McGurk Effect: The Impact of Increased Audio-Visual Perceptual Load, Saliency of Stimuli and Visual Distortion on Individual Perceptions

Erin B. Ivany

A thesis submitted to the Psychology Department in partial fulfillment of the requirements for Bachelor of Science (Honours), School of Arts and Social Science

Grenfell Campus

Memorial University of Newfoundland

April 2020

Approval

The undersigned recommend the acceptance of the thesis

"McGurk Effect: The Impact of Increased Audio-Visual Perceptual Load, Saliency of Stimuli and Visual Distortion on Individual Perceptions"

Submitted by Erin B. Ivany

in partial fulfillment of the requirements for the degree of

Bachelor of Science, Honours.

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

Peter Stewart

Thesis Supervisor

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

Mark McPhedran

Alternate Reader

Grenfell Campus

Memorial University of Newfoundland

April 2020

Acknowledgments

First and foremost, I would like to extend my utmost gratitude towards my supervisor, Dr. Peter Stewart. I am sincerely thankful for his continued encouragement and support throughout the entirety of this process. His guidance and reassurance go unmatched and his position as my mentor has proven valuable beyond belief. His undeniable commitment to the project has remained evident, especially during the last few weeks of its development. Secondly, I would like to thank my alternate reader, Dr. Mark McPhedran, for dedicating his time under such short notice, happily and with zero hesitation. Lastly, I would like to thank the entire psychology faculty for not only expanding my knowledge but allowing me to grow my love for this incredible field.

I would also like to voice my appreciation towards my family and friends, who have not only made my degree possible but enjoyable as well. To my parents, I would not be where I am today without your constant support. You have raised my spirits amidst many stressful moments and have loved me unconditionally along the way. To my sister, thanks for always keeping me on my toes and lighting a fire in me to become the best that I can be. To my friends, thank you for providing a shoulder to lean on but also for filling my life with laughter and love. And to Kelly Brown, a huge extension of my gratitude for accepting your students with open arms and always providing a sense of academic relief. Each and every one of you has played a unique role in getting me to this point and I could not be more grateful for your continued support.

Table of Contents

List of Tables and Figures

Abstract

Research has suggested that the likelihood of experiencing the McGurk Effect is impacted by attention and the visual clarity/distortion levels of the speaker. Further, research suggests that the affective saliency of the distractor stimulus is capable of altering the attentional state. In a 3 (saliency level) x 2 (distortion level) x 2 (audio-visual stimulus congruency) repeated measures design, the current study combines all of these components into a single investigation to examine the impact of these variables on the perception of the McGurk Effect. Importantly, while the majority of research has focused on one specific modality, the current study introduces multisensory (i.e. audio-visual) distractors. Results indicated that the presence of the McGurk Effect is dependent on the saliency of the audio-visual distractor and whether the visual field is blurred. Participants (n = 36) reported the correct audio at a significantly lower accuracy when the clips had neutral saliency in comparison to both positive saliency and negative saliency, as well as when the clips were blurred compared to when the clips were not blurred. Therefore, clips with positively or negatively salient audio-visual distractors and clips without blur were significantly more effective in reducing the perception of the McGurk Effect.

McGurk Effect: The Impact of Increased Audio-Visual Perceptual Load, Saliency of Stimuli and Visual Distortion on Individual Perceptions

Despite what individuals may wish to believe, the eyes and ears are found to continually play tricks and can result in experiencing an audio-visual illusion. For example, the illusory flash effect results from multiple simultaneous sounds altering the perception of a single flash of light, such that the individual believes they are witnessing multiple flashes (Shams, Kamitani, & Shimojo, 2000). A separate, but similar audio-visual illusion is exhibited when a single sound is associated with multiple flashes, leading the viewer to perceive a single visual flash despite the existence of two or more (Andersen, Tiippana, & Sams, 2004).

