Gender and the Ability to Pitch Match from Memory (December 2016)

Conor Broaders, Angela Hosbein, Eric Willms, Kaitlyn Young and Cindy Zhang

***Abstract*— The purpose of this experiment was to explore individuals' abilities to match pitches and determine if there was any relationship between how individuals matched a signal relative to another signal, and an individual's gender. In this experiment, a stimulus tone was played on a speaker for a duration of 2 seconds. The participant then tried to replicate the pitch by adjusting a potentiometer on a 555 timer circuit to match what they had heard. Based on previous studies, it was hypothesized that males would be better at matching lower frequencies, while females would be better at matching higher frequencies. The results showed that there are no statistically significant conclusions that could be made about the relationship between genders and pitch matching ability.**

I. Introduction

PITCH MATCHING is the skill of being able to distinguish differences in two tones and bring one to the same frequency as the other. It is commonly utilized by musicians to stay "in-tune" to a tuner or ensemble, and is a critical skill in the music field. Additionally, it has been found to be correlated with speech development and the ability to learn languages. Frequency resolution for the human ear, however, has been found to be dependent on the frequency that is being matched to. The objective of this study is to explore the extent to which individuals can pitch match and to determine if there are any trends relating how individuals tune a signal relative to another signal with respect to the listener's gender.

II. Background

A. Psychometric Functions

In psychoacoustics, the relationship between sensitivity of a stimulus to the magnitude of difference of the stimulus is often referred to as the "psychometric function", PF [1]. PFs obtained through one study is shown in Fig 1 and shows the non-linear relationship between minimum thresholds ($\Delta f/f$ where f is stimulus frequency) for frequency discrimination with respect to a stimulus frequency. The PF suggests that the minimum threshold for frequency discrimination (0.002 or 0.2%) occurs for the average human at around 1000 Hz [1]. Although this PF only describes the frequency discrimination ability of humans in general without regard to demographics or musical ability, it serves as a baseline and tool for comparison for acquired data in the present study.

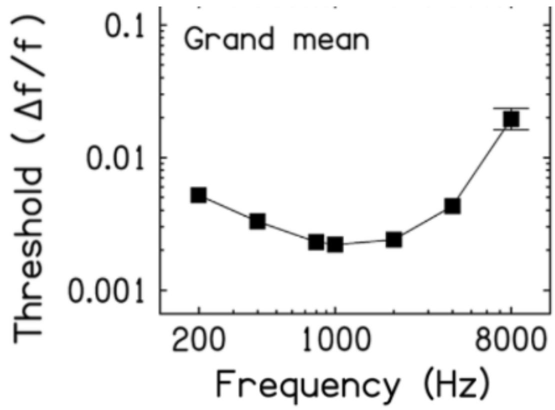


Fig. 1. Psychometric function for frequency derived from an experiment that tested discrete frequencies from 200 to 8000 Hz for 5 listeners [1]

B. Pitch Matching and Gender

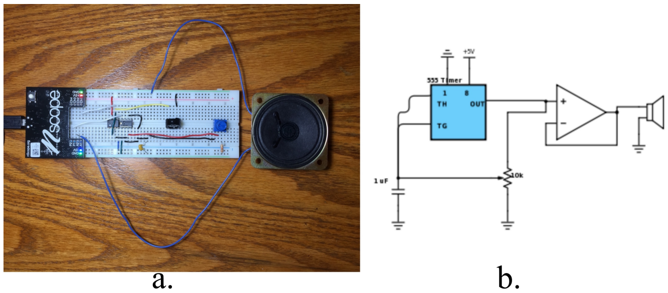
A listener's ability to pitch match in relation to his/her gender is perhaps more non-obvious than a factor like one's musical experience. A study that looked into this relationship is one performed on English-speaking preschool children [2]. This study found that boys were more accurate than girls in the low frequencies (< 275 Hz), but girls were better at matching pitch overall, with better accuracy across mid-range (275 - 375 Hz) and high frequencies (> 375 Hz) and less error overall [2]. (The experiment asked the children to sing the pitch they heard, so frequencies tested had to be within the children's vocal range). The study also found correlations for better pitch matching ability with characteristics of the child's speech such as a higher overall fundamental frequency and wider speech range [2]. These results and the study overall suggest that this may have implications to the child's speech development as he/she grows [2]. Although measuring subjects' speech characteristics is outside the scope of this present study, we do seek to obtain similar results for pitch matching ability with respect to gender.

