University of Amsterdam
Faculty of Humanities

One without the other - Can German learners of English learn vowel duration without quality?

Maxim Dauenhauer

Student number: 12993050

BA Thesis Linguistics

Supervisor: Dr. Marijn van 't Veer

26th June, 2023

Abstract

Many German learners struggle with the English vowels in pairs like bad-bed. Previous research (Bohn & Flege 1992) has shown that inexperienced L1 German learners of English are not able to distinguish the two vowels with either quality or duration, while experienced learners are able to use both. The current study expanded on these findings by including three coda conditions as opposed to only one like in Bohn and Flege (1992) and by recruiting participants with an in-between experience. Results showed that the German participants produced at least an F1 and duration contrast, whereas native speakers used F1, F2 and duration. While the learners clearly made a quality distinction, the smaller effect size suggests a non-native quality contrast. Duration was replicated native-like, both in regards to vowel and coda-vowel differences. This is interpreted as evidence that different aspects of L1 sound system can be acquired at different rates.

Acknowledgements

I would like to thank my supervisor Dr. Marijn van 't Veer for his guidance, Dr. Titia Benders for very helpful feedback regarding statistical interpretation, Prof. Paul Boersma for help with the statistical analysis and R code, Sayuri Deokar for proofreading, feedback, emotional support and enormous help with participant recruitment and Diego Ostoja-Kowalski for a lot of help with LaTeX.

Contents

1 Introduction

What do you do when you encounter a sound in a language that you are learning that does not exist in your own language? If you have ever learned a foreign language, you have likely faced this problem before. For many learners, unfamiliar sounds can certainly be quite daunting. Whenever we study a new language, the sound inventory of the that language does not usually line up entirely with the sound inventory of our native language. Ideally, we would want to replace unacceptable foreign sounds with the closest one that we can find in our own language (Paradis & LaCharité 2011). Often, however, this means losing contrastive information that distinguishes two sounds.

One such contrast that is often difficult to reproduce for learners who do not have it natively are the two English vowels found in 'bad' versus 'bed', i.e., /æ/ and /ɛ/. Standard German, for instance, only has the open-mid vowel /ɛ/, which is comparable to English /ɛ/, although there is some evidence that the German vowel is somewhat lower (Bohn & Flege 1992)[1]. The vowel /ɛ/ often also matches in cognates: compare German 'Bett' /bɛt/ with English 'bed' /bɛd/ (Harbert 2006). Crucially, there is no comparable vowel to /æ/ in German (Wiese 2000), meaning that there is no equivalent to the English minimal pair 'bad - bad' /bæd - bɛd/. So, how do German learners of English deal with this? This question will be explored in the current study.

[1] It should be noted that some German dialects have phonemic /æ/ (Flege et al. 1997), although this did not seem to be the case for any of the participants in this study.

Section 1.1 will illustrate how the contrast is produced and perceived by L1 English speakers, while section 1.2 will do the same for L1 German speakers. Lastly, Section 1.3 will explain the relevance of these findings for the present study and why the /æ/ - /ɛ/ contrast is particularly interesting to better understand the progression of L2 sound system acquisition.

1.1 The English /æ/ - /ɛ/ contrast

Section 1.1.1 will explain how native English speakers make use of vowel quality to make the /æ/ - /ɛ/ contrast; Section 1.1.2 will do the same for duration. Lastly, Section 1.1.3 will address how native English speakers perceive the /æ/ - /ɛ/ contrast, followed by a short summary in Section 1.1.4.

1.1.1 Production - Quality

In English, the /æ/ and /ɛ/ differ phonetically in both duration and quality (Bohn & Flege 1992). Hagiwara (1997) reports an average F1 difference of roughly 200 Hz for women and about 150 Hz for men (see Figure 1, left panel). In many dialects of North America, however, /æ/ is raised in some contexts, which can lead to the average F1 of the two vowels to be very similar (Hillenbrand et al. 1995).

As for F2, Hagiwara (1997) measured a difference of roughly 300 Hz for women but almost no difference for men. These results present quite a sharp contrast to Hillenbrand

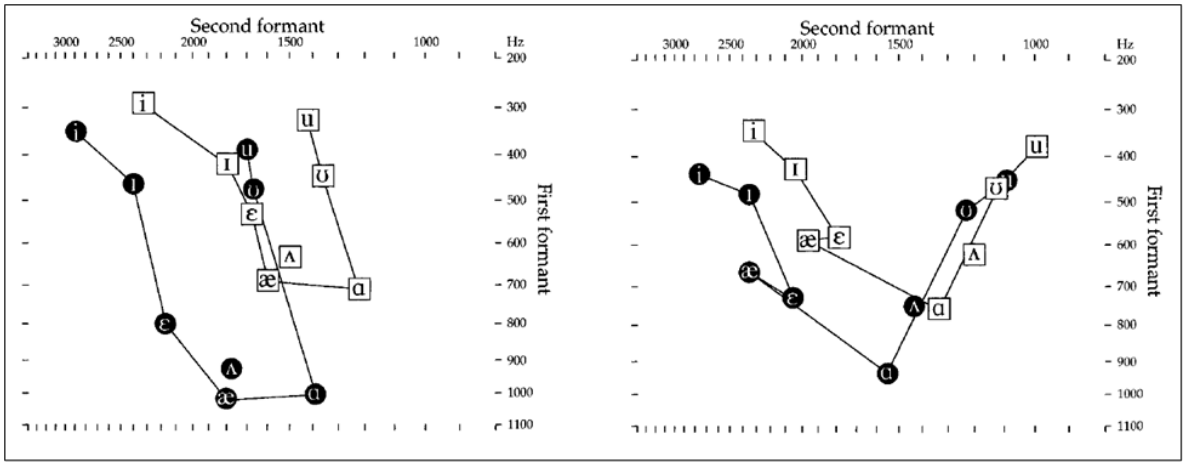

Figure 1: F1 and F2 of English vowels as measured by Hagiwara (1997) (left) and Hillenbrand et al. (1995) (right). Filled circles indicate female speakers' and open squares male speakers' vowel centers. Units are Hertz in Bark scale. Figures taken from Hagiwara (1997).

et al. (1995), who found /æ/ to actually be higher in the vowel space than /ɛ/ for women and on roughly the same height for men. The two vowels mostly differ in frontness, with /æ/ being slightly more fronted than /ɛ/, which is also opposite to what Hagiwara (1997) found. Again, the measurements by Hillenbrand et al. (1995) and Hagiwara (1997) respectively can be seen in Figure 1. Their speakers were from the Midwest, United States in the former study and California, United States in the latter.

This variation is the result of several fronting and raising processes, such as /æ/ raising before nasals, which results in /æ/ being realized with a more [eə̯] - like articulation (Mielke et al. 2017). The important fact is mainly that, depending on a speaker's dialect, the two English vowels usually differ in vowel quality, namely F1 and/or F2.

1.1.2 Production - Duration

As for duration, Umeda (1975) reports /ɛ/ to be shorter than /æ/ in all of the three conditions (prepausal, monosyllabic and polysyllabic) tested and for both speakers, who were native speakers of English from Alabama, United States. Similarly, Crystal and House (1988) measured mean durations of 106 ms for /ɛ/ and 159 ms for /æ/ in stressed positions. Their participants were also native speakers of American English.

However, there are many factors that influence vowel duration beyond their 'default' duration differences, notably the vowel's phonological environment. Onsets seem to have little effect on the vowel, with /h/ being the notable exception, as it seems to have a shortening effect on /æ/ (Umeda 1975). The effect of codas, on the other hand, is much more complicated. In the prepausal condition, i.e., when the target vowel is in the last syllable of an utterance, codas could be ordered as follows: voiceless stop « voiceless fricative « nasal « voiced stop « voiced fricative (Umeda 1975), with increasing duration of the preceding vowel. See also Figure 2.

In the monosyllabic condition, i.e., vowels in monosyllabic words that do not occur at the end of a sentence, the order of voiced stop and voiced fricative is reversed. As for the polysyllabic condition, the same trend seems to hold true as for prepausal conditions. Thus, generally speaking, vowels before voiced consonants or fricatives are longer than those before voiceless consonants and before plosives, with nasals falling in the middle of the spectrum. Importantly, voicing seems to have a larger effect on

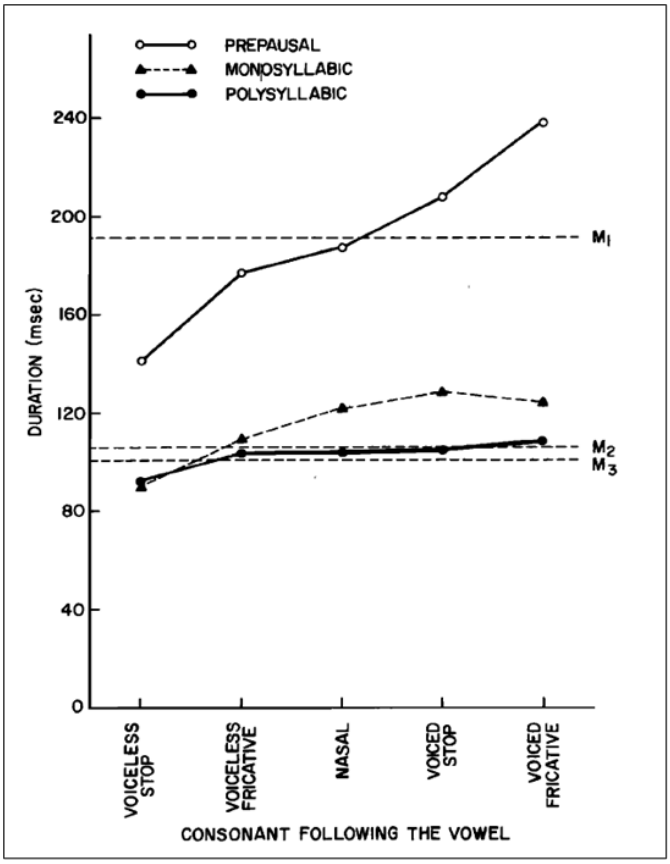

Figure 2: Average vowel duration of one speaker in prepausal condition depending on the following consonant. M1 represents the mean duration of all vowels in prepausal condition. Figure taken from Umeda (1975).

vowel duration than manner of articulation. This is not surprising since preceding vowel duration is an important cue for voicing of word-final stops in English (Raphael 1972). In other words, native English speakers use the duration of pre-coda vowels to distinguish between word-pairs like 'bad' and 'bat'.

