DURATION OF STIMULI HAS NO EFFECT ON RELATIONSHIP BETWEEN PITCH AND RATINGS OF ATTRACTIVENESS FOR MALE VOICES

DAHLSTROM, SOFIA AND LUKE, HANNAH
Psychology, College of Science and Engineering

ABSTRACT

The negative relationship between pitch and attractiveness for male voices has been firmly established throughout previous research. Lower male pitch preference is considered to have been hard-wired through evolution prior to the emergence of language and the reproductive success of males with lower-pitched voices spans beyond western culture. Despite the consensus that lower-pitched male voices are rated higher for attractiveness, research pursuing the factors that affect this relationship is limited. Yet, findings related to the vocal features that affect the relationship between lower-pitched male voices and attractiveness ratings could contribute to important discussions regarding vocal evolution. Duration, referring to the time and stimulus length of speech, has been noted to be the most reliable vocal feature for predicting attractiveness in male voices. However, research regarding duration is conflicted and does not address its effect on the relationship between pitch and attractiveness. This research investigated the effect of duration on the relationship between male voice pitch and attractiveness ratings using secondary data of female ratings for male voice recordings. Two socially relevant sentences of different durations were statistically compared against the relationship. Statistical analysis showed no significant difference in the relationship between attractiveness and pitch for either longer or shorter duration conditions. These results both align and contrast with previous research regarding duration, therefore, highlighting the complexity of vocal variability and the need for further research required to determine the importance of different vocal features, especially because of the potential in aiding the understanding of vocal evolution and the development of sexual selection.

INTRODUCTION

How we construct our first impressions of the people we meet is based upon many factors and the impressions that we construct dictate how we interact with people (Biesanz, et al., 2011; McAleer et al., 2014). From assumptions about personality to estimating physical features and aiding in partner selection throughout life, the quality of an individual's voice produces a profound amount of information to form an impression (Bruckert et al., 2006; Feinberg et al., 2008; Pisanski, et al., 2016; Mahrholz et al., 2018). Attraction is one of the qualities that is influenced by a person's vocal features (McAleer et al., 2014; Mahrholz et al., 2018). Perceived attractiveness has been linked to many vocal properties, especially pitch. In correlational and experimental settings, there is consistent reporting of increased female attractiveness ratings for male voices with lower pitch (Feinberg et al., 2005; Hughes et al., 2010; Jones et al., 2010; Hodges-Simeon et al., 2011; Xu et al., 2013). This consensus only deviates for male voice pitch below 96Hz where such a low pitch is rare and may sound unnatural to the female participant (Re et al., 2012).

Research has suggested an evolutionary aspect to the tendency of women to prefer men with lower-pitched voices (Evans et al., 2008; Hodges-Simeon et al., 2015; Jost et al., 2018). Lower pitch has been positively associated with higher testosterone levels which is suggested to be an indicator of immunocompetence, social dominance, strength, and possession of higher genetic quality (Hodges-Simeon et al., 2015; Arnocky et al., 2018). Biologically, lower voice pitch occurs when there are high levels of testosterone present during development of the vocal tract (Feinberg et al., 2008). Women may therefore rely on this vocal property to assess partner quality. This is further supported by males with lower-pitched voices reporting greater reproductive success in industrialized societies (Suire et al., 2019). This relationship seems to span beyond western cultures to isolated tribes such as the hunter-gatherers of the Hazda tribe in Tanzania where men with lower-pitched voices also report higher mating success compared to those with higher pitched voices (Apicella et al., 2007). Within the Hazda tribe it was also found that both sexes viewed lower pitched voices in the opposite sex as better at acquiring resources suggesting an evolutionary advantage to women's attraction to lower pitched male voices (Apicella et al., 2007). Further support for voice preference being an adaptation for mate choice guidance comes from Zhang et al. (2020) suggesting it was hard-wired through evolution even before the emergence of language.

Further support for the evolutionary development of voice preference comes from the theory that the information women are looking for in the male acoustics can change depending on menstrual phases and imagined mating contexts (Penton-Voak et al., 1999; Puts., 2005). When near ovulation or in short-term mating contexts women tend to be more attracted to lower pitched male voices since it is indicative of reproductive capability and security of successful offspring (Puts., 2005; Feinberg et al., 2006; Feinberg et al., 2008). Additionally, Puts (2005) found that lower male voice pitch is preferred mainly in short-term mating than in long-term mating which may be the result of males with higher testosterone levels, and therefore lower pitched voices, being considered carriers of higher genetic quality but being less invested in their partners (Penton-Voak et al., 1999; Puts., 2005). Therefore, it may have been beneficial for females to prefer such males with lower pitched voices and higher testosterone for sexual relationships compared to committed ones, especially during the fertile phase of their menstrual cycle. Hence, female preference for lower pitched male voices may have evolved due to the indication of increased heritable benefits that accompany the higher testosterone required to produce a lower pitch, and such female preference may have further acted as a selection pressure on male voice pitch.

