L1 Vowel Change as a Result of L2 Immersion

BACHELOR'S THESIS

Simon Boerstra

11699175

Department of Linguistics

University of Amsterdam

22 June 2020

Supervised by Suki Yiu, Ph.D.

Abstract

Speech plasticity and the influence of one's first language (L1) on a second language (L2) has been well attested. According to the Speech Learning Model (SLM), when speech sounds in an individual's phonological representation occupy the same space, convergence of these speech sounds occurs. This study focused on the systemic change in vowel space of Carice van Houten (CvH; L1 Dutch, L2 English), comparing spontaneous L1 speech in the periods 2006-2010 to 2013-2019. CvH spent considerable time in the US for work between 2012 and 2019. US English vowel formants F1 and F2 are on average higher than the Dutch counterparts. Therefore, we hypothesized that CvH's Dutch F1 and F2 would increase due to the assimilation predicted by the SLM. We also considered age-related F1 and F2 decrease. Linear mixed effects models with PERIOD and AGE as fixed effects and FOLLOWING SEGMENT and PREVIOUS SEGMENT as random effects were employed for analysis. Contrary to expectation, F1 and F2 exhibited a significant period-related decrease (F1: -35.35 Hz, 95% CI [-51.25 Hz, -19.45 Hz], $p < .001$; F2: -74.99, 95% CI [-115.74 Hz, -34.25 Hz], $p < .001$). Natural age-related decrease was for F1 -2.46 Hz/annum, 95% CI [-4.87 Hz/annum, -.05 Hz/annum], $p = .046$; and for F2 7.81 Hz/annum, 95% CI [1.63 Hz/annum, 13.99 Hz/annum], $p = .013$. Effects of AGE and PERIOD on individual vowels are also reported. These results indicate a post adolescent change in native speech, however they do not conform to predictions made by the widely used SLM.

Acknowledgement

First of all, I would like to thank my supervisor, Suki Yiu, for her continued support during this thesis, for our long conversations about what is good methodology and what is not, and for the fact that you spent way more than the allotted 15 hours on this thesis. I would also like to alphabetically thank my classmates and friends, Adrielli, Anastasija, Keeley, Tegan and Zack. It has been a joy to study with you, and because of our common projects and discussions, I was able to write this manuscript. Third, I want to give thanks to my parents and my brothers for their love and their patience with me finding my field of interest. Last, but (certainly) not least: I would like to thank my partner in crime, Mona. These past three years have been challenging in so many different ways and I will always be thankful for your inspiration, help, support, numerous sleepless nights spent working together, tea, walks, warmth and bearing with me, always. I cannot wait to embark on our next adventure together.

To Mona

Table of contents

1. INTRODUCTION 1

1.1. Monolingual studies in speech plasticity 1
1.2. Bilingual studies in speech plasticity 1
1.3. Focus of this paper - Vowel inventories of Dutch and English 3
1.4. Objectives of the present study 4

2. METHOD 5

2.1. Carice van Houten 5
2.2. Recordings 5
2.3. Acoustic analysis 6
2.4. Data analysis 7

3. RESULTS 7

3.1. Overall formant changes over time 7
3.2. Individual vowel F1 analyses 9
3.3. Individual vowel F2 analyses 10

4. DISCUSSION 11

5. CONCLUSION 13

6. REFERENCES 14

7. APPENDIX 16

1. Introduction

1.1. Monolingual studies in speech plasticity

Speech plasticity, or the trajectory along which speech develops and changes through childhood and adulthood, is an area of research that has not been extensively studied. Most of our current understanding is the result of studies investigating monolingual speakers. For example, Harrington (2006) presented a longitudinal analysis of vowels produced by Queen Elizabeth II during her annual Christmas broadcasts over a span of 50 years. Harrington found that sound changes that had occurred in Standard Southern British English (SSBEE), such happY-tensing, are changes to which even the Queen is not impervious. HappY-tensing is the phenomenon in which the final vowel in words like 'happy' have shifted. During the first half of the 20th century, this vowel, [ɪː] was phonetically closer to [ɪ] than to [iː] (i.e. 'tenser'), however the phonetic distance between the two has reduced in the 6 decades prior to the investigation. To test the hypothesis that an individual's accent is guided by phonetic changes that occur within society, (Harrington 2006) compared vowel qualities of the Queen in the 1950s to SSBE speakers' vowels in the 1980s, as well as comparing the former to the Queen's vowel production around the turn of the century. The long term changes that the author indeed did find include lowering of the first, second and fourth formants (F1, F2 and F4) and a raising of F3 in the schwa vowel /ə/. The change detected in F1 was greatest, with a decrease of 123 Hz between the 1950s and the 1990s, whereas the change in F2 was small (72 Hz in the same time period). Similarly, it was found that the [ɪː] produced in the 1990s was tenser than in the 1950s. Raising of vowels was thus found extensively, however fronting was found marginally. These findings corroborate the idea that vowel changes in an individual are guided by changes in one's linguistic environment.

1.2. Bilingual studies in speech plasticity

In a similar study, (Reubold & Harrington 2011)[1] investigated language production by Alistair Cooke, a British-American radio commentator, using broadcasts spanning 60 years. The aim of the investigation was to dissociate age-related changes from phonetic changes. Of major interest in this study was that the informant's accent changed over time. Alistair Cooke went from being perceived as a speaker of Received Pronunciation (RP) to being perceived as a speaker of US English, and, unexpectedly, an accent reversion: in his later years, Alistair Cooke's speech phonologically reverted back to RP, even though he had not left the US. Accent here was measured as a function of F1 and F2. Additionally, it was found that F1 changes in low vowels go in the opposite direction from those of the high vowels. It is suggested that the high vowels follow the fundamental frequency f0 more than low vowels, because high vowels have an F1 that is perceptually closer to f0. With increasing age, f0 and F1 of the low vowels decrease, possibly be due to aging factors such muscular atrophy leading to an

[1] It must be noted that this study could be classified under section *1.1 – Monolingual studies* as well, as there is no reason not to see this as a study in accent change. The question of what consists of a language is, however, out of the scope of this paper.

increasingly smaller mouth opening with increasing age, or the lowering of the larynx.

Studies on language production during or after second language acquisition often focus on the influence of L1 on L2. Whilst this is much attested, a smaller, yet growing body of evidence is showing an opposite effect: an L2 effect on an L1. In 1987, James Flege measured the voice onset time (VOT) and of French and English words spoken by highly proficient L1 French speakers of English, as well as highly proficient L1 English speakers of French. The result of this cross-sectional study indicated that proficient L1 English speakers of French produce VOTs that are shorter (more French-like) when speaking English, compared to monolingual English speakers. This investigation laid the groundwork for Flege's (1995) *Speech Learning Model*. This model states that that a bilingual's speech sounds are represented in one common phonological space. If the unfamiliar perceived speech sounds are sufficiently different from the already existing phonetic categories, new L2 phonetic categories can be formed. This formation, however, may be blocked by the mechanism "equivalence classification [which means that] a single phonetic category will be used to process perceptually linked L1 and L2 sounds [which eventually will] resemble one another in production" (Flege 1995:239). This process takes place bidirectionally between L1 and L2, and has as an explicit prediction that when a new category is not formed, the L2 sound will become part of the L1 phonetic category. As a result of this, these vowel qualities will converge. This effect has been shown cross-sectionally for French-English (Flege 1987; Fowler et al. 2008), Portuguese-English (Major 1992) and Spanish-English (Flege & Eefting 1987) bilinguals, among others.

