Dutch Sung and Spoken Vowel Space

How Vowels Are Used to Strengthen Song

Tori van Oostveen

12609870

BA thesis Linguistics

University of Amsterdam

Supervised by Suki Yiu

27-06-2022

Acknowledgements

I’d like to thank my supervisor Suki Yiu for helping me in the creation and process of this research. I’d also like to thank my family for being there for me. Especially my father, who helped me talk through the musical side of this research and who did the trial run with me. I’d also like to thank my participants for participating and the choir they are a part of for letting me recruit through them.

Abstract

This paper looks at the Dutch sung and spoken vowel space with a focus on the effects of mode and pitch. Previous research showed sung vowel space to be smaller than spoken and an effect called formant tuning where when the pitch gets high enough mostly the first formant starts to match up with the pitch. This was investigated by asking participants to say and sing at a number of notes with Dutch words containing the nine monophthongs of the language. This was recorded and the formants were extracted and analysed. Both mode and pitch caused significant changes in vowel space as was predicted from previous research. Sung vowel space was found to be smaller than spoken vowel space. There was also evidence of formant tuning.

Table of Contents

1. Introduction

When people sing they make use of a number of strategies to make it sound good, to change mood, or volume. Some of these strategies concern the manipulation of vowels, which changes the vowel space. This paper will look at vowel space in spoken and sung Dutch. Specifically it will look at the behaviour of the first and second formants. The difference in vowel space and the effect of the pitch of the note and the experience, education and range of the singer.

This research follows on from Bradley (2018), where the vowel space of spoken and sung English was compared and a significant difference was found. The vowel space tended to be smaller in sung than spoken English and the F1 tended to move towards the middle, while F2 got lower. Similar results are expected in this experiment.

The pitch range of the singer also seems to play a role in the vowel space. The higher one gets the more likely a singer is to try to match one or more formants to the pitch or one of its harmonics to amplify the sound (Henrich et al., 2011). In lower ranges there is also a tendency to lower the fourth and fifth formant to match the third, also to amplify the sound. This is called the singer's formant and is especially prevalent in lower voices (Bradley, 2018).

This leads to the prediction that vowel spaces will change based on range and pitch. When someone gets higher in pitch the F1 and F2 will start to go up as well. This behaviour is expected in higher voices like soprano's and possibly alto's and tenors. It is not expected in the lower ranges, such as basses and baritones.

Section 2 provides a more in-depth review of the literature. Section 3 contains the methodology of the experiment conducted. Section 4 shows the results of the experiment. Section 5 is a discussion of the experiment in the context of previous findings. Section 6 is the conclusion.

2. Literature Review

Henrich et al. (2011) gives a clear overview of previous research into formant tuning in the introduction. The paper then investigates the phenomenon. Formant tuning is a technique used in singing, which is mostly visible at higher pitches and thus higher ranges. The term is used when a singer, most often sopranos and some altos and tenors, matches one, or both of the formants of a vowel to the pitch at which they are singing. They can also match to one or more octaves above the pitch of the note, which happens more among singers with lower ranges. This strengthens the volume while sacrificing coherency. It is a widely described phenomenon and supported theory (Bresch & Narayanan, 2010; Garnier et al., 2010; Henrich et al., 2011; Joliveau et al., 2004; Miller & Schutte, 1990; Neumann et al., 2005; Schutte et al., 2005; Sundberg, 1975; Sundberg & Skoog, 1997; Titze et al., 1994; Yu Chan & Do, 2021)

Henrich et al. (2011) made use of token words that contained the target vowels. The participants were asked to say these words and sing them at indicated pitches. The pitches followed an ascending scale from the lowest part of the singer's comfortable range to the highest. The study found resonance tuning was consistently employed among its soprano participants, both F1 to f0 and F2 to the 2f0 at higher pitches. Among the altos it was used to a lesser extent and on a smaller part of their range, the lower part of their range they tended to tune F1 to 2f0. They also found evidence of resonance tuning to 2f0 and 3f0 among tenors and baritones in the higher part of their range. At lower pitches tuning happened to higher octaves.

Most studies have focused on the soprano range (Bresch & Narayanan, 2010; Garnier et al., 2010; Joliveau et al., 2004; Sundberg & Skoog, 1997). This is the range where resonance tuning has the most obvious and marked effect, although not all agree that it is a uniform phenomenon (Bresch & Narayanan, 2010; Joliveau et al., 2004). The reason that it is so marked in the soprano range is because the pitch comes into the range of the formants earlier and the formants are thus forced to join the f0 across a larger part of the range to get the volume bolstering effect of formant tuning.

There are less studies into other ranges although all have been looked at. Studies into male voice ranges have found that singers in those categories tend to tune their formants to higher octaves of the pitch (Miller & Schutte, 1990; Neumann et al., 2005; Schutte et al., 2005; Titze et al. 1994). Even when tenors can tune to the f0, they seem to avoid this and choose to tune to a higher octave, possibly to avoid a more feminine sound (Titze et al., 1994).

Bradley (2018) looked at the difference in vowel space between sung and spoken English. The methodology made use of a modified preexisting song. This song contained the target vowels. The participants were asked to both sing the song and say it "as if to an audience in the manner of a poem" (Bradley, 2018, 386). This is a method that focuses on mode and does not take into account the effect pitch and range can have on vowel space. Using a song with a changing pitch makes results less reliable because of the aforementioned effect. The experiment also had an unbalanced participant distribution in both range and experience which went unaddressed in the discussion.

Bradley (2018) found a significant effect of mode on both F1 and F2. When the participants were split according to gender, women had a significant effect on both and men had a significant effect on F2 and not F1. The effect of mode on vowel space was a general lowering of F2, a compression of vowel space and less variance in the vowels themselves, when sung.

3. Methods

For this research participants who have experience in choir singing were recruited. They were recorded both singing and speaking a text and words. They were also asked some questions about their singing experience, education and other background information. The recordings were analysed with Praat (Boersma & Weenink, 2021).

3.1 Participants

Ten participants aged between 66 and 78 years old were recruited. They were asked in a questionnaire about their singing experience, including their training and the style of singing they do. This information was gathered to get more insight into which factors could influence the amount of vowel change between sung and spoken language. Table 1 shows the range reported, and the lowest and the highest note recorded per participant. This paper makes use of the term octave, which indicates a doubling of frequency. So when a note (a4) has a frequency of 440 Hz, the note an octave higher (a5) will have a frequency of 880 Hz.

Table 1: Participants

Participant	Range	Lowest note	Highest note
1	Bass	e2	g4
2	Soprano	g3	e5
3	Tenor	c3	e4
4	Tenor	c3	e4
5	Alto	e3	c5
6	Soprano	g3	e5

3.2 Stimuli and Procedure

The target vowels were surrounded by two consonants as in table 2. The vowels were placed between consonants so that participants could read the words and could interpret the vowel themselves. This meant the vowel did not have to be indicated auditory and the pronunciation could not be influenced by the experimenter. The consonants were chosen so they did not differ too much between vowels so that the same amount of consonant interference could be expected across the different stimuli.

Table 2: Target vowels within lexical context.

Target vowel	Word	IPA transcription	Translation
[uː]	voet	[vuːt]	foot
[iː]	fiets	[fiːts]	bike
[ɪ]	fit	[fɪt]	fit
[yː]	fuut	[fyːt]	grebe (type of bird)
[ʏ]	fut	[fʏt]	pep/energy
[ɛ]	vet	[vɛt]	fat
[ɔ]	vod	[vɔt]	rag
[ɑ]	vat	[vɑt]	barrel
[aː]	vaat	[vaːt]	dishes

The participants were given room to warm up their voice if they wanted to. This was not recorded. After possible warming up the note to note range of the participants was approximated using their reported range and trial and error. The words were presented to the participants and they were asked to first say the words in a neutral speaking voice twice to get data on the spoken vowel space. After that they were asked to sing it on a number of pitches, the notes of a c-major chord (c, e and g), with a4 set to 440 Hz. For more information on the notes and corresponding pitches see appendix B.1. These pitches were presented using a tuning app. The indicated frequencies can be seen in appendix B.1 together with the mean of the recorded pitch. They were asked to do so with little to no vibrato (Garnier et al., 2010; Henrich et al., 2011; Joliveau et al., 2004). For the recording a Zoom H2n microphone was used and the data was analysed using Praat (Boersma & Weenink, 2021). Sampling frequency was at 96000Hz. The recording sessions took approximately twenty minutes.

3.3 Method of Analysis

The vowels each got their own file, to make analysis easier. The recordings were loaded into praat and a TextGrid was made for every vowel through a Praat script that can be found in appendix C.1. The files were annotated by hand to include only the vowels. These sound and TextGrid objects were saved by hand. Then a different Praat script was used to extract f0, F1 and F2. This script can be found in appendix C.2. This script also recorded this information and put it, together with the participant number and the information gathered from the questionnaire (excluding age), into the info window from which it was saved as a .csv file. This .csv file was used in the data analysis for which R was used. A linear model was used to analyse the data as a whole and per vowel (R Core Team, 2022).

4. Results

4.1 Effect of Mode

Figure 1 shows the effect of mode. The sung vowel space is in red and the spoken vowel space is in black. Spoken vowel space is shown to be bigger and have a higher F2. most back vowels, except [uː], move to the front and front vowels move to the back. High vowels lower and low vowels move up. Both [ɔ] and [ɛ] move down slightly as well. A linear model with F1 as a dependent variable and mode as an independent variable, shows that singing mode has a significant effect of lowering the F1 by 66.31 Hz ($p < 0.001$, 95% CI: [44.67 Hz, 87.95 Hz]). And a linear model with F2 as a dependent variable and mode as an independent variable shows that singing mode has a non-significant effect of lowering F2 by 61.38 Hz ($p = 0.070$, 95% CI: [-127.80 Hz, +5.04 Hz]).

