
voed, moed, hoed, koe, hoe, moe, groef, hoef, boef, boeg, kroeg, vroeg, boel, koel, doem, roem, doen, groen, zoen, stoer, boer, hoer, hoes, bloes, stoet, groet	sloef, stoef, gloei, boei, roei, boek, doek, bloem, stoep, groep, moet	toef, zoek, snoep, moes

Participant code: JV57				Cutoff length: 114.24ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+51 +54 +65	moet	-59 -43 -34	moed	+17 +37 +53	boek	-72 -72 -69
koe	+97 +116 +119	groet	+4 +17 +47	hoed	-39 +31 +62	zoek	-72 +17 +30
hoe	+38 +46 +83	stoet	+5 +8 +10	voed	+5 +21 +22	doek	-64 -57 -55
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-79 -71 -66	bloem	-68 -49 -37	doen	+13 +26 +29	hoes	+40 +48 +75
stoep	-79 -78 -71	roem	+29 +44 +51	groen	+23 +40 +46	bloes	+4 +28 +31
snoep	+15 +36 +38	doem	+31 +53 +103	zoen	+15 +43 +63	moes	+20 +36 +41
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+28 +57 +65	sloef	-82 -57 -48	vroeg	+42 +47 +70	boer	+53 +58 +88
boef	+17 +36 +58	stoef	-77 -71 -39	boeg	+43 +71	hoer	+45 +72 +80
groef	+23 +42 +57	toef	-63 -58 -49	kroeg	+20 +33 +40	stoer	+57 +81 +85
/l/				/j/			
Word	Distance			Word	Distance		
koel	-52 -3 +11			gloei	-61 -60 -20		
boel	-24 +2 +18			boei	-43 -25		
				roei	-55 -49 -33		

Participant code: JV57		
Consistently long	**Consistently short**	**Inconsistent**
voed, moed, koe, moe, hoe, hoef, boef, groef, kroeg, vroeg, boeg, doem, roem, doen, groen, zoen, snoep, boer, hoer, stoer, hoes, bloes, moes, stoet, groet	stoef, toef, sloef, gloei, boei, roei, boek, doek, bloem, stoep, groep, moet	hoed, zoek, boel, koel

Participant code: CG66				Cutoff length: 86.31ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+6 +44 +72	moet	-27 -13 -13	moed	+10 +21 +25	boek	-46 -41 -35
koe	+1 +5 +72	groet	+19 +24 +34	hoed	+21 +23 +59	zoek	-39 -35 -19
hoe	+24 +30 +32	stoet	-26 +16 +28	voed	+9 +18 +31	doek	-46 -40 -36
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-43 -42 -33	bloem	-37 -36 -22	doen	+32 +44 +61	hoes	+25 +34 +45
stoep	-64 -54 -43	roem	+27 +27 +58	groen	+19 +45 +48	bloes	+34 +42 +48
snoep	-34 -33 -10	doem	+38 +62 +75	zoen	+33 +33 +34	moes	-29 -18 +29
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+28 +57 +65	sloef	-36 -36 +25	vroeg	+22 +24 +30	boer	+51 +74 +86
boef	+18 +22 +34	stoef	-32 -24 -7	boeg	+6 +5 +27	hoer	+16 +23 +43
groef	+15 +17 +65	toef	-32 -30 -27	kroeg	+5 +19	stoer	+13 +36 +50
/l/				/j/			
Word	Distance			Word	Distance		
koel	+34 +53 +11			gloei	-27 -20 -4		
boel	+40 +44 +55			boei	-35 -1 -5		
				roei	-36 -17 -3		

Participant code: CG66		
Consistently long	**Consistently short**	**Inconsistent**
voed, hoed, moed, moe, hoe, koe, boef, hoef, groef, boeg, vroeg, kroeg, boel, koel, roem, doem, groen, doen, zoen, boer, hoer, stoer, bloes, hoes, groet	stoef, toef, gloei, roei, boei, zoek, boek, doek, bloem, snoep, groep, stoep, moet	sloef, moes, stoet

