Audio Engineering Society

Convention Paper 10607

Presented at the 152nd Convention
May, 2022, In-Person and Online

Watching on the Small Screen: The Relationship Between the Perception of Audio and Video Resolutions

Nicholas Henry Bartel and Song Hui Chon

Audio Engineering Technology, Mike Curb College of Entertainment & Music Business, Belmont University, 1900 Belmont Boulevard, Nashville, Tennessee 37212, USA

Correspondence should be addressed to NicholasHenryBartel@gmail.com, songhui.chon@belmont.edu

ABSTRACT

A new quality assessment test was carried out to examine the relationship between the perception of audio and video resolutions. Three video resolutions and four audio resolutions were used to answer the question: "Does lower resolution video influence the perceived quality of audio, or vice versa?" Subjects were asked to use their own equipment, which they would be likely to stream media with. They were asked to watch a short video clip of various qualities and to indicate the perceived audio and video qualities on separate 5-point Likert scales. Four unique 10-second video clips were presented in each of 12 experimental conditions. The perceived audio and video quality ratings data showed different effects of audio and video resolutions. The perceived video quality ratings showed a significant effect of audio resolutions, whereas the perceived audio quality did not show a significant effect of video resolutions. Subjects were divided into two groups based on the self-identification of whether they were visually or auditorily inclined. These groups showed slightly different response patterns in the perceived audio quality ratings.

1 Introduction

Many individuals watch films on small devices, such as their laptops or phones, streaming films or videos over the internet as convenience [1, 2]. A stipulation with streaming is buffering, which could cause a dip in video and audio quality at any time from various factors [3, 4]. This occurs when the internet connection tries to stabilise to the required resolution. Network conditions and other factors contribute to the quality of the experience when streaming videos, sometimes to the effect that streaming degradation hinders the viewing experience.

This study investigates the perceived quality differences in various audio and video resolutions and their interactions. The interest was to determine if the lower resolution audio or video affected the perceived quality when paired with higher resolutions in the opposite domain. Stimuli included high-resolution video played with lower bitrate audio, and uncompressed audio with low-resolution video.

The research questions were "Will the perceived audio quality be affected by different video resolutions?" and the opposite "Will the perceived video quality be affected by different audio resolutions?" With streaming on small devices being so popular, and with the quality complications that may come with streaming, the goal of this study was to examine how different combinations of picture and sound resolutions may affect the perception of one another.

With the media industry transitioning further into streaming, it is important to understand how viewers will perceive the quality of the audio and video, along with how they interact together. This project will help us better understand the interaction between higher and lower resolutions in audio and video. The findings from this study will be relevant to streaming media providers.

2 Prior Art

Much of the existing studies on multimedia and perceptual factors are related to the threshold of detectability of audio-visual asynchrony. They have shown that the auditory system responds significantly faster than its visual counterpart; however, visual information has been reported to override the auditory information when there are conflicts in audio and visuals [5–9]. According to these studies, the visual system will generally significantly influence the auditory system, even though visuals are processed more slowly in the brain.

The two most recent studies regarding audio and video quality were conducted in 1999 and 2001 [10, 11]. Unfortunately, they were just at the beginning of the digital era, so what they used as the standard of home media and reproduction during the period is far before high definition or even ultra-high definition media available today. Streaming was also not available during this time. Joley, Montard, and Buttin concluded that the audio significantly influenced the video [11], whereas Berrends and De Caluwe reported the opposite [10]. In 2015, a similar experiment was conducted, although solely on audio changes. The experiment used manipulations of the audio by adding noise or degrading quality to study its perceived impact when paired with the video. The results showed that the perceived video quality ratings of unaltered videos were lower when they were combined with poor-quality audio [12]. In 2014, Schoeffler and Herre conducted a study to examine if there was a relationship between audio quality and the influence it had on user ratings on the quality reduction from a file-sharing website, in this case YouTube. They reported no correlations with the user's preference ratings and the quality degradation from YouTube's compression [13]. In other words, the audio quality did not have any effect on the users' ratings.

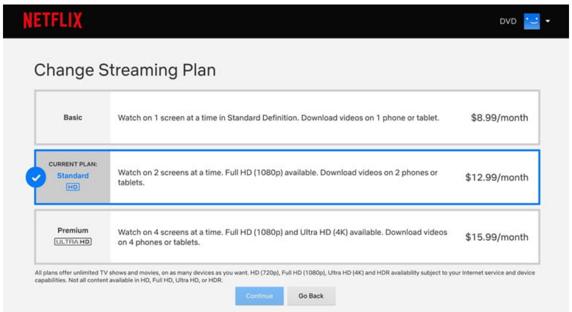

Fig. 1. Netflix streaming plan, including video resolutions per plan (captured from a NETFLIX account on October 16, 2020).