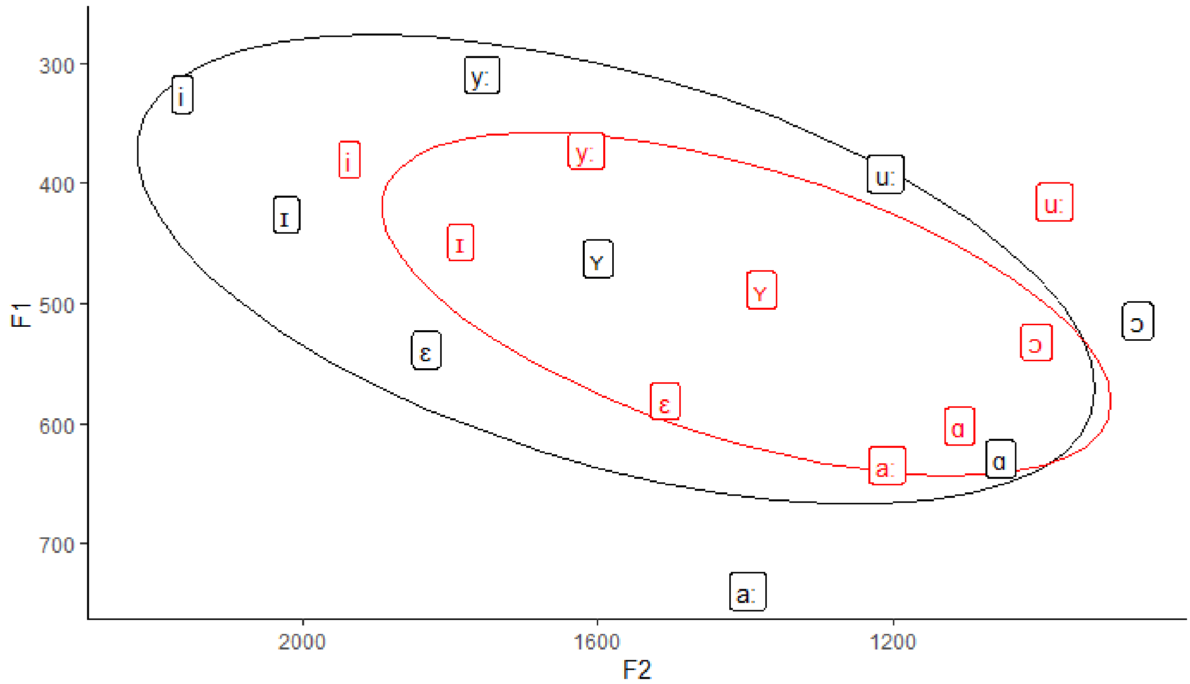


Figure 1: Spoken and Sung vowels; spoken are indicated in black and sung are red.

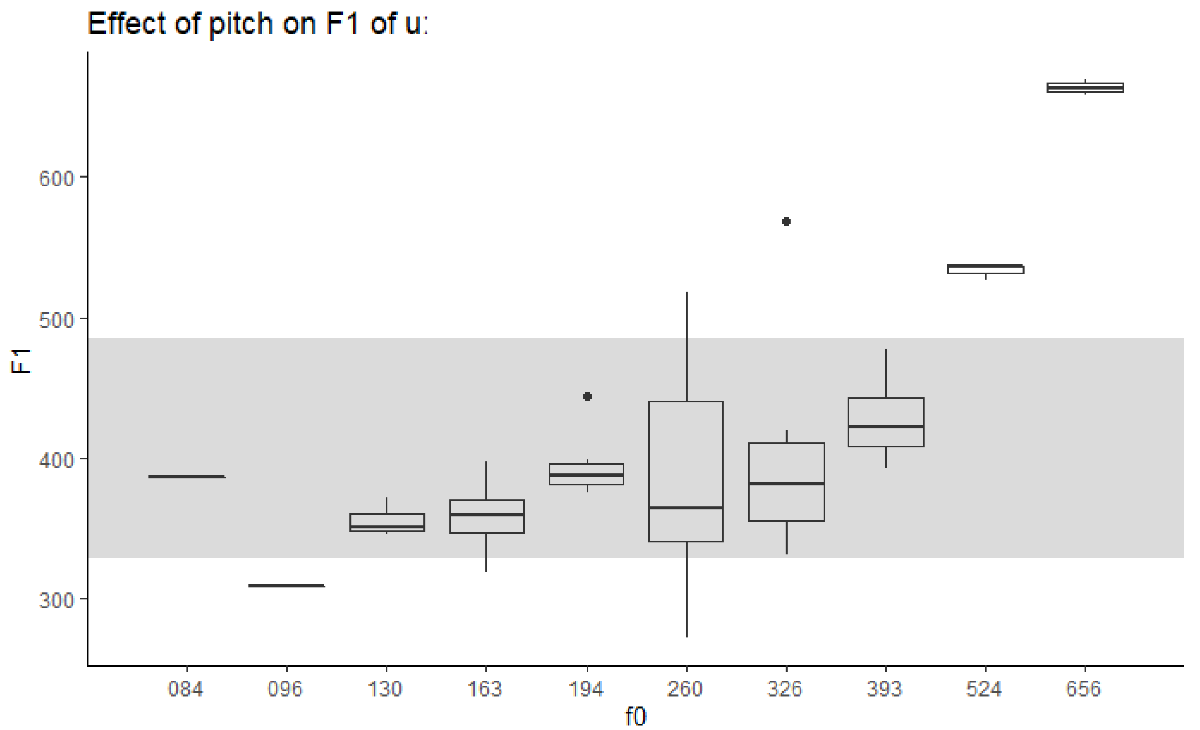
Table 3 shows the effect of mode on both the F1 and F2 per vowel. Per vowel two linear models were used. One with mode as the independent and F1 as the dependent variable and one with mode as the independent and F2 as the dependent variable. It gives the effect in Hz, the p-value and the 95% confidence interval in Hz. When 0.05 > p > 0.01 the result is marked in blue, when 0.01 > p > 0.001 the result is marked in orange and when 0.001 > p the result is marked in red. It shows a general lowering of F1 among high vowels and [a] and little significant effect on F2, except for the significant lowering of the F2 of [u], [ɛ] and [a].

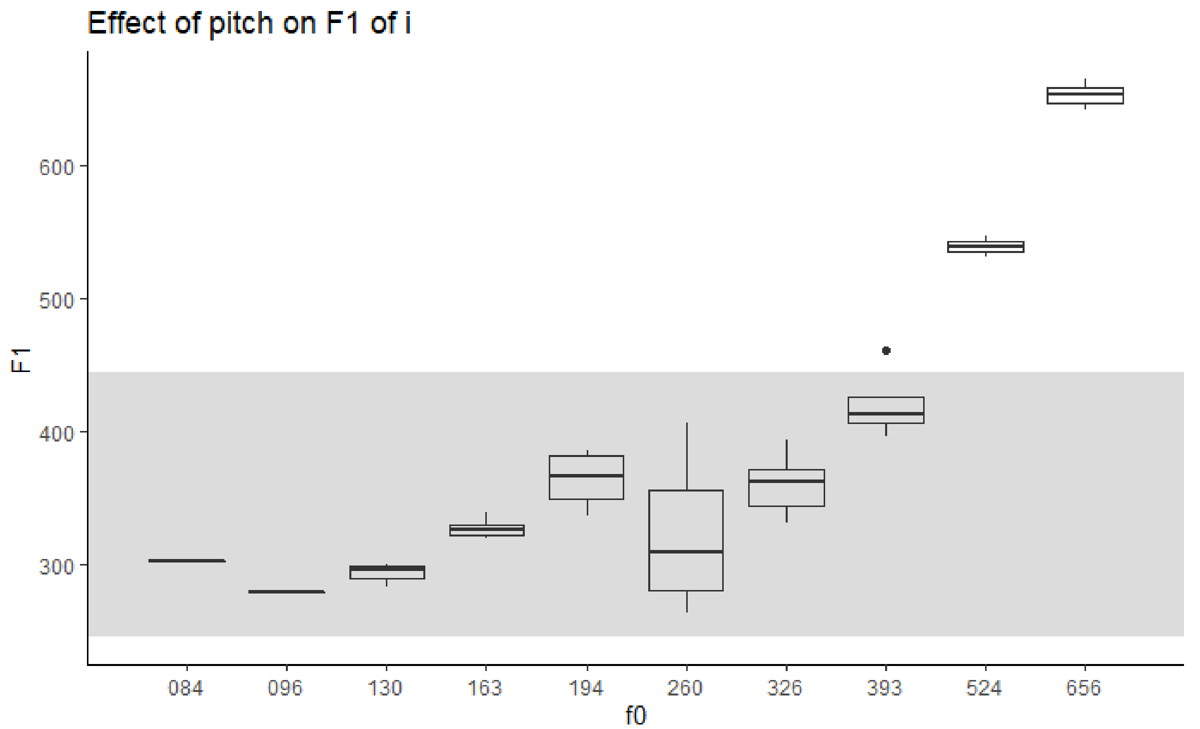
Table 3: Effect of Mode

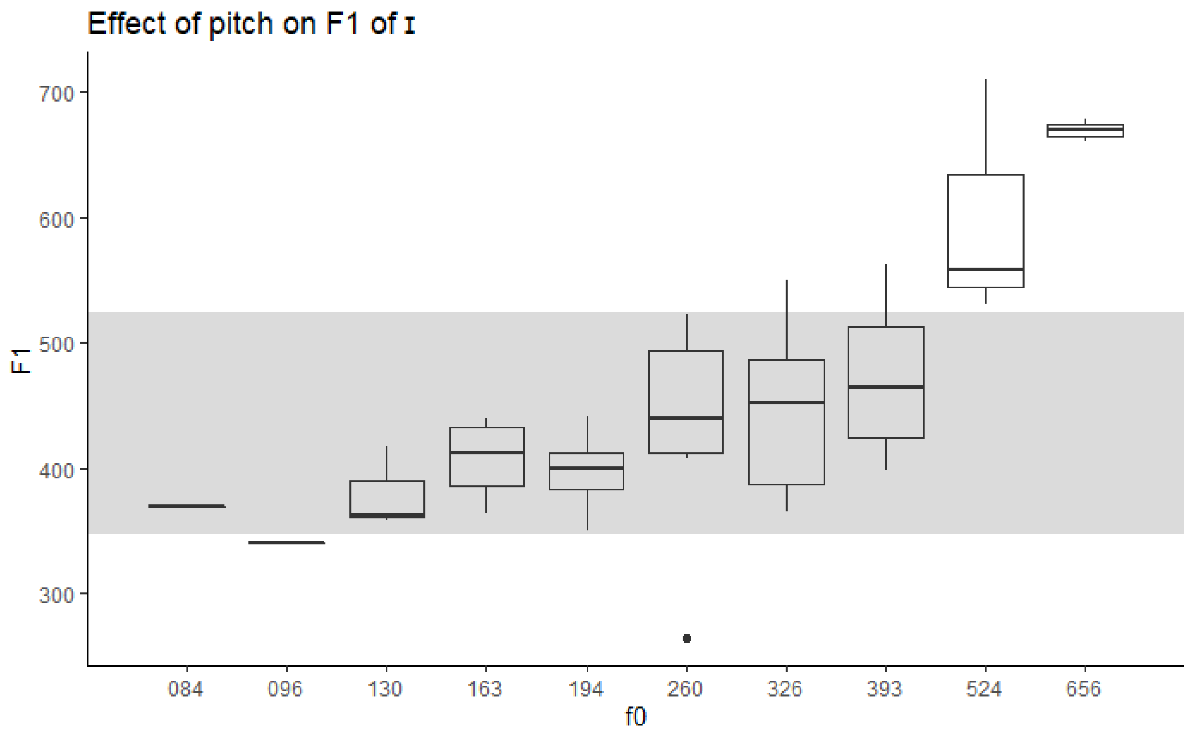
	F1			F2		
	Est. (Hz)	p-value	CI (Hz)	Est. (Hz)	p-value	CI (Hz)
[u]	-91.46	< 0.001	[-141.99, -40.93]	-395.30	0.006	[-670.98, -119.62]
[i]	-88.55	< 0.001	[-134.25, -42.86]	141.81	0.119	[-38.29, 321.92]
[ɪ]	-99.14	< 0.001	[-143.65, -54.63]	14.12	0.769	[-82.22, 110,45]
[y]	-87.85	< 0.001	[-132.04, -43.66]	92.09	0.105	[-20.20, +204.39]
[ʏ]	-66.37	0.004	[-109.93, -22.81]	-109.63	0.073	[-230.02, +10.76]
[ɛ]	-32.96	0.271	[-92.69, +26.77]	-108.12	0.029	[-204.64, -11.60]
[ɔ]	-46.42	0.108	[-103.48, +10.64]	-97.15	0.486	[-376.20, +181.90]
[ɑ]	-10.48	0.777	[-84.87, +63.91]	32.08	0.379	[-40.82, +104.98]
[a]	-76.53	0.050	[-152.92, -0.13]	-102.92	0.048	[-205.07, -0.77]