Sofia: sofiapearly@gmail.com. Hannah: hannahmluke@gmail.com

Evidently, the preferences of women in the fertile phase of their menstrual cycle shifts towards more masculine features (Puts., 2005). This can be seen in not just vocal preferences, but also in facial attractiveness and odour preference: women seem to demonstrate a higher preference for certain male scents during peak ovulation period compared to the rest of their cycle (Thornhill and Gangsted., 1999). Furthermore, Penton-Voak et al. (1999) found female preferences for more masculine or feminine male faces to be cyclical: at the peak ovulation period there was a higher preference for the more masculine faces, whereas outside of the peak ovulation period the preference was for more feminine faces. It has been suggested that this preference is based upon the desire for a short-term partner to have good, strong genetics while the woman is fertile, but for long-term partners to have more feminine traits as there is an association between more feminine male faces and more pro-social behaviours in relationships (Puts., 2005; Feinberg et al., 2006; Feinberg et al., 2008). This is further supported by Mazur and Booth (1998) reporting a higher likelihood of marital instability, divorce and domestic violence from male participants with more masculine traits.

While female menstrual cycle changes can vastly affect preferences towards male traits, there are many other factors that can further affect these preferences. One of these is duration – a very important variable in first-impression research. When investigating the relationship between duration and ratings of attractiveness for faces, Willis and Todorov (2006) concluded that only minimal exposure time was required for participants to make trait judgments about different faces, and that any additional time only served to reinforce these initial impressions. Additionally, when exposure time to the faces increased from 100 to 500 milliseconds, participants judgements become more negative (Willis and Todorov., 2006). Furthermore, Stróżak and Zielińska (2019) investigated the difference in response for different durations of exposure to both attractive and unattractive faces. They found that for shorter durations, unattractive faces would be rated as more attractive than they would be when viewed for the longer durations. This is supported by investigations from Gerger et al. (2017) as well as Rashidi et al. (2012) who both found that when the presentation duration of faces increased, the ratings of attractiveness decreased.

There is evidence that face and voice attractiveness perceptions stem from interpretations of similar underlying hormone levels that convey signals allowing the overall fitness of an individual to be suggested (Feinberg et al., 2008). Therefore, it could be inferred that the results from duration research on vocal attractiveness would echo that of facial attractiveness. In voice research, duration refers to the length of time of the utterance and is typically determined by the stimulus length or editing software. Babel et al. (2014) found that when comparing the effect of monosyllabic words and vowels, male voices with shorter durations were found more attractive. Additionally, Ferdenzi et al. (2013) reported that lengthened stimuli were significantly less attractive than shortened stimuli after digitally altering durations of the French word “bonjour” and vowels. However, voice attractiveness varied as a function of stimulus type, and the word “bonjour” was significantly more attractive than the single vowels. While contradictory, this may have been because longer duration stimuli provide potential emotional and social information which made the judgment of attractiveness easier (Jones et al., 2008; Ferdenzi et al., 2013). Moreover, it was found that the duration of non-manipulated sound samples did not affect perceived attractiveness (Ferdenzi et al., 2013; Yunji Kim., 2013).

However, Yunji Kim (2013) did find a statistically significant relationship between duration and attractiveness ratings provided that vowel space (defining vowel sounds by their position in acoustic and articulatory space) and fundamental frequency (an acoustic measure of voice pitch) were controlled. It was concluded that a 5 millisecond increase in duration would contribute to an increase in vocal attractiveness ratings. They further noted that duration was the most reliable vocal feature for predicting attractiveness in male voices (Yunji Kim., 2013). Therefore, it is not clear what effect duration has on attractiveness ratings due to the variance between and within studies. Additionally, there are further methodological differences that make it difficult to determine the role of duration.

In the current study, it is vital to consider that much of the methodology of previous research in this field fails to define or report their range of voice sample durations, and they use different programmes and manipulations to alter the duration of the recordings. Ferdenzi et al. (2013) discussed, in detail, the significant variations in the durations used in many of the studies that investigate the relationship between pitch and attractiveness. Collins and Missing (2003) used a vowel duration of 250 – 380 milliseconds, Feinberg et al. (2005) used an average duration of 640 milliseconds, Bruckert et al. (2010) used a duration of 201-477 seconds and many other studies do not disclose their duration range. Many studies in this area also use a range of manipulation techniques to digitally alter the durations of the recordings instead of using natural voice recordings of different durations. This makes comparisons across different studies in this field difficult and subjective due to the different stimuli and manipulation techniques. It also highlights queries regarding whether the preference for shorter duration stimuli in previous research is reliant on the quality and source of digital manipulation of the recording.