The specifications for various streaming platforms provide the base resolutions for the presented experiment. Netflix reports that they use adaptive bitrate streaming for the audio, the transfer rate of which ranges between 192 kbps to 640 kbps for standard 5.1 arrangements and 448 kbps to 768 kbps for Dolby Atmos arrangements [14, 15]. For video resolutions, Netflix offers three services based on user subscription tiers: "basic," "standard," and "premium" (Fig. 1). "Basic" offers standard definition (while not specified, the current norm for standard definition is 480 progressive scan, or 480p henceforth), whereas "standard" offers Full HD at 1080p. "Premium" offers videos in 1080p and 2160p, also known as 4k or Ultra HD [16, 17]. Amazon Prime Video supports "standard" definition, with a minimum of 640 x 260 pixels, "high" definition (HD), with a minimum of 720p, and "ultra-high" definition (Ultra HD), with a minimum of 2160p [18, 19]. While Amazon Prime Video reports the specifications of their surround sound capabilities, the objective quality of the audio could not be found; however, Amazon Music reports that current "standard audio" is a bitrate up to 320 kbps. Beyond the "standard" quality, Amazon Music also features HD music that has 16-bit encoding and an average bitrate up to 850 kbps, and Ultra HD tracks with 24 bit encoding and an average bitrate of 3,730 kbps [20, 21]. Music streaming providers also report a range of objective audio quality measurements. Spotify reports that transfer speeds for free users can vary among "low" (~24 kbps), "normal" (~96 kbps), and "high" (~160 kbps) conditions (Fig. 2). Spotify allows premium users to reach transfer speeds around 320 kbps [22]. Tidal, a lossless music streaming provider, reports that their streamed quality is 1,411 kbps (which appears to be the FLAC encoding rate, although not clearly mentioned), which is supposed to be equivalent to 16 bit and 44.1 kHz CD quality [23]. Other streaming sites, such as Hulu [24] and YouTube Music [25] have reported similar quality specifications.

Music quality

	Spotify free	Spotify Premium
Web player	AAC 128kbit/s	AAC 256kbit/s
Desktop, mobile, and tablet	**Automatic:** Dependent on your network connection **Low:** Equivalent to approximately 24kbit/s **Normal:** Equivalent to approximately 96kbit/s **High:** Equivalent to approximately 160kbit/s	**Automatic:** Dependent on your network connection **Low:** Equivalent to approximately 24kbit/s **Normal:** Equivalent to approximately 96kbit/s **High:** Equivalent to approximately 160kbit/s **Very high:** Equivalent to approximately 320kbit/s

Fig. 2. Spotify audio quality specifications. (captured from SPOTIFY's support website on April 14, 2022).

Many problems in streaming such as stalling, initial delay, memory requirements, and bandwidth utilization are related to buffer size and segment size; however, a small buffer of a few segment lengths has proven to be sufficient while streaming for most bandwidth conditions [26]. Streaming algorithms are proprietary between services and the specifications are unknown, although there has been an attempt to reconstruct and analyze algorithms with limited success [27]. This means that various streaming platforms are affected differently by suboptimal internet connections, which may affect the audio and video quality differently. Audio quality can be negatively affected by lower internet speeds—often the audio loses clarity and data compression causes undesirable artifacts [28]. The perception of audio and video are multisensory in that they affect each other; therefore, when there is a change in the perceived quality in one domain, it could affect the other sense even without any change in the perceived quality in the relevant domain. By utilizing various audio and video resolutions and combinations, we can study how the perception of one sense affects the other.

3 Methods

The reported resolutions of audio and video from different streaming providers form the basis for the selected resolutions for our experiment: 16 kbps, 24 kbps, 56 kbps, and 320 kbps for audio resolution, and 240p, 480p, and 1080p for video resolution. Therefore, 12 conditions (= 4 audio resolutions × 3 video resolutions) were examined based on the factorial design. We excluded 4K, as we could not expect most subjects to own a screen of that resolution.