In stressed prepausal conditions, Crystal and House (1988) report mean durations of 155 ms for short vowels before voiceless stops, 165 ms for short vowels before voiced stops and an even larger difference for long vowels with 179 ms and 216 ms respectively. The focus here is on stressed prepausal vowels since the duration differences are usually

most pronounced here and are most comparable to a wordlist-reading task, which was used in this study (see Section 2.1).

Crystal and House (1988) also present somewhat conflicting evidence to Umeda (1975) regarding coda influences on vowel duration. They report manner of articulation of coda consonants to only be a relevant factor on the duration of long vowels, with averages of 130 ms for stops versus 157 ms for fricatives. As for sonorants versus obstruents, they claim that any difference in duration would solely be the result of interactions with stress-conditions, with no underlying difference between the two. However, regarding fricatives versus stops, they also only seem to have an effect on the duration of long vowels (Crystal & House 1988). They found that fricatives lead to longer durations, which is in fact in line with Umeda (1975). As for place of articulation, stressed vowels are shorter before velars than before alveolars or labials, although this refers to non-prepausal conditions (Crystal & House 1988). While the exact effects of the coda on vowel duration are not always clear, partly due to the high number of influencing factors (place of articulation, manner of articulation, voicing, prepausality, stress, inherent/historical vowel length), it seems quite evident that the coda has at least some effect on vowel duration. This means that any study looking at duration differences between English vowels should not disregard the influence that the coda may or may not have on duration measurements. For learners of English, such as the participants in this study, this means that not only do they have to acquire the overall duration

differences between English /æ/ and /ɛ/, but also the coda effects, since there are no reports of German using duration in a similar way as a coda cue.

1.1.3 Perception

As for perception, English speakers mostly distinguish /æ/ and /ɛ/ with spectral properties, i.e., F1 and F2. Bohn & Flege (1990) report that in perception, spectral cues are more salient than duration for native speakers of American English, as they shift their classification of synthetic stimuli on a bet-bat continuum more after a change in vowel quality than a change in vowel duration. This makes sense due to the aforementioned importance of vowel duration as a cue for voicing of the following consonant (Raphael 1972).

1.1.4 Summary - English speakers

In summary, native English speakers use both duration and quality for perception and production of the /æ/ - /ɛ/ contrast, with evidence from perception studies suggesting that quality is treated as more important than duration. The exact nature of the quality contrast differs in the literature, possibly due to dialectal and/or methodological differences. However, both F1 and F2 are generally reported to be of importance for native speakers. Lastly, duration in English does not only differ between /æ/ and /ɛ/, but also depending on the coda, where the most important factor seems to be voicing.

1.2 The English /æ/ - /ɛ/ contrast by German speakers

The following sections will illustrate how L1 German learners of English produce (Section 1.2.1) and perceive (Section 1.2.2) the /æ/ - /ɛ/ contrast. Additionally, Section 1.2.3 will address the influences of the phonological context of vowels and final devoicing in German learners' production of English.

1.2.1 Production

Bohn and Flege (1992) demonstrated that inexperienced German learners of English are not able to produce a significant contrast between /æ/ and /ɛ/, while experienced learners are able to do so, albeit not in the same way as native English speakers. They recorded two groups of L1 German learners of English who had either spent little (mean = 0.6 years) or much time (mean = 7.5 years) in an English-speaking country, as well as a group of L1 English speakers from Alabama, United States.

The inexperienced group produced neither a significant difference in duration nor in quality. In other words, they merged the two English vowels into a category whose F1 was located somewhere between English /æ/ and German /ɛ/ spectrally. The experienced group, on the other hand, was able to produce both a significant duration and quality contrast. Their /æ/ was spectrally not significantly different from the native English speakers' /æ/. However, while the experienced German participants were able

to make a significant duration contrast between the two English vowels, the absolute durations of their /æ/ were shorter than those of native English speakers.

Bohn and Flege (1992) interpret these findings at least partly as evidence for their 'Equivalence Classification Hypothesis' (see also Flege 1987), which proposes that with enough experience, learners can successfully produce identical and novel L2 sounds, while struggling with L2 sounds that are similar but not identical to L1 sounds. Crucially, their study only used the English words 'bat' and 'bet' as stimuli, meaning that all other coda conditions have not been investigated yet. Considering the strong effects that different codas can have on English vowel duration, in combination with the fact that /CVt/-tokens are located towards the short end of the vowel duration spectrum, further investigation using stimuli with various codas is warranted.

Lastly, the Duden pronunciation dictionary (Kleiner et al. 2015) reports that German speakers sometimes only use duration but not quality for the /æ/ - /ɛ/ contrast in English loanwords. As an example, they give the minimal pair of 'Geck' [gɛk] ('fool') versus 'Gag' [gɛːk] ('gag') which they claim is only distinguished by vowel length, due to final devoicing also taking place. However, Kleiner et al. (2015) do not refer to any source when making this claim and I am also not aware of this ever being shown in an experimental study. Thus, the present study aims to provide evidence for this claim. This 'duration only' contrast would be natural for German speakers, as they should be used to the distinction between /ɛ/ and /ɛː/ from formal German (Flege et al. 1997); Compare 'nett' /nɛt/ ('nice') and 'wähle' /vɛːlə/ ('choose') (Kleiner et al.

2015). Theoretically, possible alternatives to this would be either faithful replication of the English vowel contrasts even in loanwords or full neutralization, like the experienced and inexperienced groups from Bohn and Flege (1992), respectively.

1.2.2 Perception

Not only do L1 German speakers produce the English /æ/ - /ɛ/ contrast differently than L1 English speakers, there is also evidence that they both use perceptual cues differently. As mentioned in Section 1.1, native English speakers seem to rely more on spectral cues than duration. Bohn & Flege (1990) suggest based on their findings that the opposite might be true for native German speakers. In one experiment, their L1 German participants judged German [ɛ] to be the most similar German vowel to English [ɛ], while they chose [ɛː ∼ eː] and [ɛa̯] to be the most similar to English [æ]. Their second experiment also corroborated this; duration was a more salient cue than quality for German speakers when classifying synthetic stimuli on a bet-bat continuum. Again, experienced German learners behaved more like native English speakers than inexperienced German learners did. Since they made more use of quality and less use of duration than inexperienced learners, experienced learners are somewhat in between the two other groups. The researchers conclude that the German speakers' "use of duration indicates [...] a general speech perception strategy that takes over whenever information conveyed by spectral differences is insufficient" (Bohn & Flege 1990: 326).

1.2.3 Phonological context and final devoicing

Section 1.1 illustrated that the coda of a syllable can influence the duration of its vowel quite significantly in English. One notable difference between German and English that might be relevant here is the fact that German features final obstruent devoicing, while English does not. While L1 German speakers have been shown to produce some cues indicating a word-final voicing contrast in their own language, the differences are quite subtle in comparison to the cues used by English speakers (Smith et al. 2009). This indicates that they transfer their own final-devoicing rule, albeit not fully, since German speakers' stop voicing is even more neutralized in German than in English (Smith et al. 2009).

Smith et al. (2009) also showed that German learners of English are able to produce a vowel duration difference between English syllables ending in a voiced or voiceless stop, in a way that was comparable to native English speakers (see also Raphael 1972). Crucially, they did not produce a vowel duration difference of the same magnitude in German, meaning that they acquired coda-dependent vowel duration specifically for English rather than simply transferring a rule from German.

Furthermore, the German participants in Smith et al. (2009) produced a much less pronounced difference between word-final voiced and voiceless stops, but maintained the duration difference. As a result, pairs like 'bet' - 'bed' were only distinguished by vowel duration. This could be interpreted as a 'precedent' for German speakers

resorting to duration when a contrast is made in a way that is not familiar to them, such as word-final stop voicing. Therefore, it seems plausible that something similar would happen with the /æ/ - /ɛ/ contrast, i.e., that the quality distinction would be lost while the duration difference is not. This would be logical because German has phonological vowel length (Wiese 2000), so differentiating vowels with duration as their main cue should be a familiar concept to German speakers. This is an important difference between English and German – English speakers seem to use duration only as a secondary cue (Bohn & Flege 1990), whereas German has actual minimal pairs that are only distinguished by phonological length, i.e., by duration phonetically (Wiese 2000). Compare, for instance, 'Bann' /ban/ ('ban') and 'Bahn' /baːn/ ('rail, train').

1.3 Present study

As outlined in Section 1.2, previous research has shown that experienced L1 German learners of English are able to produce the English /æ/ - /ɛ/ contrast more or less faithfully, i.e., using both duration and vowel quality, while inexperienced learners use neither. However, Bohn and Flege (1990, 1992) somewhat simplify the important variable of experience. They divided their participants into two groups, depending on how much time a participant has spent living in an English-speaking country. While Bohn and Flege demonstrated how learners start out using neither of the two main cues and end up using both somewhat well, they do not provide any data on how a possible intermediate stage might look like. This is worth investigating because it might give some

indication how learners progress from a rather L1-like production to a more faithful L2 production. Therefore, the present study aims to expand Bohn and Flege's notion of experience by recruiting participants who do not necessarily fit this dichotomy, namely L1 German speakers who have lived in an 'international city' (Amsterdam, the Netherlands) for a few years, where English is commonly spoken as both an L1 and L2 by expats and locals alike (Hinskens & Muysken 2007).