Reports of audio-visual illusions have also suggested that placement of sound can alter the perception of an object in motion. Initially discovered by Sekuler, Sekuler, and Lau (1997), sound inserted at a point in a visual scene is capable of altering the perception of two objects moving towards one another, inducing the perception of a collision and a bouncing back effect rather than a passing by of the objects in their original directions when no accompanying sound is presented. Further, the addition of a third nearby moving object has been found to convince the viewer that a once-believed bouncing back of the objects now appears as the passing of the objects through one another (Kawachi & Gyoba, 2006). All of these examples are evidence for a strong cross-modal interaction amongst the audio and visual domains. The current study also investigated an audio-visual illusion, emphasizing the connection and competition between vision and hearing. To further understand the misperceptions that arise from cross-modal interaction, the current study focused on the mechanisms of the McGurk Effect. Aware of the strength of audio-visual illusions, the study aimed to explore the possibility of inhibiting such misperceptions.

Although many studies have shown the influence of audition on visual perceptions, perhaps the most researched illusion involves the influence of vision on the processing of auditory events. Originally identified by McGurk and MacDonald (1976), the McGurk effect refers to the impact that a visual stimulus can have on speech perception. For example, when an auditory stimulus such as the syllable 'ba' is presented but the visual input from the speaker's mouth implies 'ga' then the resulting perception will often be 'ga'. However, very often the sound that they will hear may be a fusion, resulting in the perception of a sound that represents a compromise such as "baga" or another potentially similar, but unrepresented syllable such as "da". This particular effect has been found to remain intact regardless of both prior knowledge of the cross-modal incongruence as well as after repeated exposure to the McGurk Effect (McGurk & MacDonald, 1976), making it an extremely fascinating focus of research.

Given its high rates of replicability, generalizability with other consonants, and observability across many different age groups (McGurk & MacDonald, 1976), the McGurk Effect has sparked the curiosity of an abundance of individuals in numerous psychological fields. Whether interest is driven by the broad desire of better understanding our language systems (Hazan, Kim, & Chen, 2010), individual aspects of attention (Buchan & Munhall, 2011) or to understand the underlying function of the human sensory/perceptual systems (Paré, Richler, Hove, & Munhall, 2003), the McGurk effect has proven an excellent tool for investigation.

A common theme when studying the McGurk Effect, revolves around the attentional characteristics of the viewer and the distractor details used to disrupt this phenomenon. One study by Tippana, Sams, and Andersen (2001) attempted to weaken the multisensory integration seen during the McGurk Effect by providing attentional instructions to participants aged 19-37. Individuals instructed to pay attention to a distractor stimulus experienced a much weaker

McGurk Effect than those instructed to attend to the talking face. This data suggested that attentional state can play a huge role in audio-visual integration. In other words, shifting our attention to specific stimuli, or away from certain stimuli, can alter audiovisual speech perception by weakening the impact of visual speech. A follow-up study by Andersen, Tiippana, Laarni, Kojo, and Sams (2009) reported consistent findings, suggesting that directing of visual spatial attention to a single face when presented multiple faces is essential in determining whether the McGurk Effect will occur. Therefore, it is evident that instructions regarding attention and instructed areas of focus are key elements in determining whether the McGurk Effect will occur. The same is true when dual tasks procedures, designed to increase the perceptual load, are introduced. Evidence of this was found in a recent study by Gibney et al. (2017), who explored the importance of attention for audio-visual integration. Gibney et al. used a simultaneous McGurk and speeded detection task, which consisted of 90 trials in each of three categories: auditory only, visual only and multisensory conditions. The speeded detection task consisted of participants indicating, through the press of a button, the moment they noticed either stimulus type on the screen which allowed researchers to measure their response times. Results found the McGurk illusion to be significantly decreased when increasing perceptual load, as individual stimuli were harder to detect.

Although altering the participants' attention has been found to weaken this audio-visual integration (e.g., Andersen et al., 2009; Gibney et al., 2017; Tippana et al., 2001), the McGurk Effect is extremely robust and difficult to completely remove. When asked to direct their attention to one modality over another, Buchan and Marshall (2011) found that participants were only moderately successful at indicating the correct audio information when competing visual information was present. Other studies using similar cross-modal oddball tasks (i.e. where

participants quickly indicate on a keyboard their detection and categorization of either odd or even numbers in the presence of other stimuli that act as distractors) have concluded that unexpected auditory distractors fail to capture attention or disrupt performance (Marsja, Neely, & Ljungberg, 2018). To better understand this inconsistency, it may be useful to not only explore the attentional instructions and distractions provided by the researchers but to consider the gaze behavior and eye movements associated with participant attention.