3.1 Subjects

Participants were recruited from former and current students in Belmont University, who were 18 years and older. Many subjects were from the Mike Curb College Of Entertainment & Music Business, which encompasses the University's audio engineering, video production, motion picture, music business, and songwriting majors. Some others were from the commercial music program in the School of Music. Therefore, our subjects include individuals who are trained with critical evaluation of media, more so than the general public. Overall, 42 subjects participated in this online experiment.

3.2 Stimuli

There were 48 unique clips of 10-second duration created for the experiment, organized in four sets of 12 clips each. Each set included all 12 conditions while each condition was represented with a different clip. No videos were repeated during this study. All the clips consisted of program material from past senior capstone student films from Belmont University. They were chosen to minimize previous viewing familiarity, which could impact the subjective quality ratings. Consent was obtained from the students who created these films for the use of their work in this experiment. Each film was originally shot in 4K picture and recorded with 48 kHz and 24-bit audio. The selected segments were adjusted in video and audio resolutions for the experiment: The video resolutions consisted of 240p, 480p, and 1080p. The audio resolution consisted of 48 kHz and 24-bit audio with AAC bitrate reductions to 16 kbps, 24 kbps, 56 kbps, and 320 kbits. There were no conditions with 4K video because it was not practical to present in an online test; most subjects could not be expected to own 4K viewing screens.

The entire set of stimuli was exported through Adobe Premiere, which can export video clips at various video resolutions and audio AAC bitrates. Each of the exported clips were sorted into a timeline on VLC Media Player to be screen recorded though OBS Studios. The OBS Studio export settings were set to the highest uncompressed and lossless video quality, while the audio quality was set to 320 kbits. This consolidated the complete set of stimuli into a single video file that can be easily downloaded and watched. A second set of the same clips in a different randomized order was created so as to minimize any unexpected order effects of stimuli presented. This was a compromise necessary in the format of an online experiment, which was the only possible option at the time of this experiment due to COVID-19.

3.3 Procedure

The experiment was accessed through a Qualtrics survey. After the consent and demographic information, subjects were asked to download the appropriate version between the two stimuli files through a Google Drive link. The selection criterion was the subject's birth month; for example, subjects born between the months of January and June were asked to select version A, while those born between July and December version B. The reason for allowing a download of the video instead of embedding it on a website was to minimize any inadvertent effect of the internet connection for the experiment. Subjects were also asked to identify whether they were more auditorily or visually inclined. The purpose of this was to examine whether individuals' inclination would lead to different perception patterns.

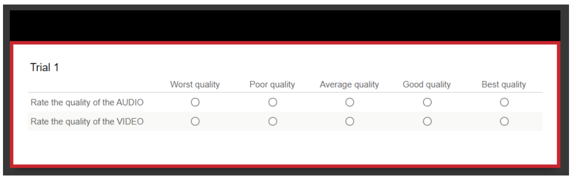

Fig. 3. A screenshot of two Likert scales for perceptual ratings for a trial (captured from Qualtrics).

One of the requirements for this study was to own a computer or laptop with a viewing screen of 1080p or higher and a secondary device, such as a smartphone or tablet. The purpose was for subjects to play the video for the study in the full screen mode without any distractions while answering the questions on their secondary device. The only usage of the primary computer device was to download the video from the Qualtrics link and play it in full screen mode; the remainder of the questions were answered on the secondary device.

The experiment consisted of separate perceptual evaluations of visual and auditory quality. Each trial presented a clip that was exactly 10 seconds long. Once the clip ended playing, subjects had seven seconds to separately rate the quality of the audio and the quality for the video. Seven seconds was selected as an adequate time between trials, so that subjects who take longer to answer had sufficient time while those who answer quickly did not spend much time waiting. Subjects responded using two five-point Likert scales to indicate the perceived audio and video quality separately, with 1 representing the lowest quality and 5 for the highest (Fig. 3). Randomly placed between every 4 to 6 trials, there was a three second clip which consisted of noise and glitch video. The purpose of this was to reset the subject's perception of the stimuli to minimize any inadvertent effects of prior trials. Lastly, there was a five-minute break at the half-point of the experiment (between the second and the third sets), therefore the entire experiment took a total of a little over thirty minutes.