C. Hypothesis

With the knowledge of the results from the study by Trollinger (2003), it was hypothesized that men would be better at pitch matching on lower frequencies (< 800 Hz) while women would be better at pitch matching at higher frequencies (> 800 Hz).

III. Methods & Materials

This section pertains to the design of the experimental setup and the procedures for testing and analysis of the data.



a. b.

Fig. 2. (a) The test circuit with the blue potentiometer on the right used for pitch modulation and (b) the diagram of the complete timer circuit.

A. *Sound Circuit*

The circuit shown in Fig 2 consists of a 555 timer in order to produce a square-wave, which is then output to a small speaker. The frequency of the output is determined by a 10k Ω potentiometer which could be rotated by the listener to alter the frequency of the timer output. A larger speaker was also attached in separate circuit that connected the nScope's P1 output to ground to deliver the stimulus tone.

B. *Experimental Design and Data Collection*

In the experiment each participant was given 5 different tones (200, 500, 800, 1100, and 1400 Hz). This range of pitches was chosen in order to observe a range of tones roughly spanning the range of the potentiometer, and also covers typical speaking range. The participants were given the tones in a random order, one at a time for 2 seconds each. The participants then attempted to match the pitch as closely as possible from their memory. This was then repeated again for each tone, for a total of 10 tones per participant. For the script that was used in the experiment please see Appendix A. Fig 3 shows the gender distribution of our participants.

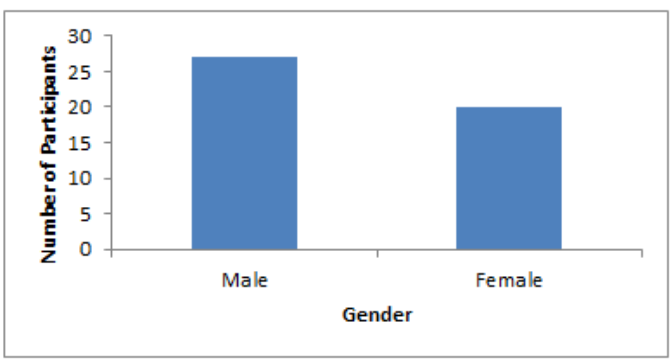


Fig. 3. The gender distribution of participants.

C. *Transposition of Frequency Data*

Because of the nature of pitch matching, some participant identified the correct note of the tone played, but were off a factor of 2^n in frequency (Hz). This is because they were likely trying to match the pitch in a different octave than was given. To account for this discrepancy, the data was scaled such that pitches that were determined to be closer to a different octave were scaled by a factor of 2^n where n is the number of octaves the pitch appeared to be displaced.

The threshold for each octave was defined to be at a ratio of 1.41, rather than the numerical halfway of 1.5, since perceived pitch is not linear to the tone's frequency. This ratio defines the tritone, or augmented 4th, which is the interval that is halfway between a pitch and its octave in the equal-temperament tuning system.

IV. Results

Performance of both males and females were directly compared for each frequency tested. For 200 Hz, males and females reported on average 290 Hz and 366 Hz respectively. The standard deviation of male's frequencies was also smaller (306 Hz) than that of females (479 Hz). The comparison boxplots can be seen in Fig 4.

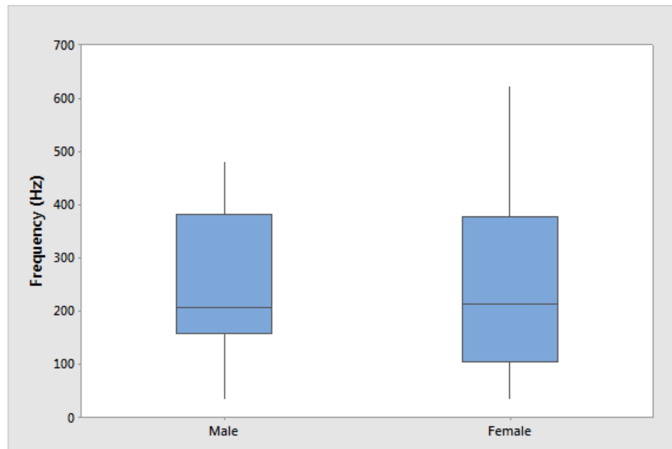


Fig. 4. Male vs. Female frequency response for 200 Hz.

