

Capstone project: Pitch change in Kazakh-Russian-English multilingual young adults

Aida Isteliyeva

SSH, Nazarbayev University

WLL 499: Languages, Linguistics, and Literatures Capstone II

Prof. Jenni Lehtinen

Advisor: Prof. Mayila Yakefu

May 3, 2023

Abstract

If you ever interacted with the same person in different languages, you might have noticed their voice sounding differently. This phenomenon is the object of this study about the change of pitch Kazakh-Russian-English trilinguals exhibit in speaking different languages. A physical representation of a human's pitch is fundamental frequency - the frequency with which vocal cords vibrate during speech. To see whether a change of pitch can be proved empirically, I conducted a series of recordings among 32 students of Nazarbayev University. This resulted in 192 recordings which were grouped depending on the gender and the first language of the speaker, as well as the elicitation method. The recordings were analyzed for each separate participant and in groupings using a t-test and linear mixed effect model. The participants also had to provide us with their self-assessed level of linguistic skills and their linguistic repertoire. The possible explanations for our results were provided by previous research in Bilingualism studies and Psycholinguistics. My study results in several findings regarding the interaction of different factors and the change of pitch. One of them is that the change of pitch seems to depend on the first language of the speaker. The change of pitch overall does not exhibit universal tendencies: separate analysis for each of the participants shows that change of pitch exists as it manifests in 28 out of 32 participants, however, there is no single reason or interaction that attests to all of the differences present in my data. Confidence level also shows a significant result for Kazakh speakers but not the others.

Introduction

A person's pitch is one of the linguistic features that is the easiest to follow and analyze in everyday life. Anyone can make a note of someone speaking in a higher or a lower pitch. One of the observations that I made about myself is that I speak with a much higher pitch when I talk in Kazakh - a language that I am supposed to know fluently yet I do not. A similar observation was often expressed to me by my peers. So this leads me to the topic of my capstone project which is the change of pitch in a multilingual. As such a study will answer my question of whether my observations are a common phenomenon, I am motivated to learn more about my topic.

Besides satisfying my curiosity, such a study will also help us learn more about the possible interaction between a person's pitch and their linguistic and extralinguistic situation which would add to the study of psycholinguistics which interests me. It also would add to the study on multilinguals as it may introduce the difference in phonetic behaviour depending on a language of a speaker.

The main analytic focus of my research is the speaker's fundamental frequency [FF] which is also commonly known as pitch. The fundamental frequency of a speaker is the frequency of vibration of vocal cords during vocalization (Reetz & Jongman, 2020). When we perceive someone as having a 'higher pitch' what it means is that their vocal cords vibrate with a higher frequency. A person whose voice is described as 'low' has vocal cords that for some reason or another vibrate with a lower frequency. The usual reason for such a difference can be the difference in physique or such destructive behaviours as drinking or smoking which are reported to affect the condition of the vocal cords.

One of my more scientific motivations for this research is that a study by Niebudek-Bogusz et al (2006) showed that fundamental frequency can be used to diagnose voice pathologies. So a person using different pitches depending on the language they are speaking could be a reason for misdiagnosis. Preventing such a case is a motivation for researching fundamental frequency in multilinguals.

This study is especially fruitful for Kazakhstan's linguistic environment because its educational system includes the obligatory acquisition of Kazakh, Russian, and English. This means that all of the people 30 years or younger have a relatively good fluency level in all three of the languages. The situation is especially so for Nazarbayev University where English is the language of instruction. The multilingual environment and constant use of all three languages among young people provide a situation in which the change of pitch is not merely an issue of acquaintance with the language as people get enough speaking practice.

Research questions

The main motivation behind this capstone project is built upon my past observations of people changing their pitch depending on the language they are using. I am interested to know whether this difference is personal or whether there can be some factors that influence a significant number of us. Following from that, my main research question "Does the language spoken by Kazakh-Russian-English multilingual young adults influence their pitch?" I have to specify the languages spoken by the multilinguals in question as I cannot generalize my findings to other combinations of languages that can be met in multilinguals around the world. If I find that the differences in pitch are personal rather than affected by some factor, I infer that these differences are a new parameter that needs to be considered in speech pathology research.

To come to the answer to the main research question, I will have to go through several secondary questions. As I am also interested in what factors affect the pitch, my question list is quite extensive and can be seen below:

1. Do Kazakh-Russian-English multilingual speakers change their pitch when speaking different languages?
2. Do different languages associate with different average fundamental frequencies?
3. What linguistic factors can possibly influence a person's pitch? Inquiry on the order of language acquisition and fluency.
4. What extralinguistic factors can possibly influence a person's pitch? Inquiry on the influence of gender and confidence level.

Literature review

The difference in pitch in monolinguals

One possible explanation for a difference in pitch in multilinguals is that languages may have different fundamental frequencies associated with them. A study by Wagner & Braun (2003) studied the difference in phonetic behaviour in German and Polish monolinguals. The study found that speakers of Polish on average had a higher fundamental frequency as compared to the frequencies of Italian and German speakers. What is important in this study is that the findings correspond to the qualities of the speakers expected by stereotypes: Germans were expected to sound harsher, and Polish speakers were expected to sound softer as associated with the cultural stereotypes connected with their languages. So it is possible here too that the multilinguals will speak Kazakh, Russian and English at different frequencies because of the fundamental frequencies associated with these languages.

The difference in pitch in multilinguals

The existing research shows that multilinguals speak at a different fundamental frequency depending on the language they are speaking. But it also shows that the presence of such a phenomenon depends on the languages under research. For example, a study by Schwab & Goldman (2016) researched bilingual speakers of three groups: English-French, English-German and French-German speakers. Their findings show that while there is a significant change in pitch in English-French speakers, there is no significant result for French-German speakers. A similar finding is present in research by Altenberg & Ferrand (2006). Their research had a similar method and investigated English/Russian and English/Cantonese bilinguals. Their results showed a significant difference in pitch for the English/Russian group with speakers using higher frequency in Russian than in English. However, the English/Cantonese group did not show a significant difference in their fundamental frequency.

Ordin & Mennen also investigated the change of FF, this time in English/Welsh bilinguals (2017). This study separately analyzed the results for female and male speakers and they found that the difference in pitch was present and consistent for the female participants while male participants either did not change their pitch or did so inconsistently. This result is what motivated me to investigate both female and male participants but do so separately.

Possible factors that can influence a person's pitch

Many studies investigated the effects of different factors on the pitch of a speaker. Some of them apply to monolinguals, and some to multilinguals. A study by Harrington, Palethorpe, & Watson (2007) and Nishio & Niimi (2008) show that age can influence the pitch of a speaker.

Research finds that to a point in life, a person speaks with a lower pitch the older they are. Another possible point similar to age is that pitch changes with the level of voice exhaustion (Niebudek-Bogusz et al, 2006). Such a point can be relevant in the application of fundamental frequency as a diagnostic tool as was mentioned in the introduction section.

As for linguistic factors, such parameter as language fluency also seems to influence a person's fundamental frequency. An article by Järvinen, Laukkanen, & Izdebski (2007) found that the more familiar the person is with a language, the lower the frequency with which they speak. These findings can also be associated with a person's confidence. Past research also shows that confidence can be a factor as to why a person talks with a particular pitch (Jiang & Pell, 2017). Research shows, that the more confident the person is, the lower they will talk. This confidence can come from the general confidence in self or the confidence in the information they are giving. These findings also indirectly agree with research done by (Zraick et al, 2006). They researched how a context of dialogue can affect the pitch with which a person speaks. They found out that such factors as the topic of the conversation, the setting and the interlocutor can affect with which pitch a person talks. The more familiar the setting and the closer the interlocutor is, the lower the fundamental frequency of a participant.

