UNIVERSITY OF ICELAND

Lokaverkefni til B.Sc. prófs

í sálfræði

Cross-Modal Correspondence

The influence of brightness on pitch perception

Alexandra Inga Alfredsdóttir
Sóley Hólm Jónsdóttir

Leiðbeinandi: Dr. Sabrina Hansmann-Roth, júní, 2024

SÁLFRÆÐIDEILD

Cross-Modal Correspondence

The influence of brightness on pitch perception

Alexandra Inga Alfredsdóttir

Sóley Hólm Jónsdóttir

Lokaverkefni til B.Sc. gráðu í sálfræði

Leiðbeinandi: Dr. Sabrina Hansmann-Roth

Sálfræðideild

Heilbrigðisvísindasvið, Háskóli íslands

Júní, 2024

Cross-Modal Correspondence

The Influence of Brightness on Pitch Perception

Ritgerð þessi er 15 eininga lokaverkefni til B.Sc. prófs við Sálfræðideild á Heilbrigðisvísindasviði Háskóla Íslands

Abstract

Cross-modal correspondence is a perception mechanism based on stimulus compatibility. In this study we explore the cross-modal correspondence between brightness and pitch. Previous research has suggested that we are more likely to choose bright colors in congruence with higher pitch and dark colors in congruence with lower pitch. In accordance with previous research our hypothesis was that brighter visual stimuli corresponds to higher pitch. We conducted two experiments to test this hypothesis. 12 observers participated in the experiments, five in Experiment 1 and seven in Experiment 2. The observers were presented with a visual stimulus combined with an auditory one. The visual stimulus was a screen showing different brightness values and the auditory stimulus were different pure tone values. This was followed by a fixation cross before showing a second stimulus combination. The task was to indicate whether the first or second tone had a higher frequency. In Experiment 1 the maximum likelihood conjoint measurement (MLCM) was used to measure how two physical dimensions (brightness and pitch) can affect perception (pitch perception). The results from Experiment 1 did not support the hypothesis but did indicate that extreme brightness values would potentially have more influence on pitch perception. In Experiment 2 the brightness values were only two, black and white, and psychometric functions were used to analyze the data. Results from Experiment 2 did support the hypothesis where most observers perceived a tone to be higher if paired with a white visual stimulus. Our results suggest that extreme values of stimulus intensity are more suitable for the evaluation of cross-modal correspondence on brightness and pitch.

Acknowledgements

We would like to thank our instructor Dr. Sabrina Hansmann-Roth for giving us a chance, trusting us with this project and helping us design it. Her incredible patience and guidance throughout the process has been invaluable and we are grateful to have been part of it. We would like to thank everyone who participated in the study. We would also like to thank our friends and family for all their help and understanding.

Table of Contents

Introduction

Perception is usually a combined experience from multiple senses. Seeing, hearing, touching or even smelling something can be unified into a meaningful individual experience that the person next to you wouldn't share even if exposed to the same combination of stimulus as you. However, as humans we tend to be more susceptible to the correspondence of certain attributes of the stimuli that make up those experiences. How does the brain repeatedly combine stimulus features from unrelated dimensions such as light or color with sounds, creating a perceptual experience that is consistent within the human population?

The brain's information process is built on a myriad of neural connections. Some of these neural connections represent low-level sensory processes such as *synaesthesia,* a phenomenon that causes perception simultaneously in two sensory modalities by stimulating only one (e.g. hearing stimulating visual perception). While this phenomenon is ultimately a rare ability it has been suggested to be a part of a much broader sensory process that affects most humans (Baron-Cohen, 1996; Ramachandran & Hubbard, 2001). *Cross-modal correspondence* is the systematic processing of a variety of inputs from multiple senses. A concept that is presumably an innate tendency we have to facilitate information processing of our environment. The feature dimensions of cross-modal correspondence that have been studied include brightness, pitch, size, weight, shapes etc. (Bernstein & Edelstein, 1971; Walker, 2017; Marks, 1989; Ramachandran & Hubbard, 2003).

Synaesthesia

Synaesthesia is an involuntary and automatic phenomenon where one sense modality is stimulated and triggers a perception experience in another modality simultaneously. Synaesthesia can happen in two unrelated modalities, an example of this is when a sound triggers the perception of a color. It could also happen intermodally e.g. vision, as seen in those who perceive colors when looking at numbers or letters (Ramachandran & Hubbard, 2001). According to Baron-Cohen et al. (1996) it seems that synaesthesia is usually an inherited trait that affects about 5% of the population, 95% of which are women. That does not mean it's entirely genetic but can also be acquired after an injury to the brain or nerves though how it develops remains a question.

A case documented by Armel and Ramachandran (1999) involved a person acquiring synaesthesia following a disease that damaged their retina leaving them blind. The person reported color perception following tactile stimulation (using Semmes Monofilaments) of the arms with increase in vividity when the arms were located in what was previously the visual field. In a way this closely resembles the experience of patients who feel tactile sensation in phantom limbs when mirroring the intact limb while touching it. Although the acquisition of synaesthesia through injury is possible the odds are it's something people are born with. The most common type of synesthesia is the grapheme-color type (Ward & Simmer, 2022). In their study, Ramachandran, and Hubbard (2001) disclose a subject who experiences number-evoked colors that do not appear in real life. So, it begs the question, how can you perceive something you have never experienced? Previous research suggests a genetic component affecting the development of neural pathways including insufficient pruning in neighboring brain areas (Baron-Cohen et. al, 1996; Ramachandran & Hubbard, 2001). The V4 is the main color center of the brain's visual system and lies in the fusiform gyrus. Using functional magnetic resonance imaging (fMRI), it has been established that the V4 becomes activated in colored hearing synaesthetes in absence of a visual stimulus. They perceive colors when they hear spoken words even though their eyes are closed. This did not occur with non-synaesthetics even with training (Nunn et. al, 2002). Ramachandran and Hubbard's (2001) subject also reported experiencing colors in accordance with facial expressions. Since the fusiform gyrus is known to have cells that respond to faces and objects and the V4 lies in the fusiform gyrus it would further support the hypothesis of cross-wiring. Synaesthetic responses are usually restricted to basic stimulus features and have been speculated to be a subtype of a more universal phenomenon called cross-modal correspondence (Sagiv & Ward, 2006; Spence, 2011).