Participant code: DH10				Cutoff length: 85.98ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+20 +37 +68	moet	-41 -31 -19	moed	+18 +33 +50	boek	-46 -41 -35
koe	+46 +58 +66	groet	+15 +24 +36	hoed	+31 +40 +45	zoek	-39 -35 -19
hoe	+24 +30 +32	stoet	-34 -28 +26	voed	+17 +38 +66	doek	-46 -40 -36
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-43 -42 -33	bloem	-37 -36 -22	doen	+32 +44 +61	hoes	+23 +33 +47
stoep	-64 -54 -43	roem	+27 +27 +58	groen	+19 +45 +48	bloes	-49 +28 +81
snoep	-34 -33 -10	doem	+38 +62 +75	zoen	+33 +33 +34	moes	+22 +53 +29
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+27 +28 +43	sloef	-35 -20	vroeg	+15 +38 +78	boer	+29 +34 +36
boef	+2 +6 +28	stoef	-50 -38 -29	boeg	+23 +47 +69	hoer	+46 +58 +67
groef	+43 +49 +61	toef	-48 -40 -18	kroeg	+35 +38 +45	stoer	-20 +23 +54
/l/				/j/			
Word	Distance			Word	Distance		
koel	+34 +53 +11			gloei	-27 -20 -4		
boel	+40 +44 +55			boei	-35 -1 -5		
				roei	-36 -17 -3		

Participant code: DH10		
Consistently long	**Consistently short**	**Inconsistent**
hoed, voed, moed, moe, hoe, koe, hoef, boef, groef, boeg, vroeg, kroeg, zoek, boel, koel, roem, doem, doen, zoen, groen, snoep, boer, hoer, hoes, moes, groet	toef, stoef, sloef, gloei, boei, roei, doek, boek, bloem, groep, stoep, moet	stoer, bloes, stoet

Participant code: NG35				Cutoff length: 73.35ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+41 +47 +52	moet	-38 -34 -34	moed	+7 +9 +34	boek	-41 -38 -34
koe	+32 +34 +45	groet	+11 +11 +39	hoed	-36 +17 +36	zoek	-18 +14 +17
hoe	+21 +34 +38	stoet	-22 -20 -17	voed	+9 +30 +44	doek	-43 -41 -40
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-39 -37 -30	bloem	-41 -40 -32	doen	+16 +28 +34	hoes	+31 +32 +40
stoep	-64 -54 -43	roem	+6 +28 +41	groen	+12 +18 +22	bloes	+1 +7 +36
snoep	-34 -33 -10	doem	+4 +11 +38	zoen	+18 +22 +28	moes	+17 +30 +35
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+26 +45 +62	sloef	-44 -41 -38	vroeg	+9 +10 +13	boer	+21 +35 +59
boef	+11 +12 +37	stoef	-41 -22 -16	boeg	+16 +20 +61	hoer	+15 +21 +45
groef	-20 +34 +42	toef	-31 -29 -5	kroeg	+10 +17 +21	stoer	+24 +38 +49
/l/				/j/			
Word	Distance			Word	Distance		
koel	-9 +34 +47			gloei	-40 -35 -13		
boel	+15 +23 +34			boei	-48 -34 -3		
				roei	-34 -28 -23		

Participant code: NG35		
Consistently long	**Consistently short**	**Inconsistent**
voed, moed, moe, koe, hoe, hoef, boef, boeg, kroeg, vroeg, boel, doem, roem, zoen, doen, groen, boer, hoer, stoer, hoes, moes, groet, bloes	stoef, toef, sloef, gloei, boei, roei, doek, boek, bloem, groep, stoep, moet, stoet	hoed, groef, zoek, koel, snoep

Participant code: FL97				Cutoff length: 95.93ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+5 +116 +119	moet	-38 -34 -34	moed	+8 +11 +38	boek	-49 -45 -43
koe	+34 +68 +79	groet	+11 +11 +39	hoed	+5 +28 +37	zoek	+4 +5 +6
hoe	+36 +63 +66	stoet	-22 -20 -17	voed	+9 +9 +37	doek	-60 -35 -12
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-39 -37 -30	bloem	-67 -60 -57	doen	+16 +28 +34	hoes	+31 +32 +40
stoep	-64 -54 -43	roem	+13 +18 +24	groen	+12 +18 +22	bloes	+1 +7 +36
snoep	-34 -33 -10	doem	+35 +43 +63	zoen	+18 +22 +28	moes	+17 +30 +35
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+10 +21 +26	sloef	-55 -49 -37	vroeg	+5 +15 +39	boer	+21 +35 +59
boef	+14 +25 +26	stoef	-52 -48 -47	boeg	-24 +23 +33	hoer	+15 +21 +45
groef	+8 +44 +57	toef	-48 -21 -9	kroeg	+5 +6 +7	stoer	+24 +38 +49
/l/				/j/			
Word	Distance			Word	Distance		
koel	+7 +35 +47			gloei	-28 -14 +85		
boel	+37 +39 +54			boei	-57 -2 -2		
				roei	-40 -5 -3		