So to make a short list of possible factors that could influence pitch, there is age, voice exhaustion, confidence, setting, interlocutor, information, as well as the emotional state of a speaker (Ghiurcau et al, 2010) and gender stereotypes of the culture connected with the language spoken (Ohara, 1999). Most of these factors will be in one way or another controlled in my experiment explained in the *Methods* section.

Methods

The main focus of my research is the change in bilinguals' fundamental frequency depending on what language they are speaking. I will investigate whether the change happens at all and, if it does, try to find out which factors influence the change.

Subjects

The participants of the current study are volunteers among undergraduate students of Nazarbayev University of various years of study and majors. Their age ranges from 18 to 22 with the average being 19.92. This general section of the population was chosen for several reasons. The first one is that students of Nazarbayev University are mostly graduates of Kazakhstani schools and all of them are expected to be fluent in Kazakh, Russian, and English by the educational system of the country. This way we are minimizing the possible effect of fluency. Although it - participants' fluency - will still be one of the parameters of the analysis - we minimize the chances of participants having drastic differences in their language skills in the three aforementioned languages. The second reason is the expected states of participants' voices. As they are already young adults, we expect that they will have well-formed stable voices. But, as they are still young, the influence of ageing as well as the possible effect of self-destructive behaviours - like drinking and smoking - are expected to have minimal to no effect on the quality of their voices.

The participants were invited to participate in this study through a university-wide email that explained the research topic as well as the experiment procedure. All participants are either native or fluent speakers of Kazakh, Russian and English. This way, we will be able to observe the biggest number of language comparisons. Many of the participants also know other

languages but most are at the earliest stages of their studies. This factor will also be considered during the analysis part of the experiment, but it is expected that it will have no significant effect.

Overall, 81 participants filled out a questionnaire about their demographic and linguistic information. Only 46 of them continued to participate in the second step of this study which is a recording session. For the sake of statistics and due to the quality of the recorded material, the number of participants was reduced to 32. The subjects are now balanced across gender and their first language and can be divided into 4 groups of eight people: female Russian native speakers, male Russian native speakers, female Kazakh native speakers, and male Kazakh native speakers. As the amount of data across groups will be equal, we will be able to utilize the t-test to see whether the difference among groups is significant and whether we can point out a factor that influences the change of pitch. This division among groups will help us to test the possible influence of the first language and the gender of the participant as well as the interaction of the two factors. Table 1 below shows demographic information and the average acquisition history for each of the groups.

None of the participants chosen for the analysis were sick during or closely before the recording sessions. Some of the 81 participants reported having a speech disorder - their results were not included in the analysis due to the possible influence of the disorder on their pitch change. Some of the participants also had overexaggerated British accents when they recorded in English - these participants were not included in the analysis as well. More than half of female participants also reported an experience with vocal training or vocal therapy. This too can be one of the parameters that influenced the presence of voice change and will be analyzed as an additional factor. However, as only two of the male participants mentioned receiving vocal training, this parameter can only be analyzed for the female portion of the participants.

Table 1. Information on demographics and linguistic acquisition of the groups of participants.

Group	Age	Average age	Dominant language	AoA of Kazakh	AoA of Russian	AoA of English	English proficiency
8 Men L1-Kazakh	18-22	20.5	6 - Kazakh 2 - Russian	L1 - 3.44	L2 - 7.75	L3 - 15.13	7
8 Women L1-Kazakh	18-21	19	5 - Kazakh 3 - Russian	L1 - 3.19	L2 - 5.25	L3 - 13.25	7.19
8 Men L1-Russian	19-22	20.63	2 - Kazakh 6 - Russian	L2 - 8.25	L1 - 3.86	L3 - 14	6.38
8 Women L1-Russian	18-22	20.38	2 - Kazakh 6 - Russian	L2 - 8.25	L1 - 2.86	L3 - 12.5	6.94

Materials

I am conducting an experiment for which I am doing recording sessions with each participant separately. The sessions consist of two parts. One of them is a reading task in which they need to read a paragraph in all three languages. The order of the languages is different depending on what is the first language of the participant - text in their first language is never first. The texts are a description of the National Museum of the Republic of Kazakhstan - the text is provided in Appendix 2. They do not hold any content that should invoke emotions in the participants. A study by Ghiurcau et al (2010) shows that the emotional state of a person and the emotion that they are trying to express through speech will affect what pitch they are going to use. By maintaining the emotion invoked by the questions in the same spectrum, we minimize the possible effect it can have on a person's pitch. The other part is a speaking task. Participants have to answer in the language in which I am asking the question. They are asked to speak for at least 10 seconds. The questions are the same for all recording sessions but differ from language to language. The questions are:

1. Which courses are you taking this semester? (in English)
2. What did you do this summer? (in Kazakh)
3. Could you please describe your room? (in Russian)

The recordings are needed to have material to analyse. The only thing that will be obtained from the analysis is the fundamental frequency of speech. The speech elicitation through questions is chosen because this will result in spontaneous speech which should yield the best results. The reading task is chosen to minimise the effect on the confidence level of the participants. Sometimes people struggle to talk when they do not really have an idea of what to say. With a ready text, they do not have to think about what to talk about. Another reason to choose more than one task was a finding by Zraick et al (2000) which shows that female participants often perform differently depending on the presented task. By finding an average of the two tasks I can obtain a result closer to reality.

For a deeper analysis of the data, I am also conducting a questionnaire. Through it, I can obtain data on participants' self-evaluated speaking skills in all three languages. I decided to use self-evaluation as Kazakh and Russian do not really have standard exams to evaluate one's speaking skills. Besides that, I also learn about their language use and obtain demographic information. The questionnaire consists of 17 questions and can be found in a link provided in Appendix 1.

The questions about the first language and dominant languages now will help evaluate their effect on a person's fundamental pitch. One possibility is that a person will talk in their lowest pitch in their first language. Or they can talk in their highest pitch in the language used in their school because they are used to talking in it to superiors (teachers). The question about their self-evaluation will give information on how confidence and acquisition degree might affect

participants' FF. One possibility is that person will talk in a higher voice in a language in which they have less confidence - evaluate their speaking with a lower score.

Procedure

Participants fill out the questionnaire online using Google Forms. After that, we will schedule a meeting. The recordings are conducted in study rooms on campus. The rooms are usually of the size to accommodate 2-4 people. During a recording, the door is closed and the only people in the room are the participant and the researcher (me). An article by Zraick et al (2006) made a conclusion that the interlocutor can also affect how a person talks. A person of higher social status or of an older age may urge a person to talk at a higher pitch. As my participants are university students, it is optimal that I am doing the recordings with them, as I am too a student. Before the recording participants are reminded that all information is 100% confidential and that, if they ever want to be pulled out of the experiment, that can be done. Then they are presented with the details of how the recording is going to go on. The tasks for the recording session were described in the Materials part. The two tasks come in random order.