Cross-modal correspondence

Even though synaesthesia is not present there seems to be a commonality in how people perceive things. A study Ramachandran and Hubbard (2003) performed, though originally by Köhler in 1929, demonstrated how everyone might experience synaesthesia on certain levels. They showed people two different pictures of shapes, one shaped like an inkblot and the other one like shattered glass. Next the participants were presented with two names, “bouba” and “kiki” and instructed to match each name with a shape. Their results showed that 98 percent of people paired the name “bouba” to the inkblot and the name “kiki” to the shattered glass. They wondered if it could stem from the way the sounds match the curves of the shapes or if the pronunciation, we make with our lips match the sharpness or curved sound of the shapes. Unlike true synaesthesia where perception of both modalities happens simultaneously, cross-modal correspondence is rather a perception mechanism for stimulus compatibility. For example, extreme attributes of a stimuli on one dimension correspond to extreme attributes on another dimension. An important factor is that the stimuli are available to many people e.g. light and sounds (Spence, 2011).

The audio-visual cross-modal correspondence has sparked the most interest in research. It has been repeatedly established that people are more likely to pair a bright color to a tone with a higher pitch and darker colors with a lower pitched tone (Marks, 1989; Spence, 2011). Chimpanzees seem to show this tendency as well which could indicate an evolutionary factor for humans (Ludwig et. al, 2011). This cross-modal matching also seems to affect where the stimulus is located, with bright, high-pitched stimulus at higher visual elevation, while darker, low pitched at lower visual elevation (Spence, 2011).

Other feature dimensions that have not received as much attention are the size and weight of objects in correspondence with perceived brightness and pitch. Walker et. al (2017) argued that heaviness of an object is clearly a feature dimension aligning the others and conducted a study exploring exactly that. Their results showed that bigger and heavier objects are perceived as dark and low pitched and smaller and lighter objects more bright and higher pitched. This correspondence occurred both when participants could not lift the objects (they only had information about size but not weight), and when participants could lift the objects (they had information about size and weight). Their results indicate that an object's weight has a stronger influence on perceived brightness and pitch than size does. When people had information about the actual weight of the objects rather than just the perceived heaviness of them, the size both influenced the judgment of brightness and pitch even though weight had a stronger influence. The results are however in reverse of how participants perceived the brightness and pitch when they could not lift the objects and not matching how this correspondence would be elsewhere, where increase in size was paired with more brightness and higher pitch (Walker et. Al, 2017).

The relevance of cross-modal correspondence to our perception experience is the real assessment. How does this impact human information processing? Tasks determining attention or response level such as the speeded classification tasks (Bernstein & Edelstein, 1971; Evans & Treisman, 2010), are performed to see if there’s a difference in people's response time to congruent stimuli opposed to incongruent. Bernstein and Edelstein (1971) conducted one of the first experiments using speeded classification task with auditory-visual stimuli. The participants had to discriminate where a visual stimulus was presented while hearing an auditory stimulus irrelevant to the task. The results showed that a cross-modal correspondence appears between visual elevation and auditory pitch where reaction time was faster when visual elevation was congruent with the pitch. While Marks' (1989) experiments have led to correspondence between brightness and pitch, the same effect has not been confirmed between pitch and contrast. Evans & Treisman (2010) demonstrated cross-modal correspondence for visual position, size, and spatial frequency in congruence with auditory features. In their task manipulating the contrast did however not confirm an association with pitch, which is surprising given that change in contrast also changes the level of brightness. Shape has also shown congruence to pitch, where the reaction time is faster in a discrimination task when a high-pitched sound is matched with an upside-down V and a low-pitched sound with an upside-down U (Marks, 1987).

Spence's (2011) summary of studies that use the speeded classification task on cross-modal correspondence has shown that when pitch is used as the auditory stimulus it corresponds with multiple feature dimensions. Along with the ones mentioned earlier, it also features lightness, and direction. When loudness is used as an auditory stimulus it only shows cross-modal correspondence with visual brightness but not lightness. There's a possibility that the speeded classification task is an insufficient measurement to the dimensions of contrast, hue and lightness corresponding to pitch or loudness. Another is that those dimensions simply do not correspond to the particular feature dimensions measured so far. These results suggest that when stimuli are congruent it facilitates information processing.

Three types of cross-modal correspondence have been suggested. First would be the structural one, based on neural mechanisms to code our sensory information. This includes how increase in stimulus intensity is accompanied by increase in neural firing, or how two sensory dimensions are coded in nearby brain areas as seen with V4/fusiform gyrus in synaesthesia (Spence, 2011; Ramachandran & Hubbard, 2001; Nunn et. al, 2002). Previous research suggests that cross-modal correspondence takes place at a low-level process, at least in the case of the audio-visual one (Marks, 1987; Zeljko et. al, 2019). Even at a low-level process inputs in one sensory modality can influence how another modality perceives a stimulus. Zeljko et. al (2019) concluded that auditory input can affect visual processing before visual sensory analysis is complete.