4 Results

Data from 42 subjects showed different effects of audio and video resolutions on the perceived audio and video qualities. First, consider the perceived video quality ratings across all subjects in Fig. 4. The horizontal axis denotes the four audio resolutions, whereas the three video resolutions are specified by three separate lines on the graph. The three lines are clearly separated, which suggests a significant main effect of the video resolution. Indeed, a two-way repeated-measures analysis of variance (ANOVA) on the perceived video quality showed a significant main effect of the video resolution ($F(2,2052) = 347.99$, $p < .001$) and a significant interaction effect between audio and video resolutions ($F(6,2052) = 10.31$, $p < .001$). The audio resolution showed an almost significant main effect, $F(3,804) = 251.27$, $p = .057$.

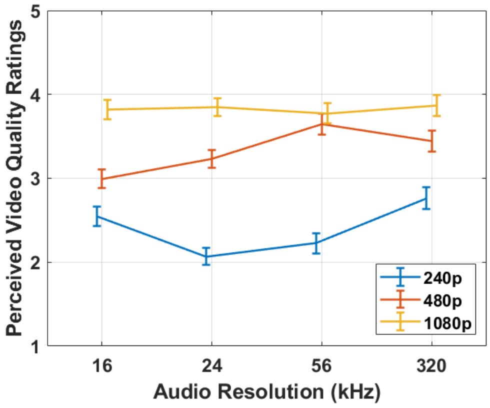

Fig. 4. The average perceived video quality ratings across all subjects. The y-axis represents the video quality ratings and the x-axis represents the audio resolution. Different video resolutions are specified in separate lines. The error bars indicate standard error.

A two-way repeated-measures ANOVA on the perceived audio quality showed a significant effect of the audio resolution ($F(3,2052) = 134.25$, $p < .001$) as did the interaction between audio and video resolutions ($F(6,2052) = 3.47$, $p < .05$). The video resolution did not have any significant main effect ($F(2,2052) = .08$, $p = .92$). This analysis result is in agreement with Fig. 5, where we can clearly see the significant main effect of the video resolution. The significant interaction effect reflects that the audio quality ratings are almost identical for 56 kHz and 320 kHz audio resolutions regardless of the video resolutions combined.

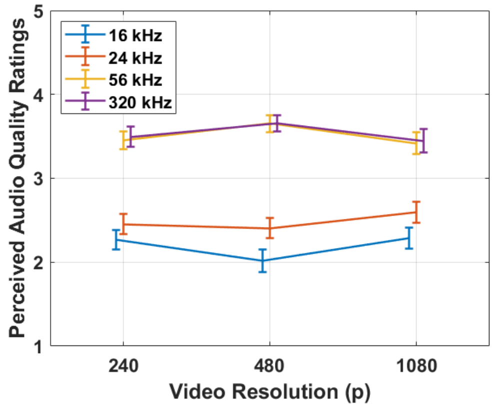

Fig. 5. The perceived audio quality ratings across all subjects. The y-axis represents the audio quality ratings and the x-axis represents the video resolution. Different audio resolutions are specified in separate lines. The error bars indicate standard error.

Since our subjects identified themselves as either visually inclined or auditorily inclined, we also analyzed the ratings including this inclination. In Fig. 6, the two groups (auditorily inclined in the left panel, and visually inclined in the right panel) show the same response patterns for perceived video quality ratings. A three-way repeated-measures ANOVA confirms that there is no significant effect of inclination, $F(1,41) = 0.04$, $p = .845$. Significant main effects were observed for the video resolution, $F(2,82) = 185.91$, $p < .001$, and for the audio resolution, $F(3,123) = 6.06$, $p < .05$. The interaction between audio and video resolutions was also significant, $F(6,246) = 10.53$, $p < .001$. No other effects were statistically significant.

Similarly, the perceived audio quality ratings by the two groups are shown in Fig. 7, which suggests some group differences. A three-way repeated-measures ANOVA on the perceived audio quality ratings show a strong effect of inclination approaching significance, $F(1,41) = 3.33$, $p = .075$. The audio resolution shows a significant main effect, $F(3,123) = 166.84$, $p < .001$. Additionally, the audio resolution shows significant interactions with two factors: $F(3,123) = 2.76$, $p < .05$ with inclination; $F(6,246) = 3.43$, $p < .05$ with the video resolution. No other effects were significant.

Notice that the auditorily inclined group showed lower perceived audio quality ratings on average for 16 kHz and 24 kHz stimuli than the visually inclined group. An independent-samples t test on the 16 kHz ratings showed a significant difference between the two groups, $t(514) = -3.44$, $p < .001$. The results for 24 kHz ratings also showed significant difference between the two groups, $t(514) = -3.41$, $p < .001$. The ratings of stimuli in higher audio resolutions did not show any significant group differences.