The recordings will later be analyzed using software called Praat (Styler, 2013). For each recording, the fundamental frequency will be calculated as an average of 10 samples. The samples will be collected from 2-3 places in the recording. The places have to be in a middle of an utterance to minimize the possible effect of intonation change at the beginning and the end of a phrase. Each sample will be of an approximate length of 0.5 seconds. The samples from a similar place will overlap by 0.2 seconds. The software will collect information on the fundamental frequency of each sample and then I will calculate their average, excluding complete outliers.

Analysis methods

After everything is recorded, all of the recordings will be analyzed using software called Praat (Styler, 2013). This software was created for analyzing speech recordings for most acoustic parameters: pitch, formants, intensity and pulses. Using this software, I will be able to calculate the average fundamental frequency or pitch of a speaker in each recording. I will be doing so by first extracting the frequency of pitch from 0.5-second samples of the recording. The samples will be overlapping by approximately 0.2 seconds, so that there are minimal sharp increases or decreases in data. I will be calculating the average based on 10 samples of each recording. The samples will be collected from the middle of uninterrupted utterances. I plan to do so because people have a tendency to speak at a higher pitch at the beginning of their utterances and to lower their pitch at the end (Gussenhoven, 2002). By analyzing only the middles, I will minimize the number of outlier-frequencies inside of the recordings which would have influenced the average. The data included outliers that had a much higher or a much lower voice than other participants. These participants were not included in the further statistical analysis.

Then, the fundamental frequencies and their averages will be analyzed in several ways. The first way will be surface-level analysis. There, I will be looking at the interactions of results for each participant separately. I will be looking at their average fundamental frequencies across all six recordings and their answers to the questionnaire. This way I will on a surface level look for possible interactions between factors like first language, dominant language and gender, and their difference in pitch. I will also make a statistical analysis for each separate participant. In it, I will be using a t-test to make six pairwise comparisons. I will compare pairings of languages

for each elicitation method separately. So I will have the following comparisons of the recordings of a single participant:

- Reading in English and reading in Kazakh;
- Reading in English and reading in Russian;
- Reading in Kazakh and reading in Russian;
- Speaking in English and speaking in Kazakh;
- Speaking in English and speaking in Russian;
- Speaking in Kazakh and speaking in Russian.

This will provide me with statistical evidence on whether or not the difference in pitch is statistically different or not. After that, I will be making group-wise analyses. The data of all participants of a group will be compiled together to make a total of 80 points of data for each of the four groups mentioned in the Participants section. The data then will be analyzed using a linear mixed-effects model using lme4 (Bates et al., 2014). That way, we will be able to statistically infer whether gender, first language, and dominant language are significant factors influencing differences in pitch. If the situation is such that none of these factors are proven to be significant, then the difference in pitch is either individual or comes from a factor outside of these ones.

Another factor that will be present in the analysis is speaking skills. It was assessed by the participants themselves and can show three factors: their real speaking skills, familiarity with the language and confidence level. As all these factors are interconnected and proven to be influential by past research (Jarvinen et al., 2007; Jiang & Pell, 2017), we cannot ignore them. However, we also cannot put too much weight on them or separate them from one another as they represent participants' own evaluation of their skills. Having a professional analyze their

speaking skills would be out of the possibilities of this project and existing examinations of speaking skills cannot be performed by a non-professional like myself. Thus, their speaking skill will be analyzed for being influential together with the possible effect of confidence level.

In addition, I will also look into comparing my data to the data of existing research on fundamental frequency in monolingual native speakers of English and Russian. This inquiry is to see whether the simple fact that these are different languages is what makes the pitch different, as it is possible that the pitch range is somehow intrinsic to the acquired language.

Coding

All of the recordings are coded by what is the content of the recording. For example, “Speech in Kazakh” is coded as [SK]; “Reading in English” is coded as [RE]. In the case of obtained fundamental frequencies, each individual will be categorized into either [CH] or [UNCH] for “changed” and “unchanged” fundamental frequency of their voice depending on the language. Besides that, in further analysis of the data in synthesis with the answers to the questionnaire, it is possible for other codings to arise. For example, the gender of the participants, where they come from, and whether they had vocal training. For example, in some of the graphs, you will see notions like [ENG] but also [EN]. These were chosen for the metadata of the data wherein three-letter abbreviations were used in describing the first language of the participant and the two-letter abbreviations were used for the language of the recording. I will be adding commentaries on all abbreviations present in the graphs in their titles.

Results

Demographic information of the participants

Figures 1-3 represent some of the demographic outcomes of the questionnaire part of the experiment. The outcomes do not really contribute to the analysis of this project. Figure 3 shows other languages known by the participants. The list is varying and does not provide anything for the scope of this project besides the presence of languages outside of the three languages under investigation. The statistical analysis did not show this parameter to be a significant variable for the change of pitch. Figures 1 and 2 may show a significant effect on the study of the pitch change, however, the number of participants for each of the categories is not equal so it cannot be determined whether the effect is true rather than random.

For example, participants' place of birth yielded a significant result. ANOVA test showed that for the female population, the significance of the place of birth was 0.006, with its interaction with first language resulting in $p\text{-value}=0.01$. A closer look at the interaction shows that the participants from North and East Kazakhstan had a significantly higher FF than those from the other regions.