The second is the Bayesian integration theory. It's a statistical cross-correspondence that reflects our brains adaptive responses to our environment resulting in learned associations of stimulus attributes in our natural surroundings. These naturally co-occurring features include size and weight, where naturally we assume bigger equals heavier based on prior knowledge or coupling priors (Walker et. al 2017; Kwak et. al, 2020). The findings of Zeljko et. al (2019) are in support of the statistical theory with lightness/pitch and elevation/pitch congruence and argue it's a learned perceptual experience. While explaining the correspondence between naturally occurring stimulus features such as size/weight with this theory seems ideal, not all correspondence have this natural co-occurrence. Therefore, it does not fully explain all cross-modal correspondence leaving out congruency like brightness and weight (Walker, 2017).

The third would be semantically mediated correspondence where linguistic terms are used to describe stimulus on other dimensions. The way high and low are used to describe pitch (Spence, 2011) or heavy to describe an olfactory stimulus (Walker, 2016). Just as the bouba/kiki effect, Marks' (1987) experiment with pitch frequency in correspondence with upside down V and U, could also be considered a semantical correspondence regarding the sounds that the letters produce when said out loud. Since there are no exact matches in different modalities, semantic mediated correspondence is purely contextual. Investigating the contribution of different physical dimensions in the environment on perception, requires precise psychophysical methods. Here, we will apply two different methods to determine the influence of two physical dimensions: brightness and pitch on perceived pitch. First, we will address the influence by using Maximum likelihood conjoint measurement to directly quantify the influence of brightness on perceived pitch. Second, we will fit psychometric functions to the data to analyze the contribution of brightness on pitch.

Measurement - MLCM and Psychometric Functions

The *maximum likelihood conjoint measurement* (MLCM) model can be used to show how more than one physical dimension can affect perception (Knoblauch & Laurence, 2012). Ho et.al (2008) used the conjoint measurement model to understand how two surface properties, glossiness, and bumpiness, are perceived. When people judge how a stimulus is perceived, if presented simultaneously with another stimulus, it is hard to focus on just one or the other. In their study they found out that people tend to perceive shinier surfaces to be more bumpy and bumpier surfaces to be shinier. When using the conjoint measurement, they could see how perception at one level influences how perception is perceived on another level. In Experiment 1 we can use the conjoint measurement to do the same, see how perception at one level affects perception at another, but we will focus on brightness and tone frequency.

The main focus for a psychometric function is to look at a person's discrimination ability (Falmagne, 1982). The method describes how intensity from one physical dimension can influence an observer's response accuracy on a task. In Experiment 2 we can use the psychometric function to understand how brightness influences the observer's response to tone frequency. It can also be used to see if the color has enough influence on the perceived pitch in a way that the response changes because of the contribution of color.

The Current Study

The objective of this study is to determine the effects of an irrelevant visual stimulus on perceived pitch. This will be done with two separate experiments. In Experiment 1 the focus is on the cross-modal correspondence effect using the MLCM method to determine the influence of visual brightness on auditory pitch. In Experiment 2 we continue our study but focus on the extremities of brightness values as visual stimulus, anticipating the extreme values show a more defining effect on pitch perception than the moderate brightness values of Experiment 1. We expect the outcome of both experiments to be that brighter colors correspond to higher pitch levels, in accordance with previous studies (Marks, 1987; Spence, 2011).

Method

Participants

A total of twelve observers participated in the experiments, five in Experiment 1 and seven in Experiment 2. The male to female ratio was 6:4 with their ages ranging from 22-39 with a mean age of 28. In Experiment 1 each participant completed two sessions on separate days. In Experiment 2 the participants completed one session. All participants were associates of the researchers and received no payment for their participation. Participants signed informed consent before participating in the experiment.

Apparatus

The experiment was run on a 27-inch ASUS ROG PG279 monitor with the resolution of 1920x1080. The experiment was designed in PsychoPy v2021.2.3 (Pierce et.al., 2022) using MacBook Pro. The conditions file for PsychoPy was done in Microsoft Excel (Microsoft Corporation, 2018). The auditory stimulus was presented through Bose QuietComfort 35 ii headphones in speech volume. Participants also used a keyboard. The experiment took place in a dark room and participants were instructed to not use any equipment that lights up e.g. phones or smartwatches.

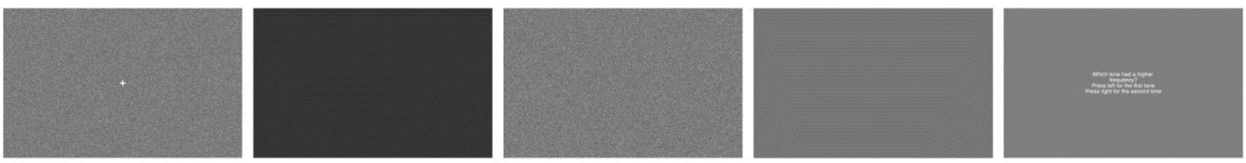
Procedure

To begin, the participants were given consent forms with information to sign before researchers explained the task at hand, followed by a brief training session to make sure they understood the procedure. When they were ready to begin, the researcher turned off the light and left the room. To indicate the start of every trial a white fixation cross appeared on a noise patch background for 500ms. Following these two stimuli were presented that consisted of a color that covered the whole screen while listening to a clear tone simultaneously. The stimuli were present for 1 second (1000ms) followed by a noise patch for 1 second as well before presenting the second stimuli for 1 second (see figure 1 for better understanding). After each trial the participants were asked to indicate with the left or right arrow key which tone frequency was higher. The instructions stated that the participant should press the left arrow key on the keyboard if they experienced the first auditory stimulus to have a higher frequency and the right arrow key on the keyboard if they experienced the second auditory stimulus to have a higher frequency. The participants answered using a keyboard that was placed in front of them and after pressing a key the next trial began. Each session took about 50-60 minutes.