For 1400 Hz, males and females reported on average 1043 Hz and 1265 Hz respectively. The standard deviation of male's frequencies was also smaller (562.6 Hz) than that of females (613.3 Hz). The comparison boxplots can be seen in Fig 5.

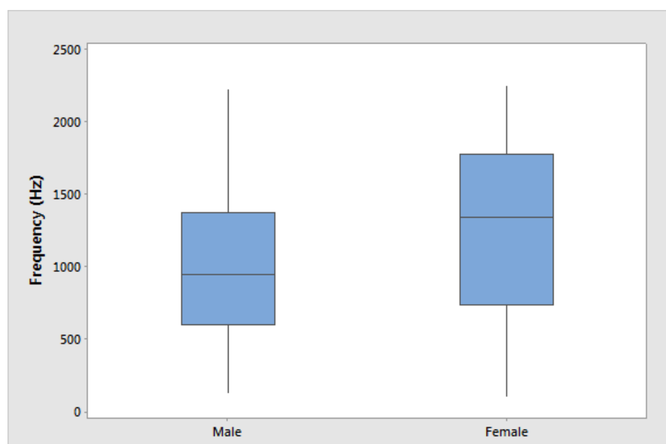


Fig. 5. Male vs. Female frequency response for 200 Hz.

Additional plots for each tested frequency can be found in Appendix B.

To address this widespread variance, tones were then scaled to their corresponding notes using the 1.41 threshold ratio discussed previously, and male and female responses for each tone were compared.

For scaled 200 Hz, males and females reported on average 205.9 Hz and 206.8 Hz respectively. The standard deviation of males' frequencies was 34.2, whereas females had a standard deviation of 38.6 Hz. The comparison boxplots can be seen in Fig 6.

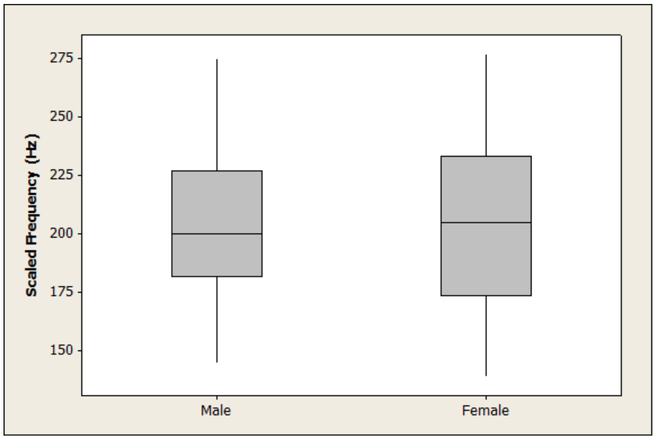


Fig. 6 Scaled Male vs. Female frequency response for 200 Hz

For scaled 1400 Hz, males and females reported on average 1464.5 Hz and 1458.9 Hz respectively. The standard deviation of males' frequencies was 333.1, whereas females had a standard deviation of 270 Hz. The comparison boxplots can be seen in Fig 7.

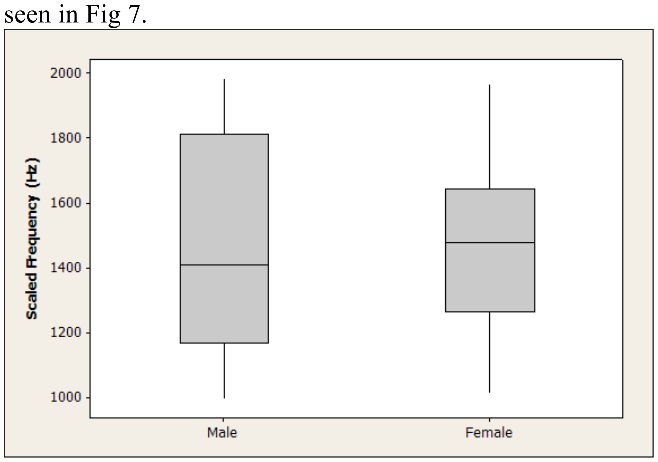


Fig. 7 Scaled Male vs. Female frequency response for 1400 Hz

Plots for all frequencies tested can be found in Appendix C.

Male and female data was directly compared and the significance of any differences were determined. The confidence intervals for male versus female means were found, and can be seen for each tone in Table 1.