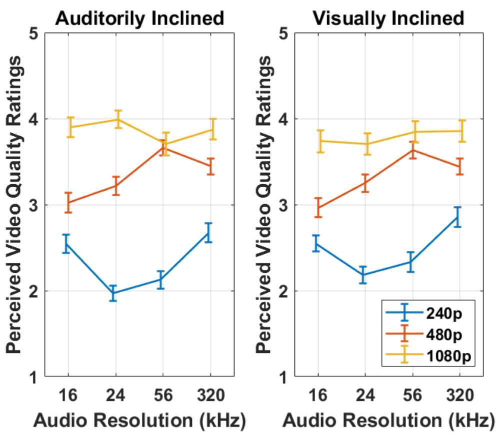

Fig. 6. The perceived video quality ratings from the subjects divided into two groups. The y-axis represents the video quality ratings and the x-axis represents the audio resolution. Different video resolutions are specified in separate lines. The left graph represents the auditorily inclined subjects, whereas the right represents the visually inclined subjects. The error bars indicate standard error.

5 Discussion

In this study, we examined the relationship between the perceived audio and video quality ratings in terms of the audio and video resolutions. Furthermore, the subject group was included in analyses as an additional factor that could impact the perceived quality ratings based on subjects' self-identification as auditorily- or visually-inclined.

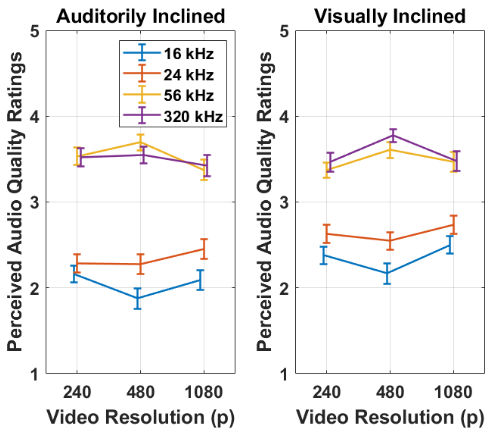

Fig. 7. The perceived audio quality ratings from the subjects divided into two groups. The y-axis represents the audio quality ratings and the x-axis represents the video resolution. Different audio resolutions are specified in separate lines. The left graph represents the auditorily inclined subjects while the right represents the visually inclined subjects. The error bars indicate standard error.

First, a two-way ANOVA on the perceived video quality ratings across all subjects (Fig. 4) showed a significant main effect of the video resolution, an *almost* significant main effect of the audio resolution, and a significant interaction of the two factors. It makes sense that different video resolutions would lead to different perceived video qualities. Note that the video quality ratings had a strong influence on the audio resolution as well. When we consider the results from a three-way ANOVA (Fig. 6), the main effect of the audio resolution becomes significant. This means that the almost-but-not-quite-significant main effect of the audio resolution from the two-way ANOVA was due to the not-significant effect of the group lumped together. Therefore, the perceived video quality is significantly impacted by both the audio resolution and the video resolution, as well as the interaction of the two. These findings are consistent with the studies by Joly, Montard, and Buttin [11] and Långvik [12] that auditory perception may influence how visuals are perceived.

In contrast, the perceived audio quality ratings across all subjects (Fig. 5) showed a significant main effect of the audio resolution and not of the video resolution. It means that different video resolutions did not influence the perceived audio quality, which is contradictory to the study reported by Berrends and De Caluwe [10] or several other studies that reported that the auditory system is dominated by the visual system [5–9]. The interaction was significant between the audio and video resolutions, probably because the mean ratings were very similar (and not statistically different) for the 56 kHz and 320 kHz audio resolutions.

When subjects are divided into two groups based on their self-identification, a significant group difference is found in the perceived audio quality ratings, where the auditorily inclined group showed lower mean ratings for stimuli in 16 kHz and 24 kHz audio than the visually inclined group (Fig. 7). These differences were indeed statistically significant, suggesting that people might differently perceive the same stimuli depending on their inclination. However, this group difference was not observed in the audio resolutions of 56 kHz and 320 kHz.