Observers partaking in Experiment 1 had to complete two sessions, each containing 650 randomized trials where the visual stimulus took on different shades of gray. In Experiment 2 the observers completed one session consisting of 44 trials presented 15 times, or a total of 660 trials. The visual stimulus consisted of two shades, black and white.

Figure 1

An example of a trial.



Note. a) Fixation cross, b) First stimulus, c) Noise patch, d) Second stimulus, e) Question.

Stimuli

The stimuli in Experiment 1 consisted of four different shades of a gray color and five clear tone frequencies. The gray colors had the following values: -0.0766, -0.5737, -0.8032 and -0.9092. Gray values here correspond to the lightness scale used in PsychoPy that ranges from -1 (back) to 1 (white). An example of the shades used can be seen in figure 1. The tone frequencies that were used were 900 Hz, 967 Hz, 1039 Hz, 1116 Hz and 1200 Hz. Brightness values and frequencies were logarithmically spaced to follow Fechner's law that the relationship between stimulus and perception is logarithmic.

The stimuli in Experiment 2 were two colors, white and black, with the values: 1 and -1 along with a wider range of tone frequencies. The frequencies that were used were 830 Hz, 880 Hz, 934 Hz, 991 Hz, 1051 Hz, 1115 Hz, 1183 Hz, 1256 Hz, 1332 Hz, 1413 Hz and 1500 Hz.

Data analysis

The data was analyzed in the software R (Team, 2018). Statistical significance was assesses using a significant level (α) of 0.05. For Experiment 1 the MLCM package was used to compare three models (Knoblauch & Maloney, 2022). The comparison of models was done with χ2-statistics (chi-square test) using ANOVA. These models were the independent model, additive model, and full model. The independent model assumes that colors have no effect on frequency perception, and only tone frequency influences perceived pitch.

$$\Delta(i, j, k, l) = \psi_i^f - \psi_k^f + \varepsilon$$

The additive model assumes that colors influence frequency perception, and the influence is additive, where the influence of color is independent of the actual frequency.

$$\Delta(i, j, k, l) = \left(\psi_i^f + \psi_j^b\right) - \left(\psi_k^f + \psi_l^b\right) + \varepsilon$$

The full model assumes that there is an interaction between the colors and the frequency. Color will influence perceived pitch and that contribution depends on the actual frequency level.

$$\Delta(i,j,k,l) = \left(\psi_{i}^{f} + \psi_{j}^{b} + \psi_{ij}^{fb}\right) - \left(\psi_{k}^{f} + \psi_{l}^{b} + \psi_{kl}^{fb}\right) + \varepsilon$$

In Experiment 2 we used psychometric functions to analyze the data. Psychometric functions are fitted separately for the two colors (black and white). A shift between the two functions would indicate that one color induces the observer to select the stimulus more often than the other stimulus.