Furthermore, the studies by Bohn and Flege (e.g. 1992) only ever recorded tokens of English bVt words, i.e. 'bet' and 'bat'. The voiceless stop /t/ is usually preceded by relatively short vowels, meaning that it is unclear how German speakers would treat vowel durations in words with codas that have a lengthening effect on the preceding vowel. As described in Section 1.3, German speakers seem to weight duration as more important than English speakers do, meaning that it is possible that they would make use of duration more when distinguishing the two vowels in words with overall longer vowel durations.

Combining these two expansions on previous work (coda and experience), the present paper will try to address the following general research question:

> "How do native speakers of German living in Amsterdam produce the English /æ/ - /ɛ/ contrast in different phonological contexts?"

The following paragraphs illustrate the hypotheses for this study. First, the participants in this study are meant to represent a somewhat intermediate state in the learning of the

two English vowels. This can be justified because they are in between the inexperienced and experienced group from Bohn and Flege (1992). In the present experiment, the participants had been living in Amsterdam for a mean of 2.50 years (see Section 2.1), as opposed to 0.6 and 7.5 years for the inexperienced and experienced group, respectively in Bohn and Flege (1992). In this 'intermediate state', the participants are hypothesized to use duration to distinguish /æ/ and /ɛ/ due to familiarity with length contrasts in German, while using the same vowel quality, as claimed by Kleiner et al. (2015). Such a finding would be interesting insofar as it would show that language learners are able to use more advanced techniques when dealing with novel sounds than simply replacing them with an L1 category. It should be noted that Bohn and Flege (1992) did not consider such an 'intermediate state' and hence provided no evidence for or against it. The present study aims at filling this gap.

Second, in order to expand on their study, the variable 'phonological context' is specifically included and is a more exploratory aspect of this paper. Since there is no way of knowing whether the participants in this study would behave more like the inexperienced or experienced groups from Bohn and Flege (1992) and there is no data on CVd and CVn tokens, formulating any hypotheses regarding coda influence on vowel quality is difficult.

Lastly, I hypothesize that that German speakers will have the same pattern of increasing vowel durations from /t/ to /d/ codas as native English speakers, as has been shown by Smith et al. (2009). This is logical because replicating the coda influences on

English vowel duration is an easy and familiar way for them to compensate for loss of information due to final devoicing. The hypotheses are summarized below.

i. The English vowels /æ/ - /ɛ/ will be distinguished with duration by both the German and the English group, while it will be distinguished with F1 and F2 only by the English group. There might additionally be an effect of coda on F1 and/or F2 for both groups, which is a more exploratory aspect of this study.

ii. The German participants will have increasing vowel durations for the codas /t – n – d/ in the same way as the English-speaking participants will.

2 Methods

In order to test the hypotheses, a wordlist reading task was designed, which was approved by the ethics committee of the University of Amsterdam, Faculty of Humanities (project number: FGW-873).

2.1 Participants

10 native speakers of German (GE) and 7 native speakers of American English (EN) were recruited in the Amsterdam region. One participant from the GE group was excluded because he reported to be dyslexic after finishing the experiment and two more participants from the same group had to be excluded because the audio files containing their recordings were corrupted. This left 7 participants per group for the analysis. A closer breakdown of the participants can be found in Tables 1 and 2.

The GE participants had been living in Amsterdam for a mean of 2.35 years. No participant from the GE group had more than one L1. They were all raised in Germany and had learned English through instruction at school in Germany for a mean of 9.70 years. The GE participants grew up in Baden-Württemberg (1), Hamburg (1), Brandenburg (1), Lower Saxony (1), Rhineland-Palatinate (1), North Rhine-Westphalia (1) and Bavaria (1). It should be noted that in this group, 3 participants indicated that they had lived in an English-speaking country, namely USA (10 months), Australia (6

	Mean	SD	Min	Max
Age	22.29	1.98	20	26
Years living in Amsterdam	2.50	1.26	1	5
Years of English instruction	10.57	1.51	9	12
English proficiency	8.29	0.76	7	9

Table 1: GE Participants' age, how much time they have spent living in Amsterdam, how many years of formal (school) English instruction they have had in Germany and how they would rate their own English proficiency on a scale of 0 - 10.

	Mean	SD	Min	Max
Age	29.00	9.12	20	44
Years living in Amsterdam	9.92	7.11	2	20

Table 2: EN Participants' age and how much time they have spent living in Amsterdam.

months) and Canada (did not specify). Due to the relatively small participant number and the overall short durations of their stay, it was decided not to exclude these participants. This will be addressed in Section 4.3. Nevertheless, the self-rated proficiency in the GE group was incredibly homogeneous. Most participants reported a proficiency of 8 (very good) or 9 (excellent); only one reported a proficiency of 7 (good) (see also Table 1).

Two speakers in the EN group grew up in Europe and not in the United States, but had gone to an international school, where they acquired an American English accent. One of them reported to be bilingual with Dutch and English. The other 5 speakers in the EN group had all grown up in the United States, namely in Massachusetts (2), California (1), Indiana (1) and Washington, D.C. (1).

Participants from both groups had only moved to Amsterdam after reaching adulthood. The male-to-female ratio was 5:2 in the GE group and 4:3 in the EN group.

2.2 Stimuli

The stimuli for this production experiment consisted of 18 monosyllabic target words of the structure CVC (see Table 3) and 18 filler words of the same structure, but with different vowels. This simple syllable structure was chosen for all stimuli to minimize effects of prosody. Most target and filler words were nouns, while a few others were verbs and adjectives. A full list of stimuli can be found in appendix A.

	CVt	**CVn**	**CVd**
/**æ**/	bat	ban	bad
/**ɛ**/	bet	Ben	bed

Table 3: Illustration of the six conditions with bVC tokens.

For the target words, there were three coda conditions and two vowel conditions, namely /æ/ and /ɛ/. The three possible codas were /t/, /n/ and /d/. The codas were chosen because the literature showed voicing to be the most important factor, with /n/ as the supposed middle of the spectrum (Umeda 1975). Since manner and place of articulation were shown to have less of an influence (see Section 1.1) and their influence is less clear, fricatives and non-alveolars were not included as codas of the target words. Alveolar stops were chosen for no specific reason other than that they make stimuli

selection easier due to their high frequency. Onsets were matched across conditions where possible, meaning that minimal pairs such as 'bad – 'bed' were favored. /h/ was avoided as an onset due to the shortening effects reported by Umeda (1975).

2.3 Procedure

The participants' speech was recorded in a soundproof booth at the speech lab of the University of Amsterdam. The stimuli were presented in two blocks, i.e., each individual stimulus was presented twice in total. In each block, all target and filler words were presented one by one on a screen in a different randomized order. Participants had to press the space bar in order to proceed to the next stimulus. They were instructed to read every word embedded in the carrier phrase "I will say _____", analogously to Bohn and Flege (1992). This mimics the prepausal conditions of Umeda (1975) and Crystal and House (1988), i.e., the target vowel occurred in the last syllable of an utterance (see Section 1.1). Prepausal utterances were chosen because, as mentioned previously, the vowel duration differences between coda conditions are most pronounced in this context. Participants were told to read the words at a normal pace, i.e., neither too fast nor overly slow, and to say the words as they would in any day-to-day conversation.

After recording, participants were asked to fill in an adapted version of the LEAP-Questionnaire (Marian et al. 2007) in order to control for any individual language background related or sociolinguistic influences. The questionnaire can be found in Appendix B. Adaptation consisted mostly of simplifying the LEAP-Q, since there is

no need for an in-depth background analysis in this study and it would be difficult to justify collecting a large amount of data for post-hoc analyses only.

2.4 Data Analysis

After all recordings had been made, all target vowels were manually annotated and their average first and second formants and total durations were extracted using the computer program Praat (Boersma & Weenink 2023). The script which was created in order to do this can be found in Appendix C. In order to extract the formants, the Burg algorithm was used, which is built-in in Praat (see also de Waele & Broersen 2000). All durations were log-transformed in order to account for speech rate. Therefore, any time that duration is mentioned in the following paragraphs, the log-transformed values are being referred to (unless specified otherwise). A conscious decision was made not to apply any formant normalization such as Lobanov (1971) in order to not risk introducing any artifacts. Furthermore, normalization would not really have been possible due to a lack of data of each participant's full vowel space.

These three dependent variables could then be used to determine whether and how speakers made use of vowel quality (F1 and F2) and duration. In order to test the first hypothesis, namely that the German participants will distinguish the two vowels only with duration but not with quality, several linear mixed-effects models were fitted, using the R-package 'lmerTest' (Kuznetsova 2017). First, three models (one for each dependent variable) were fitted for each participant group, with vowel as a fixed effect

and participant as a random effect. The resulting six models could then be used to check whether there is any evidence that either participant group distinguished the /æ/ and /ɛ/ with F1, F2 or duration, respectively. Since any effects are only really meaningful if compared between groups, these models are mostly just intended to get a rough understanding of the within-group results and not as the main analysis.

Furthermore, three more linear mixed-effects models were fitted (again one for each dependent variable) in order to check for any interactions between the fixed effects: vowel, participant group and coda. Again, participant was included as a random effect. Lastly, other independent variables that were collected via the questionnaire were tested for significant influences on all the dependent variables, using three linear mixed-effects models. These independent variables were age, gender, self-reported English proficiency, years spent living in the Netherlands, years of formal English instruction, age when starting to learn English and age of reaching fluency in English.

For all linear mixed-effects models, the contrasts for the variable 'coda' were coded in a way that the model checked a) whether there is any effect of /t/ vs /d/ (the alleged endpoints of the duration spectrum) and b) whether /n/ lies exactly in the middle of these two endpoints.