It is known that audio-visual illusions have a significant influence on saccadic eye movements and other related visuo-motor behaviors (Fracasso, Targher, Zampini, & Melcher, 2013). This influence may in part be due to the fixation point an individual elicits on the McGurk effect illusion, which has been found to interrupt integrations and perceptions (Buchan & Munhall, 2011). Using infrared eye tracking technology, results show that gaze fixation on the talker's mouth is sufficient in increasing the frequency of this extremely robust audio-visual illusion (Gurler, Doyle, Walker, Magnotti, & Beauchamp, 2015). Subsequently, when the gaze of the individual is fixated beyond the mouth there is a significant reduction in the McGurk effect response. Such a result was replicated by Buchan and Munhall (2011). Using the Eyelink II eye-tracking system, Buchan and Munhall found a reduction in the McGurk Effect when visual attention was directed beyond the speaker's mouth. Consistent with other studies, this reduction only occurred when fixation was 10-20 degrees past the mouth of the speaker. Otherwise, the McGurk effect remained intact (Paré et al., 2003; Buchan & Munhall, 2011). Again, the strength of the McGurk effect is extremely apparent. Although visuo-motor behaviors and gaze fixations can be used to lessen the McGurk effect, the conditions and degrees under which this can occur appear limited.

In another attempt to fully break the audio-visual integration seen during the McGurk Effect, researchers turned to distorting the visual clip, by altering visual components including gender of the speaker (Mallick, Magnotti, & Beauchamp, 2015), gender incompatibility of the speaker (Green, Kuhl, Meltzoff, & Stevens, 1991) speaker familiarity (Walker, Bruce, & O'Malley, 1995), color (Jordan, McCotter, & Thomas, 2000), proportion of the speaker's face shown (Jordan & Thomas, 2011) and of personal interest to the current study, blur (Thomas & Jordan, 2002; Buchan & Munhall, 2011). Through examination of blurring and inversion on the McGurk response, Thomas and Jordan found that blurring did not affect audio-visual integration of congruent trials but did affect the auditory perception of sounds incongruent with the visual speech. In addition, faces with extreme blurring on top of inverted orientation remained powerful enough to alter perception of the auditory component. In general, degradation through blurring in one domain (i.e. auditory or visual) was found to increase the influence of the other un-blurred modality (Hazan et al., 2010). In contrast, other findings report that when the visual image was blurred but still recognizable, the McGurk effect remained equally as strong when compared to the unfiltered image (Buchan & Munhall, 2011). Inconsistencies in findings regarding blurring of the visual component serve excellent purpose for future examination of blurring effects and their abilities to either break or sustain the McGurk response. The current study aimed to examine the effects of visual distortion (i.e. blurring) on audio-visual perceptions by recreating similar experiments. To further explore the inconsistencies in the literature, the current study analyzed the effects of blurring on the likelihood of a McGurk response through comparing both blurred and un-blurred conditions across congruent and incongruent trials.

In addition, saliency of stimuli has also been explored in audiovisual integration, surprisingly without using the McGurk Effect. Research suggests that both pleasant and

unpleasant emotional stimuli are significantly more effective at drawing attention and altering fixations when compared to stimuli with neutral emotional associations (Calvo & Lang, 2004). Other studies show that images with affective saliency, particularly negative stimuli, have the biggest effect on eye movements and fixations of the visual field (Niu, Todd & Anderson, 2012). Individuals were found to quickly and more frequently fixate on negatively salient stimuli as opposed to those with neutral or positive affect (Humphrey, Underwood, & Lambert, 2012). This is consistent with several findings concluding that saliency of audio and visual stimuli can drastically affect attention, and therefore impact multisensory integration (Coutrot, Guyader, Ionescu & Caplier, 2014). With affective salience as a predictive variable of eye movements and attention, the question then becomes whether saliency effects can continue to predict such behaviors whilst faced with an audio-visual illusion.