In their seminal case study, Sancier & Fowler (1997) investigated language plasticity of a Brazilian-Portuguese – English 27 year old late bilingual woman. The voice onset times (VOTs) of plosives were measured at three time moments; after a 4.5-month stay in the US before leaving for Brazil, after her stay in Brazil for 2.5 months upon her return, and finally after a 4 month stay in the US before leaving for Brazil. The study found that, in line with the prediction made by the SLM, the VOT values of both English and Brazillian Portuguese drifted: they were both shorter after a stay in Brazil, than after a stay in the US. In a follow-up study, these results were replicated with Spanish-English bilinguals (Tobin, Nam & Fowler 2017). This study, however, yielded a large amount of interspeaker variability.

Whereas Sancier & Fowler (1997) discussed changes in L1 production as a result of recency (the temporal closeness to being immersed in a different language environment), a similar study was performed investigating speech plasticity as a result of recency as well as novelty (the relative unfamiliarity with a new vowel system, which by default must also be recent). In this study, 19 adult US English L1 novice learners of Korean were investigated during a 6-week intensive course of Korean (Chang 2011; Chang 2012; Chang 2013). Vowel formants as well as VOT values were measured weekly, starting at the second week. English VOTs were measured to be significantly longer during the last recording session when compared to the first. Moreover, a systemic and uniform phonetic vowel drift was reported as well. To measure this vowel drift, Chang measured change in formants 1 and 2 (F1 and F2) of non-spontaneous speech, i.e. the participants in his study were recorded reading sentences which embedded words that contained the target vowels. Results indicated a systemic phonetic drift, i.e. a global shift of the entire vowel space. The female participants in his study, especially, exhibited such a large decrease of F1 that the mean values of English F1 went from being statistically significantly different from the Korean baseline in week 1 to not significantly different in week 5. Regarding F2, no change was found, with English F2 means staying approximately 160 Hz higher than their Korean counterparts.

In a recent longitudinal study, De Leeuw (2019) investigated 40 years of existing recordings, obtained from online media resources, of the L1 production of a German informant Stefanie Graf) after the start of her exposure to L2-English at the age of 17. The study covered a total of 1250 segments consisting of voiced lateral approximant /l/ and front vowels /a/ and /i/. The investigator hypothesized that, in line with the SLM, one phonetic category for /a/ and /i/ would host both phonemes in German and English, and the F1 and F2 of the German vowels would converge towards the English vowels' values. The results of this investigation show significant lowering of the F2 and F1 in /l/, F2 increase in /i/, and F1 increase in /a/, confirming the hypothesis in these cases, but not in the F1 of /i/ or F2 of /a/. Segments /r/ and /u/ were also insignificant, yet remained unreported, hindering comparison. These findings contrast with Chang's (2012): whereas he did find a systemic change in the learner's L1, De Leeuw did not and whereas De Leeuw did find evidence of the F2 change in her investigation, Chang did not. De Leeuw interprets her results as potentially a result of age-related and socio-indexical changes.

In order to expand on these findings, the present study aims to add to the knowledge by investigating the language plasticity of a native Dutch speaking actress Carice van Houten (CvH), who spent considerable time in the US Similarly to De Leeuw (2019), this investigation investigates existing recordings of spontaneous speech covering over a decade, however as the investigation by de Leeuw (2019) reported two vowels and one lateral only, this investigation covers the vowel system in its entirety, similar to Chang (2012).

1.3. Focus of this paper - Vowel inventories of Dutch and English

This paper focuses on the effect that an L2 has on an L1. In order to hypothesize direction of change according to the SLM, the vowel systems of both languages must be compared. If the vowel systems have overlapping phonemes, covergence of the L1 is expected. The Standard Dutch vowel system contains 12 monophthongs /ɑ, a, ɛ, e, ɪ, i, o, ɔ, u, ø, y, ʏ/ and schwa /ə/. Diphthongs in Dutch were not taken into consideration in this study. The US-English vowel system contains 13 monophthongs: /ɑ, a, ɛ, e, ɪ, i, o, ɔ, u, ʊ, ʌ, æ, ɚ/ and schwa /ə/. Table 1 shows the reported average values of F1 and F2 in Dutch (Adank, Van Hout & Smits 2004) and English (Hillenbrand et al. 1995; Yang

1996; Hagiwara 1997). Of note are the discrepancies shown in the high vowels /i/ and /u/, as well as back vowels /o/ and /ɔ/ and front vowel /ɛ/, which all have appreciably higher values in English when compared to Dutch. The only values that are higher in Dutch than in any regional variety of US English are the F1 and F2 of /a/. The only similar formant value is F1 of /e/ in Dutch and Southern California, which differ only 2 Hz (442 Hz in Dutch cf. 440 Hz in English). Generally speaking, US English high vowels are lower, i.e. have higher formant values than Dutch vowels and US English front vowels are more frontal, i.e. have higher F2 values, when compared to Dutch.

Table 1 - Table 1 - Average female F1 and F2 vowels in Standard Dutch and US English, adapted from Adank, Van Hout & Smits 2004 (Northern Dutch, Nor Dutch); Hillenbrand et al. 1995 (MidWest); Yang 1996 (South & Southwest, S&SW; Hagiwara (1997) *Southern California, SoCal. All vowels except the bottom are shared by both vowel systems. Values marked with n/a have no reported value.*

	Dutch		English					
	Nor Dutch		MidWest		S&SW		SoCal	
Vowel	F1	F2	F1	F2	F1	F2	F1	F2
/ɑ/	758	1280	936	1551	n/a	n/a	997	1390
/a/	912	1572	n/a	n/a	782	1287	n/a	n/a
/ɛ/	535	1990	731	2058	631	2244	808	2163
/e/	442	2343	536	2530	521	2536	440	2655
/ɪ/	399	2276	483	2365	466	2373	467	2400
/i/	294	2524	437	2761	390	2826	362	2897
/o/	445	964	555	1035	528	1206	516	1391
/ɔ/	419	918	781	1136	777	1140	n/a	n/a
/u/	286	938	459	1105	417	1511	395	1700
/y/	305	1918	n/a	n/a	n/a	n/a	n/a	n/a
/ʏ/	417	1830	n/a	n/a	n/a	n/a	n/a	n/a
/ø/	445	1713	n/a	n/a	n/a	n/a	n/a	n/a
/ʊ/	n/a	n/a	519	1225	491	1486	486	1665
/ʌ/	n/a	n/a	753	1426	701	1641	847	1753
/æ/	n/a	n/a	669	2349	825	2059	1017	1810
/ɚ/	n/a	n/a	523	1588	523	1550	n/a	n/a