3 Results

All statistical analysis was carried out in RStudio (R Core Team 2023). Plots were created using the online tool Visible Vowels (Heeringa & van de Velde 2018) and RStudio.

First, it should be noted that some participants exhibited various degrees of vocal fry. However, the program still seemed to be able to extract reasonable formant values from all of them. For this reason, none of the participants in question were excluded. Table 4 illustrates the overall results with some descriptive statistics of the dependent variables.

Group	Vowel	F1	F2	Duration	Log Duration
GE	/æ/	631.7 (26.6)	1571.9 (139.5)	0.19 (0.04)	-1.71 (0.2)
	/ɛ/	581.8 (26.8)	1620.2 (111.2)	0.12 (0.03)	-2.09 (0.2)
EN	/æ/	825.1 (117.2)	1755.1 (207.9)	0.22 (0.07)	-1.55 (0.3)
	/ɛ/	694.0 (81.3)	1851.1 (147.4)	0.15 (0.04)	-1.96 (0.3)

Table 4: Mean F1, F2 in Hertz, duration in seconds and duration (log-transformed) per group and vowel. Values in brackets indicate standard deviations. GE and EN speakers have a significant F1 and duration difference between /æ/ and /ɛ/. The F1 difference is smaller in the GE group, while duration differences are comparable. EN speakers also have a significant F2 difference.

Section 3.1 will address the results of the six linear mixed-effects models that were fit for each dependent variable in each group, while Section 3.2 will address the three linear mixed-effects models that were fit with the entire data, with subsections dedicated to each dependent variable and one to inter-speaker variation. Section 3.3 will shortly address the results of the models fitted with questionnaire data for post-hoc analysis.

3.1 Within-group effects

The GE speakers' F1s were 50.0 Hz higher on average for /æ/ than for /ɛ/ ($p = 0.0069$) and their duration was 0.38 s (log) higher on average for /æ/ than for /ɛ/ ($p < 0.001$). The model for F2 in the GE group yielded no significant effect for vowel, although the confidence interval includes relatively large effects where the F2 for /ɛ/ would be higher than for /æ/ (95% CI = -118.68 ... 22.07). Therefore, there is no evidence for or against an effect of F2 on vowel in the GE group. Figure 3 illustrates that an F2 effect is not clearly present in the GE group, but an F1 effect is, at least in /t/- and /d/-conditions. Since /n/-conditions had somewhat different results, a similar graph (Figure 7) displaying those results can be found Section 3.2.1.

As for the EN group, their F1s were 131.1 Hz higher on average for /æ/ than for /ɛ/ ($p < 0.001$), while their F2s were 96.0 Hz lower on average for /æ/ than for /ɛ/ ($p = 0.0020$). Both F1 and F2 effects for this group are clearly visible in Figure 3. Furthermore, their duration was 0.41 s (log) higher on average for /æ/ than for /ɛ/ ($p < 0.001$).

This means that, if the coda is not taken into account, both groups produced a significant F1 and duration difference between the target vowels, although the effect size for F1 is much larger in the EN group. The effect size of duration is roughly the same in both groups. As was to be expected, the EN group also produced a significant F2 difference, whereas no clear statement can be made about an effect of F2 in the GE group due to the aforementioned large confidence interval. The overall differences of all

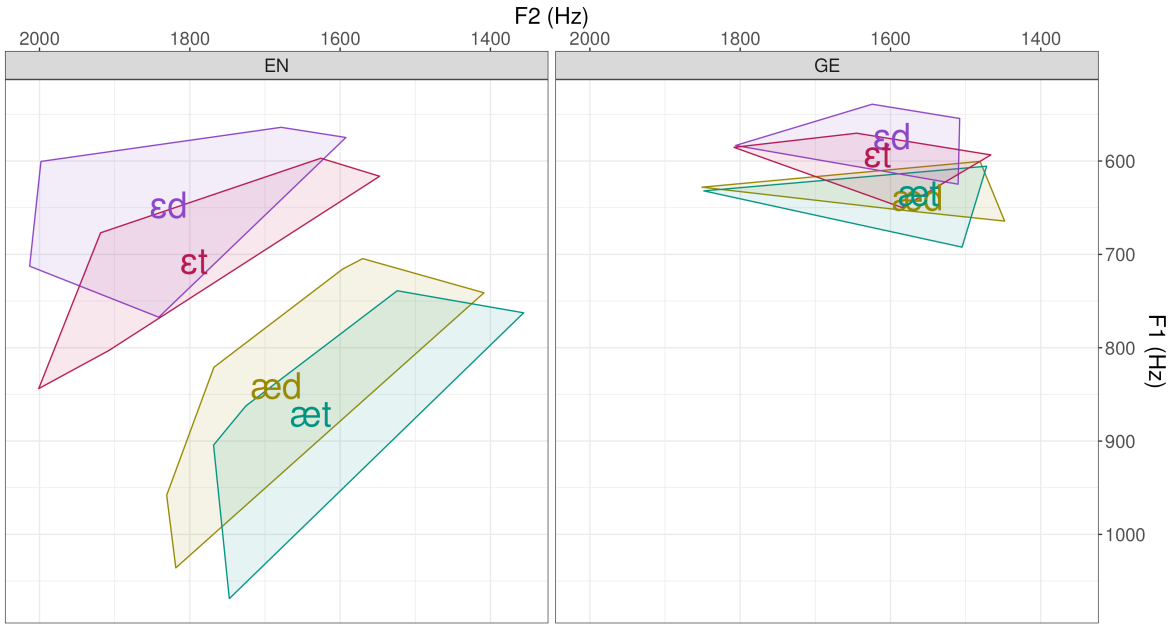

Figure 3: Range of the English vowels /æ/ and /ɛ/ in the coda conditions /d/ and /t/, as produced by the EN group (left) and GE group (right). In these conditions, EN speakers seem to make the /æ/ - /ɛ/ contrast with both F1 and F2, while GE speakers appear to only use F1.

three target variables and for both groups are illustrated on the left side of Figures 4 through 6.

3.2 Between-group effects

Before evaluating the results of the linear mixed-effects models that were fitted with the entire data, it should be noted that direct group comparisons are somewhat unreliable since the EN group had a slightly higher female-to-male ratio than the GE group and no formant normalization had been applied. These sex differences[2] might at least partly

[2] The questionnaire administered before the experiment asked for gender. However, any phonetic differences between formants would be due to the participants' sex.

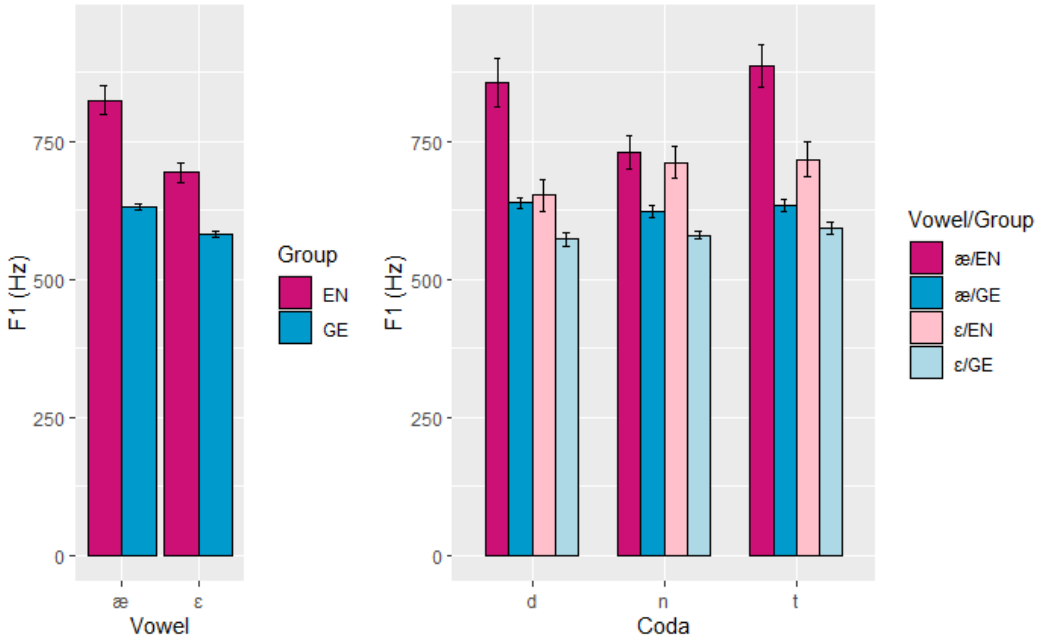

Figure 4: Mean F1 in Hertz of the English vowels /æ/ and /ɛ/ as produced by both groups overall (left) and for each coda (right). Error bars indicate standard error. Both groups make the contrast using F1, with the EN group having a larger effect size. EN speakers clearly exhibit some /æ/-raising before /n/ as a group.