Furthermore, there is a lack of consensus regarding longer or shorter durations being perceived as more attractive, and a gap in research specifically investigating attractiveness ratings for pitch in different durations. The current study, therefore, explored the relatedness of attractiveness ratings based on pitch across two stimulus lengths. This research used secondary data from Mahrholz et al. (2018), with socially relevant stimulus conditions that applied to the participant population of students and were created specifically for Mahrholz et al. (2018), to compare the relationship between pitch and attractiveness in different conditions. The shorter duration condition of “Hello” and the longer duration condition of “I urge you to submit your essay by the end of the week” were used. The differing duration stimuli were chosen instead of manipulating the duration of stimuli to avoid participants selecting the more natural sounding voice recording. Previous research found that male listeners have less experience ranking male voices in terms of attractiveness and are unwilling to give male voices high attractiveness ratings, which has the potential to skew results (Pisanski and Rendall., 2011; Babel et al., 2014). Therefore, this report focused on the attractiveness ratings of only female participants for male voices. Additionally, age range was restricted to 17-30 years for both speakers and listeners to reduce the potential effects of age-related positivity bias. Based on previous research, this research hypothesised a significant difference between the relationship of vocal pitch (in Hz) and attractiveness ratings, as measured on a visual analogue scale (VAS), when judgments were based upon hearing one word or one sentence.

MATERIALS AND METHODS

Ethics

All procedures used by Mahrholz et al. (2018) were approved by the University of Glasgow Ethics Committee and were in accordance with the ethical standards of the 1964 Declaration of Helsinki (World Medical Association., 2001). All

participants provided either written consent or pressed an online confirmation of consent button. This indicated their agreement for their data to be stored, treated and for their voice recordings to be stored in an open-access database (Mahrholz et al., 2018).

Participants

Voice Recording

Sixty native English speakers (30 females: 20.2 ± 2.95 years (range: 17-27 years) and 30 males: 23.2 ± 3.75 years (range: 17-30 years)) were recruited for stimuli recording via the University of Glasgow School of Psychology Subject Pool (Mahrholz et al., 2018). Mahrholz et al. (2018) advertised for Scottish participants, between 17 and 30 years of age without speech impairments.

Online rating experiment

Additionally, 181 new participants, 138 females: 20.1 ± 2.45 years (range: 18-30 years) and 43 males: 21.3 ± 2.78 years (range: 18-27 years) took part in the original online voice rating experiment (Mahrholz et al., 2018). Each participant could only contribute to the voice recordings or the voice ratings. Participant recruitment was via the same means and criteria as for the voice recording participants.

The current study selected the data from 24 female participants (19.88 ± 3.19 years (range: 18-33 years)) that rated male voices for attractiveness and the 30 male participants' voice recordings (23.2 ± 3.75 years (range: 17–30 years)).

Materials

Mahrholz et al. (2018) recorded each participant's voice individually in a custom-made sound-attenuated chamber. Participants were instructed to read the two unfamiliar texts five times in a neutral voice. They did this for content relevant and content ambiguous related words and sentences of similar lengths. Further information on the materials can be found in the original study (Mahrholz et al., 2018) and the Open Science Framework (OSF) depository (osf.io/s3cxy). The current study selected the content relevant conditions which included the socially relevant word “Hello” and the socially relevant sentence “I urge you to submit your essay by the end of the week”.

Procedure

Mahrholz et al. (2018) conducted their experiment online through the Experiment webpages of the School of Psychology, University of Glasgow (https://experiments.psy.gla.ac.uk/). Participants were asked to complete it in a quiet environment and were randomly assigned to a personality type (trustworthiness, dominance, or attractiveness) for female or male voices. The stimulus was rated using a visual analogue scale (VAS) slider ranging from “not at all [trait]” (left) to “extremely [trait]” (right). The ambiguous relevant and content relevant words and sentences were presented in counterbalanced order with one variable changing at a time. Each stimulus was presented twice. This resulted in 240 ratings for each participant. The current study only used the data collected on female ratings of attractiveness for male voices.

RESULTS

Before completing any data analysis, the dataset was filtered for the variables of this report, any incomplete responses were removed and the mean response for each voice recording for both the word and the sentence conditions were calculated. All data wrangling, visualisation and analysis was conducted using the R programming environment (Version 4.0.3, R Core Team, (2020)) using the tidyverse library (Version 1.3.0), cocor (Version 1.1.3), pwr (Version 1.3.0) and attached packages (Diedenhofen and Musch., 2015; Wickham et al., 2019; Champely., 2020).

We hypothesised a significant difference between the individual correlations of vocal pitch (in Hz) and attractiveness ratings, as measured on a VAS, when judgements are based upon hearing either one word or one sentence. To determine this, we compared the correlational values from the two duration conditions. A correlational analysis measures the strength of a linear association between two variables. Based on a significance level (alpha) of 0.05, power of 0.8 and a sample size of 29, we could reliably detect a difference between the correlations of r_diff = 0.494. Alpha is the probability of rejecting the null hypothesis when it’s true and power is the probability of rejecting the null hypothesis when it’s false. The assumptions for the Pearson correlation analysis are that the data is continuous, the data is paired, there is an absence of outliers, there is a linear relationship between points, the data is relatively normally distributed, and the data is relatively homoscedastic (all data points must be around the same distance from the line of best fit). We checked and confirmed these assumptions through visual inspection of the data, and we performed a z-score test to confirm that our data follows a normal distribution and to highlight any outliers. It was decided in advance that we would remove outliers if it fell outside of ± 2.5 standard deviation on the z-score test. Through this method we found one outlier based on pitch: one voice recording had an unusually high pitch for a male voice. This voice recording participant was removed from both conditions before the correlational analysis was run. After removing the outlier, the voice recording participant demographics were 29 male voices (23.3 ± 3.79, a range of 17-30).