1.4. Objectives of the present study

The objective of this study is to answer the question *How does the vowel space measured by F1 and F2 of late Dutch-English bilingual CvH change over time after being immersed in an US-English language environment?* We hypothesize that, in line with findings by Chang (2011, 2012, 2013) as well as findings by de Leeuw (2019) and conform Flege's Speech Learning Model, L1 vowel formants F1 and F2 will adjust over time as they are influenced by exposure to English vowels; specifically, we expect that Dutch vowels will have higher formant values for the period after CvH's prolonged exposure to US English. This will mean lowering of the high vowels and fronting of the front vowels This study aims to add to the existing knowledge on long-term effects of L2 acquisition on L1 phonology, as no longitudinal research has been done on the effects of L2 English immersion on L1 Dutch. Taking previous longitudinal research in language change (e.g. Reubold & Harrington 2011; de Leeuw 2019; Harrington 2006) into consideration, a potential effect of age on formant change must be taken into consideration

The rest of this thesis is outlined as follows: Section 2 presents the design of this study and the methods used to analyze the changes found. Section 3 compares spontaneous speech produced by CvH, before and after long term immersion in a second language community in the United States. Section 4 discusses these findings and finally in Section 5 the main conclusions are summarized.

2. Method

2.1. Carice van Houten

Carice van Houten (CvH) was born on September 5th 1976 in Leiderdorp, the Netherlands and raised in a monolingual household. CvH had her professional acting education at the *Kleinkunstacademie* in Amsterdam. In 1999, at age 22, CvH played her first leading role in a TV film. After a number of movies, CvH played the leading role in Blackbook (2006) which catapulted her career internationally, after which she began starring in English speaking movies, such as *Valkyrie* (2008), *Repo Men* (2010). This culminated in her being cast for the TV series *Game of Thrones* from 2012-2019.

As a child in the Dutch education system, CvH has received formal education in English starting from grade 5 (age 11) of elementary school until year 6 of high school (age 18). In 2015, at 38 years of age, CvH met her current partner, Guy Pearce, who is a monolingual Australian actor.

2.2. Recordings

Table 2 - Number of recordings, minutes and segments per year of data

Year	Number of recordings	Total minutes of speech (mm:ss)	No. of segments
2006	2	3:01	665
2007	2	4:37	1068
2009	3	5:47	1353
2010	1	1:45	428
2013	1	17:48	4051
2016	1	0:18	71
2019	2	1:46	452

Recordings of CvH were obtained through various channels (e.g. YouTube, *Uitzending Gemist* (Dutch Television) and *Linda.TV*). These segments were recorded in various contexts, for example an acceptance speech after winning an award, an interview after the conclusion of Game of Thrones or a discussion at a Dutch late night television show. 16 recordings were selected for inspection, after which 12 recording were used in this investigation. The recordings that were not used for analysis were rejected for background noise, low quality recording or because the recording was interspersed with interjetion by another speaker, yielding too few segments available for analysis. In all, 12 recordings were analysed. The earliest recordings used were captured in 2006 (CvH is 29 years old here), the latest in 2019 (CvH is 42 years old here). Table 2 shows the total length of recordings per year, as well as the number of recordings and analyzed vowel segments by year.

The sampling frequencies of the original sound files were not obtained. The reason for this is that the downloading process instantly converted audio into a .wav file with sampling frequency 44.1 kHz. Videos uploaded with the resolution 360p to YouTube, prior to 2011 had audio converted into 96 kbps AAC quality. This has gone up since then to maximally 256 kbps AAC. The data before 2011, therefore, is of lower quality than that of after 2013. All recordings used for this study consisted of spontaneous speech (i.e. not intended for linguistic analysis). Manual preparation was done in PRAAT acoustic analysis software (Boersma 2001). From the recordings various parts of speech were removed, such as unintelligible speech, false starts, overlapping speech or speech with applause in the background. After this initial cleaning of the data, all audio files were uploaded to the Bavarian Archive for Speech Signals (BAS) (Schiel 1999; Schiel, Draxler & Harrington 2011; Kisler, Reichel & Schiel 2017) for automatic speech recognition (ASR). The ASR output was raw text, which was then used to create a rough annotation of the sound files in PRAAT. The resulting pairs of files, consisting of annotation files and sound files, were once again uploaded to the BAS, using their 'pipeline' function, in which a series of services were performed. These were in order: grapheme-to-phoneme conversion, phonemic segmentation and syllabification. The output of the pipeline included a broad phonemic transcription that was used for analysis. Where needed, the automatically transcribed segments were corrected by the author.

2.3. Acoustic analysis

The recordings' vowel data was automatically analysed in PRAAT for the first and second formant frequency (F1 and F2). As mentioned, for the purpose of this study, only monophthongs were analyzed. Segments /øː/ and /œː/ were omitted from analysis because these segments were not present in available years. The distribution of all segments can be found in Table 2. Analyses were done at the center of the vowel (from 25% of the length of the vowel to 75% of the length) to reduce influences from surrounding segments.

Table 3 - Segmental distribution

Vowel	No. of segments	Vowel	No. of segments
/ɪ/	709	/ə/	2280
/i/	287	/ɔ/	448
/ʏ/	124	/oː/	487
/y/	85	/ɑ/	1172
/ɛ/	960	/u/	181
/eː/	666	/œː/	51
/aː/	689	/øː/	34

2.4. *Data analysis*

A PRAAT script was written to automatically extract the relevant information (i.e. formant values, following and preceding segments and parent-word) from all audio-transcript pairs that were available, and export these in a single CSV file. This file was subsequently exported to R software (R Core Team 2019). For linear mixed effect models, the lme4 package (Bates et al. 2015), as well as the lmerTest package (Kuznetsova, Brockhoff & Christensen 2017) were used, with as dependent variable either F1 or F2, and with as fixed effects (1) PERIOD (before or after 2012; binary) and (2) AGE of CvH (as a function of YEAR; continuous). Random effects were preceding and following segments. The fixed effect PERIOD was chosen as the general aim of this research is to establish whether there is an effect of L2 immersion on L1, therefore the YEAR data was divided into pre- and post-2012. AGE will nevertheless be assessed for use as a second fixed effect. This choice is motivated by previous findings (e.g. Reubold & Harrington 2011) which state that there is a potential effect of age on F1 and F2.

3. Results

3.1. *Overall formant changes over time*

In order to build the most parsimonious model, three models were compared: in the null model, only PERIOD was used as a fixed effect, with preceding and following segments as random effects, whereas in the comparing models AGE was added as an additional fixed effect, or as an additional random efefct. AIC model selection was used, in which the best-fit model is one with the addidtion of AGE as a fixed factor.