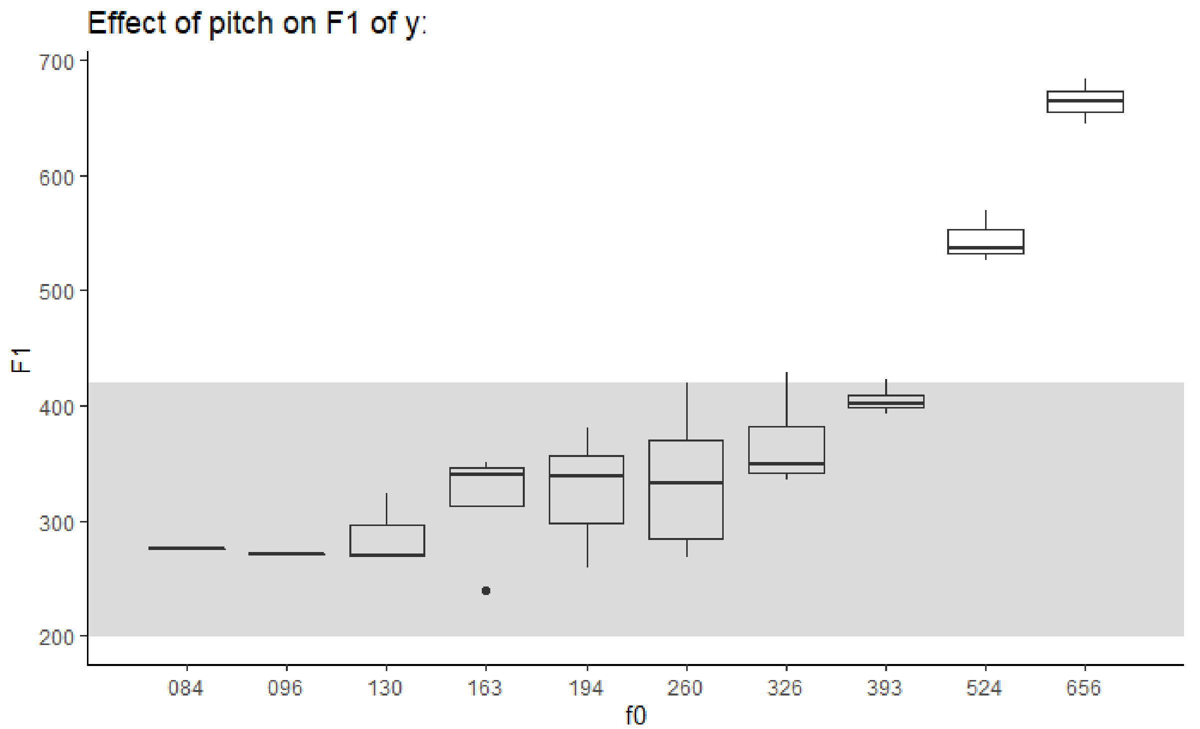
4.2 Effect of Pitch

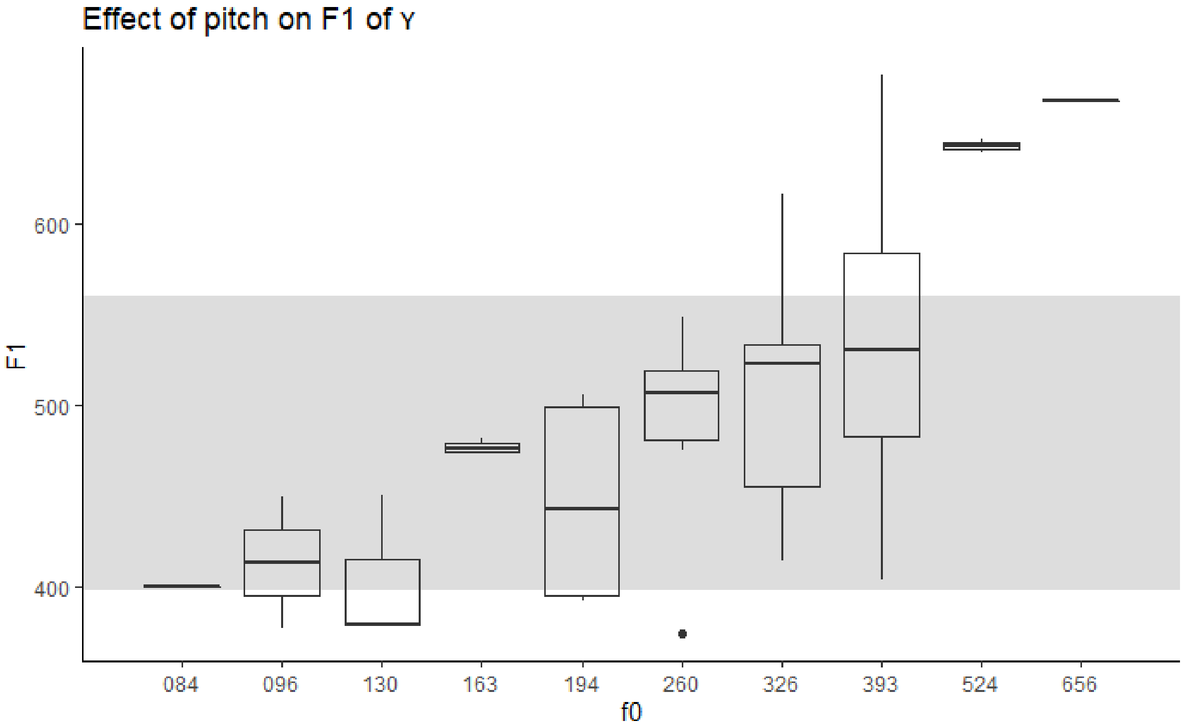
Figure 2 shows the effect of pitch on F1 per sung vowel. No adjustment of the x-axis has been done to account for the exponential nature of pitch. The grey area indicates the range of the F1 in spoken Dutch as a point of comparison. In the high vowels a clear departure from the spoken range can be seen. The formant frequency matches up with the pitch of the note. This can be seen in to a lesser extend in [ɔ] and [ɛ] and not at all in the low vowels.

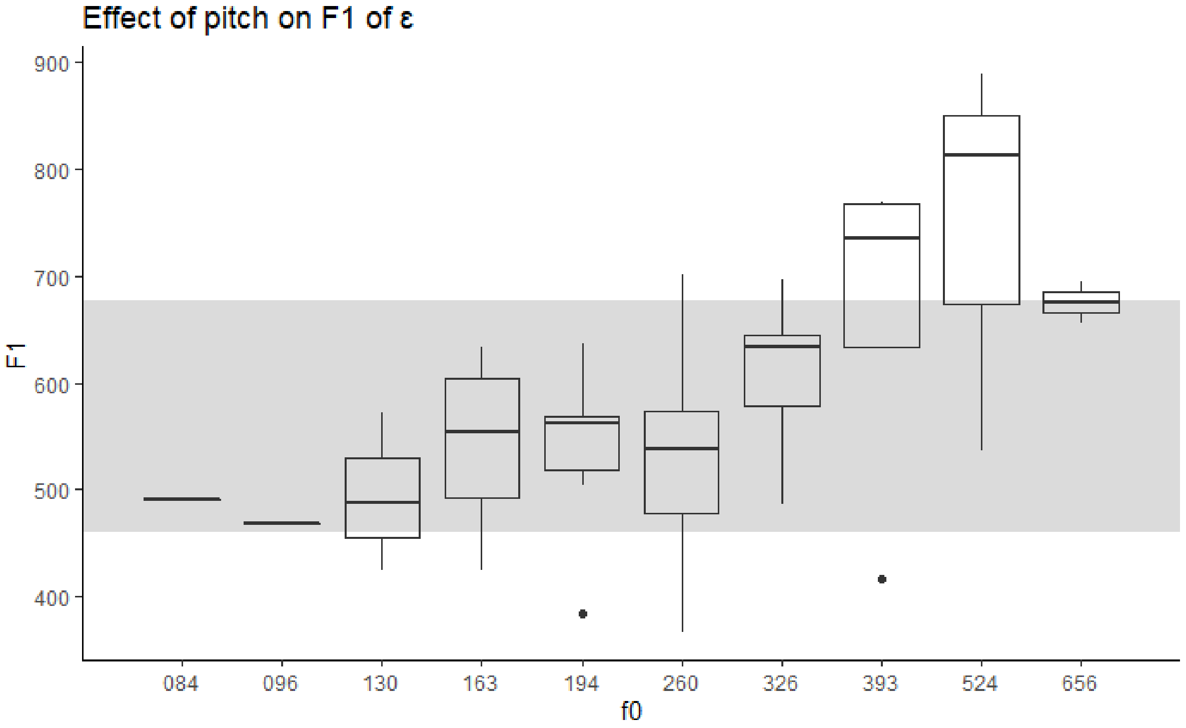

Effect of pitch on F1 of u:
F1
600
500
400
300
084
096
130
163
194
260
326
393
524
656
f0