Table 1 Voice Recording Participants Raw Data

Voice Code	Age (years)	Gender	Sentence Condition Mean Response (VAS)	Sentence Condition Mean Pitch (Hz)	Word Condition Mean Response (VAS)	Word Condition Mean Pitch (Hz)
AMS16	30	M	259.56250	118.36038	197.7917	123.26568
APW26	18	M	115.35417	117.80874	206.2083	125.53026
DLS16	21	M	302.68750	97.90013	282.2917	97.56558
gm_M01	27	M	221.66667	105.93733	279.5208	103.23495
gm_M03	20	M	146.35417	142.03800	132.7083	144.20216
gm_M04	23	M	162.35417	138.36158	220.8125	146.64469
gm_M06	17	M	210.93750	101.62637	306.0625	107.23603
gm_M07	23	M	267.87500	118.84998	177.7708	117.83197
gm_M08	27	M	198.64583	120.44308	224.5417	104.01856
gm_M10	21	M	176.33333	107.34482	203.8750	116.14500
gm_M11	21	M	170.43750	125.95632	213.3542	157.88704
gm_M12	20	M	177.79167	132.33550	235.1250	121.13907
gm_M13	25	M	274.43750	110.07491	243.5833	115.45019
gm_M14	25	M	224.97917	95.59592	272.6667	102.14747
gm_M15	21	M	270.27083	106.80286	294.2708	115.38686
gm_M16	20	M	292.72917	127.98802	258.7083	111.64883
gm_M17	25	M	185.52083	126.66411	246.1875	136.52932
gm_M18	18	M	240.27083	119.01121	285.5000	115.60392
gm_M19	22	M	234.16667	120.50673	258.5417	93.56724
gm_M21	23	M	213.70833	90.70458	214.0417	113.93174
gm_M22	24	M	228.77083	91.28815	221.2708	95.72063
gm_M23	30	M	219.43750	123.21037	241.3542	125.71635
MLS27	27	M	81.14583	148.40884	120.0833	151.00864
MON02	25	M	172.39583	113.57855	228.0000	108.40334
PMG30	28	M	267.02083	146.93444	294.5833	123.25623
PPD02	30	M	182.89583	89.17183	194.6667	93.88287
SHD27	21	M	267.70833	136.48822	218.8333	143.78478
SSN18	26	M	266.10417	111.71432	295.4375	125.41236
WMS16	18	M	159.47917	103.82187	236.2083	101.12556