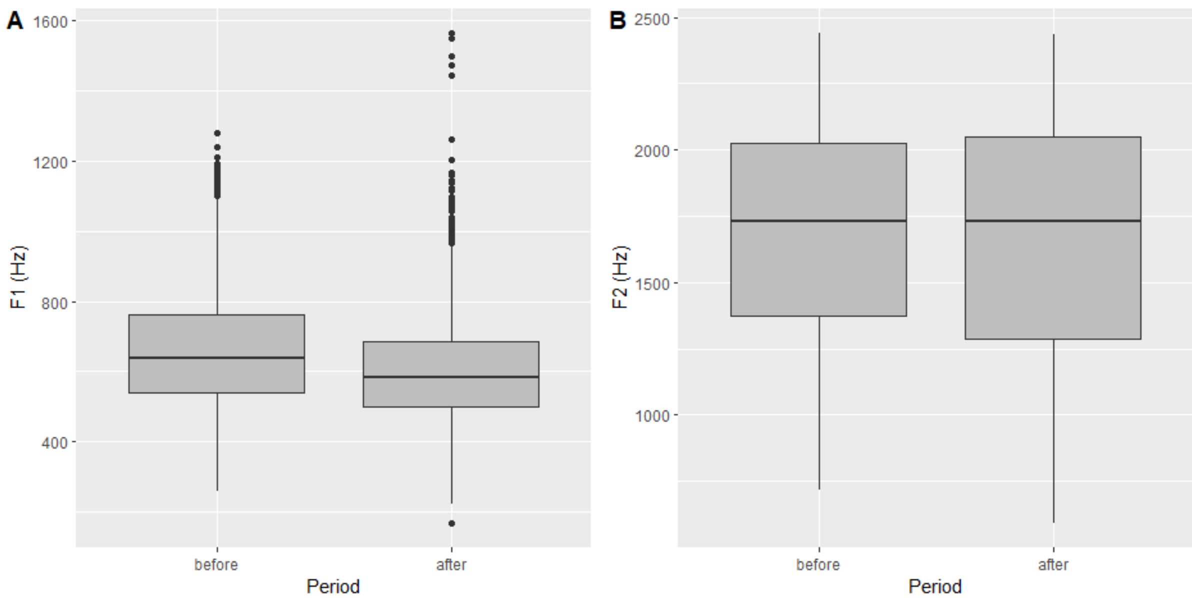

*Figure 1 - Mean formant values by period. N = 5798 vowels (2519 before; 3289 after). See **A** for F1 and **B***

First, linear mixed effects regression models were built with F1 as the dependent variable in order to see if, in line with the hypothesis, an F1 increase would be detected. All vowels except schwa /ə/ were incorporated into the model (5798 vowels of which 2519 tokens belonged to the *before* bin, and

3289 tokens belonged to the *after* bin). The fixed factors AGE and PERIOD both had a significant effect on F1 (AGE: t(5761.6) = -2.00, p = .046; PERIOD t(5765.3) = -4.36, p < .001). The estimated effect of AGE on F1 was -2.46 Hz/annum, 95% CI [-4.87 Hz/annum, -.05 Hz/annum], and the estimated effect of PERIOD on F1 was -35.35 Hz, 95% CI [-51.25 Hz, -19.45 Hz]. Second, linear mixed effects regression models were built with F2 as the dependent variable, with the same motivations as the first model discussed above. Again, the fixed factors AGE and PERIOD both had a significant effect on F2 (AGE: t(5766.1) = 2.48 Hz, p = .013; PERIOD t(5769.0) = -3.61, p < .001). The estimated effect of AGE on F2 was 7.81 H/annum z, 95% CI [1.63 Hz/annum, 13.99 Hz/annum], and the estimated effect of PERIOD on F2 was -74.99 Hz, 95% CI [-115.74 Hz, -34.25 Hz]. Fig. 1 shows the overall trend of F1 and F2 value decreasing significantly when comparing F1 and F2 before and after 2012. In Fig. 2A, an F1 decline is visible between 2010 and 2013. This period corresponds to the period during which CvH started working on the set of Game of Thrones. All individual vowels' findings are summarized in Table 4 in the appendix, and relevant results are discussed in sections 3.2 and 3.3. CvH's phonetic drift is visualized in Figure 3, in which a noticable shift of vowel space is visible.

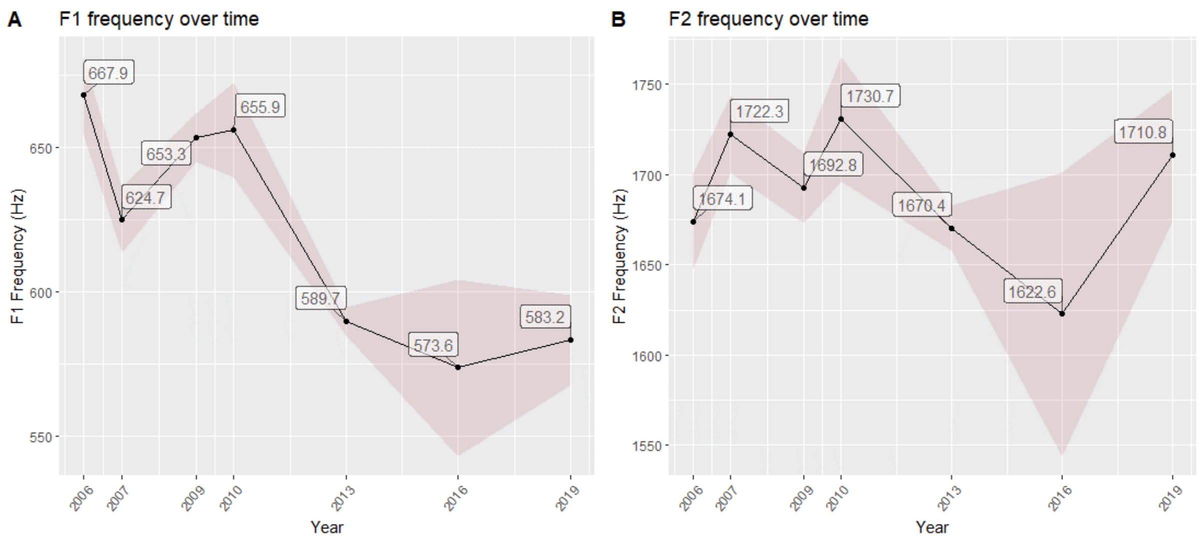

Figure 2 - F1 and F2 of all vowels over time, red bands indicate 95% CI.