Results

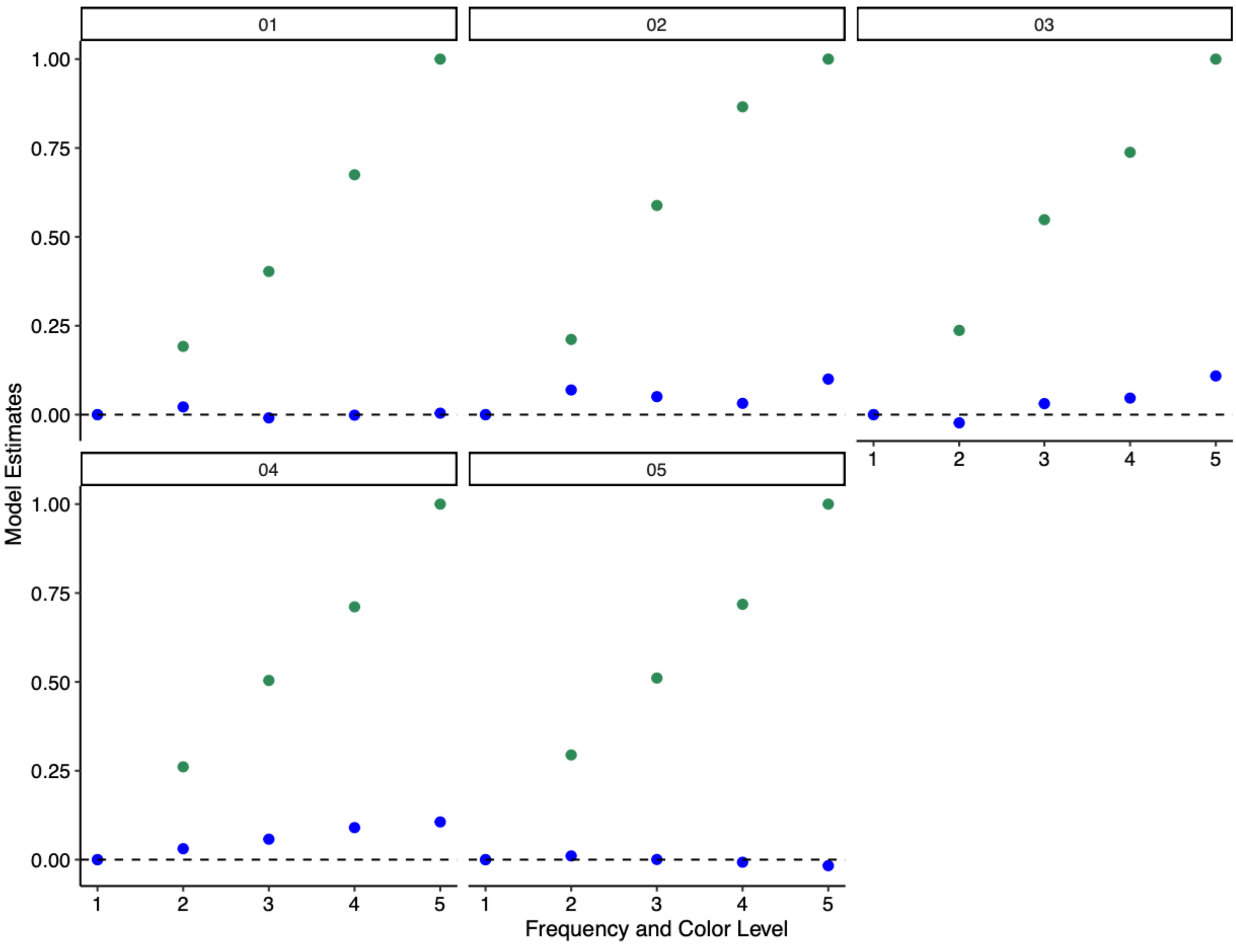
Experiment 1

In Experiment 1 observers were presented with two combinations in succession of stimuli consisting of a brightness level presented as a gray value and a frequency value of a pure tone simultaneously. They were asked to pick which of the frequency they perceived having higher pitch. The goal was to assess if the brightness level influenced the observers perceived pitch. The expectations were that the observer would perceive the frequency to be higher if combined with a brighter visual stimulus. This was measured using the maximum

likelihood conjoint measurement. Figure 2 plots the results for all participants in Experiment 1.

Figure 2

The influence of color on pitch perception



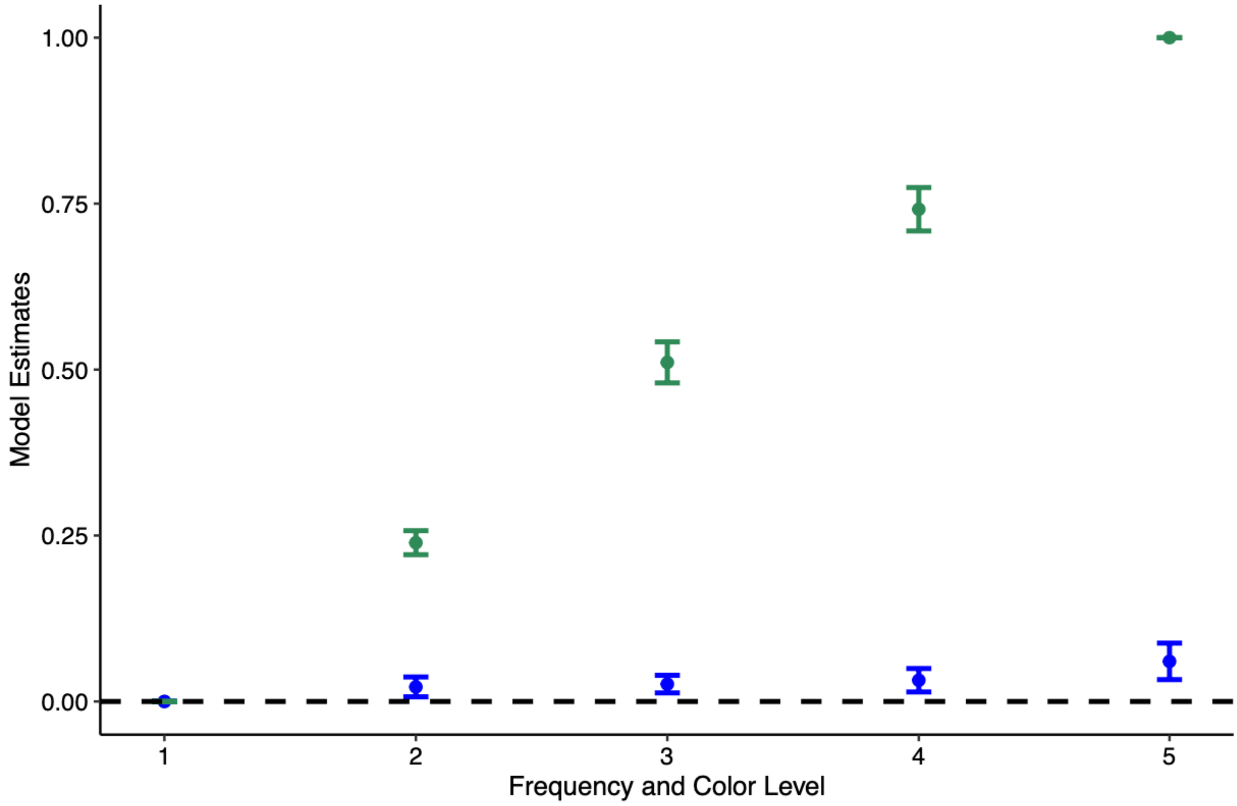
Note. The green data points correspond to the frequency of the stimulus, the blue data points correspond to the contribution of color. The x-axis tells us how frequency of pitch is influenced by the contribution of color, the y-axis corresponds to the normalized parameter estimates of the model.

For participant three $\chi 2(4, N = 5) = 25.13$, $p < 0.01$ and participant four $\chi 2(4, N = 5) = 27.27$, $p < 0.01$, the blue line increases from left to right and shows contribution from the color

($p < 0.01$). The effect color had on perceived pitch for those participants was about 10%. For the other three participants, the blue line does not increase from left to right like it does with participants three and four and was not significant. The color does not significantly contribute anything to the perception of pitch for participant one $\chi2(4, N = 5) = 1.81$, $p = 0.77$ with the effect from color 0,4%, two $\chi2(4, N = 5) = 4.43$, $p = 0.35$ with effect from color 0,099% and five $\chi2(4, N = 5) = 1.73$, $p = 0.79$ with effect from color -0,007%. Participant two did however show an increase for the fifth value of color (the brightest value) but not for any other color values.

Figure 3

The summary of all five observers using the MLCM in Experiment 1.