Participant code: FL97		
Consistently long	**Consistently short**	**Inconsistent**
hoed, moed, voed, moe, hoe, koe, boef, hoef, groef, vroeg, kroeg, zoek, boel, koel, doem, roem, groen, zoen, doen, hoer, stoer, boer, hoes, bloes, moes, groet	toef, sloef, stoef, boei, roei, boek, doek, bloem, groep, snoep, stoep, moet	boeg, gloei, stoet

Participant code: GF91				Cutoff length: 76.58ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+5 +116 +119	moet	-38 -34 -34	moed	+8 +11 +38	boek	-49 -45 -43
koe	+34 +68 +79	groet	+11 +11 +39	hoed	+5 +28 +37	zoek	+4 +5 +6
hoe	+36 +63 +66	stoet	-22 -20 -17	voed	+9 +9 +37	doek	-60 -35 -12
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-39 -37 -30	bloem	-67 -60 -57	doen	+16 +28 +34	hoes	+31 +32 +40
stoep	-64 -54 -43	roem	+13 +18 +24	groen	+12 +18 +22	bloes	+1 +7 +36
snoep	-34 -33 -10	doem	+35 +43 +63	zoen	+18 +22 +28	moes	+17 +30 +35
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+10 +21 +262	sloef	-55 -49 -37	vroeg	+5 +15 +39	boer	+21 +35 +59
boef	+14 +25 +26	stoef	-52 -48 -47	boeg	-24 +23 +33	hoer	+15 +21 +45
groef	+8 +44 +57	toef	-48 -21 -9	kroeg	+5 +6 +7	stoer	+24 +38 +49
/l/				/j/			
Word	Distance			Word	Distance		
koel	+7 +35 +47			gloei	-28 -14 +85		
boel	+37 +39 +54			boei	-57 -2 -2		
				roei	-40 -5 -3		

Participant code: GF91		
Consistently long	**Consistently short**	**Inconsistent**
hoed, voed, koe, hoe, moe, boef, groef, boeg, kroeg, vroeg, boel, koel, roem, doem, doen, groen, zoen, stoer, hoer, boer, moes, hoes, stoet, groet	toef, sloef, stoef, gloei, roei, boei, boek, doek, bloem, groep, snoep, stoep, moet	moed, hoef, zoek, bloes

Participant code: GL98				Cutoff length: 95.93ms			
∅		/t/		/d/		/k/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
moe	+21 +41 +49	moet	-30 -20 -15	moed	+4 +11 +20	boek	-38 -27 -23
koe	+29 +42 +48	groet	+9 +19 +20	hoed	+12 +22 +25	zoek	-24 -22 -21
hoe	+29 +36 +55	stoet	+4 +20 +25	voed	+3 +11 +15	doek	-36 -36 -31
/p/		/m/		/n/		/z/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
groep	-31 -27 -24	bloem	-38 -36 -27	doen	+34 +35 +36	hoes	+11 +16 +26
stoep	-40 -31 -21	roem	+13 +29 +31	groen	+16 +32 +36	bloes	+1 +6 +18
snoep	-40 -31 -20	doem	+12 +34 +46	zoen	+26 +27 +29	moes	-12 -5 -3
/v/		/f/		/ɣ/		/r/	
Word	Distance	Word	Distance	Word	Distance	Word	Distance
hoef	+8 +16 +22	sloef	-41 -40 -37	vroeg	+9 +11 +24	boer	+40 +51 +68
boef	+15 +30 +34	stoef	-32 -24 -22	boeg	+22 +23 +30	hoer	+42 +44 +63
groef	+7 +20 +44	toef	-32 -20 -16	kroeg	-33 +12 +31	stoer	+41 +48 +67
/l/				/j/			
Word	Distance			Word	Distance		
koel	-15 +27 +61			gloei	-34 -29 -28		
boel	+18 +36 +38			boei	-32 -20 -3		
				roei	-35 -35 -23		