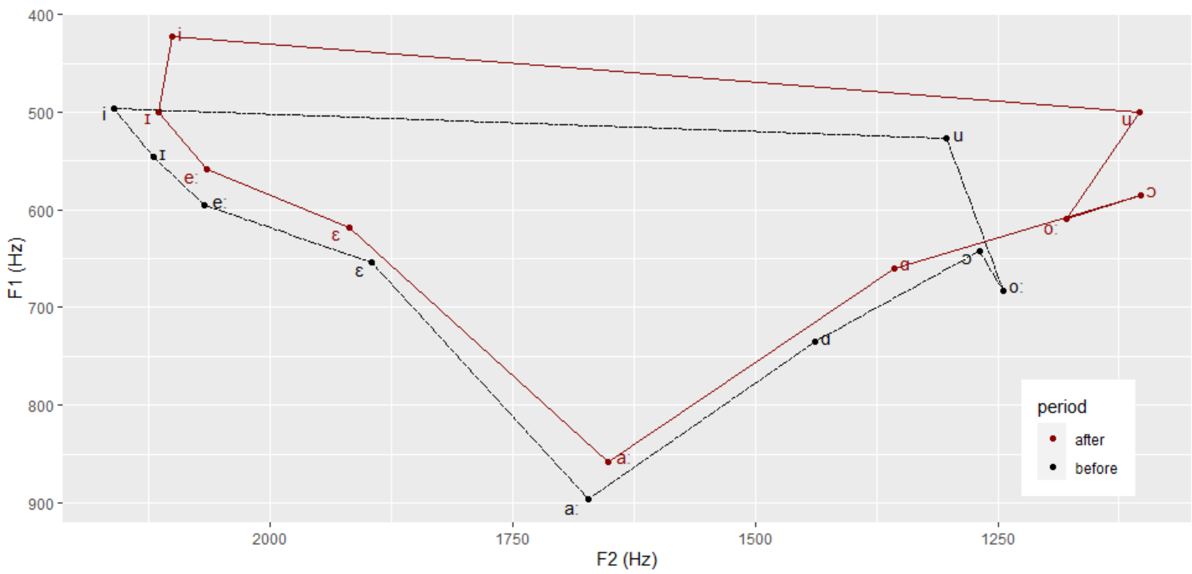

Figure 3 - Phonetic drift of CvH's vowel space in Dutch. The uninterrupted line depicts the period after 2012, the interrupted line depicts the period before 2012.

3.2. Individual vowel F1 analyses

The values of F1 can be seen on the vertical axis of Fig. 3. By convention, the axis' values are reversed, so that a decrease in F1 is translated by an upward movement. For the analyses of each individual vowel's F1 (with the exception of /ə/), models were fitted using lmerTest in R with PERIOD and AGE as fixed effects and following and preceding segments as random effects, and using Bonferroni corrections to set α = .0045 (.05 / 11 vowels), yielded the following results. For all individual vowels' F1, AGE was not a significant fixed effect in any vowel, whereas PERIOD only was a significant factor in some of the vowels. This was the case for five vowels: /ɑ/, /ɪ/, /eː/, /oː/ and /y/. The details are as follows: /ɑ/ (t(1151.8) = -3.94, *p* < .001, with an estimate of -61.4 Hz, 95% CI [-92.00 Hz, -30.87 Hz], /ɪ/ (t(683.38) = -3.64, *p* < .001, with an estimate of -48.5 Hz, 95% CI [-74.55 Hz, -22.39 Hz]; /eː/ (t(650.98) = -3.26, *p* = .0012, with an estimate of -45.65 Hz 95% CI [-73.14 Hz, -18.23 Hz]; /oː/ t(472.85) = -3.60, *p* < .001, with an estimate of -67.95 Hz, 95% CI [-105.15 Hz, -31.10 Hz] and finally /y/ t(82) = -3.06, *p* = .0030, with an estimate of -108.06 Hz, 95% CI [-176.79 Hz, 39.32 Hz]. All other vowels did not show individual significant differences over PERIOD. Fig. 4 shows the boxplots for all vowels in this F1 analysis. Boxplots that are marked in green report on significant *p*-values and estimates for PERIOD as a fixed effect.

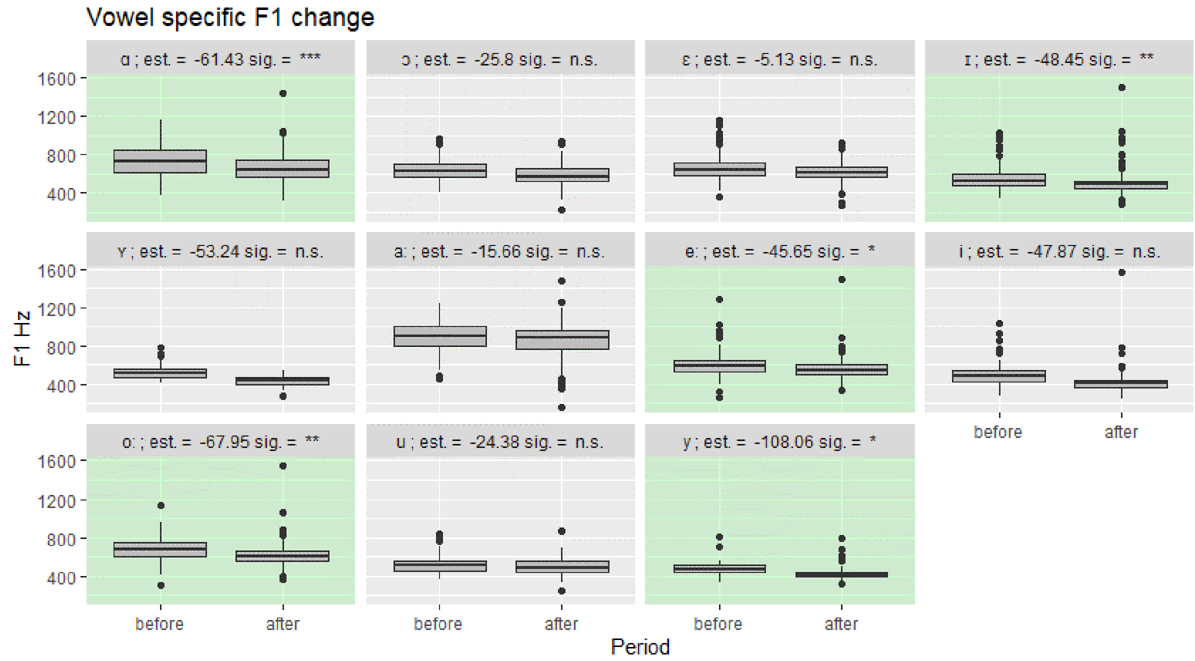

Figure 4 - F1 changes per vowel. p-*values are of* PERIOD *as fixed effect. In green are marked those vowels that are significant for* $\alpha = .0045 = \frac{.05}{11}$. *Significance codes:* $* < \frac{.05}{11}$, $** < \frac{.01}{11}$, $*** < \frac{.001}{11}$, . *See appendix for a more detailed list of results. Estimates are in Hz.*