Figure 2 Ratings of Attractiveness compared to Pitch in the Word Condition

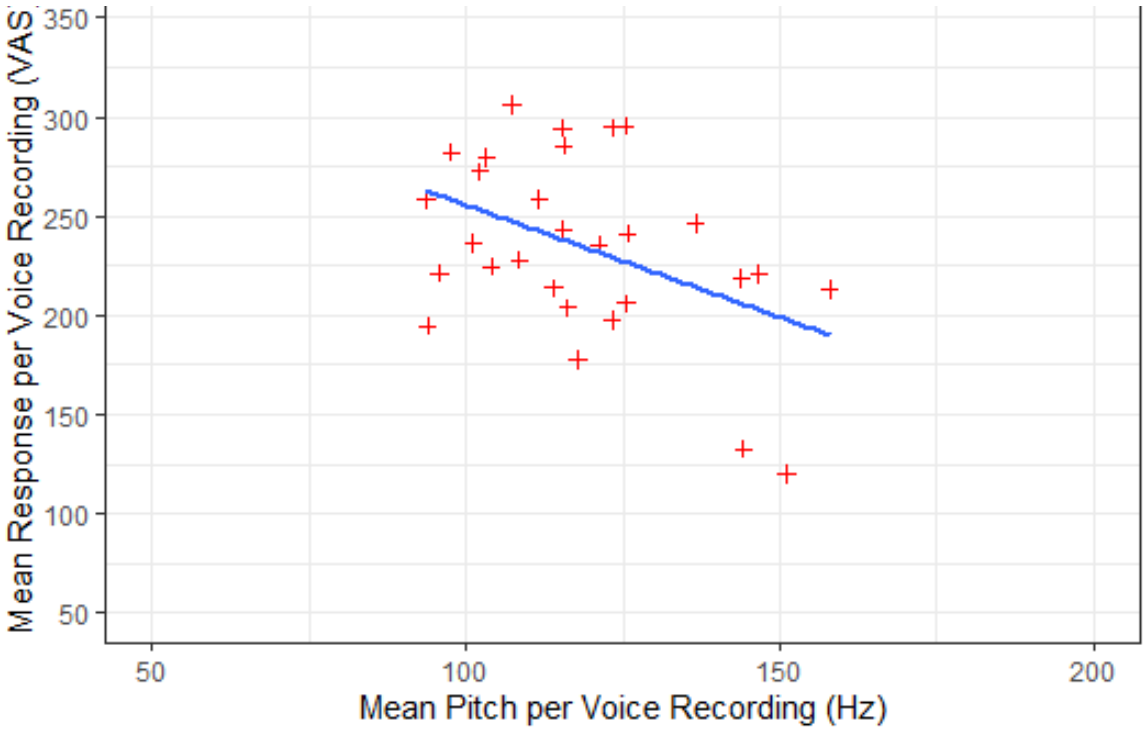
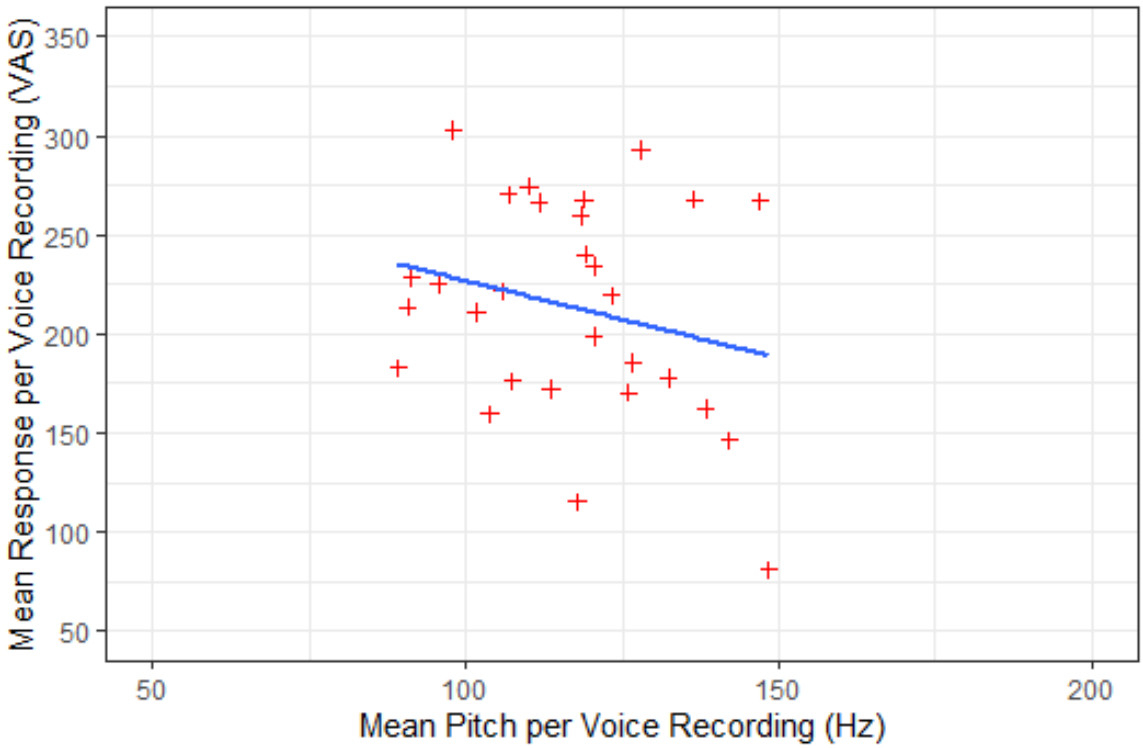


Figure 1 Ratings of Attractiveness compared to Pitch in the Sentence Condition



Based on the scatterplots above, the correlation between the mean pitch (Hz) and attractiveness (VAS) (Figure 1) suggested a weak negative correlation for the sentence condition. The correlation between the mean pitch (Hz) and attractiveness (VAS) (Figure 2) suggested a medium negative correlation for the word condition. This was supported by the Pearson correlation coefficients, denoted by r, obtained for the word condition (r = -0.438) and the sentence condition (r = -0.240).

DISCUSSION

The data followed the negative relationship trend between pitch and attractiveness that has formed the consensus of several studies and the negative r values from the Pearson correlations aligned with the sentence condition having a weaker negative correlation than the word condition (Feinberg et al., 2005; Hughes et al., 2010; Jones et al., 2010; Hodges-Simeon et al., 2011; Xu et al., 2013; Babel et al., 2014). Additionally, comparison of the two correlations using a z-transformation showed that there was no significant difference in attractiveness ratings based on the pitch for the shorter duration word condition and the longer duration sentence condition for male voices judged by females (Pearson and Filon., 1898). The results did not align with the hypothesis that there would a significant difference between the relationship of vocal pitch (in Hz) and attractiveness ratings, as measured on a VAS, when judgements are based upon hearing either one word or one sentence. Therefore, the experimental hypothesis was rejected, and the null hypothesis was retained suggesting there is no effect of duration on the relationship between pitch and attractiveness.