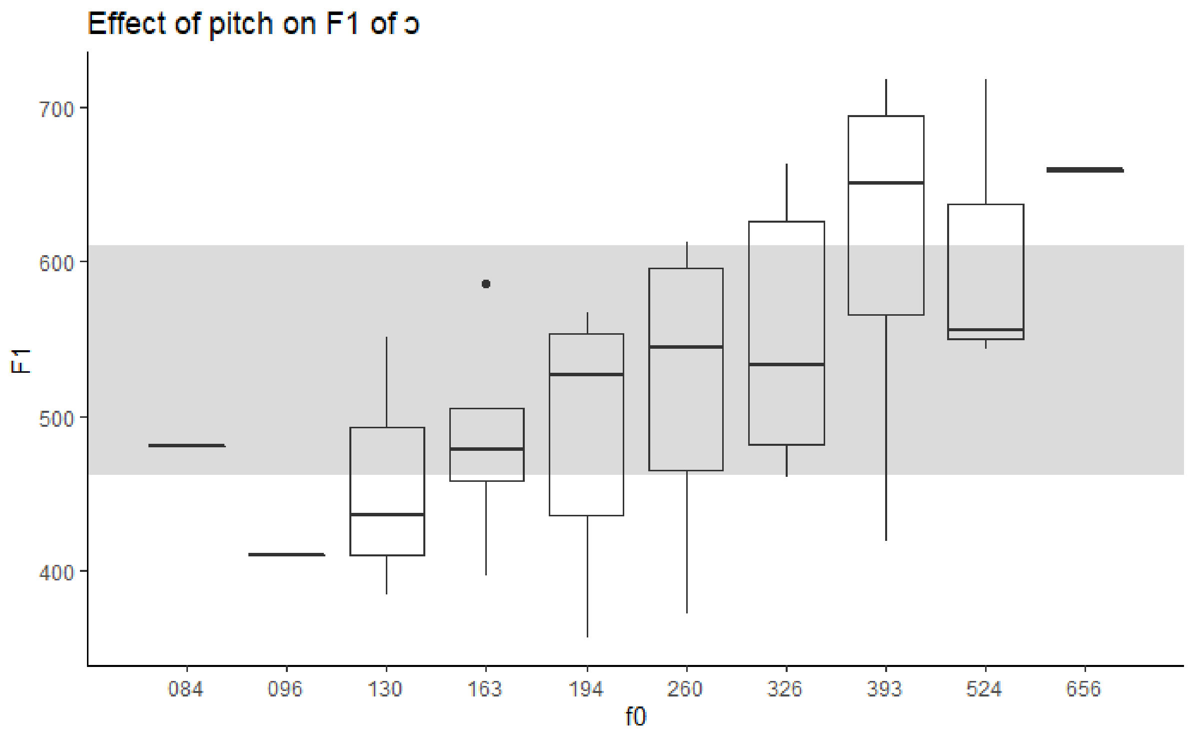

Effect of pitch on F1 of i
F1
600
500
400
300
084
096
130
163
194
260
326
393
524
656
f0

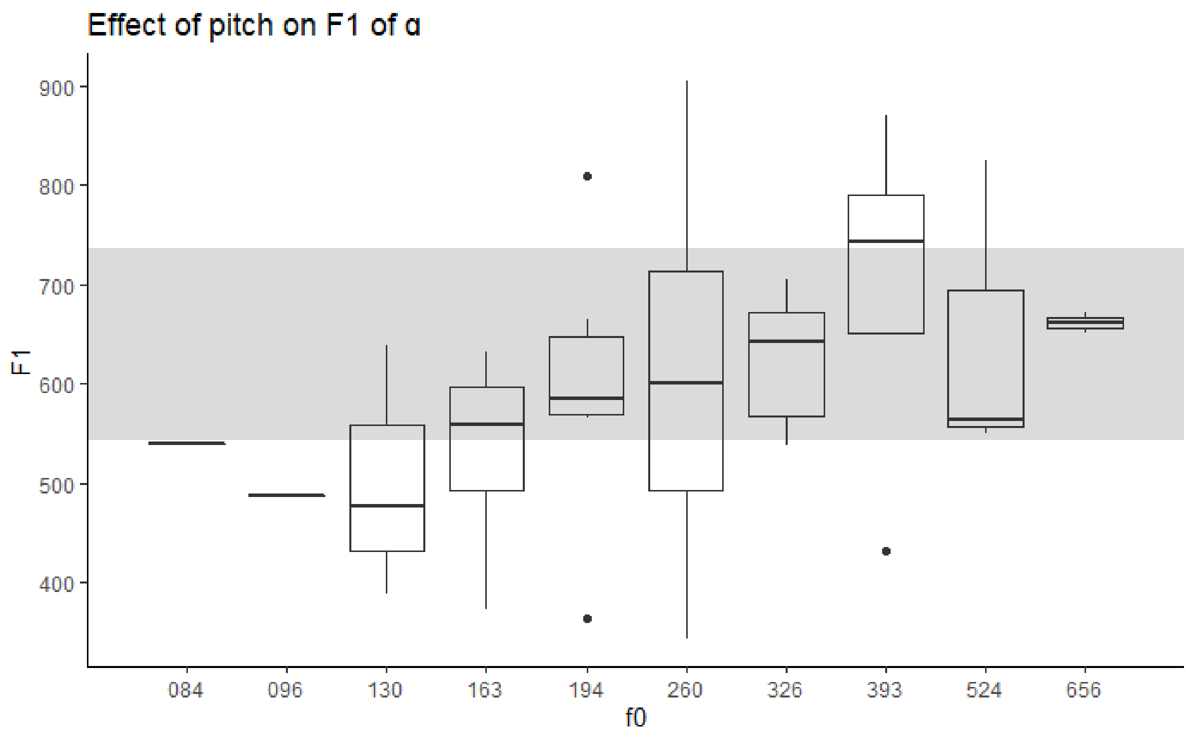

Effect of pitch on F1 of ɪ
F1
700
600
500
400
300
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F1 of yː
F1
700
600
500
400
300
200
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F1 of ɣ
F1
600
500
400
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F1 of ε
F1
900
800
700
600
500
400
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F1 of ɔ
700
600
500
400
F1
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F1 of ɑ
900
800
700
600
500
400
F1
084
096
130
163
194
260
326
393
524
656
f0