3.3. Individual vowel F2 analyses

The values of F2 can be seen on the horizontal axis of Fig. 3. By convention, the axis' values are reversed, so that a decrease in F2 is translated by a leftward movement. An analysis similar to that of F1 was performed for F2, using the same fixed and random effects, as well as the same α, yielded the following results. Similar to F1, AGE showed no significant interaction with any vowel, whereas PERIOD was significant in the following vowels: /a/, /ɔ/, /ɛ/, /i/ and /oː/. The results of these vowels are as follows: /ɑ/, (t(1149.80) = -4.73, *p* < .001, with an estimate of -124.37 Hz 95% CI [-176.00 Hz, -72.94 Hz]; /ɔ/, (t(439.36) = -3.36, *p* < .001, with an estimate of -181.21 Hz 95% CI [-288.45 Hz, -75.51 Hz]; /ɛ/, t(934.72) = 3.35, *p* < .001, with an estimate of 70.42 Hz, 95% CI [29.28 Hz, 111.79 Hz]; /i/, t(273.65) = -3.12, *p* = .0020, with an estimate of -137.67 Hz, 95% CI [-224.12 Hz, -51.18 Hz]; and finally /oː/, t(443.91) -4.61, *p* < .001, with an estimate of -160.48 Hz, 95% CI [-228.52 Hz, -92.12 Hz] All other vowels did not show individual significant differences over PERIOD. Fig. 5 shows the boxplots for all vowels in this F2 analysis. Boxplots that are presented in green report on significant *p*-values and estimates for PERIOD as a fixed effect. Fig. 3 shows the F1xF2 shift that is described in sections 3.2 and 3.3.

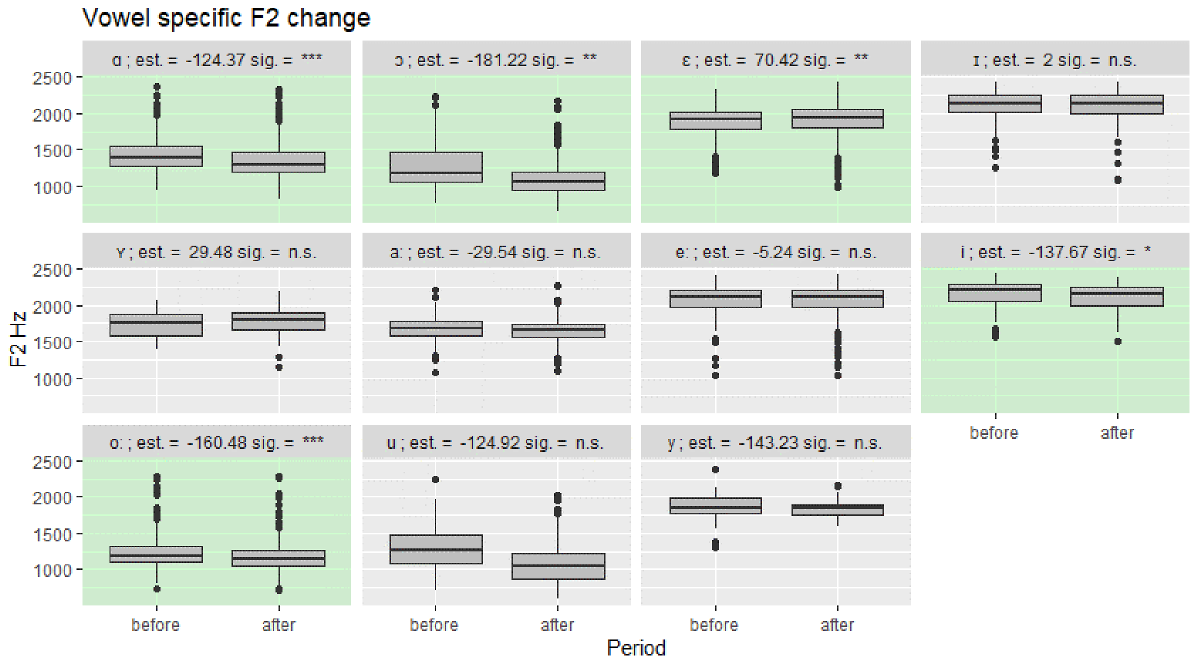

Figure 5 - F1 changes per vowel. p*-values are of* PERIOD *as fixed effect. In green are marked those vowels that are significant for* $\alpha = .0045$*. Significance codes:* $* < \frac{.05}{11}$, $** < \frac{.01}{11}$, $*** < \frac{.001}{11}$*. See appendix for a more detailed list of results.*

4. Discussion

The aim of this research was to establish whether a detectable change in CvH's L1 Dutch occurred as a result of increased exposure to a different language community, L2 US English. It was hypothesized that as a result of this linguistic exposure, an L1 change would be observed in the general direction of the L2 phonemes that overlap, e.g. if the English /y/ is lower and further back, that is the direction that the Dutch phonemic counterpart would shift towards. It was also hypothesized that as a natural result of aging, a decrease in F1 and F2 values would be observed. As seen in Fig. 5, the direction of change was in fact the opposite direction: away from the English counterparts. Theories such as Flege's Speech Learning Model (1995), a widely used framework in L2 speech analysis, as well as previous research (e.g. de Leeuw 2019 and Chang 2011) indicate an interaction between the L2 of interest and the speaker's L1. This raises the question of why do we then find results that not only fail to find this result, but also directly oppose it. A number of considerations must be made.

First, the biological effects of aging need to be taken into consideration. Harrington's Christmas broadcast study (2006) reports a significant effect of aging on speech sounds. It must be noted, however, that the corpora used in this investigation spanned multiple decades, comparing the Queen when she was 26-34 years old to when she was 69-76 years old, whereas in this investigation a mere 13 years of data was analysed, during which CvH was 30-43 years old. Furthermore, when adding AGE as a fixed effect to the linear mixed effects model, a very small effect was noted, of -2.46 Hz/annum for F1 and 7.81 Hz/annum for F2. More longitudinal research is needed in order to determine the effects of aging on speech sounds, and specifically on when in an individual's life these changes have the most effect, in order to be able to improve statistical models analyzing change of speech.