The results of the current study coincide with the research by Ferdenzi et al. (2013) and Yunji Kim (2013) who also found that duration did not affect ratings of attractiveness for different vowel sounds. The results align most with the results of Yunji Kim (2013), whereby the duration of the non-manipulated stimuli did not affect attractiveness ratings. However, both Ferdenzi et al. (2013) and Yunji Kim (2013) did find a significant difference between ratings of attractiveness for manipulated stimuli duration. With results indicating that shortened stimuli duration was preferred compared to lengthened stimuli (Ferdenzi et al., 2013; Yunji Kim., 2013). This aligns with research done on facial attractiveness, whereby shorter duration exposure to faces resulted in higher attractiveness ratings (Willis and Todorov., 2006; Rashidi et al., 2012; Gerger et al., 2017; Stróżak and Zielińska., 2019). Pell and Kotz (2011) theorised that longer duration recordings would be rated as more attractive because they convey more emotional and social information; this is especially applicable in the current study using socially relevant stimuli. However, these findings are contradicted by the results of the current study and other research that has demonstrated a lack of relationship between duration and ratings of attractiveness (Ferdenzi et al., 2013; Yunji Kim., 2013). Due to the nature of the current study using non-manipulated stimuli, it is understandable that the results showed no effect of duration as seen in Yunji Kim (2013).

Some of the research that suggests duration of stimuli plays a role in the assessment of attractiveness (especially studies that suggest shorter durations are considered more attractive) could have achieved such results due to the practice of manipulating the stimuli. Ferdenzi et al. (2013) and Babel et al. (2014) found that the lengthened stimuli were less attractive than the original shorter stimuli, however, this could have been caused by the PSOLA (Pitch Synchronous Overlap and Add) manipulation algorithm used to lengthen the stimuli. Lengthening the stimuli in such a way could have caused it to sound unnatural which may have biased the attractiveness ratings towards the original shorter, more natural sounding, duration stimuli. Furthermore, the majority of previous research does not include the use of sentences as stimuli, which the current study does. The current study also uses socially relevant stimuli to assess the relationship between pitch and attractiveness, which has not been investigated before to our knowledge. These methodological differences and the limited literature regarding the effect of stimulus duration on attractiveness ratings for pitch makes it difficult to compare the current study to previous research, although the results do align with the findings of Ferdenzi et al. (2013) and Yunji Kim (2013).

Although using natural, non-manipulated stimuli increased the ecological validity of the current study and avoided the unnatural sounding stimuli of the PSOLA algorithm, it may also have caused some limitations. Yunji Kim (2013) found that when certain vocal features were controlled, duration showed a statistically significant relationship with attractiveness ratings, whereby an increase in duration contributed to an increase in attractiveness ratings. Since the current study used secondary data, we were not able to control any vocal features and there may have been other elements that could have affected the attractiveness ratings. Higher intensity of speech rate, for example, has been found to have a positive impact on voice attractiveness, increased male speech tempo (the number of speech units of a given type produces within a given time) may also result in higher attractiveness ratings from females, and even men speaking “with a smile” may sound less-masculine which could lead to lower voice attractiveness ratings (Avery and Liss., 1996; Quené et al., 2016; Lewandowksi et al., 2018). Furthermore, there are factors related to the female participants that may have influenced the relationship between ratings of attractiveness and speech variables. Women using oral contraceptives show weaker vocal masculinity preferences, and ovulating women have a stronger preference for low-pitched male voices (Feinberg et al., 2006; Feinberg et al., 2008). Controlling for the menstrual phase of female participants could therefore provide further insights into the effect of duration on attractiveness ratings throughout cycles.

The results of the research in this field are difficult to generalise across languages and cultures due to participant selection. Mahrholz et al. (2018) advertised for Scottish participants, however, previous literature uses Korean, French, Californian English, and other languages (Yunji Kim., 2013; Ferdenzi et al., 2013; Babel et al., 2014). It has been found that there is variability in attractiveness ratings in different accents of English and that the vocal properties that listeners pay attention to vary across languages (Bresnahan et al., 2002; Baus et al., 2019). This restricts the generalisability of the findings in this field and makes it difficult to determine if duration may have an effect in one language but not another. Furthermore, while some studies ask participants to report their sexuality, other studies, and the current study, assume that participants are heterosexual (Ferdenzi et al., 2013). Most studies within this field fail to disclose any sexual orientation information. This furthers the challenge in concluding the effect of duration on attractiveness ratings since all the participants may not have been heterosexual, and the results may only apply to heterosexual individuals and heterosexual attractiveness ratings. When considering the methodological differences and the challenges in generalising studies in this area with vastly different participant criteria, the effect of duration of stimuli on the relationship between pitch and attractiveness becomes inconclusive.

CONCLUSION

In the current study, we did not find a significant difference between the conditions possibly due to it being underpowered. As an underpowered study, there was not a sufficiently large enough sample size to answer the research question. Further research could therefore aim to repeat this analysis using a larger sample size with the potential to find a different result. It could also aim to do so with greater control of the vocal features presented in the voice recordings as seen in Yunji Kim (2013). This report highlights a gap in voice research regarding stimuli durations and the use of sentences to assess this. Since previous research has suggested an evolutionary link to voice attractiveness that may have become hard-wired before the advent of language, it would be interesting to explore where longer duration stimuli of full sentences and the effect of context that conveys emotional and social information fit into this signal of reproductive status (Pell and Kotz., 2011; Zhang et al., 2020). Furthermore, including a wider variety of participants with different sexual orientations would allow consideration of sexual selection in participant voices and participant ratings that do not identify as heterosexual.