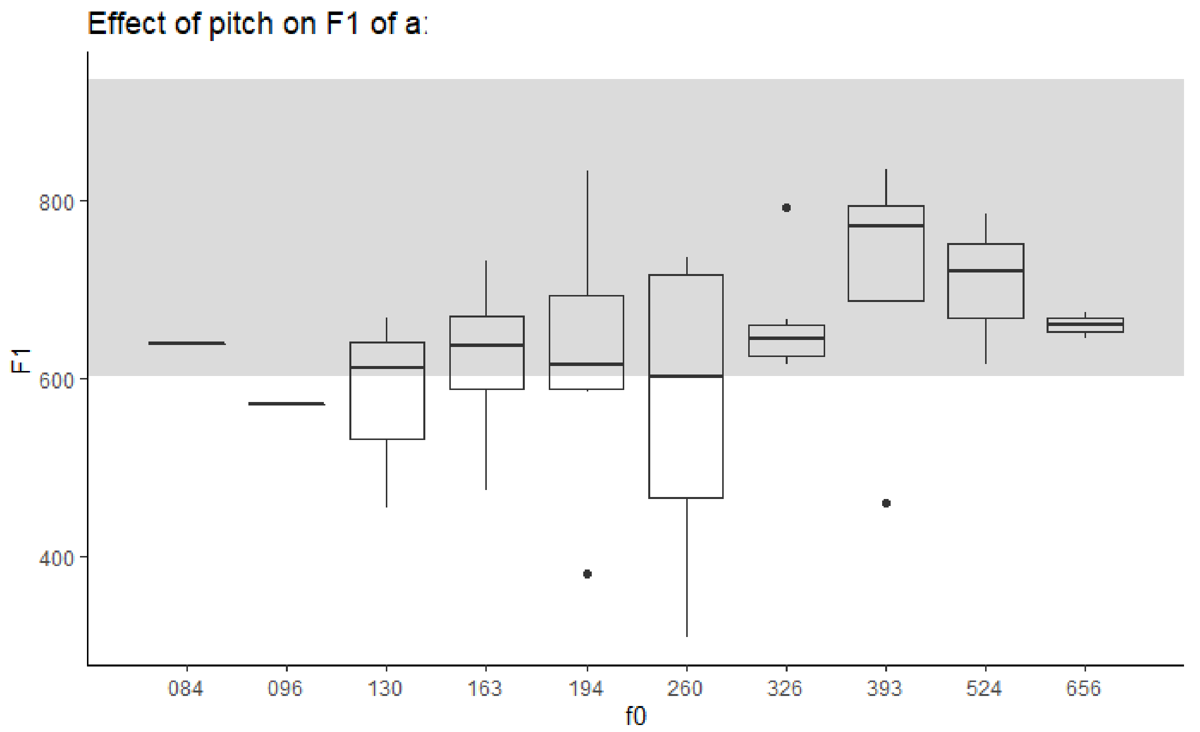
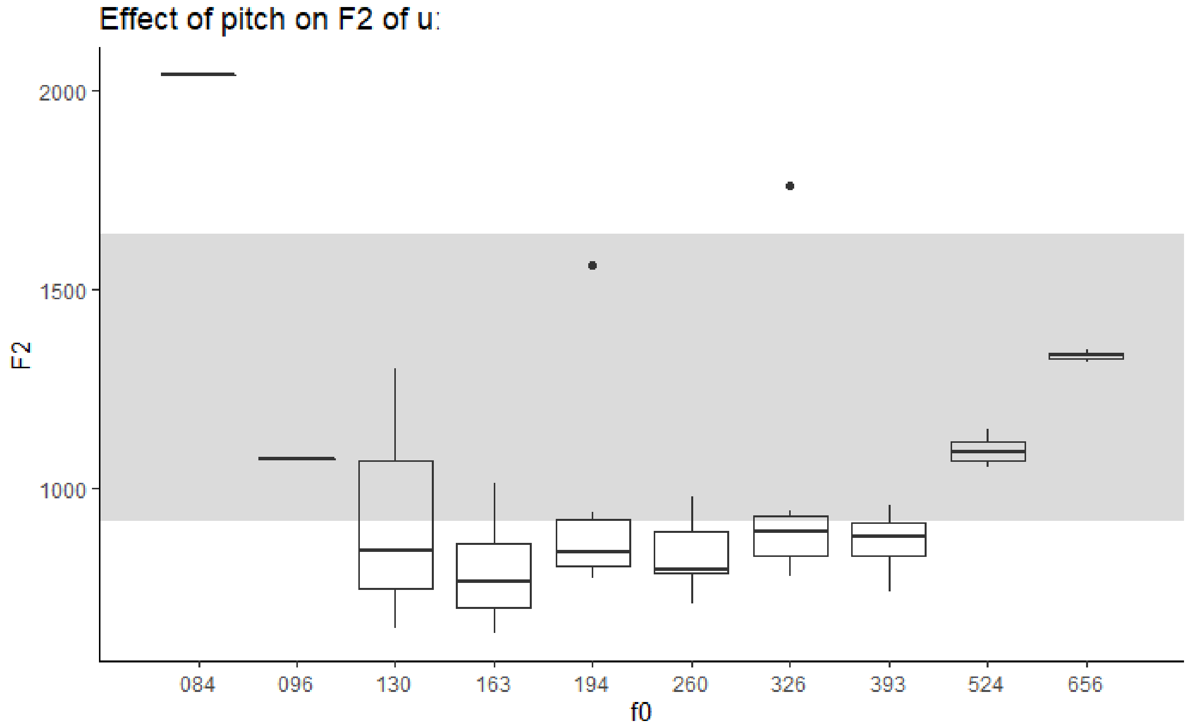
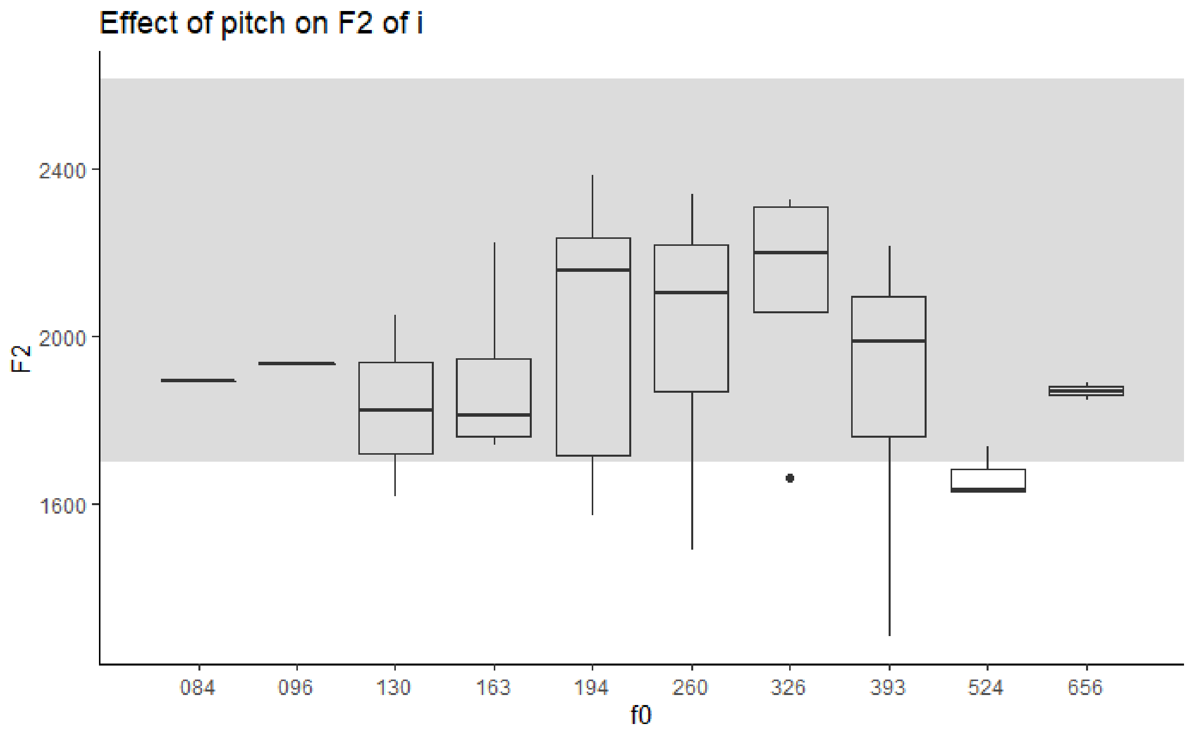
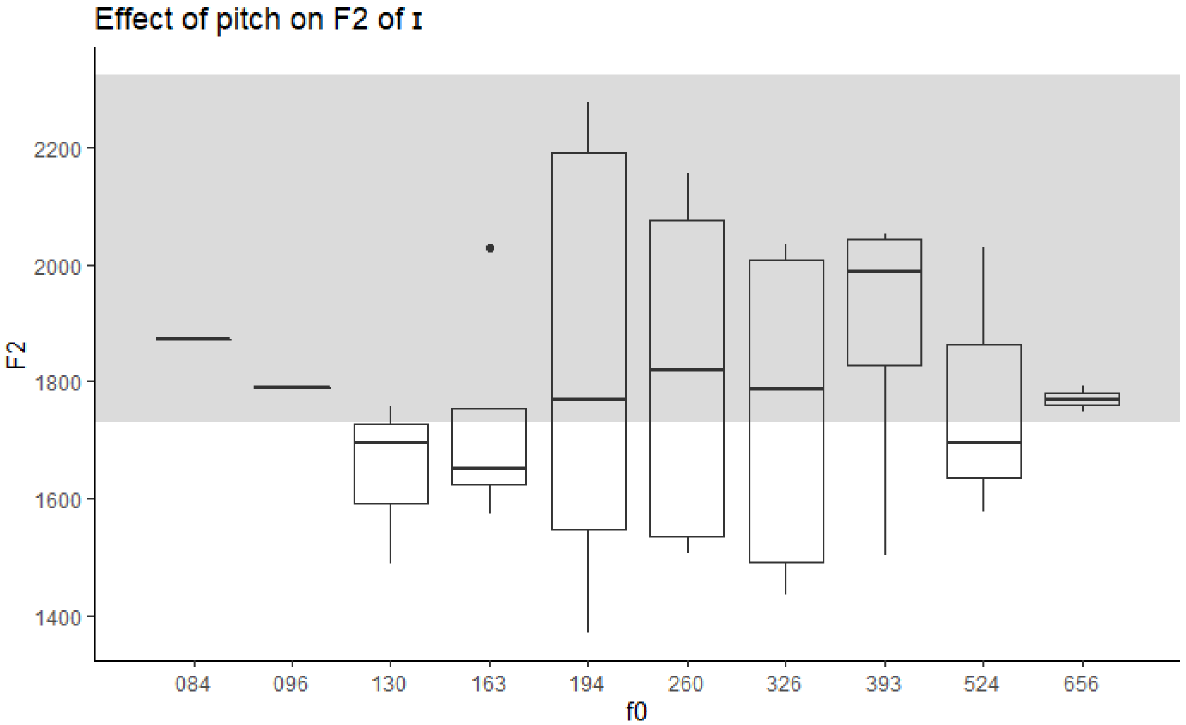


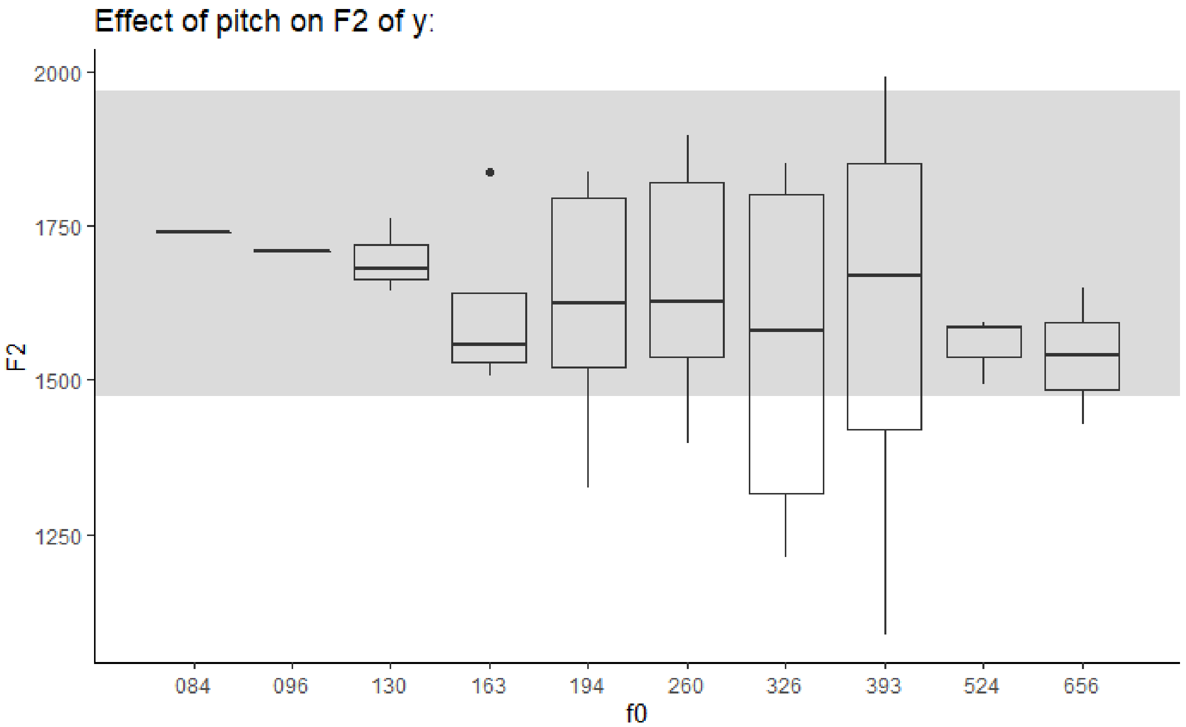
Figure 2: Effect of pitch on F1 in song.

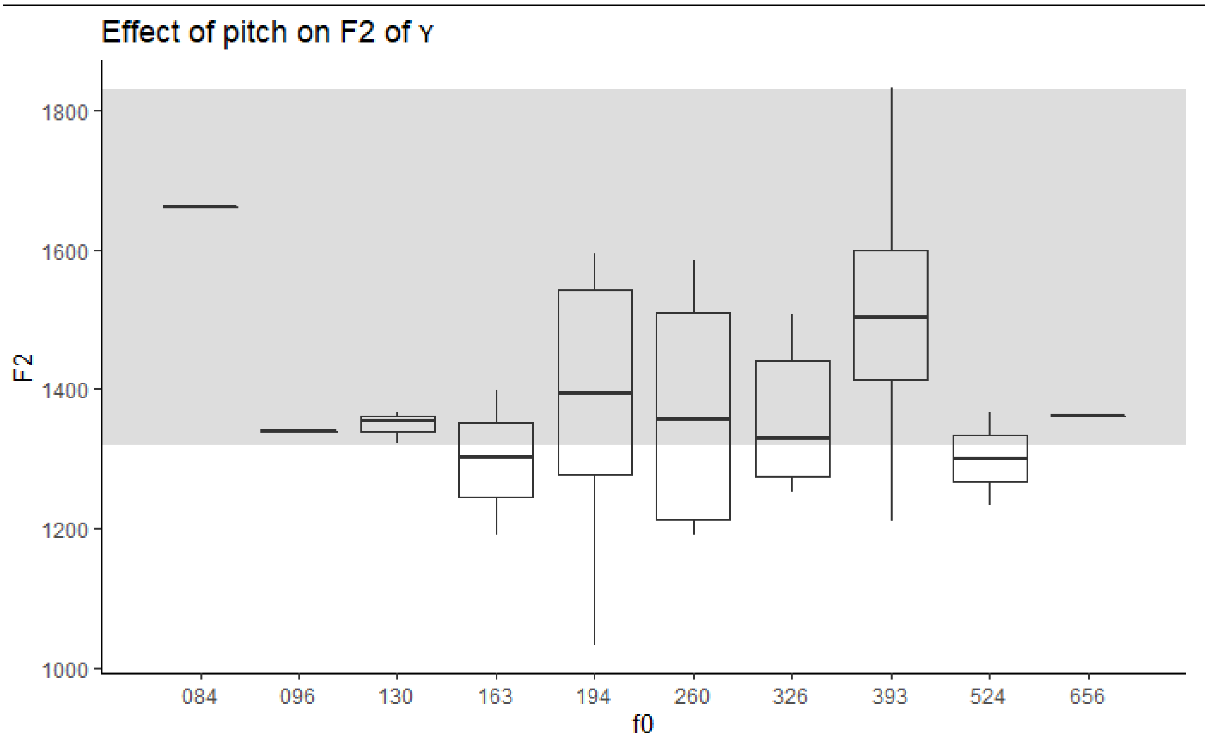
Figure 3 shows the effect of pitch on F2 per sung vowel. No adjustment of the x-axis has been done to account for the exponential nature of pitch. The grey area indicates the range of the F2 in spoken Dutch as a point of comparison. Front vowels move back and back vowels move to the front. Back vowels also show a tendency to match up with the pitch, although only the [ɔ] moves out of the spoken range area. Front vowels also show more variation in the F2 than back vowels.