Table 1. Scaled Frequency Response Male vs. Female Confidence Intervals

Tone	Sample Mean Difference	Confidence Interval
200 Hz	-0.90	(16.17, 14.36)
500 Hz	26.7	(-9.5, 62.9)
800 Hz	30.7	(-31.2, 92.6)
1100 Hz	-27.8	(-112.3, 56.6)
1400 Hz	5.6	(-118.1, 129.3)

Accuracy of both genders at each tone was determined by looking at how close responses were to the desired value. Distances away from desired values determined these error values, therefore all errors are positive.

Interval plots show how typical values of error ranged for each gender at each frequency. At 200 Hz, males had an average error of 26.6 Hz while female's error was 32.2 Hz. Fig 8 shows this expected error for respondents.

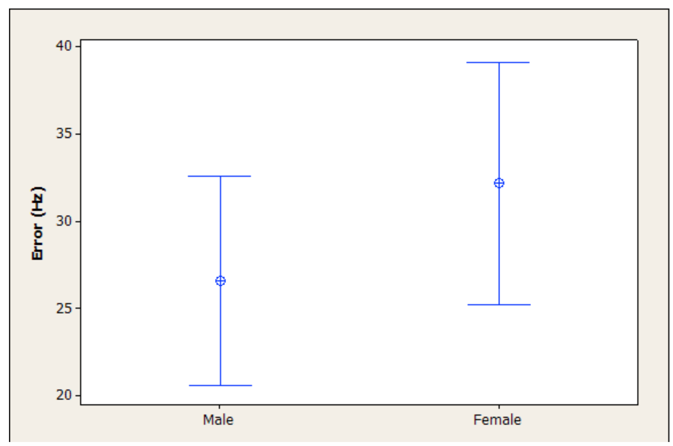


Fig. 8 Scaled Male vs. Female error for 200 Hz

At 1400 Hz, males had an average error of 279 Hz while female's error was 231 Hz. Fig 9 shows this expected error for respondents.

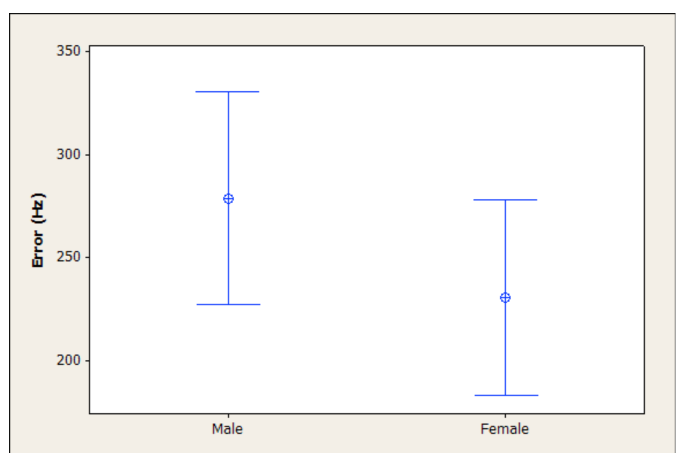


Fig. 9 Scaled Male vs. Female error for 1400 Hz

Plots for all frequencies tested can be found in Appendix D.

Confidence intervals directly comparing these error values were found, looking at male means versus female means. Intervals for each tone can be seen in Table 2.

Table 2. Scaled Error Male vs. Female Confidence Intervals

Tone	Error Mean Difference	Confidence Interval
200 Hz	-5.59	(-14.64, 3.46)
500 Hz	3.7	(-19.5, 26.8)
800 Hz	5.4	(-33.3, 44)
1100 Hz	-33.1	(-80.6, 14.3)
1400 Hz	48.4	(-20.8, 117.6)

Though male respondents tended to have less error than female respondents at 200 Hz (26.6 Hz versus 32.2 Hz), this trend was not statistically significant. In addition, though female respondents had less mean error than male respondents at 1400 Hz (231 Hz versus 279 Hz), this was also not a statistically significant difference.

V. Discussion And Conclusion

There are several components of the experimental design that could be improved in the future to obtain more accurate results. The speaker playing the given tone and the speaker that the participant would use to match the tones should have been the same, in order to ensure that the input and output frequencies of both are consistent and the same volume. Additionally, the experiment could have used headphones or conducted all trials in a soundproof room in order to eliminate ambient noise that may have affected the results of the trials. Due to the fact that the potentiometer used to control the pitch adjusted the frequency linearly, we noticed that it was more difficult for participants to pitch match the higher notes since pitch increases exponentially to frequency. In order to offset this effect in future experiments, it would be important to make the potentiometer used by participants increase in frequency exponentially, so that the perceived pitch is linear.