Note that both groups of subjects showed similar (and statistically not different) quality ratings for 56 kHz and 320 kHz audio resolutions (Figs. 5 and 7). This suggests that the objective resolution increase from 56 kHz to 320 kHz does not lead to a perceived quality improvement. This pattern appears to stay constant regardless of the combined video resolution, which hints at a possible *ceiling* effect. Overall, these patterns indicate that increasing audio resolutions from 16 kHz to 24 kHz or 24 kHz to 56 kHz leads to improvements in perceived audio qualities; however, there appears to be little to no perceptual merit of higher audio resolutions over 56 kHz.

Based on the limited number of subjects, all these findings should be interpreted with caution. Further studies are called for to validate the effect of inclination as well as the perceptual difference (or no difference) between the audio resolutions of 56 kHz and 320 kHz.

6 Conclusion

In this study, perceptual asymmetry was observed. The audio resolution had a significant effect on the perceived video quality, whereas the video resolution did not influence the perceived audio quality. The video resolution increase from 240p to 480p then to 1080p led to improvement in perceived video quality. An increase in audio resolution from 16 kHz to 24 kHz or from 24 kHz to 56 kHz also resulted in the improved perceived audio quality. However, the audio resolution change from 56 kHz to 320 kHz did not lead to any significant difference in perceived audio quality, despite a large increase in objective audio resolution. Additionally, the two subject groups showed different audio quality perception of lower resolution audio; the auditorily-inclined subjects were more critical of lower resolution audio.

As this study is our first endeavor on this topic, further studies are required. An in-person version of this experiment is planned to be conducted in a controlled theatre, which is equipped with a 4K projector and surround sound speakers. In this current design of an online experiment, 4K video condition was excluded as subjects could not be expected to own such high resolution monitors. We will be able to include the 4K video condition to better understand the relationship of the perceived audio and video qualities when the video resolution is very high.

The findings of this study are directly relevant to the streaming media industry such as Netflix, Amazon, and Hulu. Our results suggest that, within the video and audio resolutions considered, the audio resolution influences both the perceived audio and video qualities, whereas the video resolution influences only the perceived video quality. Interestingly, the audio resolution appears to reach a perceptual ceiling at 56 kHz. Therefore, the streaming media providers may want to make sure to provide audio at 56 kHz at least before increasing the video resolution as the bandwidth allows. However, these findings should be re-examined with more diverse stimuli and more subjects and more conditions before making a conclusive interpretation.

References

[1] S. Kalla, "Majority of TV Watchers Prefer Phone or Laptop to the Big Screen - Cord Nevers," *Forbes*, 25 Sep. 2014. [Online], Available: https://www.forbes.com/sites/susankalla/2014/09/25/small-screens-mean-big-bucksthey-dominate-tv-watching/ [Accessed 14 Apr. 2022].

[2] M. Harris, "What If the Movie Studios Decide They Don't Need Theaters After All?" *Vulture*, 1 Sep. 2020. [Online], Available: https://www.vulture.com/2020/09/what-if-studios-decide-they-dont-need-movie-theaters.html [Accessed 12 Oct. 2020].

[3] International Telecommunication Union, "ITU-T P.1203," *International Telecommunication Union*, (October 2017). [Online], Available: https://www.itu.int/rec/T-REC-P.1203/ [Accessed Apr. 14, 2022].

[4] J.K. Willcox, "Is Your Internet Fast Enough for Streaming?" *Consumer Reports*, 5 Apr. 2019. [Online], Available: https://www.consumerreports.org/broadband/internet-fast-enough-for-streaming-broadband-speed/ [Accessed 12 Oct. 2020].

[5] J.P. Makela, M. Hamalainen, R. Hari, and L. McEvoy, "Whole-head mapping of middle-latency auditory evoked magnetic fields," Helsinki University of Technology, Mar. 1994.

[6] K. Amano, N. Goda, S. Nishida, Y. Ejima, R. Rakeda, and Y. Ohtani "Estimation of the Timing of Human Visual Perception from Magnetoencephalography," *The Journal of Neuroscience*, pp. 381-391, Apr. 2006.

[7] B. Yvert, A. Crouzeix, O. Bertrand, A. Seither-Preisler, and C. Pantev, "Multiple Supratemporal Sources of Magnetic and Electric Auditory Evoked Middle Latency Components in Humans," *Institute for Experimental Audiology*, May 2001.

[8] A.J. Lotto and L. Holt, "Psychology of Auditory Perception," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol 2, pp. 479-489 Sep. 2011.