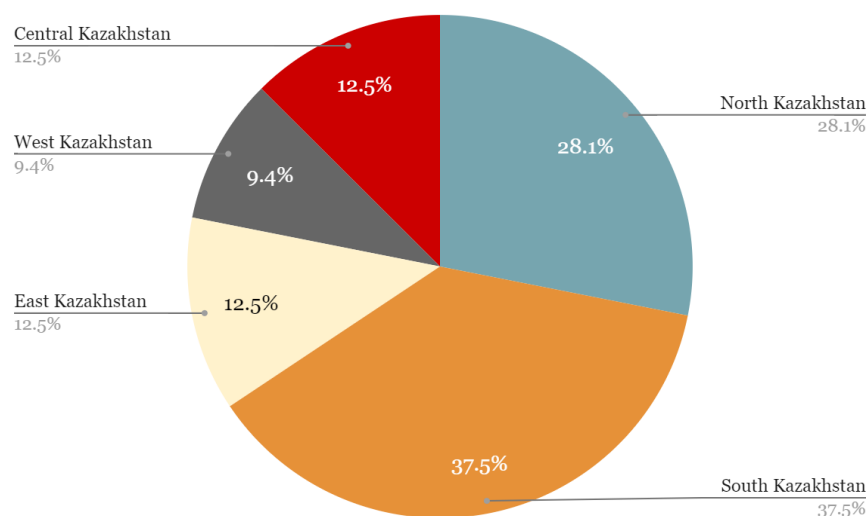


Figure 1. Participants' place of birth

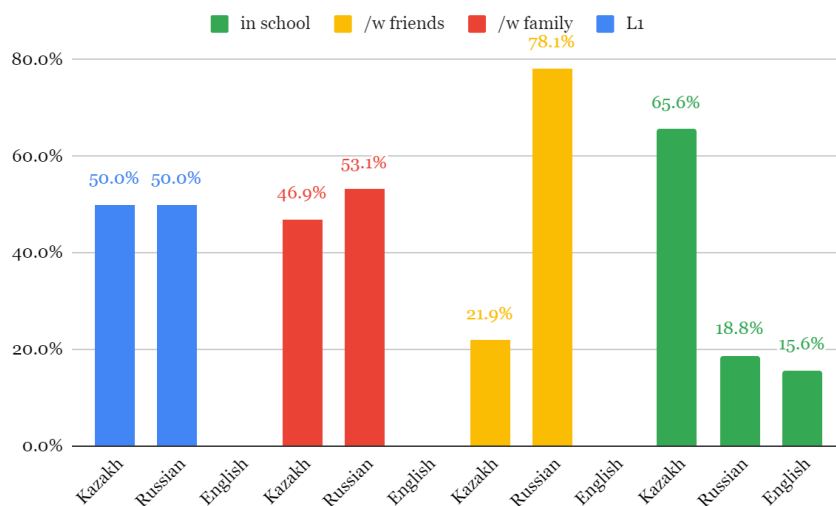


Figure 2. Participants' dominant languages in different linguistic domains

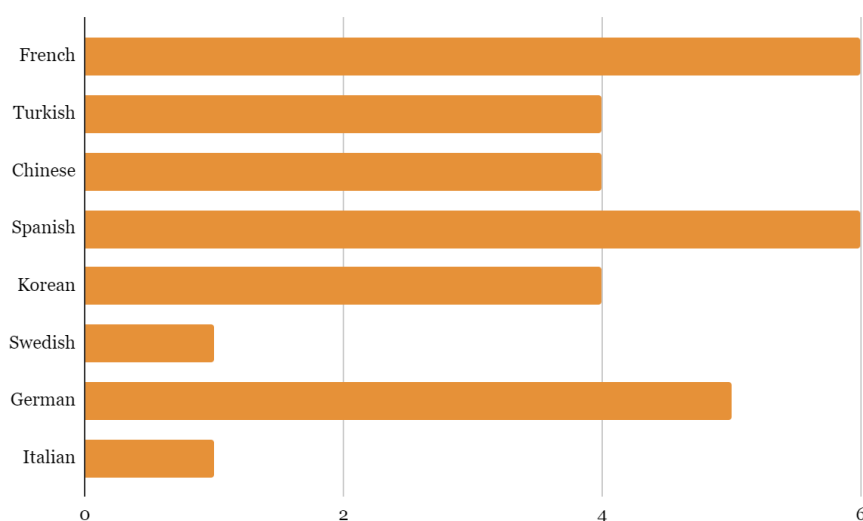


Figure 3. Other languages known by the participants

Association of particular fundamental frequencies with particular languages

Figure 4 and Figure 5 present boxplots showing how the language of the recording affected the average pitch of the participants. As we can see, the range of the data does not vary significantly across the three languages of the recording for both genders. I also calculated the mean values for each of the boxplots. For the male participants, the average fundamental

frequencies are: 108.16 for English, 109.12 for Kazakh, 107.12 for Russian. For the female participants, they are: 219.07 for English, 217.69 for Kazakh, and 217.09 for Russian.

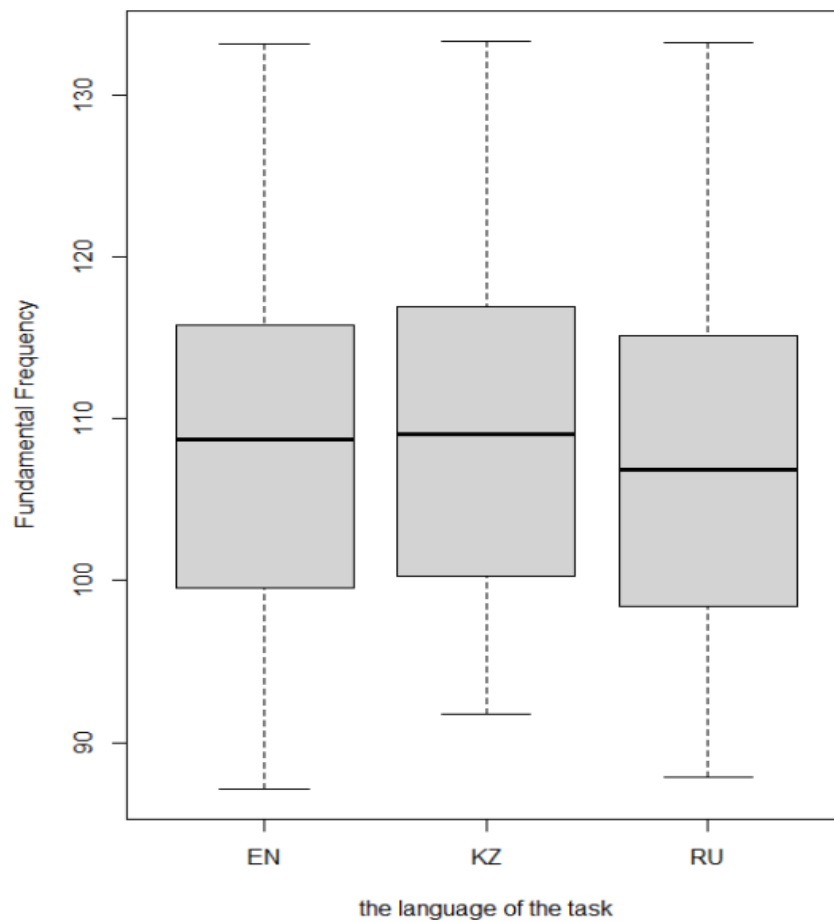


Figure 4. Fundamental frequencies of male participants in different languages of recording. EN, KZ and RU stand for the language of the recording being English, Kazakh and Russian respectively.

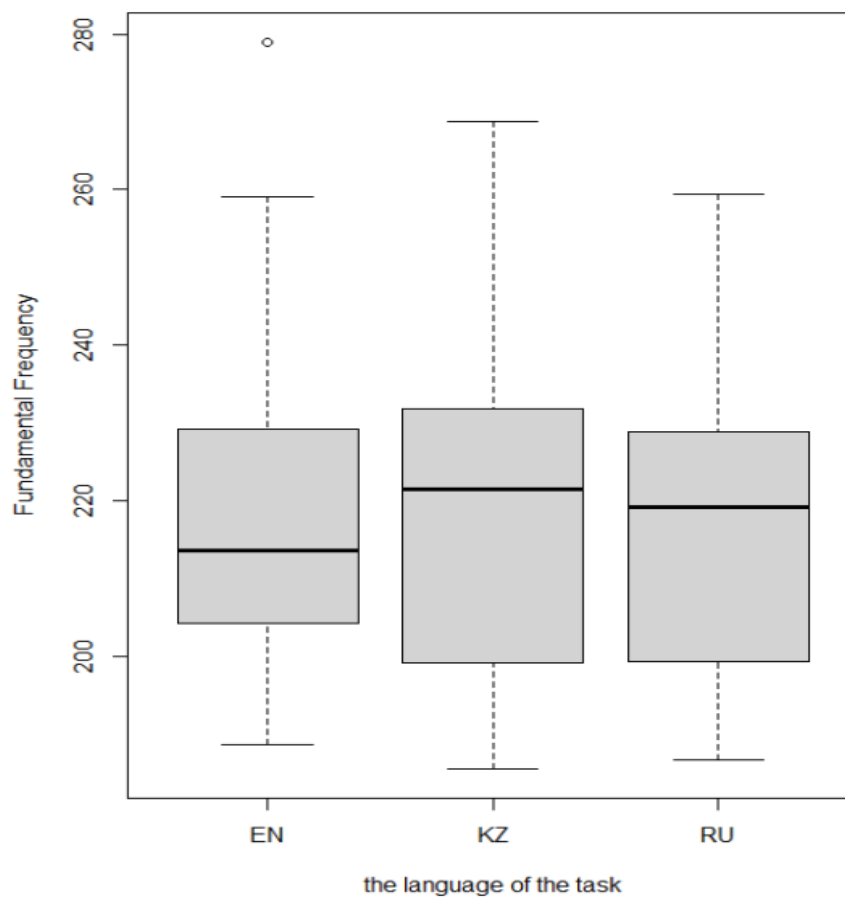


Figure 5. Fundamental frequencies of female participants in different languages of recording. EN, KZ and RU stand for the language of the recording being English, Kazakh and Russian respectively