Participant code: GL98		
Consistently long	**Consistently short**	**Inconsistent**
hoed, moed, voed, koe, hoe, moe, boef, hoef, groef, boeg, vroeg, boel, doem, roem, groen, zoen, doen, stoer, boer, hoer, hoes, bloes, groet, stoet	stoef, toef, sloef, gloei, boei, roei, boek, doek, zoek, bloem, stoep, snoep, groep, moes, moet	kroeg, koel

Appendix I - Etymology and cognates

Loanwords	
Word	Origin[11]
hoes	French, ~12th century
boer	Eastern Low Franconian dialect, ~15th century
stoer	Eastern Low Franconian dialect, ~18th century
kroeg	unsure
groep	Possibly French, ~16th century
stoef	Possibly French, ~16th century
boei	French, ~16th century
toef	French, ~14th century

Tussentaal	Antwerp[12]	Thorn[13] & Venlo[14]	Den Bosch & Tilburg[15]	Proto-WGerm.[16]
moe	[my:x]	[mø:x]	[my:x]	*mōþī
koe	[kuj]	[ku:]	[kuj]	*kū
hoe	[(h)u:]	?	[hu:]	*hwō
moet	[mut]	[mot]	[mɔt]	*mōtan
groet	[ɣry:t]	?	[ɣry:t]	*grōtijan
stoet	?	?	[stu:t]	?
moed	[mu:t]	[mo:t]	[mu:t]	*mōd
hoed	[(h)u:t]	[ho:t]	[hu:t]	*hōd
voed	[vy:t]	[vø:t]	[vy:t]	*fōdijan
boek	[buk]	[bo:k]	[bu:k]	*bōk
zoek	[zy:k]	[zø:k]	[zy:k]	*sōkijan
doek	[duk]	[do:k]	[du:k]	*dōk
groep	[ɣrup]	?	?	/
stoep	?	[stu:p]	[stu:p]	*stōp
snoep	[sny:p]	?	[sny:p]	?
bloem	[blum]	[blo:m]	[blɔm]	*blōmō[17]
roem	?	?	?	*hrōm
doem	?	?	?	*dōm
doen	[duŋ]	[do:n]	[du:n]	*dōn
groen	[ɣry:n]	[ɣrø:n]	[ɣry:n]	*grōnī
zoen	?	?	?	*s(w)ōnō

[11] van der Sijs, 2010. etymologiebank.nl
[12] Camerman, 2009
[13] Bokken et al., 2011
[14] van der Sijs, 2015. ewnd.ivdnt.org
[15] Swanenberg, 2011
[16] van der Sijs, 2010. etymologiebank.nl; de Vries et al., 1971
[17] Brabantian reflex is most likely from an alternative form /blumme/ (Paul Boersma, p.c.)

hoes	?	?	?	/
bloes	[blu:s]	?	?	/
moes	[mu:s]	[mo:s]	[mu:s]	*mōs
hoef	?	[ho:f]	[hu:f]	?
boef	?	?	?	*bōbō
groef	[ɣru:f]	[ɣro:f]	[ɣru:f]	*grōbō
sloef	[sluf]	?	[slɔf]	?
stoef	[stuf]	[stof]	[stɔf]	/
toef	[tuf]	?	?	/
vroeg	[vry:x]	TR [vrø:x] / VL [vry:x]	[vry:x]	*frōwī
boeg	?	?	?	*bōg
kroeg	?	?	?	?
boer	[bu:r][18]	[bu:r]	?	*būr
hoer	?	[ho:r]	[hu:r]	*hōrō
stoer	?	?	?	?
koel	[ky:l]	[kø:l]	[ky:l]	*kōlī
boel	[bu:l]	[bu:l]	[bu:l]	*bōþl
gloei	[ɣly:j]	[ɣlø:j]	[ɣly:j]	*glōan
boei	?	?	?	/
roei	[ry:j]	?	[ry:j]	*rōan

[18] Most likely a borrowing from an eastern dialect

Appendix J - Textgrid-isolating Praat script

```
# Init
tgrid = selected("TextGrid")

selectObject: tgrid
nTiers = Get number of tiers
for i from 1 to nTiers
	nIntervals = Get number of intervals: i
	for j from 1 to nIntervals
		str$ = Get label of interval: i, j
		strLen = length(str$)
		if strLen > 0
			intBegin = Get start time of interval: i, j
			intEnd = Get end time of interval: i, j
			appendInfoLine: str$, " ", intEnd-intBegin
		endif
	endfor
	appendInfoLine: ""
endfor
```