The findings of this report show no significant difference in the relationship between attractiveness and pitch for longer or shorter duration conditions. However, voice research is complex and multidimensional. Hence, it would require a larger sample with greater control of acoustic variability to investigate the effect of duration in the detail necessary to form an evidence-based conclusion. This report does, however, highlight an area of voice research that could benefit from further study, especially due to its potential in aiding the understanding of vocal feature evolution and the development of sexual selection.

ACKNOWLEDGMENTS

Many thanks for Dr. McAleer for his support and guidance with this project.

REFERENCES

Apicella, C. L., Feinberg, D. R., Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. Biology Letters, 3 (6), 682-684.

Arnocky, S., Hodges-Simeon, C. R., Ouellette, D., Albert, G. (2018). Do men with more masculine voices have better immunocompetence? Evolution and Human Behavior, 36, 602–610. doi: 10.1016/j.evolhumbehav.2018.06.003

Audacity Team (2020). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 2.4.2 retrieved Sep 20th 2020 from https://audacityteam.org/

Avery, J.,and Liss, J. (1996). Acoustic characteristics of less masculine-sounding male speech. JASA, 99 (6), 3738–3748.

Babel, M., McGuire, G., King, J. (2014). Towards a More Nuanced View of Vocal Attractiveness. PLoS ONE 9(2) https://doi.org/10.1371/journal.pone.0088616

Baus, C., McAleer, P., Marcoux, K., Belin, P., Costa, A. (2019). Forming social impressions from voices in native and foreign languages. Sci Rep 9, 414 https://doi.org/10.1038/s41598-018-36518-6

Biesanz, J., Human, L., Paquin, A., Chan, M., Parisotto, K., Sarracino, J. (2011). Do We Know When Our Impressions of Others Are Valid? Evidence for Realistic Accuracy Awareness in First Impression of Personality. Social Psychology and Personality Science, 2(5), 452-459.

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y., Shearman, M. S. (2002). Attitudinal and affective response toward accented English. Language and Communication, 22(2), 171–185.

Bruckert, L., Bestelmeyer, L. M., Rouger, J., Charest, I., Rousselet, G., Kawahara, H. (2010). Vocal attractiveness increases by averaging. Current Biology, 20(2), 116-120

Bruckert, L., Liénard, J. S., Lacroix, A., Kreutzer, M., Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. Proceedings. Biological sciences, 273(1582), 83–89. https://doi.org/10.1098/rspb.2005.3265

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0.

Collins, S., and Missing, C. (2003). Vocal and visual attractiveness are related in women. Animal Behaviour, 65(5), 997-1004.

Diedenhofen, B., and Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. PLoS ONE, 10(4).

Evans, S., Neave, N., Wakelin, D., Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. Physiology and Behavior, 93(4-5), 783–788. https://doi.org/10.1016/j.physbeh.2007.11.033

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. Animal Behaviour, 69(3), 561-568. https://doi.org/10.1016/j.anbehav.2004.06.012

Feinberg, D. R., Jones, B. C., Law-Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., Hillier, S. G., Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. Hormones and Behavior. 49, 215–222. doi:10.1016/j.yhbeh.2005.07.004

Feinberg, D. R., DeBruine, L. M., Jones, B. C., Little, A. C. (2008). Correlated preferences for men's facial and vocal masculinity. Evolution and Human Behavior, 29, 233–241.

Ferdenzi, C., Patel, S., Mehu-Blantar, I., Khidasheli, M., Sander, D., Delplanque, S. (2013). Voice attractiveness: influence of stimulus duration and type. Behav Res Methods, 45(2), 405-13. doi: 10.3758/s13428-012-0275-0. PMID: 23239065.

Gerger, G., Forster, M., Leder, H. (2017). It felt fluent but I did not like it: Fluency effects in faces versus patterns. The Quarterly Journal of Experimental Psychology, 70(4), 637–648. https://doi.org/10.1080/17470218.2016.1145705

Hodges-Simeon, C. R., Gaulin, S. J. C., Puts, D. A. (2011). Voice Correlates of Mating Success in Men: Examining "Contests" Versus "Mate Choice" Modes of Sexual Selection. Arch Sex Behav 40, 551–557. https://doi.org/10.1007/s10508-010-9625-0

Hodges-Simeon, C. R., Gurven, M., Gaulin, S. J. C. (2015). The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. Evolution and Human Behavior, 36 (4), 294-302. 10.1016/j.evolhumbehav.2015.01.002

Hughes, S. M., Farley, S. D., Rhodes, B. C. (2010). Vocal and Physiological Changes in Response to the Physical Attractiveness of Conversational Partners. J Nonverbal Behav 34, 155–167. https://doi.org/10.1007/s10919-010-0087-9