The influence of linguistic repertoire on a person's pitch

Figures 6 and 7 below show the box plots for the fundamental frequencies different speakers produced during the recordings. Figure 6 represents the data for the male participants and Figure 7 represents the data for the female participants. The parameters by which the data was divided are the first language of the speaker (three-letter abbreviations) and the language of the recordings.

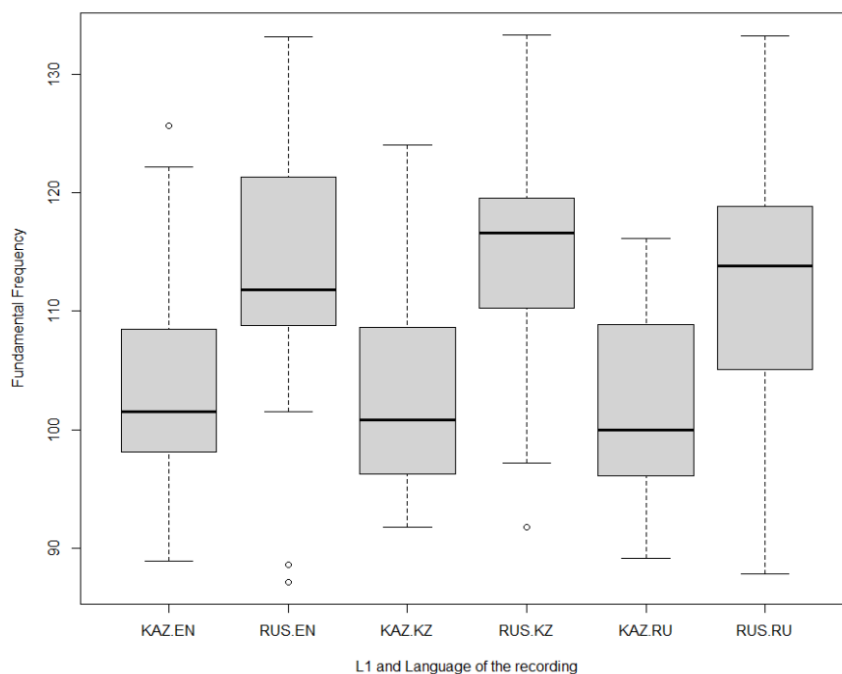


Figure 6. Fundamental frequencies across speakers with different first language - Male version. Three-letter abbreviations stand for the first language of the speaker; two-letter abbreviations stand for the language of the recording

As we can see from Figure 4, there is a clear difference between the speakers depending on their first language: men with Russian as their first language exhibited a much higher fundamental frequency across all three languages of the recordings than the men with Kazakh as their native language. The t-test shows that the difference is significant ($p\text{-value} < 0.001$). ANOVA test analyzing the significance of the first language shows a $p\text{-value}$ of 0.07.

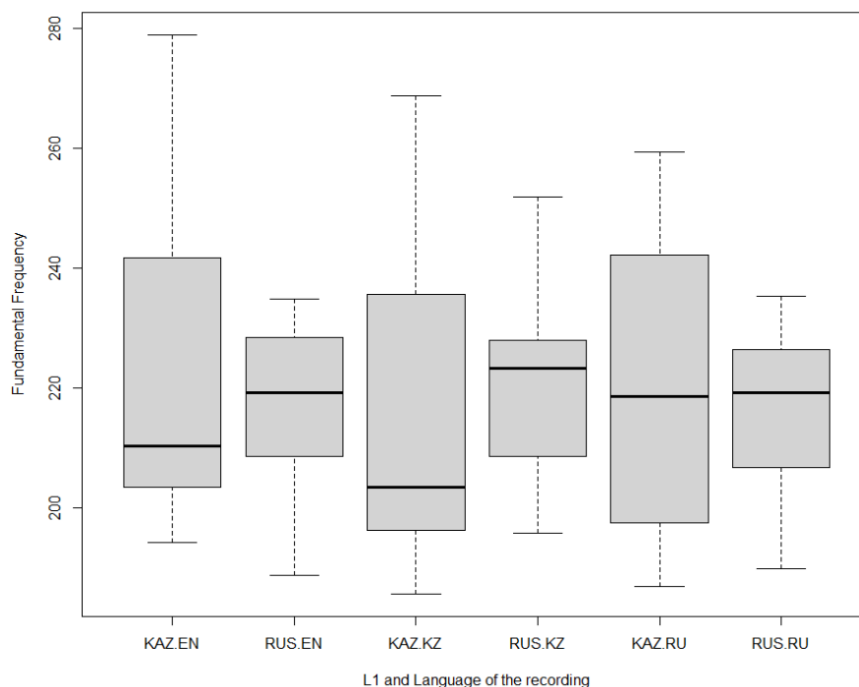


Figure 7. Fundamental frequencies across speakers with different first language - Female version. Three-letter abbreviations stand for the first language of the speaker; two-letter abbreviations stand for the language of the recording.

Figure 4 shows that there is no significant difference between the means of fundamental frequencies across the participants with different first languages or the recordings in different languages. One thing needs to be noted and it is that the participants with Kazakh as their native language had a much larger interquartile range than those of the Russian language. The t-test also shows that the difference among women with different first languages can be random ($p\text{-value} > 0.8$).

I was unable to collect data for reliable statistical research on whether the dominant language is what influenced the fundamental frequency due to the difference between the first language of the participant and the languages they most frequently use in different domains. The results that do exist, when ran through ANOVA, show little significance: $p\text{-value}$ for the influence of the language used with family = 0.67; $p\text{-value}$ for the language used with friends = 0.83; $p\text{-value}$ for the language used in school = 0.76.