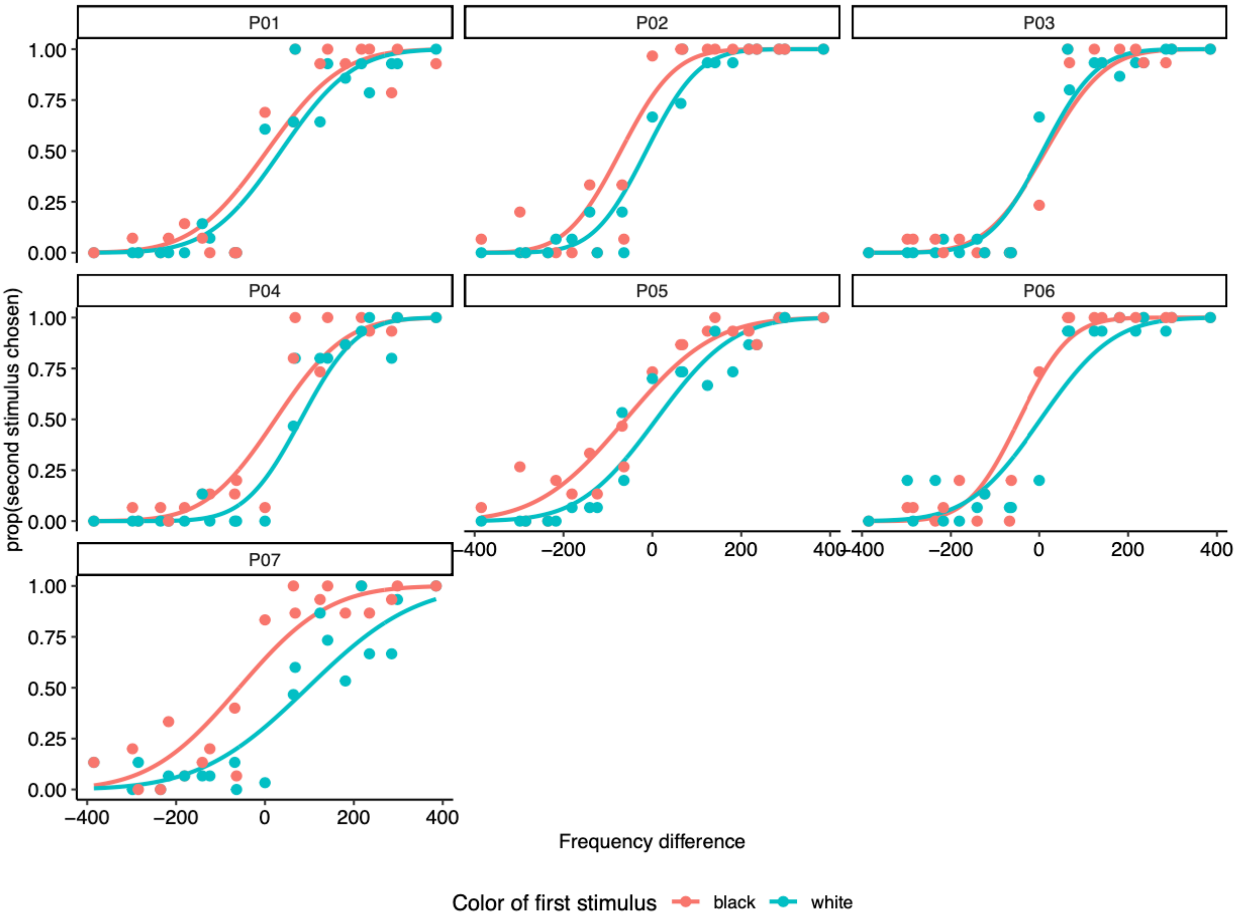
Note. The green data points correspond to the frequency of the stimulus, the blue data points correspond to the contribution of color.

Experiment 2

With the results from Experiment 1 it was decided to follow up with Experiment 2 to assess whether the influence would be more defined using extreme values of brightness instead of the various shades of gray. The observers were asked which of the frequency they perceived as a higher pitch in combination with either a black or a white visual stimulus.

Figure 4

The influence of color on pitch perception using psychometric functions.