Second, the basis of the hypothesis must be considered. In this paper it was hypothesized, based on the SLM, that any difference in vowel space detected that was not directly attributalble to aging effects would be the result of her immersion in an English speaking environment. The hypothesized direction of change, a higher F1 and a higher F2, was based on (1) the interaction between the prediction in the SLM that when speech sounds occupy the same phonological space, they will converge; and (2) there is a difference in average formant values of Dutch on the one hand and English on the other. Regarding point (1), the possibility cannot be excluded that CvH's Dutch speech sounds in fact did not ocupy the same phonological space as did her English speech sounds. According to the SLM, in the case that speech sounds are sufficiently different, new categories may be formed. As there are now two categories of speech sounds that are in relatively close proximity, a diverging phonological shift may be required to provide the speaker with the required contrast. Regarding point (2), it is important to note the methodological differences of data acquisition in the studies that provided the average formant values for Dutch and English. For Dutch, Adank, Van Hout & Smits (2004) report formant values based on 10 female speakers, consisting of 5 speakers aged 20 to 44 years old, and 5 speakers aged 45 to 60 years old. For English, both Yang (1996) and Hagiwara (1997) investigated speech produced by college-aged women, with 10 participants for the former and 9 participants for the latter. Hillenbrand et al. (1995) did not report the age of their 48 female participants. The average formant for Dutch women of around CvH's age may, assuming a systemic formant decrease as a result of biological aging, be higher than the reported values in Table 1. This difference, however, may not be enough to bridge the gap between the findings of this study and the discrepancy it has with prior research: as the average Dutch vowel formant values are averaged between "younger women" and "older women" (Adank, Van Hout & Smits 2004:1730), in order for the average formant values for younger women to be higher than their US English counterparts, a large age-related decrease is required. For example, for /ɑ/, an average Dutch F1 of 758 Hz is given, whereas for Midwest US English an F1 of 936 Hz is given. If 'younger' Dutch women have an F1 that is higher than 936 Hz, it would require for 'older' Dutch women to have an average F1 value of 580 Hz. This would in turn result in an F1 decrease of 356 Hz, a 38% decrease, which is unlikely to say the least. Future research may be needed to clarify the difference in acoustics between age groups of Dutch.

Third, while our findings do not corroborate the findings reported by de Leeuw (2019) and Harrington (2006), it must be considered that the present paper provided a much more comprehensive analysis (i.e. the analysis of the phonological space consisting of 11 vowels), whereas de Leeuw (2019) reported vowels /a/ and /i/, and Harrington's 2006 study mainly covered /ə/, /i/ and /ɪ/. The limited scope of these previous investigations may lead to limitations in generalizability.

Finally, due to the nature of this research into the spontaneous speech over a number of years, it must be taken into account that the variance in sound quality in general and the relatively low sound quality of the earlier years in particular, may have led to slightly different data than otherwise would have been found, if all speech segments were of high quality. There is a slight possibility that audiofiles generated in the period 2006-2010 will have a slightly lower quality and that this lower sound quality translates into higher formant values. Using higher quality audio for research might remedy this potential issue.

5. Conclusion

Previous empirical research have corroborated the Speech Learning Model, showing convergence of L1 and L2 vowels that share a phonological space. In the present paper, evidence is shown for the opposite effect: whereas the US English vowel space is in general more frontal and lower than the Dutch counterpart, a significant and systemic raising and backing of the vowel space was observerd, and not the expected fronting and lowering. These results must of course be taken with caution, as they are the result of the analysis of speech sounds generated by a single individual. Moreover, biological effects must not be understated and continued research is needed to disentangle environmental effects from aging effects. Further longitudinal studies into the speech plasticity of individuals experiencing L2 immersion will be able to enhance our insight into adult speech development.

6. References

Adank, Patti, Roeland Van Hout & Roel Smits. 2004. An acoustic description of the vowels of Northern and Southern Standard Dutch. *Citation: The Journal of the Acoustical Society of America* 116. 1130. doi:10.1121/1.1779271.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). doi:10.18637/jss.v067.i01.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10). 341–347. doi:10.1097/AUD.0b013e31821473f7.

Chang, Charles B. 2011. Systemic Drift of L1 Vowels in Novice L2 Learners. *Proceedings of the International Congress of Phonetic Sciences*(August). Hong Kong. 428–431.

Chang, Charles B. 2012. First Language Phonetic Drift During Second Language Acquisition. *PhD Thesis*. doi:http://dx.doi.org/10.1016/j.physbeh.2008.04.026.

Chang, Charles B. 2013. A novelty effect in phonetic drift of the native language. *Journal of Phonetics* 41(6). Academic Press. 520–533. doi:10.1016/j.wocn.2013.09.006.

Flege, James 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. doi:0095a4470/87/010047! http://www.citeulike.org/group/278/article/520469.

Flege, James E. 1995. Second Language Speech Learning: Theory, Findings, and Problems. In Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233–277. Baltimore: York Press. doi:10.1111/j.1600-0404.1995.tb01710.x. http://jimflege.com/files/Flege_in_Strange_1995.pdf (21 June, 2018).

Flege, James E. & W. Eefting. 1987. Production and perception of English stops by native Spanish speakers. *Journal of Phonetics* 15. 67–83.

Fowler, Carol A., Valery Sramko, David J. Ostry, Sarah A. Rowland & Pierre Hallé. 2008. Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics* 36(4). 649–663. doi:10.1016/j.wocn.2008.04.001.

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

Harrington, Jonathan. 2006. An acoustic analysis of “happy-tensing” in the Queen’s Christmas broadcasts. *Journal of Phonetics* 34(4). Academic Press. 439–457. doi:10.1016/j.wocn.2005.08.001.

Hillenbrand, James, Laura A Getty, Michael J Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5). 3099–3111. doi:10.1121/1.411872.

Kisler, Thomas, Uwe Reichel & Florian Schiel. 2017. Multilingual Processing of Speech via Web Services. *Comput. Speech Lang.* 45(C). GBR: Academic Press Ltd. 326–347. doi:10.1016/j.csl.2017.01.005.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13). doi:10.18637/jss.v082.i13. http://www.jstatsoft.org/v82/i13/.

Leeuw, Esther de. 2019. Native speech plasticity in the German-English late bilingual Stefanie Graf: A longitudinal study over four decades. *Journal of Phonetics* 73. Academic Press. 24–39. doi:10.1016/j.wocn.2018.12.002.

Major, Roy C. 1992. Losing English as a First Language. *The Modern Language Journal* 76(2). 190–208. doi:10.2307/329772.

R Core team. 2019. R: A language and environment for statistical computing. Vienna, Austria. https://www.r-project.org/.

Reubold, Ulrich & Jonathan Harrington. 2011. Disassociating the effects of age from phonetic change. 9–38. doi:10.1075/impact.37.02reu. https://benjamins.com/catalog/impact.37.02reu (18 March, 2020).

Sancier, Michele L. & Carol A. Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25(4). 421–436. doi:10.1006/jpho.1997.0051.

Schiel, Florian. 1999. Automatic phonetic transcription of nonprompted speech. *Proceedings of ICPhS*. 607–610.

Schiel, Florian, Christoph Draxler & Jonathan Harrington. 2011. Phonemic segmentation and labelling using the MAUS technique. *New Tools and Methods for Very-Large-Scale Phonetics Research*. http://epub.ub.uni-muenchen.de/13684/1/schiel_13684.pdf%5Cnhttp://epub.ub.uni-muenchen.de/13684/ (19 March, 2020).