The influence of extra-linguistic factors on a person's pitch

Vocal training: As was mentioned in the Methods section, half of the female participants reported receiving vocal training. The period of training varied from 3 months to 6 years so the data is not ideally reliable, however, all of the participants that reported receiving vocal lessons also reported that they sing on a daily basis. The results can be seen in Figures 8 and 9 below:

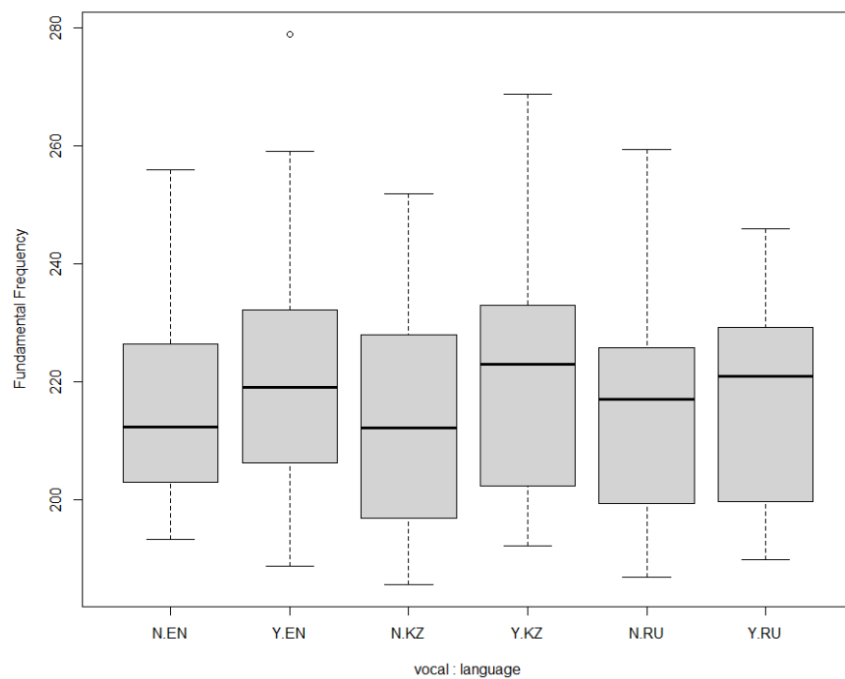


Figure 8. Fundamental frequencies of participants who did and did not receive vocal training across languages of recordings. Y and N stand for the presence or the absence of the vocal training experience; EN, KZ and RU stand for the language of the recording.

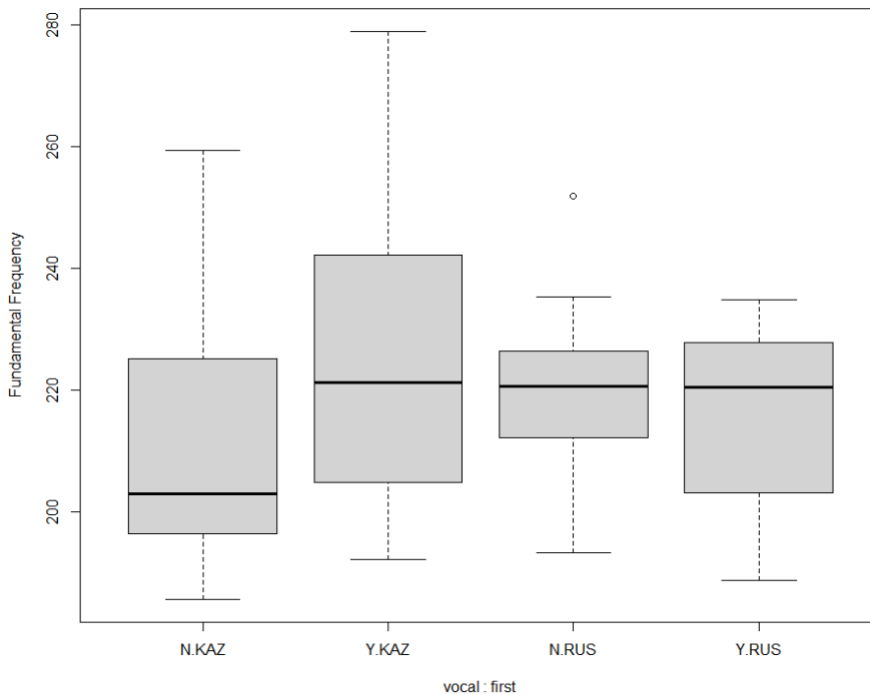


Figure 9. Fundamental frequencies of participants who did and did not receive vocal training across their native languages. Y and N stand for the presence or the absence of the vocal training experience; KAZ and RUS stand for native language of the participants.

The box plots show that across both the languages of the recordings and the native languages, the presence of vocal training resulted in a higher fundamental frequency. One important point is that the only relatively statistically significant parameter was the vocal training's influence on native Kazakh speakers (p -value=0.06). Statistical analysis also shows that the participants with vocal training had a larger difference in their pitch depending on the language they spoke. Meaning that a Kazakh native woman with some experience with vocal training had a larger difference in pitch across different languages than a woman with no experience with vocal training.

Confidence level: Another test considered the influence of participants' self-reported confidence levels. A surface-level analysis shows that 16 participants who reported differing levels of self-confidence followed this tendency: the lower the confidence level the higher the pitch in which they speak that language. At the same time, 10 participants presented the opposite

tendency wherein they spoke at a lower pitch if they were less confident in their level. ANOVA tests for each of the languages showed that only the Kazakh language yielded significant results: the lower the confidence level, the higher the voice. ANOVA showed a p-value of 0.046 for reading tasks and 0.008 for speaking tasks. All of the other tests did not yield a p-value lower than 0.37.

Individual differences

The data for each separate participant was analyzed across different recordings using a t-test. I looked into whether the difference of the pitch in different languages was significant and I accepted any p-value below 0.05 as significant. Out of the 32 participants, six showed no significant difference among all of the recordings: two men and two women with Kazakh as their first language and one man and one woman with Russian as their first language. As this phenomenon was present pretty equally among the four groups of participants, I do not find this significant for the study. 81% of the participants (26) showed a significant change of pitch depending on which language they were speaking. Overall, 56 comparisons yielded a significant difference. Table 2 below shows the p-values of the t-tests of the comparisons.

Table 2. Results of a pair-wise t-test comparison of the six recordings for each participant; cells in green represent tests with p-value < 0.05; cells in blue represent tests with p-value between 0.05 and 0.1.

№	RE/RK p-value	RE/RR p-value	RK/RR p-value	SE/SK p-value	SE/SR p-value	SK/SR p-value
Men; L1 - Kazakh						
1	0.7527	0.0149	0.0107	0.0903	0.0950	0.9608
2	0.0018	0.1961	0.1994	0.5742	0.0379	0.1702
3	0.2363	0.1527	0.9927	0.3819	0.2046	0.3665
4	0.000002	0.00008	0.1224	0.8659	0.8436	0.6999
5	0.0278	0.5947	0.0252	0.0150	0.0680	0.1757
6	0.0983	0.0851	0.9961	0.2799	0.5748	0.1906
7	0.3434	0.0107	0.1209	0.5103	0.1016	0.1546
8	0.0130	0.9620	0.0038	0.7112	0.2298	0.0982