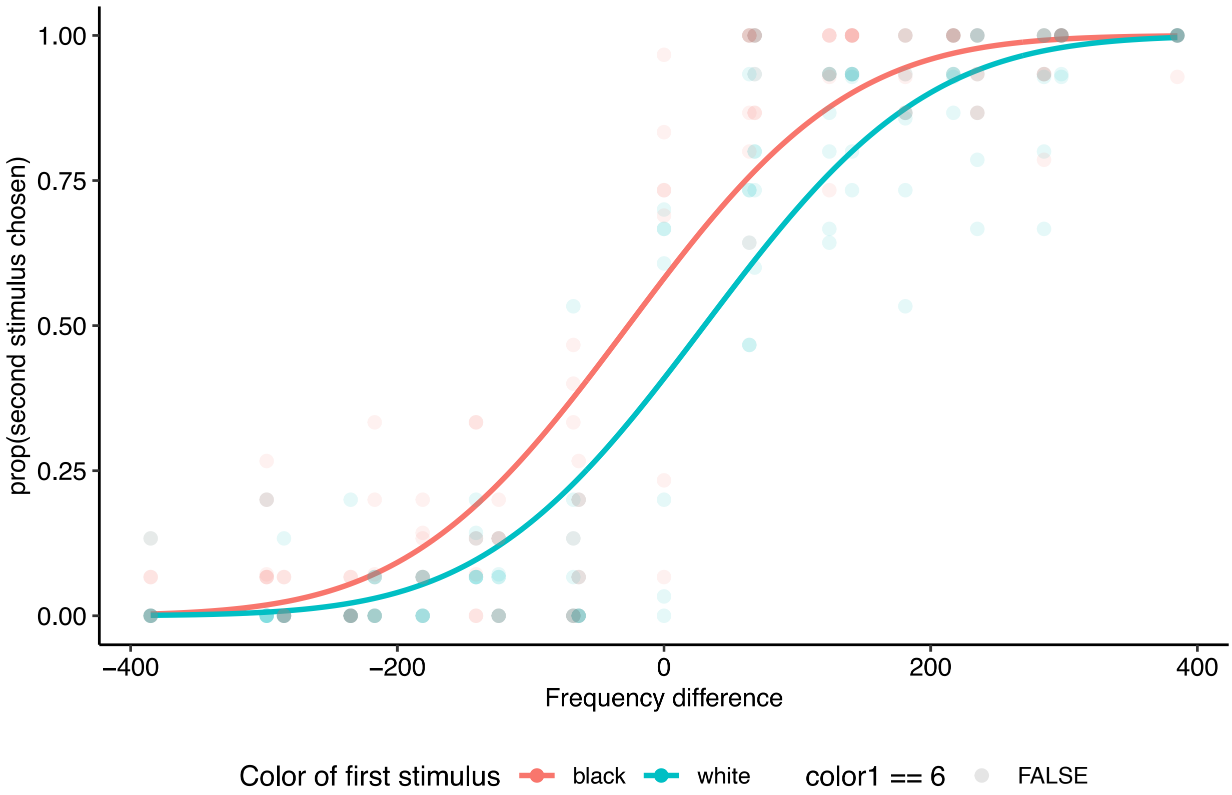


Note. The red curve corresponds to all trials where the first stimulus was black, when shifted to the left it represents participants' response that the second stimulus had higher pitch. The blue curve corresponds to all trials where the first stimulus was white, when shifted to the right it represents the participants' response that the first stimulus had higher pitch.

As shown in figure 4, the color influenced the response of all participants except for participant three. For most participants, the psychometric functions between the white and black stimulus are shifted and do not overlap. When the curves are close together it represents little influence from color, the further they are apart the greater the influence. The curves for participant three are closely aligned with small reverse compared to other participants showing little to no effect of the color on perceived pitch. Participant one showed little influence of color to their response. Participant seven showed the greatest effect so the color has more influence on perceived pitch. A shift in the psychometric function refers to a shift in the response criterion. When the first stimulus was black (red curve), observer more likely chose the second stimulus has having a higher pitch, which is the stimulus that had a white color. When the first stimulus was white, it was more likely chosen than the black stimulus. We extracted the threshold (x-value at y = 0.5) from each participant and each psychometric function. A t-test between the thresholds of the two psychometric functions revealed a significant difference $t(6) = -3.1993$, $p = 0.01862$.

Figure 5

Summary of all observers using psychometric functions in Experiment 2



Note. The red curve corresponds to all trials where the first stimulus was black, when shifted to the left it represents participants' response that the second stimulus had higher pitch. The blue curve corresponds to all trials where the first stimulus was white, when shifted to the right it represents the participants' response that the first stimulus had higher pitch.

Figure 5 shows the combined results of the observers from Experiment 2. The curves do not overlap indicating that the extreme values of brightness do influence the observers perceived pitch.

Discussion

The aim of this study was to determine whether visual brightness influences auditory pitch perception, something that has been previously established in cross-modal correspondence. The speeded classification task is one of the most commonly used ways to determine the effects of cross-modal correspondence. From studies using the speeded classification task, where results show faster response time with stimulus congruent intensities, this study’s expected outcome was that brighter colors would correspond to higher pitch and darker colors would correspond with lower pitch (Bernstein & Edelstein, 1971; Marks, 1987). This study used the maximum likelihood conjoint measurement (MLCM) to determine the effects of an irrelevant stimulus to the one the observers are asked to respond to, in this case pitch perception. To our knowledge the MLCM method has not been used to measure the effects of brightness on pitch perception.

In Experiment 1, the assessment of the full model of the MLCM showed no effect from the colors on pitch perception, rejecting the study’s hypothesis that brightness value influences pitch perception. However, out of five participants the visual brightness of colors seemed to influence two of the participants' response to the pitch and one observer showed only a similar contribution of brightness for only the brightest stimulus. This could be the result of the stimulus intensity of the gray values not contributing to pitch perception. It’s possible that the gray colors were rather perceived as a change in contrast than brightness supporting Evans & Treisman’s (2010) results where brightness levels need to be more defined. Perhaps the first step should have been conducting a baseline experiment asking people to match a range of tones to certain colors and/or different brightness levels to see if there was some sort of universal congruence. Maximum likelihood conjoint measurement assumes additive contribution of all values of the secondary dimension (here brightness). However, the results from the MLCM method to our surprise, gave us the opportunity to explore further the effects of extreme dimensions of feature stimulus. The responses of one participant showed influence to the brightest value used on pitch perception, suggesting that matching stimulus intensity might be relevant in cross-modal correspondence. This was a motivation to conduct Experiment 2.

The color values chosen in Experiment 2 were black and white instead of different brightness values of gray to see if the difference in response was more apparent. All but one participant showed influence from the colors. Participant three was only influenced by the frequency level of the pitch but showed no influence from the contribution of color. For the rest of the participants, when the first stimulus was black, they were more likely to choose the second stimulus to have a higher pitch, knowing the following color would be white. The outcome remained consistent when the first visual stimulus was white resulting in the response corresponding to the prior auditory stimulus. Therefore, suggesting that not only did the color influence the response supporting the hypothesis, but the color already influenced the response before the second tone was presented. This is in support of Zeljko et. al (2019) with the exception that this study showed a reverse effect. In the current study the visual input influenced auditory perception before the observer completed the trial. Overall Experiment 2 provided a more distinct portrayal of how brightness influences perceived pitch, suggesting that extreme values are more suitable to evaluate cross-modal correspondence of brightness on pitch perception. The effect may very well be a product of the semantically mediated correspondence as described by Spence (2011) and Walker (2016). That happens when using the same linguistic terms for different feature dimensions, in this case the increase/decrease of tone frequency corresponds to increase/decrease in brightness: higher=brighter, lower=darker. This also relates to the bouba/kiki effect and Marks' (1987) upside down V and U, where the sound corresponds to the curves of the shapes (Ramachandran & Hubbard, 2003). Because there is no semantical correspondence such as high or low in different gray values to tone frequency the results of Experiment 1 might therefore not have produced the desired effect.