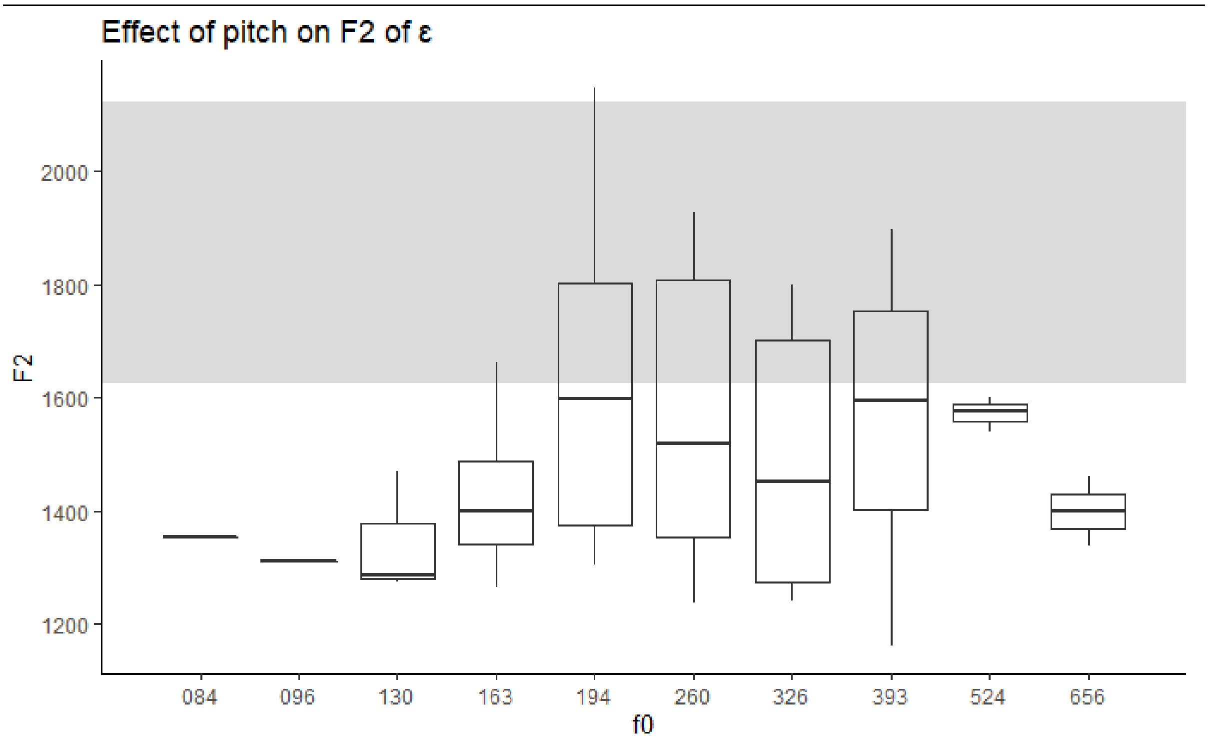

Effect of pitch on F2 of uː
2000
1500
1000
F2
084
096
130
163
194
260
326
393
524
656
f0

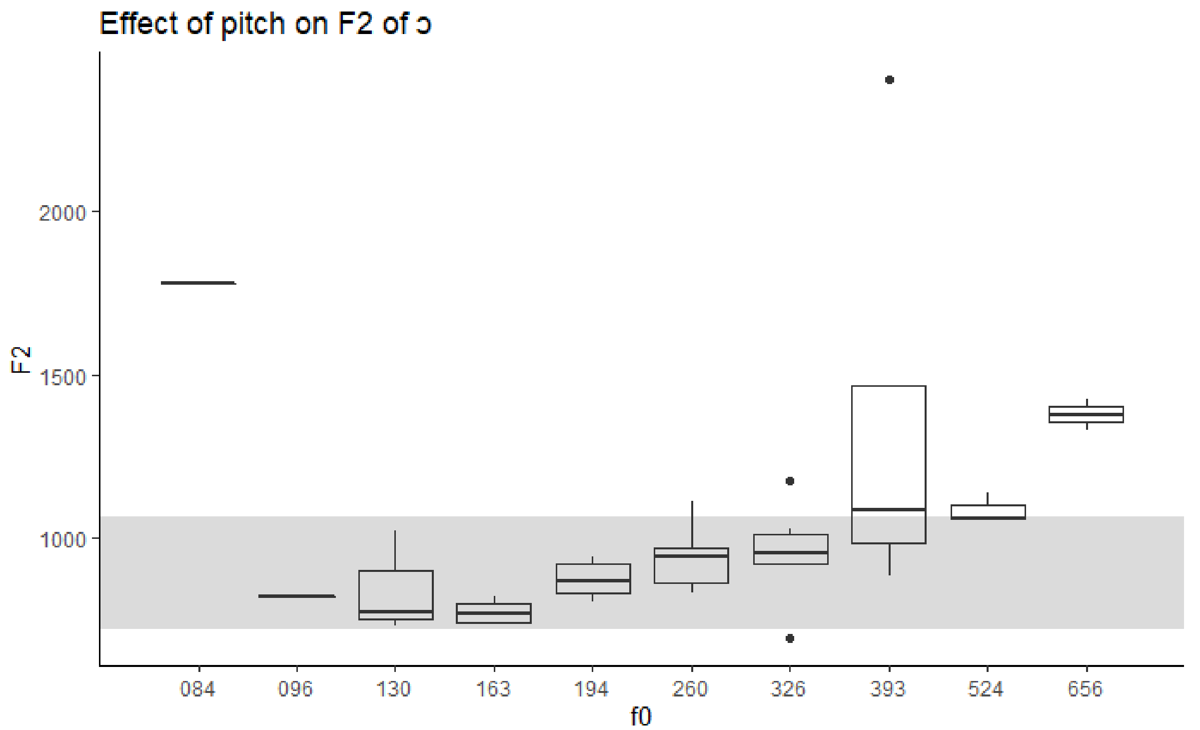

Effect of pitch on F2 of i
2400
2000
1600
F2
084
096
130
163
194
260
326
393
524
656
f0

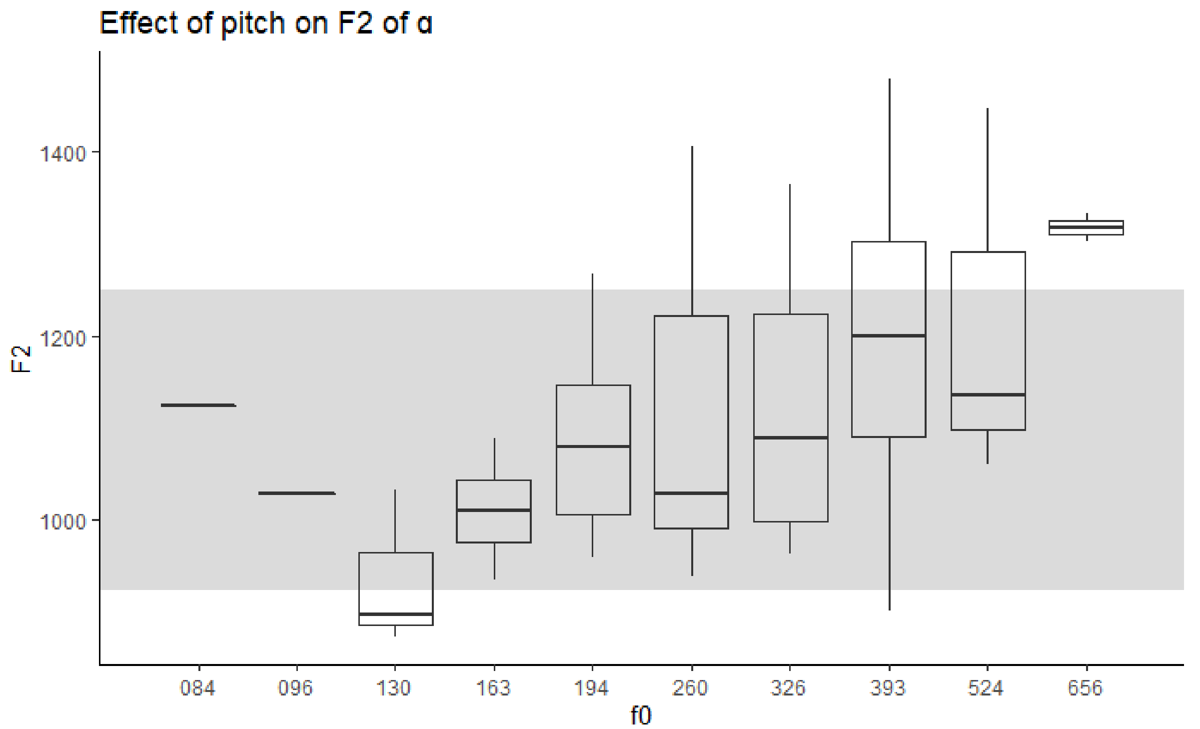

Effect of pitch on F2 of ɪ
F2
2200
2000
1800
1600
1400
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F2 of yː
F2
2000
1750
1500
1250
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F2 of ɣ
F2
1800
1600
1400
1200
1000
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F2 of ɛ
F2
2000
1800
1600
1400
1200
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F2 of ɔ
F2
2000
1500
1000
084
096
130
163
194
260
326
393
524
656
f0