Tobin, Stephen J., Hosung Nam & Carol A. Fowler. 2017. Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *Journal of Phonetics*. doi:10.1016/j.wocn.2017.05.006.

Yang, Byunggon. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24(2). 245–261. doi:10.1006/jpho.1996.0013.

7. Appendix

Table 4 - Linear mixed effect model output per vowel. Vowel, formants, estimates, standard error, t-value, p-value and 95% confidence interval is given.

Vowel		Effect	Estimate	Std. Error	*t*-value	dF	*p*-value	Low CI	High CI
ɑ	F1	Period	-61.43	15.59	-3.94	1151.79	8.64E-05	-91.99	-30.87
		(Intercept)	2841.43	4873.94	0.58	1148.34	0.56	-6711.32	12394.18
		Age	-1.02	2.43	-0.42	1148.18	0.67	-5.78	3.73
	F2	Period	-124.37	26.28	-4.73	1149.80	2.50E-06	-175.89	-72.85
		Age	7.64	4.09	1.87	1142.28	0.06	-0.38	15.66
		(Intercept)	-13896.88	8221.14	-1.69	1142.53	0.09	-30010.02	2216.27
a:	F1	(Intercept)	8161.88	7016.32	1.16	668.76	0.25	-5589.86	21913.61
		Age	-3.62	3.49	-1.04	668.74	0.30	-10.47	3.23
		Period	-15.66	22.78	-0.69	670.84	0.49	-60.31	29.00
	F2	Period	-29.54	19.30	-1.53	662.69	0.13	-67.37	8.28
		Age	3.76	2.96	1.27	665.62	0.20	-2.04	9.56
		(Intercept)	-5903.94	5947.12	-0.99	665.66	0.32	-17560.08	5752.21
ɔ	F1	(Intercept)	10008.44	5561.94	1.80	428.73	0.07	-892.76	20909.64
		Age	-4.67	2.77	-1.69	428.73	0.09	-10.10	0.76
		Period	-25.80	18.80	-1.37	427.70	0.17	-62.64	11.04
	F2	Period	-181.22	53.89	-3.36	439.36	8.39E-04	-286.83	-75.60
		Age	2.69	7.94	0.34	439.74	0.74	-12.87	18.24
		(Intercept)	-4126.22	15938.28	-0.26	439.76	0.80	-35364.68	27112.24
e:	F1	Period	-45.65	14.00	-3.26	650.98	1.17E-03	-73.08	-18.21
		Age	1.14	2.03	0.56	654.46	0.57	-2.83	5.11
		(Intercept)	-1689.93	4067.93	-0.42	654.50	0.68	-9662.92	6283.07
	F2	Age	1.34	4.00	0.33	647.61	0.74	-6.51	9.19
		Period	-5.24	27.64	-0.19	645.49	0.85	-59.40	48.93
		(Intercept)	-594.97	8038.76	-0.07	647.64	0.94	-16350.65	15160.71
ɛ	F1	(Intercept)	9617.10	3327.29	2.89	927.72	3.94E-03	3095.73	16138.48
		Age	-4.46	1.66	-2.69	927.68	0.01	-7.71	-1.21
		Period	-5.13	10.78	-0.48	926.01	0.63	-26.25	15.99
	F2	Period	70.42	21.03	3.35	934.72	8.46E-04	29.20	111.65
		(Intercept)	16111.55	6510.82	2.47	934.70	0.01	3350.57	28872.54
		Age	-7.08	3.24	-2.18	934.72	0.03	-13.44	-0.73
i	F1	Period	-47.87	24.72	-1.94	268.40	0.05	-96.31	0.58
		(Intercept)	6654.70	7947.72	0.84	270.34	0.40	-8922.53	22231.94
		Age	-3.06	3.96	-0.77	270.37	0.44	-10.82	4.69
	F2	Period	-137.67	44.11	-3.12	273.65	1.99E-03	-224.12	-51.22
		Age	14.27	7.05	2.02	276.89	0.04	0.45	28.09
		(Intercept)	-26512.64	14159.97	-1.87	276.92	0.06	-54265.68	1240.39
ɪ	F1	Period	-48.45	13.31	-3.64	683.38	2.92E-04	-74.53	-22.37
		Age	0.35	1.99	0.18	686.97	0.86	-3.55	4.25
		(Intercept)	-156.89	3997.93	-0.04	686.97	0.97	-7992.69	7678.91
	F2	Age	0.89	3.95	0.23	685.41	0.82	-6.85	8.63
		Period	2.00	26.34	0.08	678.51	0.94	-49.62	53.63
		(Intercept)	322.86	7929.54	0.04	685.42	0.97	-15218.76	15864.47
o:	F1	Period	-67.95	18.87	-3.60	472.85	3.49E-04	-104.93	-30.98
		(Intercept)	2227.31	5429.59	0.41	472.06	0.68	-8414.49	12869.10
		Age	-0.77	2.70	-0.29	472.03	0.77	-6.07	4.53

	F2	Period	-160.48	34.76	-4.62	443.91	5.10E-06	-228.61	-92.36
		Age	14.10	4.99	2.83	443.79	4.90E-03	4.33	23.87
		(Intercept)	-27067.33	10011.25	-2.70	443.85	0.01	-46689.01	-7445.64
u	F1	Period	-24.38	27.95	-0.87	177.58	0.38	-79.16	30.39
		(Intercept)	1728.62	9312.82	0.19	177.86	0.85	-16524.18	19981.41
		Age	-0.60	4.64	-0.13	177.86	0.90	-9.69	8.49
	F2	Period	-124.92	95.91	-1.30	171.91	0.19	-312.90	63.05
		(Intercept)	26652.54	31993.45	0.83	172.47	0.41	-36053.47	89358.55
		Age	-12.61	15.93	-0.79	172.46	0.43	-43.84	18.61
y	F1	Period	-108.06	35.30	-3.06	82.00	2.98E-03	-177.25	-38.87
		Age	12.61	5.37	2.35	82.00	0.02	2.08	23.14
		(Intercept)	-24853.54	10789.03	-2.30	82.00	0.02	-45999.65	-3707.43
	F2	Period	-143.23	59.29	-2.42	69.08	0.02	-259.44	-27.02
		Age	18.99	9.25	2.05	72.17	0.04	0.85	37.12
		(Intercept)	-36361.36	18581.53	-1.96	72.18	0.05	-72780.48	57.77
ʏ	F1	Period	-53.24	26.83	-1.98	118.32	0.05	-105.83	-0.65
		(Intercept)	11539.25	8521.12	1.35	118.23	0.18	-5161.83	28240.33
		Age	-5.48	4.24	-1.29	118.24	0.20	-13.80	2.84
	F2	(Intercept)	18168.93	16323.29	1.11	107.24	0.27	-13824.13	50161.98
		Age	-8.20	8.13	-1.01	107.25	0.32	-24.14	7.73
		Period	29.48	51.50	0.57	104.22	0.57	-71.45	130.41