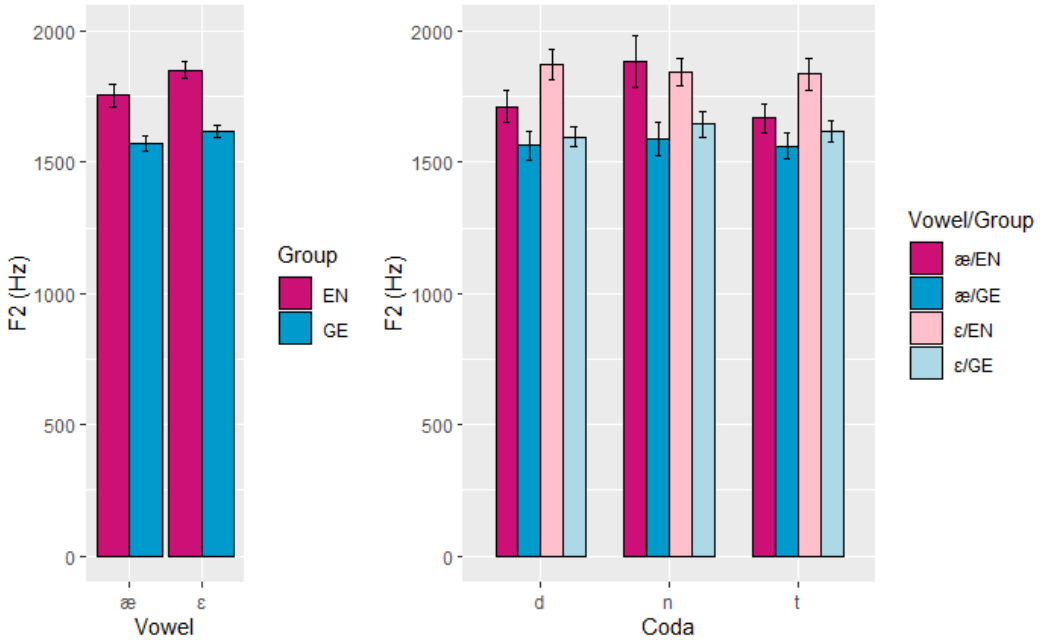

Figure 5: Mean F2 in Hertz of the English vowels /æ/ and /ɛ/ as produced by both groups overall (left) and for each coda (right). Error bars indicate standard error. Only EN speakers produced a significant F2 difference between /æ/ and /ɛ/.

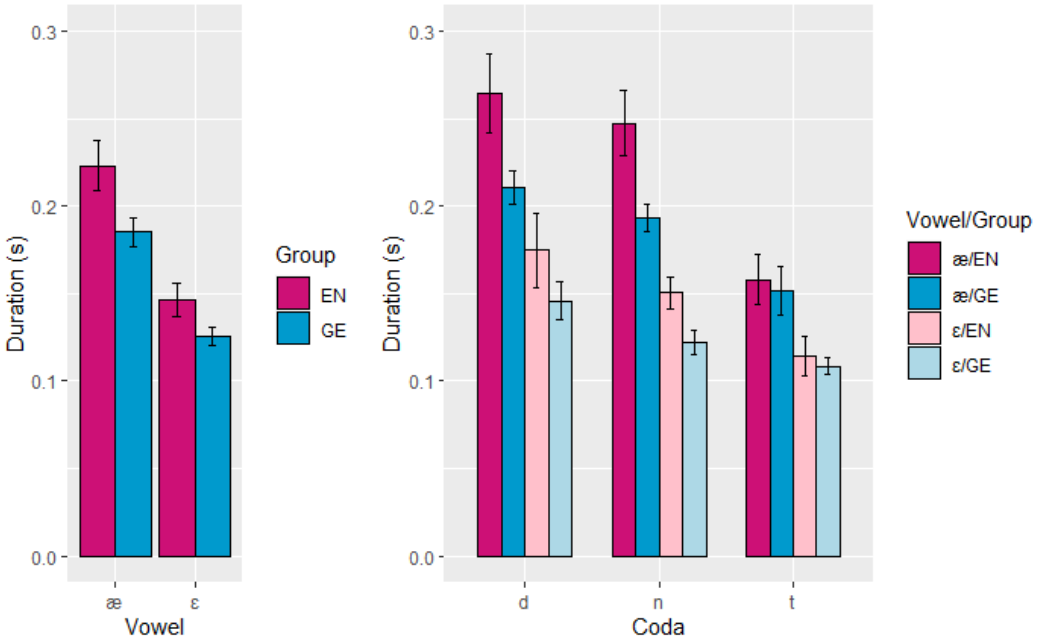

Figure 6: Mean duration in seconds of the English vowels /æ/ and /ɛ/ as produced by both groups overall (left) and for each coda (right). Note that the durations in the graph are not transformed (for readability), while in the actual analysis, they were log-transformed. Error bars indicate standard error. Both groups exhibit a native-like duration pattern.

explain the fact that the EN group overall had much higher formants than the GE group (see again Table 4). These models were fitted with data from both groups; therefore, only significant group effects and interactions with group will be presented in the following sections, since any effects across both groups are relatively meaningless.

3.2.1 F1

Figure 4 illustrates the results regarding F1. There was a significant group effect for F1, with the EN group having a 152.75 Hz higher F1 than the GE group ($p < 0.001$). There was a significant interaction between vowel and group; the F1 difference between /æ/ and /ɛ/ was 81.14 Hz higher in the EN group ($p < 0.0023$). Furthermore, there

was an unexpected interaction between coda /t/ - /d/ and group; the degree to which /t/-tokens have a lower F1 than /d/-tokens is 41.56 Hz larger in the EN group. There were no hypotheses made about coda effects on F1 or F2, meaning that this result was somewhat surprising. It seems that for both GE and EN speakers, the coda (i.e., whether the coda is /t/ or /d/) influences the F1, but more so in the EN group. Interestingly, there was a significant three-way interaction between vowel, group and coda /n/. For both groups, the F1 difference between /æ/ and /ɛ/ was larger in /n/-conditions compared to /t/- and /d/-conditions. However, this difference was 158.56 Hz larger in the GE group than in the EN group. This could have a relatively straightforward explanation, though. Many of the speakers in the EN group exhibited /æ/-raising in front of /n/, leading to a more [eə̯]-like articulation, which is common among speakers of North-American varieties of English (Mielke et al. 2017). In this study, however, only the average formant over the entire duration of the vowel was measured. This led to the vowels having similar average formants (as reported by Hillenbrand et al. 1995), even though a more sophisticated analysis would have revealed that the EN speakers clearly distinguished them, albeit with a 'monophthong versus diphthong' contrast. Since many, but not all, speakers from this group exhibited this behavior, with some still using the monophthong [æ] for /æ/, the spread of this vowel was very large before /n/ (see Figure 7, left panel).

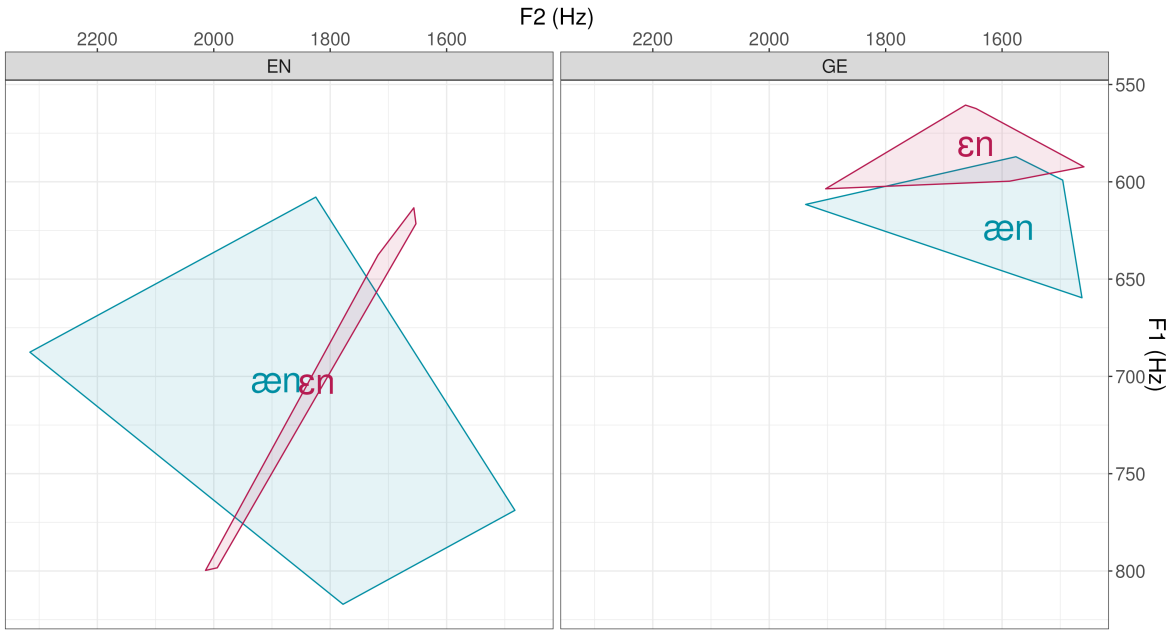

Figure 7: Range of the English vowels /æ/ and /ɛ/ in the /n/ coda condition, as produced by the EN group (left) and GE group (right). Due to /æ/-raising, EN speakers' range of the /æn/ condition is very spread out.

3.2.2 F2

There was also a significant group effect for F2, with the EN group having a 207.08 Hz higher F2 than the GE group ($p = 0.017$). Interestingly, there was no group effect for vowel, meaning that there is no evidence that any group has a larger F2 difference between /æ/ and /ɛ/ than the other (see also Figure 5). While this vowel-group interaction was not significant for F2 ($p = 0.29$), the direction of the estimated effect size follows the same pattern of the English group making a larger F2 difference, namely 47.71 Hz.

The same surprising coda interaction (/t/ - /d/) that was found for F1 was also significant for F2, meaning that the degree to which /t/-tokens are larger than /d/-tokens is 50.39 Hz larger in the EN group. Again, there was no hypothesis made about this. Lastly, the same three-way interaction that was found for F1 was again found for F2. In this case, however, the 'n-enhancing' effect on the vowel contrast is 217.80 Hz larger in the EN group. This basically means that there is quite a strong fronting effect for /n/-tokens in this group.