Since some of the trials included the same auditory value of stimulus, observers were forced to choose which pitch was higher even though they were the same. The hope was that responses would even out, but there is no guarantee. This could have resulted in the observer continually preferring one response over the other when forced to choose between the same tones.

When doing the experiment for a while, observers might possibly notice the tone variety and learn what tone is highest leading to an automatic response without giving notice to the latter tone. This could be especially true for people who have studied any form of musicology. Since the task emphasized which tone was higher, if the observer learned the tone variety and began to ignore the visual stimulus the influence of color would not be observed, making the task overall inconclusive. To counter this possible complication, it might be more productive to use virtual reality goggles, or a closed space with a screen that fills the visual field, making it harder to ignore the visual stimulus. Since not many studies have used the MLCM method to demonstrate the correspondence between brightness and pitch, it might be advantageous to use the method for further research on cross-modal correspondence. Even though our results using that method did not confirm the hypothesis, it gave us an opportunity to explore further with the findings from Experiment 1. Giving our findings in Experiment 2 it would be interesting to study further the effect of a prior stimulus on perception of a second stimulus, possibly in accordance with priming. The results from these studies were a further contribution to the research on cross-modal correspondence supporting the theory of stimulus intensity matching.

References

Armel, K. C., & Ramachandran, V. S. (1999). Acquired synesthesia in retinitis pigmentosa. *Neurocase, 5*(4), 293-296. https://doi.org/10.1080/13554799908411982

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: Prevalence and Familiality. *Perception*, *25*(9), 1073-1079. https://doi.org/10.1068/p251073

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, *87*(2), 241-247. https://doi.org/10.1037/h0030524

Evans, K. K., & Treisman, A. (2010). Natural ross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1), 1-12. https://doi.org/10.1167/10.1.6

Falmagne, J. C. (1982). Psychometric functions theory. *Journal of Mathematical Psychology*, *25*(1), 1–50. https://doi.org/10.1016/0022-2496(82)90045-1

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint Measurement of Gloss and Surface Texture. *Psychological Science*, *19*(2), 196-204. https://doi.org/10.1111/j.1467-9280.2008.02067.x

Knoblauch, K., & Maloney, L. T. (2022). MLCM: Maximum likelihood conjoint measurement. R package version 0.4.3. Retrieved from https://cran.r-project.org/web/packages/MLCM/MLCM.pdf

Kwak, Y., Nam, H., Kim, H. W., & Kim, C. Y. (2020). Cross-Modal Correspondence Between Speech Sound and Visual Shape Influencing Perceptual Representation of Shape: the Role of Articulation and Pitch. *Multisensory Research*, *33*(6), 569-598. https://doi.org/10.1163/22134808-20191330

Ludwig, V., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans.

Proceedings of the National Academy of Sciences of the United States of America. *108*(51), 20661-20665. https://doi.org/10.1073/pnas.1112605108.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 384-394. https://doi.org/10.1037/0096-1523.13.3.384

Marks, E. Lawrence. (1989). On Cross-Modal Similarity: The Perceptual Structure of Pitch, Loudness, and Brightness. *Journal of experimental psychology. Human perception and performance*. *15*(3), 586-602. https://doi.org/10.1037//0096-1523.15.3.586.

Microsoft Corporation. (2018). Microsoft Excel. Retrieved from https://office.microsoft.com/excel

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., Morris, R. G., Bullmore, E. T., Baron-Cohen, S. and Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience, 5*(4), 371–375. https://doi.org/10.1038/nn818

Peirce, J. W., Hirst, R. J. & MacAskill, M. R. (2022). Building Experiments in PsychoPy. 2nd Edn London: Sage.

Ramachandran V. S., & Hubbard E. M. (2001). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings. Biological sciences*, *268*(1470), 979-983. https://doi.org/10.1098/rspb.2000.1576

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing Colors, Tasting Shapes. S*cientific American*, *288*(5), 52-59. https://doi.org/10.1038/scientificamerican0503-52.

Sagiv, N., & Ward, J. (2006). Crossmodal interactions: lessons from synesthesia. *Progress in brain research*, *155*, 259-271. https://doi.org/10.1016/S0079-6123(06)55015-0

Team, R. C. (2018). A language and environment for statistical computing. 2013. *R Foundation for Statistical Computing: Vienna, Austria*.

Walker, P., Scallon, G., & Francis, B. (2017). Cross-Sensory Correspondences: Heaviness is Dark and Low-Pitched. *Perception*, *46*(7), 772-792. https://doi.org/10.1177/0301006616684369

Ward, J. & Simner, J. (2022). How do Different Types of Synesthesia Cluster Together? Implication for Causal Mechanisms. *Perception,* *5*(2), 91-113. https://doi.org/10.1177/03010066211070761

Zeljko, M., Kritikos, A., & Grove, P. M. (2019). Lightness/pitch and elevation/pitch crossmodal correspondences are low-level sensory effects. *Attention, Perception, & Psychophysics*, *81*(6), 1609-1623. https://doi.org/10.3758/s13414-019-01668-w