Women; L1 - Kazakh						
9	0.8928	0.1488	0.0003	0.3823	0.0169	0.0695
10	0.1165	0.2616	0.7637	0.0189	0.5121	0.0429
11	0.0254	0.2029	0.0136	0.1808	0.0095	0.0017
12	0.0769	0.0918	0.0002	0.0061	0.8988	0.0141
13	0.7725	0.7831	0.5683	0.0785	0.1715	0.5115
14	0.6634	0.6691	0.4891	0.8460	0.8515	0.5960
15	0.6560	0.3408	0.5785	0.0104	0.0054	0.6715
16	0.1905	0.5794	0.5076	0.0031	0.0199	0.7567
Men; L1 - Russian						
17	0.8595	0.3273	0.5206	0.0484	0.6606	0.0512
18	0.3957	0.5488	0.5743	0.2517	0.8410	0.0188
19	0.6378	0.4618	0.9599	0.1903	0.4067	0.9649
20	0.00004	0.0913	0.00009	0.0587	0.1576	0.0076
21	0.9474	0.9648	0.9291	0.0479	0.0435	0.6643
22	0.2118	0.2769	0.8496	0.4702	0.0128	0.0042
23	0.6373	0.9720	0.7036	0.1094	0.1606	0.8070
24	0.1060	0.5088	0.0842	0.0168	0.6121	0.0062
Women; L1 - Russian						
25	0.00004	0.0137	0.0736	0.4109	0.5075	0.7456
26	0.3118	0.7826	0.3399	0.4585	0.0888	0.1028
27	0.2679	0.3492	0.8419	0.7388	0.0545	0.0563
28	0.0382	0.0533	0.0034	0.7368	0.5814	0.7399
29	0.1211	0.0106	0.0001	0.0420	0.0676	0.1675
30	0.5764	0.0425	0.3147	0.3422	0.8682	0.2745
31	0.2732	0.3092	0.6375	0.5068	0.0003	0.0167
32	0.0099	0.0014	0.9079	0.1647	0.5145	0.2413

Discussion

My study exists to answer the question of whether the pitch of a person changes when they speak different languages, and if it does, is there a reliable way to claim what the change comes from. Answering the second secondary question of my study, the data does not support that particular fundamental frequencies are associated with particular languages. As seen from our results, when we changed the languages of the tasks, the average fundamental frequency and

its range stayed approximately the same. Figures 4 and 5 show that data for different languages of recording show similar results. The answer to the second question would have been positive if the boxplots significantly differed in their placements.

Figures 6 and 7 show that the factor that does seem to be affecting the pitch level overall is the first language. The male participants who have Kazakh as their L1 have a significantly lower voice than those with Russian as their L1. At the same time, this difference does not significantly affect the change of pitch in either group. I also need to mention that FF does not seem to be particular for the Russian or Kazakh language. I make such a conclusion from the following logic: Speakers with Kazakh as their L1 spoke Russian at a lower pitch than speakers with Russian as L1. If the pitch came as a particularity of a language, then, considering that Russian natives spoke Russian with a higher pitch than Kazakh natives spoke Kazakh, Kazakh natives should have spoken at a higher pitch when speaking Russian. Such a trend is not present - or, at least, is not consistent. So FF is not a feature of a language.

Female participants did not show the same tendency in general but we will also see the influence of L1 in discussion of the effect of the presence of vocal training. Person-by-person analysis - the results of which are present in Table 1 - shows that there is a significant change in pitch. This in conjunction with the results in Figures 4-5 show that the change of pitch is present but is not universal for these groups of multilinguals. The trend is mostly that the language that the person is more acquainted with is the language in which the person speaks at a lower pitch. However, this trend is not universal as well, as many participants spoke at a lower pitch in a less familiar language. My understanding of this phenomenon is that when people spoke in a less familiar language, to appear more fluent, they paid more attention to

what they were speaking thus speaking with more control: at a lower pace and with a lower pitch.

Vocal training is the first extra-linguistic factor analyzed in this study. It was only applied to the female groups of the participants and showed a significant result in two regards. The first one is that vocal training affected the participants with Kazakh as their first language but not those with Russian as their L1. It affected in a way that those with training had a higher average pitch than those who did not. The second one is that those affected by training had a larger pitch difference across the languages.

Confidence level is another factor that was investigated in this study. As was mentioned in the results, only the Kazakh language showed a significant relation with the confidence level of the participants. Specifically, the lower the level of confidence in Kazakh, the higher the pitch with which participants spoke it. Confidence level did not elicit a similar behaviour in other languages.

The results mostly agree with findings in previous works in that there was a difference in pitch when a person spoke a different language. At the same time, the experiments showed no stable significant difference in pitch across participants which does not agree with other findings. In other works, for example, if participants showed differences in pitch, they did so consistently: they spoke one language at a higher pitch and another at a lower pitch. In our findings, from participant to participant, we saw different changes in pitch regardless of their L1. As for the cultural differences proposed by Ohara (1999), we could not make a judgement regarding the possible effect of Kazakhstani culture on the pitch of people of different genders. This came from the fact that the participants came from different regions of Kazakhstan all of which have a varying degree of being influenced by Russian/European culture.

Conclusion

In this project, I attempted to learn whether multilinguals changed their pitch depending on the language they were speaking and if they did so, what affects that change. My results showed that the change of pitch is present in multilinguals but that it is neither constant nor is there a singular significant cause for the change. The change varied with each participant. The work analyzed the possible influence of such factors as a person's first language, the language of utterance, the confidence level, and the fluency level in English. We also analyzed such factors as place of birth and presence of vocal training experience. All these factors showed to be unreliable determiners of what the change of pitch of a person will be like.

We made a definite conclusion regarding Russian and Kazakh that a particular FF is not a feature associated with a particular language. As we saw in our results when Kazakh natives spoke Russian, they did not exhibit much of a higher pitch even though Russian natives spoke Russian in a higher pitch. The same applies to Kazakh. Therefore, FF is not an acoustic feature particular to a specific language, rather the change varies depending on the speaker.

In general, we saw that L1 affects the pitch of a person overall rather than the change of pitch: Kazakh natives speak with a lower pitch in all languages - the average magnitude of change does not depend on L1.

As for extralinguistic differences, only the presence of vocal training showed a relatively significant effect on the participants but did so only for women with Kazakh as their first language. It needs to be noted here, that male groups were not analyzed for the effect of vocal training.

This study contributes to the studies of psycholinguistics, multilingualism and the effects of vocal training. It introduces the existing variability of the change of pitch in multilinguals,

showing that professionals have to account for the possible effect of vocal training, L1 and confidence level as well as fluency level when using FF in the diagnosis of speech disorder.

Limitations

There are several points on limitations to this research. One comes from the procedure of the experiment. Namely that there is a possibility of outside noise interfering with the extraction of pitch from the recordings. One possible solution is to make the recordings in a soundproof room with a microphone that is more professional and would not pick up such interference as much. However, this would make the environment of the procedure much more scientific and would remove the natural aspect of the recording. I hypothesize that such a change of setting would change the obtained data, as the participants would feel more pressure in an environment like that and their speech might be much more controlled than what it would be in a natural setting. Another point of limitation is the mains hum - the fundamental frequency produced by the alternating current in electronic technology (Gorman, 2010). Such interference can be removed by editing the audio files before the analysis, however, we are not sure whether this frequency interferes with our readings at all.

Another point comes from the participants. One is that they all come from different parts of Kazakhstan. Kazakhstan has a variety of dialects, each of which has its own acoustic specifics recognizable by native speakers. The statistical analysis of our data showed the place of birth to be a significant influence on the pitch of a person. People from Northern and Eastern Kazakhstan showed a much higher average fundamental frequency and greater inter-language differences in pitch. I, however, cannot make any significant conclusions regarding this issue, as the number of participants from each region is varying and for some regions, the number is only three. I cannot

make definite conclusions on whether the place of birth, upbringing traditions or dialect is at issue. Besides, the factor of dialect might have interfered with the general analysis of pitch differences based on the first language.