3.2.3 Duration

The effect of group on duration was barely not significant ($p = 0.058$). This means that while there is no evidence for a group effect, one can also not really be ruled out. Importantly, there is a significant interaction between both coda contrasts and vowel across both groups ($p = 0.038$ for vowel and coda /t/-/d/, $p = 0.0026$ for vowel and coda /n/). If both groups are taken together, there is evidence that a) the durations for /t/- and /d/-tokens are not the same and b) /n/ does not lie in between these endpoints, but closer to /d/. Figure 6 corroborates this, indicating that the pattern of a steady increase in duration from /t/ to /d/ is present in both groups.

The GE speakers produced a difference of 0.13s (log) between /d/ and /n/ and 0.19s (log) between /n/ and /t/, while the EN speakers produced a difference of 0.09s (log) between /d/ and /n/ and 0.38s (log) between /n/ and /t/.

3.2.4 Inter-speaker variation

It should furthermore be noted that there was considerable inter-speaker variation within both participant groups. Figures 8 and 9 illustrate how the formant centers of each vowel differ per participant. In the EN group, many speakers have similar locations and distances between the vowel centers, with some exceptions like speaker EN04, whose /æ/ is more fronted than that of the other speakers. However, in the GE group, some speakers like GE02, GE04 or GE06 appear to have very similar locations for both vowels, while some speakers like GE09 or GE10 clearly have two different vowel positions.

Figure 8: Mean F1 and F2 in Hz of the English vowels /æ/ and /ɛ/ as produced by each speaker in the EN group across all six conditions. The large symbol indicates the center of the six conditions.

Figure 9: Mean F1 and F2 in Hz of the English vowels /æ/ and /ɛ/ as produced by each speaker in the GE group across all six conditions. The large symbol indicates the center of the six conditions.

3.3 Post-hoc analysis of questionnaire data

Neither of the three linear mixed-effects models that were fitted with the questionnaire data yielded any significant results. Therefore, there is no evidence that any of the individual differences between the German participants influenced their performance, at least in terms of the collected background information.

4 Discussion

This Section is dedicated to interpreting the results from the previous section, regarding the first hypothesis (Section 4.1) and the second hypothesis (Section 4.2). This is followed by a section explaining the limitations of this study, together with possibilities for future research (Section 4.3).

4.1 Hypothesis 1 - /æ/ and /ɛ/ with quality and duration

The results were overall mixed, but they generally supported the first hypothesis, albeit not in the exact way laid out and less extreme. For convenience, it is stated here again:

i. The English vowels /æ/ - /ɛ/ will be distinguished with duration by both the German and the English group, while it will be distinguished with F1 and F2 only by the English group. There might additionally be an effect of coda on F1 and/or F2 for both groups, which is a more exploratory aspect of this study.

The GE participants in this study were – taken as a group – clearly advanced enough to have acquired at least some quality difference. The initial 'sanity-check', i.e., the within-group models, resulted in a significant F1 contrast across all codas. On the other hand, no F2 effects were found, although the absence of a significant effect does not mean evidence for no effect. As was to be expected, the EN group made a clear F1

and F2 difference between /æ/ and /ɛ/ averaged across coda conditions, meaning that they did indeed produce the contrasts in the same way as the literature reported that native speakers of North American English behave.

More importantly, these effects need to be put into perspective using between-group comparisons, which is why the three main models had been fitted using all of the data. The first important insight from these models was the fact that the EN participants' formants were both significantly higher across the board, indicating that the GE participants did not successfully emulate the native speakers' formants. However, this was already addressed in Section 3.2 as being at least partly due to the gender-ratio differences between the two groups, which led to disproportionately high formants in the EN group. Another finding was the significant interaction between vowel and group for F1, which suggests that the EN participants made use of a much larger height contrast between /æ/ and /ɛ/. While the group-vowel interaction for F2 suggests a similar pattern for F1, this effect is not significant, meaning that nothing can be concluded group-differences regarding the use of F2 to distinguish the vowels.

These results taken together suggest loosely that the GE speakers are making their quality contrast differently from the native speakers in this study, namely by possibly focusing more on vowel height rather than frontness (see figures 3 and 7 again), while still having less of a height difference than the native speakers. Therefore, the GE group seems to be able to make a quality contrast, albeit one that is different from the EN group and therefore arguably not native-like (or at least unlike North American

English). This would mean that the German speaking participants in this experiment can indeed be placed in between Bohn and Flege's (1992) inexperienced and experienced groups. The participants in their study either failed to produce a significant quality contrast or produced it native-like. In the present study, the GE participants produced a non native-like quality contrast. Thus, it seems like the learners in this study are well on their way to acquiring the quality aspect of the English /æ/ - /ɛ/ contrast, while still not behaving fully native-like.

As for duration, the absence of any group effects in the main models combined with significant duration effects on vowel within both groups suggests that GE and EN participants alike use duration to distinguish /æ/ and /ɛ/.

In summary, the prediction that the GE participants would not make any quality contrast whatsoever clearly did not hold true. However, the general idea that in production, they rely more on duration than English native speakers would, is generally supported. The reason for this conclusion is the overall smaller quality difference that GE speakers made for F1 while also possibly disregarding F2 as a differentiating factor, combined with the overall faithful replication of the duration contrast.

It should be noted that effects of coda on vowel quality were found, which was initially stated to be a more exploratory aspect of this study. From these effects two general conclusions can be drawn. First, some but not all EN participants exhibited /æ/-raising before /n/, while none of the GE participants did. This led to some very significant group interactions, which do not necessarily give any insight about the acquisition of the

quality contrast by the GE group other than the fact that they have not acquired any /æ/-raising. This is again illustrated in figure 7. Second, it can generally be concluded that the coda has come effect on quality which is different in the two groups, although there is not necessarily a discernable pattern. Future research could try to focus more on the coda effects on the two vowels by native speakers of different varieties of English and learners of English, possibly by controlling for any raising or similar effects.

4.2 Hypothesis 2 – Influence of the coda on duration

The second hypothesis, which is also stated again below, was supported quite clearly.

ii. The German participants will have increasing vowel durations for the codas /t – n – d/ in the same way as the English-speaking participants will.

The GE participants did indeed seem to be able to faithfully replicate not only the duration difference between /æ/ and /ɛ/, but also the steady increase in overall duration from the codas /t/ to /d/. This was shown by the absence of a significant duration difference between groups combined with significant coda effects on duration across both groups. Figure 6 illustrates very clearly that this pattern is present in both groups. This means that the participants in this study were advanced enough to replicate this feature of English phonetics, even though it does not exist in their own L1. This is in line with the findings from Smith et al. (2009), where their German participants were shown to preserve the vowel duration differences before word-final voiced and voiceless

stops while mostly neutralizing voicing differences (see Section 1.4). Thus, it seems likely that German native speakers pick up on these small duration differences as a way to compensate for the loss of a voicing contrast in the coda that they might exhibit themselves in perception and/or production (see Section 1.4). Being more familiar with vowel length as a contrastive feature but not with obstruent voicing in the coda, perceiving and producing these small duration differences would be a very plausible compensation strategy. This idea could be used as a topic for further research.

It should also be noted that the absolute durations of all vowels were consistently lower in the GE group than in the EN group. While this difference was not significant (possibly due to log-transformation), this overall trend lines up with the findings by Bohn and Flege (1992). It is not clear why German speakers would have shorter overall durations for these vowels, but this difference could be a last remnant of their non-nativeness. It should be noted here that speech rate was not measured since log-transformation of durations was specifically applied to counterbalance any speech rate effects. Thus, the overall duraiton differences could simply be the result of a faster speech rate in the GE group.

4.3 Limitations

This study treated the participants as homogeneous groups, even though this was clearly not the case. While all participants in the GE group were monolinguals, the exact path that they took learning English is naturally different for each individual. For instance,

three participants had indicated that they spent some time living in an English-speaking country, which could have naturally impacted their acquisition of English. As figures 6 and 7 illustrated, all of these individual differences led to participants being at different levels in their knowledge of English, which means that all insights gained from this study will always be a generalization that individuals might deviate from (as is the case for most studies). This is exacerbated by the fact that this study did not distinguish different levels of experience, unlike previous studies. While the overall goal of recruiting people in-between Bohn and Flege's (1992) two levels of experience was overall achieved, the individual participants in this study clearly still had different levels of experience that was not accounted for.

As for the analysis, no formant normalization was applied, which made direct comparisons between groups rather difficult to interpret. Nevertheless, the within-group differences and group interactions are still noteworthy results. Furthermore, all formants were measured by averaging over the entire duration of the vowel. Measuring formants at different timestamps, say the 25th, 50th and 75th quantile, might have led to more accurate measurements. Specifically, this method would have been more suitable for those instances where EN participants exhibited /æ/ raising before /n/, most likely clarifying the nature of the contrasts that EN speakers made between /æ/ and /ɛ/ in /n/ conditions.

5 Conclusion

The main insight from this study is that two features of a languages like vowel quality and vowel duration do not really behave detached from each other. Rather, progression is likely gradual with both aspects of the L2 being acquired simultaneously. Importantly, it does seem to be possible for both of them to progress *at different speeds*. In the case of this experiment, the German participants' ability to replicate the phonetic patterns of vowel duration of English /æ/ and /ɛ/ seems to exceed their ability to replicate the intricacies of the spectral differences between the same vowels. This is most likely related to the fact that the German language does have phonological length while lacking a contrast between two open front vowels. Therefore, parallels between learners' L1 and L2 might be able to 'boost' acquisition in the relevant areas.

6 References

Boersma, P., & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.09, retrieved 2 March 2023 from http://www.praat.org.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied psycholinguistics*, *11*(3), 303–328. https://doi.org/10.1017/S0142716400008912

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in second language acquisition*, *14*(2), 131–158. https://doi.org/10.1017/s0272263100010792

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

de Waele, S., & Broersen, P. (2000). The Burg algorithm for segments. *IEEE Transactions on Signal Processing*, *48*(10), 2876–2880. https://doi.org/10.1109/78.869039

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of phonetics*, *15*(1), 47–65. https://doi.org/10.1016/s0095-4470(19)30537-6

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of phonetics*, *25*(4), 437–470. https://doi.org/10.1006/jpho.1997.0052

Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America*, *102*(1), 655–658. https://doi.org/https://doi.org/10.1121/1.419712

Harbert, W. (2006). *The Germanic Languages*. Cambridge University Press.