Effect of pitch on F2 of ɑ
F2
1400
1200
1000
084
096
130
163
194
260
326
393
524
656
f0

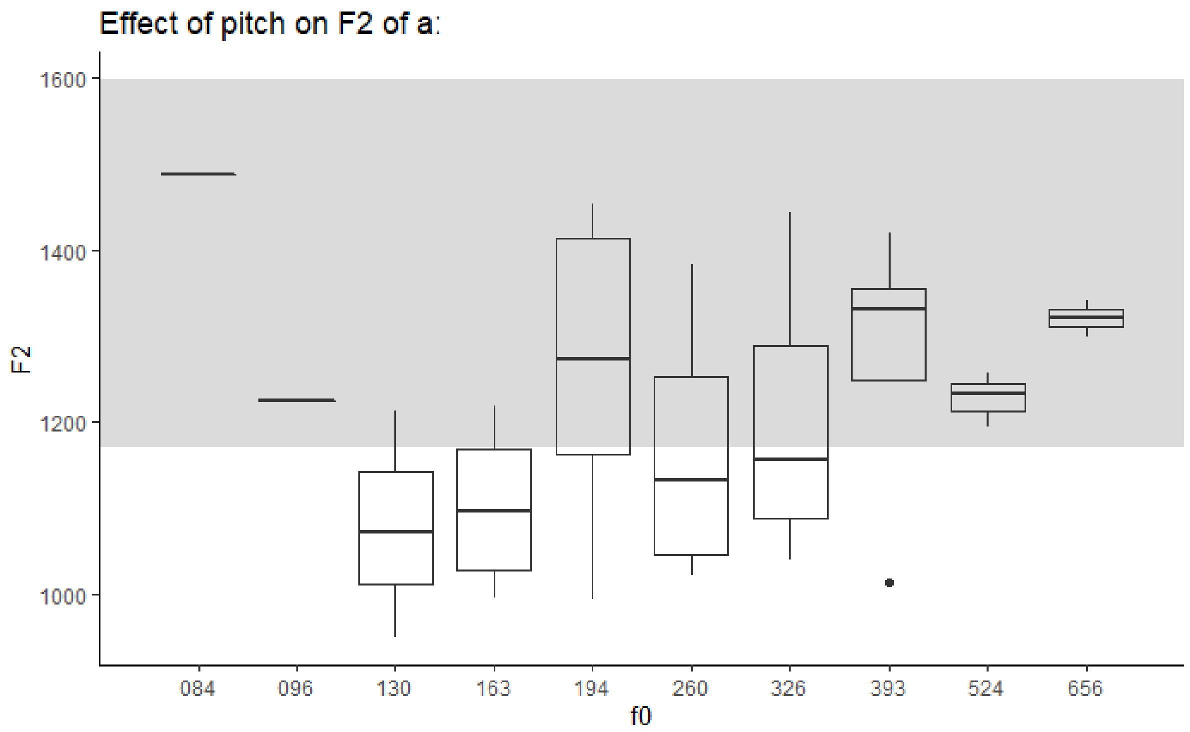


Figure 3: Effect of pitch on F2 in song.

A linear model with F1 as a dependent and pitch as an independent variable shows that pitch has a significant positive effect of 0.29 Hz/Hz on F1 in general ($p < 0.001$; 95% CI: [0.23 Hz/Hz, 0.35 Hz/Hz]). And a linear model with F2 as a dependent variable and mode as an independent variable shows that pitch has a significant negative effect of 0.33 Hz/Hz on F2 ($p < 0.001$, 95% CI: [0.16 Hz/Hz, 0.50 Hz/Hz]). Table 4 shows the effect of pitch on both the F1 and F2 per vowel. Per vowel two linear models were used. One with pitch as the independent and F1 as the dependent variable and one with pitch as the independent and F2 as the dependent variable. It gives the effect in Hz/Hz, the p-value, and the 95% confidence interval in Hz/Hz. When $0.05 > p > 0.01$ the result is marked in blue, when $0.01 > p > 0.001$ the result is marked in orange and when $0.001 > p$ the result is marked in red. It shows a significant rise in F1 among high and mid vowels, a significant lowering of F2 in front vowels and a significant rising in the F2 of [ɔ].

Table 4: Effect of Pitch

	F1			F2		
	Est. (Hz/Hz)	p-value	CI (Hz/Hz)	Est. (Hz/Hz)	p-value	CI (Hz/Hz)
u	0.40	< 0.001	[0.27, 0.53]	0.50	0.160	[-0.21, +1.21]
i	0.50	< 0.001	[0.38, 0.62]	-1.28	< 0.001	[-1.75, -0.81]
ɪ	0.43	< 0.001	[0.32, 0.55]	-0.88	< 0.001	[-1.13, -0.63]
y	0.53	< 0.001	[0.42, 0.64]	-0.81	< 0.001	[-1.10, -0.52]
ʏ	0.36	< 0.001	[0.23, 0.48]	-0.43	0.015	[-0.78, -0.09]
ɛ	0.28	< 0.001	[0.12, 0.44]	-0.73	< 0.001	[-0.98, -0.48]
ɔ	0.24	0.003	[0.08, 0.38]	0.78	0.035	[0.06, 1.51]
ɑ	-0.04	0.680	[-0.23, +0.15]	0.08	0.388	[-0.11, +0.27]
a	-0.08	0.435	[-0.28, +0.12]	-0.30	0.030	[-0.56, -0.03]

4.3 Other factors

4.3.1 Range

Two linear models were used. One with F1 as a dependent variable and one with F2 as a dependent variable, both used range as an independent variable. They show that range has a significant effect on F1 and F2. The bass has a lower F1 than the alto by 737.43 Hz ($p < 0.001$; 95% CI: [547.84 Hz, 927.03 Hz]) and F2 is 647.82 Hz lower for the bass ($p = 0.029$; 95% CI: [65.89 Hz, 1229.74 Hz]). Tenors have a lower F1 by 895.64 Hz ($p < 0.001$; 95% CI: [642,89 Hz, 1148.39 Hz]) and a lower F2 by 871.57 Hz ($p = 0.028$; 95% CI: [95.80 Hz, 1647.35 Hz]) than altos. Sopranos have a lower F1 by 372.19 Hz ($p < 0.001$; 95% CI:

[266.14 Hz, 478.24 Hz]) and a non significant lowering of F2 by 217.76 Hz ($p = 0.189$; 95% CI: [-543.26 Hz, +107.74Hz]).

4.3.2 Education and Experience

Two linear models were used. One with F1 as a dependent variable and one with F2 as a dependent variable, both had education and experience as independent variables. They showed that both education and experience have a significant effect on F1 and a non-significant effect on F2. Educated singers have an F1 that is lower by 866.55 Hz ($p < 0.001$; 95% CI: [592.35 Hz, 1140.74 Hz]) and F2 by 643.35 Hz ($p = 0.134$; 95% CI: [-1485.95 Hz, +198.25 Hz]). Experience lowers F1 by 23.37 Hz/year ($p < 0.001$; 95% CI: [16.18 Hz/year, 30.56 Hz/year]) and F2 by 17.34 Hz/year ($p = 0.123$; 95% CI: [-39.41 Hz/year, +4.73 Hz/year]).

5. Discussion

5.1 Mode and Pitch

For mode a significant negative effect on F2 was expected based on Bradley (2018). And a lowering of high vowels and rising of low vowels was expected in F1. A non-significant lowering of F2 was found, meaning no conclusions can be drawn. F1 was also found to lower. Of the vowels themselves a significant effect of mode was found mostly in F1. The [u], [i], [ɪ], [y] and [ʏ] all lowered significantly in F1. These vowels were expected to have a higher F1 when sung, as was found in Bradley (2018). This difference could be because pitch was not considered a factor in that research. Only [a], [u] and [ɛ] had a significant lowering of F2, matching up with Bradley (2018) found.

Based on previous research a significant effect of pitch on vowel space was expected (Bresch & Narayanan, 2010; Garnier et al., 2010; Henrich et al., 2011; Joliveau et al., 2004; Miller & Schutte, 1990; Neumann et al., 2005; Schutte et al., 2005; Sundberg, 1975; Sundberg & Skoog, 1997; Titze et al., 1994; Yu Chan & Do, 2021). And a significant effect was found. A gradual lowering of high vowels was expected and found as well. And a subtle fronting of back vowels. There are clear signs of formant tuning in the results of this experiment in all vowels except the ɑ, where no significant effects were found.

Previous research shows that range has a significant effect on both F1 and F2. This is also what was found in this research. However, this effect can be partly conflated with the effect of pitch. Sopranos will divert more from the spoken norm than basses and tenors, because they sing at higher pitches and thus have less choice of frequencies to tune their formants to. Basses will mostly follow the spoken norm.

Both experience and education have a significant effect on F1 and non-significant effect on F2. Probably because mode change has more effect on F1, than on F2 in general. It should also be taken into account when looking at the results that the sample population of this study has an unbalanced distribution of education and range. This and other limitations will be discussed further in the next section.

5.2 Limitations

One of the issues that could make the results less accurate is sample size. Only six participants were recruited, which may not be a large enough sample to draw conclusions. Although the participants do represent a range of ranges, both basses and altos are under-represented in the data. In the case of the bass, this also means that lower pitches had less data points and were less reliable when it comes to analysis and conclusions. The bass has also sung as a tenor most of his life, which may have caused him to apply tenor strategies to his singing despite his current range. There is also the fact that some of the participants had a smaller range than others. This makes some singers more dominant in the results. For more information into the amount of tokens per participant see appendix B.2.

Another issue is the difference in education between lower and higher ranges. The participants with a tenor or higher range had all received some kind of singing lessons, while one of the tenors and bass both had not. This may have caused a skewing of results regarding the effect of education and experience. This is especially the case because these effects are mostly observed among higher pitches and ranges. The fact that education has strong overlap with higher voice-types may have confounded the results. Knowledge of formant tuning could also play a role. Participants who have had singing lessons are more likely to know of the phenomenon.

The range of experience is also quite big, from the least experienced at 19 years of experience and the most experienced at 66 years. Of course the idea of singing experience itself is already a murky area, because of the nature and cultural prominence of singing. This issue has been partly solved by asking for experience when it comes to singing with the goal of performing. Although, again, this isn’t foolproof, seeing as this does not give room for breaks. Some of the participants may have only 10 years of singing ‘experience’ total, even though they started 60 years ago. The question asked was “when did you start?”. Accurate information on the exact number of years and intensity and how that affects vowel manipulation can thus not be answered here.

The issue of vocabulary could have had an effect on the results. One participant mentioned not knowing the word fuut (grebe), which could have affected their pronunciation of the word. Dutch is relatively phonetic in its orthography, but not completely. Some of the words are more used than others and thus better known and more likely to follow the intended pronunciation.