The last point I wanted to mention is the lack of native English speakers. If they were recruited, they would have only been recorded in English - as there are presumably few English natives that can fluently speak Kazakh. The use of their data will be described in the following section on Further research. Generally, as of now, the research does not have reliable data on the comparison of the participants' pitch in English to the pitch of a native speaker. This would be another point of analysis as maybe the participants' pitch in English differs from their pitch in Kazakh or Russian due to the English language as a whole being associated with this different fundamental frequency.

Future research

One point regarding further research was just mentioned in the Limitation section of this work. This research could benefit from having data from native English speakers. This data would act as the role of the bottom line for a comparison of the pitch in English of a native speaker and that of Kazakh/Russian native speakers. As of now, we do not know if the acoustic features of English itself interfered with our results as we do not have reliable data to compare our data to. To elaborate, let's say that the average FF for a male native Kazakh speaker in Kazakh is 100Hz, and for English, it is 115Hz. This creates a significant difference, and we can assume that this comes from the fact that English is not a native language for Kazakh speakers. However, let's say that FF for a male native English speaker in English is 116Hz. This entails that a higher frequency - 115-116Hz - is an acoustic feature of the English language, and the

Kazakh natives speak English at a higher frequency than Kazakh because the English language requires them to speak at that frequency.

Another possible point is having professionals evaluate people's fluency levels. As of now, the questionnaire of this experiment asks the participants to evaluate their speaking skills on a 1-10 scale for the three languages under investigation and to report their IELTS scores. Neither Kazakh nor Russian has a reliable system that evaluates speaking skills, so there is no way for the participants to already have a more or less objective evaluation of their skills like they have for English. And as for their self-evaluation, this includes both the skill as well as the confidence level of the participants. We do not know whether the overall confidence of the person affects their perception and evaluation of their skill, so we cannot take their evaluation to be objective. For example, a confident or delusional person may evaluate their skills as 9/10 when on an objective scale their skill is closer to a 6/10. At the same time, a self-conscious person may evaluate their skills as 7/10 when their objective result is 9/10. A separate professional evaluation of the speaking skills of all of the participants would allow the analysis to separate confidence level and fluency into two separate categories.

Overall there is a variety of ways of how this research can be taken further. There is also a possibility to repeat this experiment with a controlled number of participants from each of the regions. Or a more personal - and more risk-infused - parameter of asking the participants about the type of upbringing they had. One of the initial hypotheses of this research was that Kazakh-native women will have a larger difference between their pitch in Kazakh and Russian due to the qualities Kazakh girls are taught. By Kazakh traditions, girls are supposed to be on the quieter side with a melodious and high voice (Argynbayev, 2005). So it was expected that Kazakh-native women will have a higher FF in Kazakh than in English/Russian or a higher pitch

overall than Russian-native speakers. However, the results did not prove to be significant and the questions on the upbringing were considered to be too personal and out of the scope of this research project.

References

- Altenberg, E. P., & Ferrand, C. T. (2006). Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice, 20(1)*, 89-96.
- Argynbayev, H. (2005). Qazaqtyn otbasylyq dasturleri [Traditions of Kazakh families]. *Almaty: Qainar*.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Ghiurcau, M. V., Lodin, A., & Rusu, C. (2010). A study of the effect of emotional state upon the variation of the fundamental frequency of a speaker. *Journal of Applied Computer Science & Mathematics, 4(1)*, 79-82.
- Gorman, P. (2010). Mains Hum. *Junctures: The Journal for Thematic Dialogue, (13)*.
- Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. *In Speech Prosody 2002, International Conference*.
- Harrington, J., Palethorpe, S., Watson, C. I. (2007). Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers. *Interspeech, 2007*, 2753-2756.
- Järvinen, K., Laukkanen, A. M., & Izdebski, K. (2007). Voice Fundamental Frequency Changes as a Function of Foreign Languages Familiarity: An Emotional Effect?. *Emotions in the Human Voice, Volume 1: Foundations, 1*, 203-213.
- Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication, 88*, 106-126.

- Niebudek-Bogusz, E., Fiszer, M., Kotylo, P., & Sliwinska-Kowalska, M. (2006). Diagnostic value of voice acoustic analysis in assessment of occupational voice pathologies in teachers. *Logopedics Phoniatrics Vocology, 31*(3), 100-106.
- Nishio, M., & Niimi, S. (2008). Changes in speaking fundamental frequency characteristics with aging. *Folia phoniatrica et logopaedica, 60*(3), 120-127.
- Ohara, Y. (1999). Performing gender through voice pitch: A cross-cultural analysis of Japanese and American English. In *Wahrnehmung und Herstellung von Geschlecht* (pp. 105-116). VS Verlag für Sozialwissenschaften.
- Ordin, M., & Mennen, I. (2017). Cross-linguistic differences in bilinguals' fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research, 60*(6), 1493-1506.
- Reetz, H., & Jongman, A. (2020). *Phonetics: Transcription, production, acoustics, and perception*. John Wiley & Sons.
- Schwab, S., & Goldman, J. P. (2016). Do speakers show different F0 when they speak in different languages? The case of English, French and German?.
- Styler, W. (2013). Using Praat for linguistic research. University of Colorado at Boulder Phonetics Lab.
- Wagner, A., & Braun, A. (2003). Is voice quality language-dependent? Acoustic analyses based on speakers of three different languages. *Language, 6*(4), 2.
- Zraick, R. I., Gentry, M. A., Smith-Olinde, L., & Gregg, B. A. (2006). The effect of speaking context on elicitation of habitual pitch. *Journal of Voice, 20*(4), 545-554.
- Zraick, R. I., Skaggs, S. D., & Montague, J. C. (2000). The effect of task on determination of habitual pitch. *Journal of Voice, 14*(4), 484-489.

Appendix

Appendix 1: Link to the questionnaire part of the experiment

<https://forms.gle/mXmZvWMHBbwzhHgu5>

Appendix 2: Texts used in the reading task

In Russian:

Национальный музей Республики Казахстан самый молодой и самый крупный музей в Центральной Азии. Музей был создан в рамках реализации Государственной программы «Культурное наследие» по поручению Президента Республики Казахстан Н.А.Назарбаева. Музей находится на главной площади страны — на площади Независимости, гармонично вписываясь в единый архитектурный ансамбль с Дворцом Независимости, Дворцом мира и согласия и Национальным университетом искусств.

In Kazakh:

Қазақстан Республикасының Ұлттық музейі - Орталық Азиядағы ең жас әрі ең ірі музей. Музей «Мәдени мұра» мемлекеттік бағдарламасын іске асыру шеңберінде Қазақстан Республикасының Тұңғыш Президенті-Елбасы Н.Ә. Назарбаевтың тапсырмасы бойынша құрылды. Музей еліміздің бас алаңы – Тәуелсіздік алаңында орналасқан. Оның асқақ, бірегей ғимараты осы маңдағы «Қазақ елі» монументімен, Тәуелсіздік сарайымен, Бейбітшілік және Келісім сарайымен, «Хазірет Сұлтан» мешітімен және Ұлттық өнер университетімен үйлесім тапқан.

In English:

The National Museum of the Republic of Kazakhstan is the youngest and largest museum in Central Asia. The museum was created within the framework of the State Program "Cultural Heritage" on behalf of the President of the Republic of Kazakhstan N.A. Nazarbayev. The

museum is located on the main square of the country - on Independence Square, harmoniously fitting into a single architectural ensemble with the Palace of Independence, the Palace of Peace and Reconciliation, and the National University of Arts.