Another concern was that many participants matched the note instead of the exact pitch. This was especially true of participants with prior musical experience, possibly because these people would have been trained to tune to a transposed note naturally. In order to account for this effect, the halfway point was calculated, using equal-temperament tuning, between each tone and the corresponding note in higher and lower octaves, and multiplied or divided the response until it aligned with the proper octave if it was closer to the note in a different octave. It is for this reason that there are two box plots (Fig 5,6) for each frequency, one representing the data when the responses were scaled, and one that shows the unscaled data.

It was noted that many participants hummed the note that they were supposed to match accurately, but still had difficulty tuning the potentiometer to the correct note. It was also noted that the longer it took for a respondent to match the pitch, the more likely they were to forget the original tone entirely. This leads to the conclusion that the test was not solely a measure of pitch matching ability, but of also being able to quickly and accurately control the testing apparatus.

It was expected that men have better pitch matching ability at lower frequencies and women to have better pitch matching at higher frequencies. The sample data does not support this hypothesis as the differences in pitch matching performance by men and women were not found to be statistically significant.

References

[1] H. Dai and C. Micheyl, "Psychometric functions for pure-tone frequency discrimination," *The Journal of the Acoustical Society of America,* vol. 130, pp. 263-272, 2011.

[2] V. L. Trollinger, "Relationships between Pitch-Matching Accuracy, Speech Fundamental Frequency, Speech Range, Age, and Gender in American English-Speaking Preschool Children," *Journal of Research in Music Education,* vol. 51, p. 78, Spring2003

Appendix A: Testing Script

Pre-Test Script

This experiment will test your ability to match tones.

- What is your name?
- What kind of prior experience do you have that you think is relevant to this experiment? For example, with playing musical instruments or tuning.
- Do you have any hearing issues, such as hearing loss in one ear? Please explain if so.
- On a scale of 1 to 7 (7 being the highest score and 1 being the lowest), how would you rate your musical ability?
- How do you identify? As male or female or prefer not to say?
- How old are you?

You will hear a total of fifteen tones. You will hear each tone for a duration of 2 seconds. Once a tone has finished playing, your job is to turn the knob on the given breadboard to match the pitch you hear without transposing. Take some time to match the pitch and let me know when you have finished. Then, I will play the next tone. You will not receive feedback during or after the experiment.
You will have two trial tones to try to practice matching the tones. Then we will begin with the experiment.

Do you have any questions?

Test 1 low (300 Hz) pitch and 1 high (1300 Hz)

Do you need any clarifications?

Appendix B: Additional Unscaled Graphs

Fig 10-12 show boxplots for unscaled frequency responses for 500, 800, and 1100 Hz respectively.

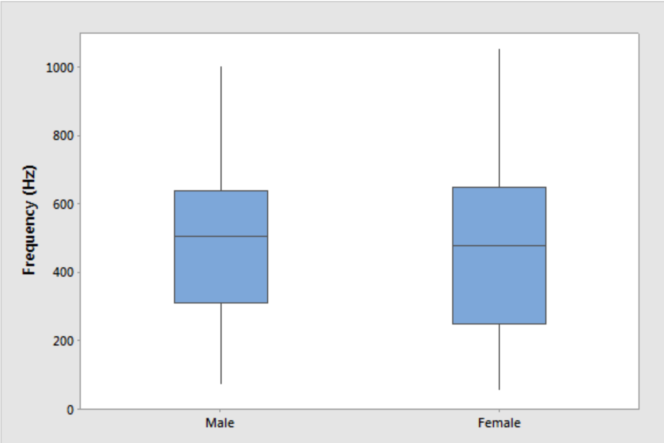


Fig. 10 Male vs. Female frequency response for 500 Hz.