Heeringa, W., & Van de Velde, H. (2018). Visible vowels: A tool for the visualization of vowel variation. *Proceedings CLARIN annual conference*, 8–10.

Hillenbrand, J., Getty, L. A., Wheeler, K., & Clark, M. J. (1994). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *95*(5), 2875–2875. https://doi.org/10.1121/1.411872

Hinskens, F., & Muysken, P. (2007). The talk of the town: Languages in Amsterdam 1507–2007. *Zeitschrift für Literaturwissenschaft und Linguistik*, *37*, 7–23. https://doi.org/10.1007/BF03379769

Kleiner, S., Knöbl, R., & Dudenredaktion (Bibliographisches Institut), I. (2015). *Duden - Das Aussprachewörterbuch*. Dudenverlag.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of statistical software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, *49*(2B), 606–608. https://doi.org/10.1121/1.1912396

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. https://doi.org/https://doi.org/10.1044/1092-4388(2007/067)

Mielke, J., Carignan, C., & Thomas, E. R. (2017). The articulatory dynamics of pre-velar and pre-nasal/æ/-raising in English: An ultrasound study. *The Journal of the Acoustical Society of America*, *142*(1), 332–349. https://doi.org/https://doi.org/10.1121/1.4991348

Paradis, C., & LaCharité, D. (2011). Structure preservation: the resilience of distinctive information. *The Blackwell Companion to Phonology*, 1–24. https://doi.org/10.1002/9781444335262.wbctp0076

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

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

Smith, B. L., Hayes-Harb, R., Bruss, M., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, *37*(3), 257–275. https://doi.org/10.1016/j.wocn.2009.03.001

Umeda, N. (1975). Vowel duration in american english. *The Journal of the Acoustical Society of America*, *58*(2), 434–445. https://doi.org/10.1121/1.380688

Wiese, R. (2000). *The phonology of German*. Oxford University Press on Demand.

Appendix A - Stimuli

	CVt	**CVn**	**CVd**
	bat	ban	bad
/**æ**/	mat	man	mad
	cat	can	pad
	bet	Ben	bed
/**ɛ**/	pet	men	red
	set	ten	Ted

Appendix A1 - Target words across all six conditions.

beer	call	cup	four	full	hear
house	leaf	life	light	love	nose
pill	pin	roll	sun	take	war

Appendix A2 - Filler words in alphabetical order.

Appendix B - Questionnaire

English as spoken by Germans in Amsterdam - Background Questionnaire

Dear participant,

You will be taking part in the "English as spoken by Germans in Amsterdam" research project conducted by Maxim Dauenhauer under supervision of Marijn van 't Veer at the University of Amsterdam, department of linguistics. Before the research project can begin, it is important that you read about the procedures we will be applying. Make sure to read this brochure carefully.

Purpose of the research project

As we know from prior research, the time that we spend living in an English-speaking country has a big impact on how our English sounds. People living in Amsterdam, however, live in a unique environment in which they use English daily, but often interact with people that do not speak English natively. This presents a great opportunity to study how German speaking people here speak English, compared to ones that lived in an English-speaking country only shortly or for a very long time. At this stage of the project, we cannot provide any further information on the factors we will be examining. You will receive further details after the experiment has ended.

Who can take part in this research?

We are inviting adult speakers of the German language to take part in this research. Before the experiment begins, we will be asking you some questions about your language background. You can take part in this research project if you one of the following is true about you:

1. German is your mother tongue and you were not brought up in a bilingual household. You also currently live in Amsterdam and have learned English as a second language in school in Germany. English is your mother tongue and you speak a North American variety of English. You also currently live in Amsterdam.

2. We also need to make sure that you do not, to the best of your knowledge, have any language related problems such as dyslexia or developmental language disorder.´

Instructions and procedure

During the experiment, you will be sitting in front of a microphone and you will see one word at a time on a screen. You will be asked to say the sentence "I will say _____" followed by the word on the screen, for a total of 34 words. After a short break, the same procedure will be performed a second time. The second repetition does not need to be more or less emphasized than the first. Please read the words at a normal pace, i.e., neither too fast nor overly slow, and say the words as you would in any day-to-day conversation. While you are speaking, your audio will be recorded. The total duration of the experiment will be approximately 10 minutes. After the actual experiment, you will be asked to fill out a short questionnaire about your language background (and possibly your experience with the English language).

Voluntary participation

You will be participating in this research project on a voluntary basis. This means you are free to stop taking part at any stage. This will not have any consequences and you will not be obliged to finish the procedures described above. You can always decide to withdraw your consent later on. If you decide to stop or withdraw your consent prior to publication of the research results, all the information gathered up until then will be permanently deleted. However, if information has been anonymized, it cannot be deleted because it is not possible to trace back the information to individual participants.

Discomfort, Risks & Insurance

The risks of participating in this research are no greater than in everyday situations at home. Previous experience in similar research has shown that no or hardly any discomfort is to be expected for participants. For all research at the University of Amsterdam, a standard liability insurance applies.

Confidential treatment of your details

The information gathered over the course of this research will be used for further analysis and publication in scientific journals only. Your personal details will not be used in these publications, and we guarantee that you will remain anonymous under all circumstances, unless you explicitly provide consent to share your personal information. Audio recordings will also never be shown in public. The data gathered during the research will be encrypted and stored separately from the personal details. These per-

sonal details and the encryption key are only accessible to members of the research staff. Anonymous data will be stored for a period of 10 years. The personal data will only be stored as long as is necessary for the research and will be deleted as soon as possible.

Further information

For further information on the research project, please contact Marijn van 't Veer (phone number: +31 20 – 525 38 72; e-mail: b.m.vantveer@uva.nl; Spuistraat 134, 1012VB Amsterdam, room 6.38). If you have any complaints regarding this research project, you can contact the secretary of the Ethics Committee of the Faculty of Humanities of the University of Amsterdam, commissie-ethiek-fgw@uva.nl, phone number: +31 20 – 525 3054; Binnengasthuisstraat 9, 1012 ZA Amsterdam.

Informed consent form

'I hereby declare that I have been clearly informed about the research project "English as spoken by Germans in Amsterdam" at the University of Amsterdam, department of linguistics, conducted by Maxim Dauenhauer under supervision of Marijn van 't Veer as described in the information brochure. My questions have been answered to my satisfaction.

I realise that participation in this research is on an entirely voluntary basis. I retain the right to revoke this consent without having to provide any reasons for my decision. I am aware that I am entitled to discontinue the research at any time, and that I can always withdraw my consent after the research has ended. If I decide to stop or withdraw my consent, all the information gathered up until then will be permanently deleted.

If my research results are used in scientific publications or made public in any other way, they will be fully anonymised. My personal information may not be viewed by third parties without my express permission.

If I need any further information on the research, now or in the future, I can contact Marijn van 't Veer (phone number: +31 20 – 525 38 72; e-mail: b.m.vantveer@uva.nl; Spuistraat 134, 1012VB Amsterdam, room 6.38).

If I have any complaints regarding this research, I can contact the secretary of the Ethics Committee of the Faculty of Humanities of the University of Amsterdam; email: commissie-ethiek-fgw@uva.nl; phone number: +31 20 – 525 3054; Binnengasthuisstraat 9, 1012 ZA Amsterdam.

Q1 I consent to participate in this research, audio recordings being made and my personal details to be stored until August 2023.

(a) Yes

(b) No

Q2 What is your e-mail address?

Q3 How old are you (in years)?

Q4 In which country did you complete the majority of your secondary education (i.e. after elementary school)?

(a) Germany

(b) USA

(c) Other

Q5 What is your gender?

(a) Male

(b) Female

(c) Other

Q6 Where did you spend the majority of the first 13 years of your life? You can mention a region, state, city, etc.

Q7 Please list all the languages you know or have learned:

(a)

(b)

(c)

(d)

(e)

Q8 What is your level for each language? (Choose one from the options below)

	Beginner	Intermediate	Advanced	Native
Language 1	○	○	○	○
Language 2	○	○	○	○
Language 3	○	○	○	○
Language 4	○	○	○	○
Language 5	○	○	○	○

Q9 Please list what percentage of time you have been exposed to each of the languages you listed over the past two years.

(a) Language 1:

(b) Language 2:

(c) Language 3:

(d) Language 4:

(e) Language 5:

Q10 For how many years have you been living in the Netherlands?

Q11 If you have ever migrated to another country, please provide name of country and the time you spent there.

Only for participants who indicated that they went to school in Germany:

Q12 How many years of formal education of learning English do you have (school, university, vocational training, etc.)?

Q13 Please check your highest education level (or the approximate equivalent to a degree obtained in another country):

(a) Hauptschule

(b) Realschule (Mittlere Reife)

(c) Abitur/Fachabitur

(d) Professional Training (Ausbildung)

(e) Bachelor's Degree

(f) Master's Degree

(g) Ph.D./M.D./J.D

(h) Other

Q14 Q11 Age when you...