Jones, B. C., Feinberg, D. R., Debruine, L. M., Little, A. C., Vukovic, J. (2008). Integrating cues of social interest and voice pitch in men's preferences for women's voices. Biology letters, 4(2), 192–194. https://doi.org/10.1098/rsbl.2007.0626

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. Animal Behaviour, 79, 57–62. doi:10.1016/j.anbehav.2009.10.003

Jost, L., Fuchs, M., Loeffler, M., Thiery, J., Kratzsch, J., Berger, T., Engel, C. (2018). Associations of sex hormones and anthropometry with the speaking voice profile in the adult general population. Journal of Voice, 32, 261–272. doi:10.1016/j.jvoice.2017.06.011

Lewandowski, N., Duran, D., Bruni, J. (2018). Acoustic and Social Correlates of Perceived Voice Attractiveness. Proceedings of the Conference on Phonetics and Phonology in German-speaking countries, 113-116. http://dx.doi.org/10.18452/18805

Mahrholz, G., Belin, P., McAleer, P. (2018). Judgements of a speaker's personality are correlated across differing content and stimulus type. PLoS ONE 13(10): e0204991. https://doi.org/10.1371/journal.pone.0204991

Mazur, A., and Booth, A. (1998). Testosterone and dominance in men. Behavioral and Brain Sciences, 21(3), 353–397. https://doi.org/10.1017/S0140525X98001228

McAleer, P., Todorov, A., Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. PLoS ONE 9(3): e90779. https://doi.org/10.1371/journal.pone.0090779

Pearson, K., and Filon, L. N. G. (1898). Mathematical contributions to theory of evolution: IV. On the probable errors of frequency constants and on the influence of random selection and correlation. Philosophical Transactions of the Royal Society of London, Series A, 191, 229-311. doi:10.1098/rsta.1898.0007

Pell, M., and Kotz, S. (2011). On the time of course of vocal emotion recognition. PloS One, 6(11).

Penton-Voak, I., Perrett, D., Castles, D., Burt, M., Koyabashi, T., Murray, L. (1999). Female preferences for male faces change cyclically. *Nature, 399*, 741-742.

Pisanski, K., and Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. J Acoust Soc Am 129, 2201–2212.

Pisanski, K., Jones, B., Fink, B., O'Connor, J., DeBruine, L., Röder, S. (2016). Voice parameters predict sex-specific body morphology in men and women. Animal Behaviour, 112, 13-22.

Puts, D. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior, 26*, 388-397.

Quené, H., Boomsma, G., Erning, R.V. (2016). Attractiveness of male speakers: Effects of voice pitch and of speech tempo. Speech prosody, 1086-1089.

Rashidi, M., Pazhoohi, F., Hosseinchari, M. (2012). Effect of facial stimuli exposure time on evaluation of facial attractiveness. Australian Journal of Psychology, 64(3), 164–168. https://doi.org/10.1111/j.1742-9536.2011.00050.x

Re, D.E., O'Connor, J.J., Bennett, P.J., Feinberg, D.R. (2012). Preferences for very low and very high voice pitch in humans. PLoS One 7(3):e32719. doi: 10.1371/journal.pone.0032719

Stróżak, P., and Zielińska, M. (2019). Different processes in attractiveness assessments for unattractive and highly attractive faces—The role of presentation duration and rotation, Acta Psychologica, Volume 200, 102946, ISSN 0001-6918, https://doi.org/10.1016/j.actpsy.2019.102946.

Suire, A., Raymond, M., Barkat-Defradas, M. (2019). Male Vocal Quality and Its Relation to Females' Preferences. Evol Psychol 17(3). doi: 10.1177/1474704919874675. PMID: 31564128.

Team, R. C. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Thornhill, R., and Gangestad, S. W. (1999). The scent of symmetry: A human sex pheromone that signals fitness? Evolution and Human Behavior, 20(3), 175–201. https://doi.org/10.1016/S1090-5138(99)00005-7

Wickham et al. (2019). Welcome to tidyverse. Journal of Open Source Software, 4(43), 1686.

Willis, J., and Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. Psychological Science (Wiley-Blackwell);17(7):592-8. doi: 10.1111/j.1467-9280. 01750.x 16866745.

World Medical Association. (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bulletin of the World Health Organization, 79 (4), 373 - 374. World Health Organization. https://apps.who.int/iris/handle/10665/268312

Xu, Y., Lee, A., Wu, W.L., Liu, X., Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. PLoS One 8(4), e62397. doi: 10.1371/journal.pone.0062397. PMID: 23638065; PMCID: PMC3634748.

Yunji Kim. (2013). Acoustic Variables as Predictors of Vocal Attractiveness. Doctoral dissertation. Seoul National University Graduate School. Accessed via https://s-space.snu.ac.kr/handle/10371/131938

Zhang, H., Liu, M., Li, W., Sommer, W. (2020). Human voice attractiveness processing: Electrophysiological evidence. Biological Psychology, 150, 107827. DOI: 10.1016/j.biopsycho.2019.107827.