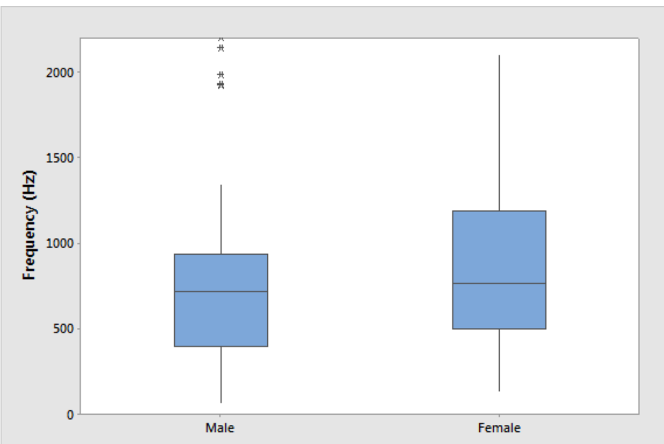


Fig. 11 Male vs. Female frequency response for 800 Hz

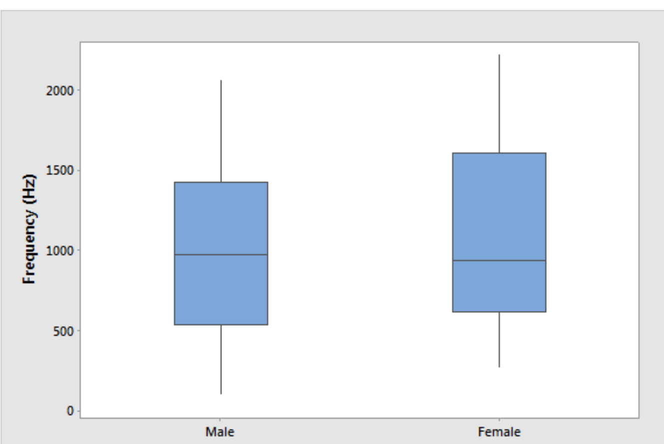


Fig. 12 Male vs. Female frequency response for 1100 Hz

APPENDIX C: ADDITIONAL SCALED GRAPHS

Fig 13-15 show boxplots for scaled frequency responses for 500, 800, and 1100 Hz respectively.