(a) began acquiring/learning English

(b) became fluent in English

Q15 On a scale from zero to ten, please select your level of proficiency in English:

(a) 0 - none

(b) 1 - very low

(c) 2 - low

(d) 3 - fair

(e) 4 - slightly less than adequate

(f) 5 - adequate

(g) 6 - slightly more than adequate

(h) 7 - good

(i) 8 - very good

(j) 9 - excellent

(k) 10 - perfect

Q16 On a scale from zero to ten, please select how much of the following factors contributed to you learning English:

	1	2	3	4	5	6	7	8	9	10
Interacting with friends	○	○	○	○	○	○	○	○	○	○
Interacting with family	○	○	○	○	○	○	○	○	○	○
Reading	○	○	○	○	○	○	○	○	○	○
Language apps or websites/self-instruction	○	○	○	○	○	○	○	○	○	○
Watching TV	○	○	○	○	○	○	○	○	○	○
Listening to radio/music	○	○	○	○	○	○	○	○	○	○

Q17 Please rate to what extent you are currently exposed to English in the following contexts:

	1	2	3	4	5	6	7	8	9	10
Interacting with friends	○	○	○	○	○	○	○	○	○	○
Interacting with family	○	○	○	○	○	○	○	○	○	○
Reading	○	○	○	○	○	○	○	○	○	○
Language apps or websites/self-instruction	○	○	○	○	○	○	○	○	○	○
Watching TV	○	○	○	○	○	○	○	○	○	○
Listening to radio/music	○	○	○	○	○	○	○	○	○	○

Appendix C - Praat Script

```
files = Create Strings as file list: "list", "Thesis Recordings/*.TextGrid"
writeInfoLine: "Durations and formants:"
numberOfFiles = Get number of strings

form Analyze which files?
    positive Number_of_first_file 1
    integer: "Number_of_last_file", "0 (= last file in the folder)"
endform

if number_of_last_file = 0
    number_of_last_file = numberOfFiles
endif

for fileNumber from number_of_first_file to number_of_last_file
    selectObject: files
    textgridFile$ = Get string: fileNumber
    textgrid = Read from file: "Thesis Recordings/" + textgridFile$
    numberOfIntervals = Get number of intervals: 1
    soundFile$ = textgridFile$ - ".TextGrid" + ".wav"
    participantName$ = textgridFile$ - ".TextGrid"
    sound = Read from file: "Thesis Recordings/" + soundFile$

    selectObject: textgrid
    gender$ = Get label of interval: 1, 2
    if gender$ = "m"
        genderHz = 5000
    elsif gender$ = "f"
        genderHz = 5500
    else
        exitScript("Gender not specified.")
    endif

    selectObject: sound
    formantObject = To Formant (burg): 0, 5, genderHz, 0.025, 50

        i_ae_t = 1
        durations_ae_t# = zero# (6)
        f1_ae_t# = zero# (6)
        f2_ae_t# = zero# (6)
        i_ae_n = 1
```

```
durations_ae_n# = zero# (6)
f1_ae_n# = zero# (6)
f2_ae_n# = zero# (6)
i_ae_d = 1
durations_ae_d# = zero# (6)
f1_ae_d# = zero# (6)
f2_ae_d# = zero# (6)
i_e_t = 1
durations_e_t# = zero# (6)
f1_e_t# = zero# (6)
f2_e_t# = zero# (6)
i_e_n = 1
durations_e_n# = zero# (6)
f1_e_n# = zero# (6)
f2_e_n# = zero# (6)
i_e_d = 1
durations_e_d# = zero# (6)
f1_e_d# = zero# (6)
f2_e_d# = zero# (6)

for interval to numberOfIntervals
    selectObject: textgrid
    text1$ = Get label of interval: 1, interval
    text2$ = Get label of interval: 2, interval
    #appendInfoLine: text1$, text2$

    #ae, t
    if text1$ = "ae" and text2$ = "t"
        startingTime = Get start time of interval: 1, interval
        endTime = Get end time of interval: 1, interval
        duration = endTime - startingTime
        selectObject: formantObject
        f1 = Get quantile: 1, startingTime, endTime, "hertz", 0.50
        f2 = Get quantile: 2, startingTime, endTime, "hertz", 0.50
        durations_ae_t#[i_ae_t] = duration
        f1_ae_t#[i_ae_t] = f1
        f2_ae_t#[i_ae_t] = f2
        i_ae_t += 1
    endif

    #ae, n
    if text1$ = "ae" and text2$ = "n"
        startingTime = Get start time of interval: 1, interval
```

```
    endTime = Get end time of interval: 1, interval
    duration = endTime - startingTime
    selectObject: formantObject
    f1 = Get quantile: 1, startingTime, endTime, "hertz", 0.50
    f2 = Get quantile: 2, startingTime, endTime, "hertz", 0.50
    durations_ae_n#[i_ae_n] = duration
    f1_ae_n#[i_ae_n] = f1
    f2_ae_n#[i_ae_n] = f2
    i_ae_n += 1
endif

#ae, d
if text1$ = "ae" and text2$ = "d"
    startingTime = Get start time of interval: 1, interval
    endTime = Get end time of interval: 1, interval
    duration = endTime - startingTime
    selectObject: formantObject
    f1 = Get quantile: 1, startingTime, endTime, "hertz", 0.50
    f2 = Get quantile: 2, startingTime, endTime, "hertz", 0.50
    durations_ae_d#[i_ae_d] = duration
    f1_ae_d#[i_ae_d] = f1
    f2_ae_d#[i_ae_d] = f2
    i_ae_d += 1
endif

#e, t
if text1$ = "e" and text2$ = "t"
    startingTime = Get start time of interval: 1, interval
    endTime = Get end time of interval: 1, interval
    duration = endTime - startingTime
    selectObject: formantObject
    f1 = Get quantile: 1, startingTime, endTime, "hertz", 0.50
    f2 = Get quantile: 2, startingTime, endTime, "hertz", 0.50
    durations_e_t#[i_e_t] = duration
    f1_e_t#[i_e_t] = f1
    f2_e_t#[i_e_t] = f2
    i_e_t += 1
endif

#e, n
if text1$ = "e" and text2$ = "n"
    startingTime = Get start time of interval: 1, interval
    endTime = Get end time of interval: 1, interval
```

```
        duration = endTime - startingTime
        selectObject: formantObject
        f1 = Get quantile: 1, startingTime, endTime, "hertz", 0.50
        f2 = Get quantile: 2, startingTime, endTime, "hertz", 0.50
        durations_e_n#[i_e_n] = duration
        f1_e_n#[i_e_n] = f1
        f2_e_n#[i_e_n] = f2
        i_e_n += 1
    endif

    #e, d
    if text1$ = "e" and text2$ = "d"
        startingTime = Get start time of interval: 1, interval
        endTime = Get end time of interval: 1, interval
        duration = endTime - startingTime
        selectObject: formantObject
        f1 = Get quantile: 1, startingTime, endTime, "hertz", 0.50
        f2 = Get quantile: 2, startingTime, endTime, "hertz", 0.50
        durations_e_d#[i_e_d] = duration
        f1_e_d#[i_e_d] = f1
        f2_e_d#[i_e_d] = f2
        i_e_d += 1
    endif

endfor

appendInfoLine: "Values for ", participantName$, ":"
appendInfoLine: ""

appendInfoLine: "Condition ae/t "
appendInfoLine: "Durations: ", durations_ae_t#
appendInfoLine: "F1s: ", f1_ae_t#
appendInfoLine: "F2s: ",  f2_ae_t#
appendInfoLine: "Mean duration: ",  mean(durations_ae_t#)
appendInfoLine: "Mean F1: ",  mean(f1_ae_t#)
appendInfoLine: "Mean F2: ", mean(f2_ae_t#)
appendInfoLine: ""

appendInfoLine: "Condition ae/n "
appendInfoLine: "Durations: ", durations_ae_n#
appendInfoLine: "F1s: ", f1_ae_n#
appendInfoLine: "F2s: ",  f2_ae_n#
appendInfoLine: "Mean duration: ",  mean(durations_ae_n#)
```

```
        appendInfoLine: "Mean F1: ",  mean(f1_ae_n#)
        appendInfoLine: "Mean F2: ", mean(f2_ae_n#)
        appendInfoLine: ""

        appendInfoLine: "Condition ae/d "
        appendInfoLine: "Durations: ", durations_ae_d#
        appendInfoLine: "F1s: ", f1_ae_d#
        appendInfoLine: "F2s: ",  f2_ae_d#
        appendInfoLine: "Mean duration: ",  mean(durations_ae_d#)
        appendInfoLine: "Mean F1: ",  mean(f1_ae_d#)
        appendInfoLine: "Mean F2: ", mean(f2_ae_d#)
        appendInfoLine: ""

        appendInfoLine: "Condition e/t "
        appendInfoLine: "Durations: ", durations_e_t#
        appendInfoLine: "F1s: ", f1_e_t#
        appendInfoLine: "F2s: ",  f2_e_t#
        appendInfoLine: "Mean duration: ",  mean(durations_e_t#)
        appendInfoLine: "Mean F1: ",  mean(f1_e_t#)
        appendInfoLine: "Mean F2: ", mean(f2_e_t#)
        appendInfoLine: ""

        appendInfoLine: "Condition e/n "
        appendInfoLine: "Durations: ", durations_e_n#
        appendInfoLine: "F1s: ", f1_e_n#
        appendInfoLine: "F2s: ",  f2_e_n#
        appendInfoLine: "Mean duration: ",  mean(durations_e_n#)
        appendInfoLine: "Mean F1: ",  mean(f1_e_n#)
        appendInfoLine: "Mean F2: ", mean(f2_e_n#)
        appendInfoLine: ""

        appendInfoLine: "Condition e/d "
        appendInfoLine: "Durations: ", durations_e_d#
        appendInfoLine: "F1s: ", f1_e_d#
        appendInfoLine: "F2s: ",  f2_e_d#
        appendInfoLine: "Mean duration: ",  mean(durations_e_d#)
        appendInfoLine: "Mean F1: ",  mean(f1_e_d#)
        appendInfoLine: "Mean F2: ", mean(f2_e_d#)
        appendInfoLine: ""
        appendInfoLine: ""

    removeObject: sound, textgrid, formantObject
endfor
```

```
removeObject: "Strings list"
```