[9] Wikipedia, "McGurk effect," Wikipedia [Online document], Available: https://en.wikipedia.org/wiki/Technical_details_of_Netflix/ [Accessed 15 Apr. 2020].

[10] J.G. Beerends and F.E. De Caluwe, "The Influence of Video Quality on Perceived Audio Quality and Vice Versa," *JAES*, vol. 47, pp. 355-362, May 1999.

[11] A. Joly, N. Montard, and M. Buttin, "Audio-visual quality and interactions between television audio and video," Proceedings of the Sixth International Symposium on Signal Processing and its Applications (Cat.No.01EX467), Kuala Lumpur, Malaysia, 2001, pp. 438-441 vol.2, Jan. 2001.

[12] S. Långvik, "Influence of audio on perception and comprehension of video sequences," *KTH School of Computer Science and Communication*, 25 Sep. 2015.

[13] M. Schoeffler, J. Herre, "The Influence of Audio Quality on the Popularity of Music Videos: A YouTube Case Study," *International Audio Laboratories Erlangen*, Nov. 2014.

[14] Netflix.com, "Bringing Studio Quality Sound to Netflix," *Netflix*, 1 May 2019. [Online], Available: https://about.netflix.com/en/news/bringing-studio-quality-sound-to-netflix/ [Accessed 7 Nov. 2020].

[15] Wikipedia, "Technical details of Netflix," Wikipedia [Online document], Available: https://en.wikipedia.org/wiki/Technical_details_of_Netflix/ [Accessed 7 Nov. 2020].

[16] Netflix.com, "Plans and Pricing," *Netflix*. [Online], Available: https://help.netflix.com/en/node/24926/ [Accessed 12 Oct. 2020].

[17] J. Supan, "Our Netflix review for 2020: Is it worth it?" *Allconnect*, 17 Jan. 2020. [Online], Available: https://www.allconnect.com/streaming/netflix-review/ [Accessed Oct. 9, 2020].
[18] Amazon.com, "Video Central - Technical Asset Requirements," *Amazon*. [Online], Available: https://videocentral.amazon.com/home/help?topicId=GDDDXTDKXPSEY7JF/ [Accessed 12 Oct. 2020].
[19] Amazon.com, "Video Direct - Asset Requirements," *Amazon*. [Online], Available: https://videodirect.amazon.com/home/help?topicId=G202129880/ [Accessed 12 Oct. 2020].
[20] M. Snider, "Amazon is newest music streamer to tempt subscribers with higher-resolution sound," *USA Today*, 17 Sep. 2019. [Online], Available: https://www.usatoday.com/story/tech/talkingtech/2019/09/17/amazon-music-streaming-service-gets-high-resolution-sound-upgrades/2350055001/ [Accessed 12 Oct. 2020].
[21] Amazon.com, "Amazon Music Unlimited HD FAQ," *Amazon*. [Online], Available: https://www.amazon.com/b?ie=UTF8&node=14070322011/ [Accessed 12 Oct. 2020].
[22] Spotify.com, "Audio Settings," *Spotify*, 22 Jun. 2020. [Online], Available: https://support.spotify.com/us/article/audio-quality/ [Accessed 12 Oct. 2020].
[23] Tidal.com, "What Audio Quality Does TIDAL HiFi Offer?" *Tidal*. [Online], Available: https://support.tidal.com/hc/en-us/articles/360002599997/ [Accessed 12 Oct. 2020].
[24] Hulu.com, "Audio quality on Hulu," *Hulu*, 10 Sep. 2020. [Online], Available: https://help.hulu.com/s/article/sound-quality/ [Accessed 9 Oct. 2020].
[25] Google.com, "Select or Change your Audio Quality," *Google*. [Online], Available: Google.com. https://support.google.com/youtubemusic/answer/9076559 [Accessed 7 Nov. 2020].
[26] L, Yitong, S. Yun, M. Yinian, L. Jing, L. Qi, and Y. Dacheng, "A study on Quality of Experience for adaptive streaming service," IEEE International Conference on Communications Workshops (ICC), pp. 682-686, Apr. 2014.
[27] M. Gruner, M. Licciardello, and A. Singla, "Reconstruction Proprietary Video Streaming Algorithms," USENIX Annual Technical Conference, Jul. 2020.
[28] I. Corbett, "What Data Compression Does to Your Music," *Sound on Sound*, Apr. 2012. [Online], Available: https://www.soundonsound.com/techniques/what-data-compression-does-your-music/ [Accessed 7 Nov. 2020].