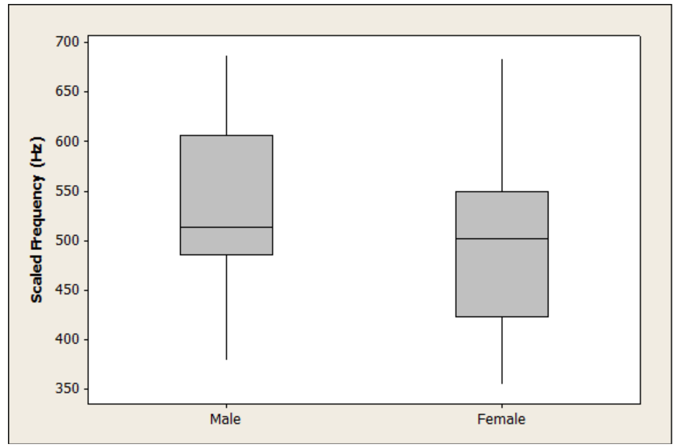


Fig. 13 Scaled Male vs. Female frequency response for 500 Hz

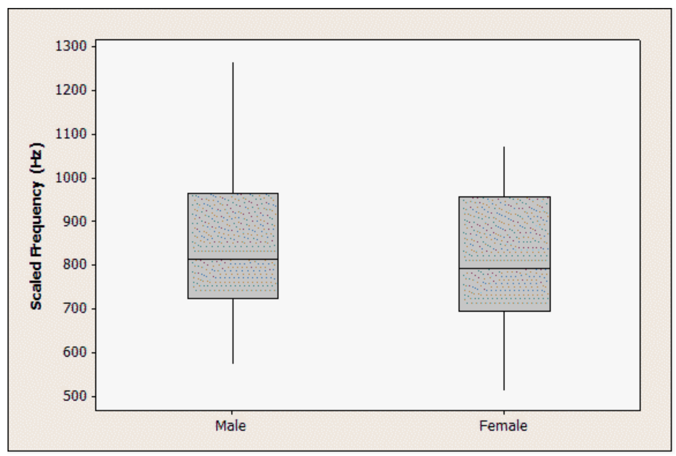


Fig. 14 Scaled Male vs. Female frequency response for 800 Hz

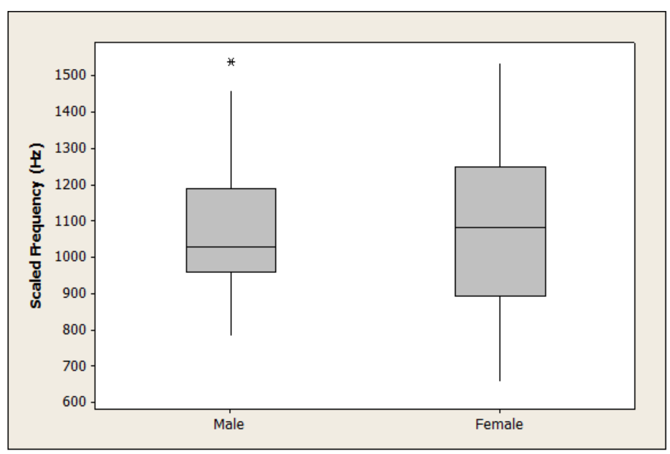


Fig. 15 Scaled Male vs. Female frequency response for 1100 Hz

APPENDIX D: ADDITIONAL ERROR GRAPHS

Fig 16-18 show plots error responses for 500, 800, and 1100 Hz respectively.

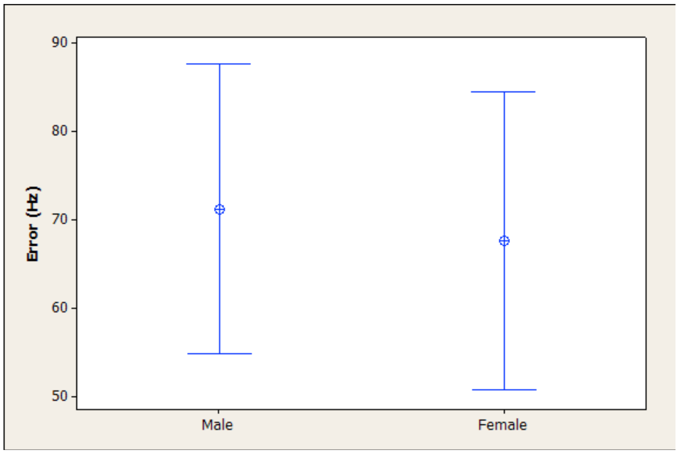


Fig. 16 Scaled Male vs. Female error for 500 Hz

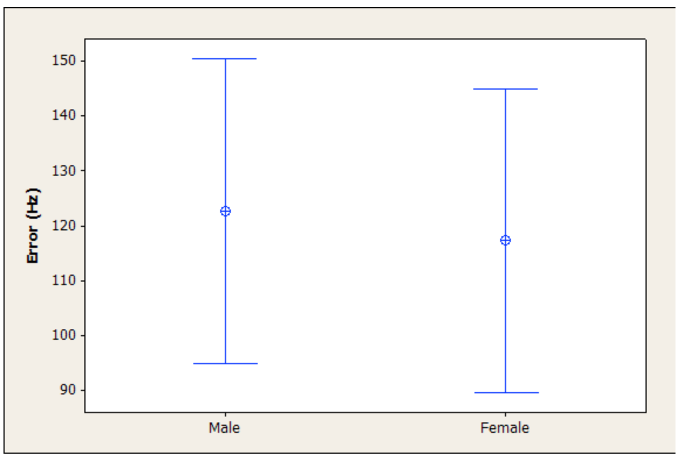


Fig. 17 Scaled Male vs. Female error for 800 Hz

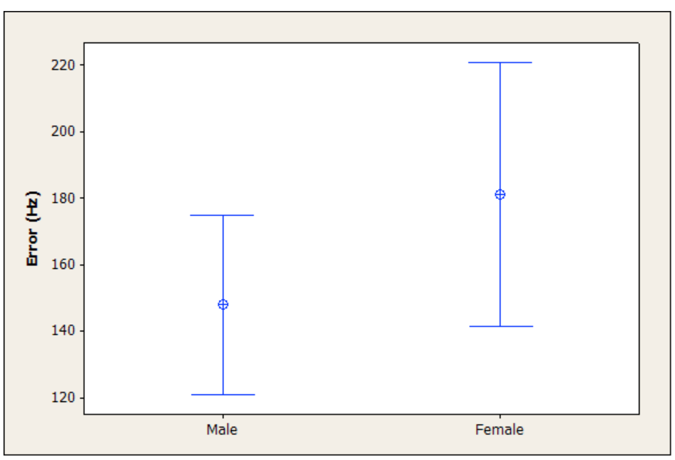


Fig. 18 Scaled Male vs. Female error for 1100 Hz