5.3 Future Research

Future research could look into expanding into other languages. The effect of both education and experience are also interesting questions that could be looked into and expanded on. Formant tuning at lower pitches and in lower ranges could also be an interesting topic.

6. Conclusion

This paper has looked into the vowel space of sung and, to a lesser extent spoken, Dutch. Past research had found a significant effect on vowel space based on both mode, singer range and pitch. Mode lowers F2 and compresses vowel space in general (Bradley, 2018). Singers with a higher range generally adapt their vowels more extremely than those with a lower range (Bradley, 2018; Garnier, 2010; Henrich et al., 2011; Joliveau, 2004) and higher pitches predict more extreme formant change than lower pitches (Garnier, 2010; Henrich et al., 2011; Joliveau, 2004). These phenomena were investigated by asking the participants to sing target words based on the nine monophthongs in Dutch. They were asked to sing these on ascending chords. On the notes of a c-major chord from the bottom to the top of their range. The results of this paper largely support previous findings, strengthening claims of change based on mode and the existence of formant tuning.

References

Boersma, P., & Weenink, D. (2021). *Praat: Doing Phonetics by Computer* (6.1.38) [Computer software]. https://www.fon.hum.uva.nl/praat/

Bradley, E. D. (2018). A comparison of the acoustic vowel spaces of speech and song. *Linguistic Research*, *35*(2), 381–394. Scopus. https://doi.org/10.17250/khisli.35.2.201806.006

Bresch, E., & Narayanan, S. (2010). Real-time magnetic resonance imaging investigation of resonance tuning in soprano singing. *Journal of the Acoustical Society of America*, *128*(5), EL335–EL341. Scopus. https://doi.org/10.1121/1.3499700

Garnier, M., Henrich, N., Smith, J., & Wolfe, J. (2010). Vocal tract adjustments in the high soprano range. *The Journal of the Acoustical Society of America*, *127*(6), 3771–3780. https://doi.org/10.1121/1.3419907

Henrich, N., Smith, J., & Wolfe, J. (2011). Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones. *Journal of the Acoustical Society of America*, *129*(2), 1024–1035. Scopus. https://doi.org/10.1121/1.3518766

Joliveau, E., Smith, J., & Wolfe, J. (2004). Vocal tract resonances in singing: The soprano voice. *The Journal of the Acoustical Society of America*, *116*(4), 2434–2439. https://doi.org/10.1121/1.1791717

Miller, D. G., & Schutte, H. K. (1990). Formant tuning in a professional baritone. *Journal of Voice*, *4*(3), 231–237. https://doi.org/10.1016/S0892-1997(05)80018-9

Neumann, K., Schunda, P., Hoth, S., & Euler, H. A. (2005). The interplay between glottis and vocal tract during the male passaggio. *Folia Phoniatrica et Logopaedica*, *57*(5–6), 308–327. https://doi.org/10.1159/000087084

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schutte, H. K., Miller, D. G., & Duijnstee, M. (2005). Resonance strategies revealed in recorded tenor high notes. *Folia Phoniatrica et Logopaedica*, *57*(5–6), 292–307. https://doi.org/10.1159/000087082

Sundberg, J. (1975). Formant technique in a professional female singer. *Acta Acustica United with Acustica*, *32*(2), 89–96.

Sundberg, J., & Skoog, J. (1997). Dependence of jaw opening on pitch and vowel in singers. *Journal of Voice*, *11*(3), 301–306. https://doi.org/10.1016/S0892-1997(97)80008-2

Titze, I. R., Mapes, S., & Story, B. (1994). Acoustics of the tenor high voice. *The Journal of the Acoustical Society of America*, *95*(2), 1133–1142. https://doi.org/10.1121/1.408461

Yu Chan, M. P., & Do, Y. (2021). Vowel modification (aggiustamento) in soprano voices. *Music and Science*, *4*. Scopus. https://doi.org/10.1177/20592043211055168

Appendix A: Questionnaire

- Hoe oud bent u?
  - How old are you?
- Hoe lang zingt u (met het doel om op te treden)?
  - How long have you been singing (with the intent to perform)?
- Heeft u zanglessen gehad?
  - Have you had singing lessons?
- Wat is uw dominante zangstijl?
  - What is your dominant singing style?
- Welke stem zingt u voornamelijk (bas, bariton, tenor, alt, mezzosopraan, sopraan, anders)?
  - What is your range (bass, baritone, tenor, alto, mezzo soprano, soprano, other)?

Appendix B: Stimuli Information

B.1: Note and Pitch Information

Note	Indicated Pitch (Hz)	Mean Recorded Pitch (Hz)
e2	82.4	84
g2	98	96
c3	130.8	130
e3	164.8	163
g3	196	194
c4	261.6	260
e4	329.6	326
g4	392	393
c5	523.3	524
e5	659.3	656

B.2: Token per Participant

P1 is a bass, P2 and P6 are sopranos, P3 and P4 are tenors and P5 is an alto

Note	P1	P2	P3	P4	P5	P6	Total
e2	✓						1
g2	✓						1
c3	✓		✓	✓			3
e3	✓		✓	✓	✓		4
g3	✓	✓	✓	✓	✓	✓	6
c4	✓	✓	✓	✓	✓	✓	6
e4	✓	✓	✓	✓	✓	✓	6
g4	✓	✓			✓	✓	4
c5		✓			✓	✓	3
e5		✓				✓	2
Total	8 (72)	6 (54)	5 (45)	5 (45)	6 (54)	6 (54)	

Appendix C: Praat scripts

C.1: From sound file to TextGrid

```
targets = Read from file: "targets.Strings"

ptcpNr$ = ""

folderPath$ = "/media/t/USB2/Thesis/Praat/sounds/" + ptcpNr$
createFolder: folderPath$

folder$ = "/media/t/USB2/Thesis/Praat/originalSounds/"

string = Create Strings as file list: "fileList", folder$ + "/*.WAV"
nrStrings = Get number of strings
for nr from 1 to nrStrings
   selectObject: string
   file$ = Get string: nr
   original = Read from file: folder$ + file$
   mono = Convert to mono
   selectObject: targets
   word$ = Get string: nr
   selectObject: mono
   Save as WAV file: folderPath$  + "/" + word$ + ptcpNr$ + ".wav"
   selectObject: mono
   textGrid = To TextGrid: "pitches", ""
   Save as text file: folderPath$  + "/" + word$ + ptcpNr$ + ".TextGrid"
   removeObject: original, mono, textGrid
endfor
removeObject: string, targets

editing = Create Strings as file list: "editing", folderPath$  + "/*"
nrEditing = Get number of strings
for ed to nrEditing
   selectObject: editing
```

```
   file$ = Get string: ed
   original = Read from file: folderPath$ + "/" + file$
endfor
removeObject: editing
```

C1.1 target.Strings

```
File type = "ooTextFile"
Object class = "Strings"

numberOfStrings = 9
strings []:
        strings [1] = "fiets"
        strings [2] = "fuut"
        strings [3] = "voet"
        strings [4] = "vaat"
        strings [5] = "fut"
        strings [6] = "vat"
        strings [7] = "vet"
        strings [8] = "vod"
        strings [9] = "fit"
```

C.2: From annotated TextGrid to .csv

```
participant =
participant$ = ""
gender =
range$ = ""
highPitch =
experience$ = ""
education$ = ""

directory$ = "sounds/" + participant$ + "/"
data$ = "data/" + participant$ + "/"

wav$ = participant$ + ".wav"
writeInfoLine:
"participant,experience_(years),education,range,word,vowel,vowel_(IPA),pitch,f0,F1,F2,stan
dard_deviation_F1,standard_deviation _F2"

files = Create Strings as file list: "fileList", directory$ + "*.wav"
noFiles = Get number of strings

for no to noFiles
  selectObject: files
  soundString$ = Get string: no
  gridString$ = soundString$ - ".wav" + ".TextGrid"
  sound = Read from file: directory$ + soundString$
  textGrid = Read from file: directory$ + gridString$
  noIntervals = Get number of intervals: 1
  selectObject: sound, textGrid
  Extract non-empty intervals: 1, "no"

  for interval to noIntervals
        selectObject: textGrid
        intervalLabel$ = Get label of interval: 1, interval
```

```
if intervalLabel$ = ""
elsif intervalLabel$ = " "
else
	selectObject: "Sound "+ intervalLabel$
	pitchOb = To Pitch: 0, 75, highPitch
	pitch = Get mean: 0, 0, "Hertz"
	removeObject: pitchOb

	selectObject: "Sound "+ intervalLabel$
	formantOb = To Formant (burg): 0, 5, gender, 0.025, 50
	for formant to 2
		selectObject: formantOb
		matrix = To Matrix: formant
		mean = Get mean: 0, 0, 0, 0
		sDev = Get standard deviation: 0, 0, 0, 0

		vowel = Read from file: "vowels.Strings"
		vowel$ = Get string: no
		vowelIPA = Read from file: "vowelsIPA.Strings"
		vowelIPA$ = Get string: no

		if formant = 1
			formant1 = round (mean)
			formant1SDev = round (sDev)
		elsif formant = 2
			formant2 = round (mean)
			formant2SDev = round (sDev)
		endif

		removeObject: matrix, vowel, vowelIPA
	endfor

	if intervalLabel$ = "spoken"
```

```
                    appendInfoLine: participant$, ",", experience$, ",", education$, ",",
range$, ",", soundString$ - wav$, ",", vowel$, ",", vowelIPA$, ",", intervalLabel$, ",spoken,",
formant1 , ",", formant2, ",", formant1SDev, ",", formant2SDev
            else
                    appendInfoLine: participant$, ",", experience$, ",", education$, ",",
range$, ",", soundString$ - wav$, ",", vowel$, ",", vowelIPA$, ",", intervalLabel$, ",", round
(pitch), ",", formant1 , ",", formant2, ",", formant1SDev, ",", formant2SDev
            endif

            removeObject: "Sound "+ intervalLabel$, formantOb
        endif

    endfor

    removeObject: sound, textGrid
endfor

removeObject: files
```

C.2.1 vowels.Strings

```
File type = "ooTextFile"
Object class = "Strings"

numberOfStrings = 9
strings []:
    strings [1] = "ie"
    strings [2] = "i"
    strings [3] = "u"
    strings [4] = "uu"
    strings [5] = "aa"
    strings [6] = "a"
    strings [7] = "e"
    strings [8] = "o"
    strings [9] = "oe"
```

C.2.2 vowelsIPA.Strings

```
File type = "ooTextFile"
Object class = "Strings"

numberOfStrings = 9
strings []:
    strings [1] = "i"
    strings [2] = "ɪ"
    strings [3] = "ʏ"
    strings [4] = "yː"
    strings [5] = "aː"
    strings [6] = "ɑ"
    strings [7] = "ɛ"
    strings [8] = "ɔ"
    strings [9] = "uː"
```