Although previous studies looked at saliency in one modality (Coutrot et al., 2014), blurring (Thomas & Jordan, 2002) and attentional instructions on capturing attention during audiovisual tasks (Tippana et al., 2001), to the best of my understanding no study has combined the three in a single investigation. Furthermore, past researchers have stressed a need to incorporate an auditory component into models of visual saliency (Coutrot et al., 2014), while others have emphasized the importance of future exploration into the significance of audiovisual stimuli, with context-sensitive information (Coutrot et al., 2014) and the reliability of such affective salience on attentional tasks (Niu et al., 2012). Taken together, a gap in the literature regarding saliency of audiovisual distractors and corresponding blurring effects becomes apparent. The current study sought to examine the effect of affective saliency (i.e., positive, neutral, and negative) and the distortion (i.e. blurring) of audio-visual distractors on the perception of the McGurk Effect. That is, we examined whether a salient or distorted perceptual

load can weaken the strength of the visual component, reducing the frequency of the McGurk response.

Considering the previous findings, it was hypothesized that audio-visual distractors with both negative and positive affective saliency would significantly weaken the overall perception of the McGurk Effect in comparison to neutral stimuli. In addition, it was hypothesized that the effect of audio-visual distractors with negative affective saliency would have a more robust weakening effect than audio-visual distractors with positive affect. Based on evidence showing that salient stimuli, especially those with negative affect, are best at capturing attention (Niu et al., 2012), and that lack of attention to the speaker reduces McGurk responses (Gibney et al., 2017), this hypothesis would be an extremely reasonable result for the combination of the two. In regard to visual distortion, it was hypothesized that blurring of the visual video component would significantly weaken the overall perception of the McGurk Effect, as blurring may cause salient background stimuli to become more noticeable and therefore more distracting. Knowing that displacement of focus from the talker's mouth leads to less McGurk perceptions (Gurler et al., 2015), this transformation of attention to the salient clip could easily alter one's susceptibility and therefore frequency of this illusion.

Method

Participants

Sixteen participants between the ages of 19-68 years (M = 37.31, SD = 19.28) voluntarily completed a categorization questionnaire in order to determine accurate degrees of saliency for each audio-visual clip. Of these 16 participants, there were 5 males (M = 45.20, SD = 22.08) and 11 females (M = 33.73, SD = 17.81). A separate sample of 36 participants between the ages of 18-59 years (M = 22.33, SD = 10.21) then completed a cross-modal task and corresponding questionnaire. Of these 36 participants, there were 8 males (M = 28.88, SD = 17.75) and 28 females (M = 20.46, SD = 6.10).

Materials

Recruitment

To obtain the projected sample size, posters and announcements in university classes were used to recruit participants. Both the poster and script utilized for recruitment purposes, as well as the email issued to interested participants, can be found in Appendix A.

Categorization Study

This was a stimulus selection and validation study designed to determine which of 60 video clips were the best exemplars for positive, negative, and neutral saliency.

Stimulus Preparation. Sixty non-copyright audio-visual clips were obtained from YouTube using screen recording on an iPhone 11. All clips were then trimmed to four-seconds using iPhone 11 built-in editing software and were then combined into a single four-minute-long video using iMovie version 10.1.14, such that each clip played one after the other.

Informed Consent. All participants completed an informed consent process. A copy of the informed consent form for the pre-test can be found in Appendix B.

Questionnaire. To determine the degree of saliency of the distractors (i.e. positive, neutral and negative) used in the current experiment, a questionnaire was completed by a separate population. This questionnaire was completed in response to the four-minute video, consisting of all possible audio-visual distractors. A copy of this questionnaire can be found in Appendix B.

Audio-Visual Study

Stimulus Preparation. Audio-visual recordings of a single female speaker were recorded on an iPhone 11 and used in the study. In one recording the speaker is pronouncing the two-letter syllable "la" three consecutive times, while in the second recording the speaker is pronouncing the two-letter syllable "fa" three consecutive times. All edits and manipulations of audio and visual stimuli were completed using iMovie (version 10.1.14). The volume of both initial recordings was set to 400% and although both were used for the congruent trials, incongruent trials were also created. For these trials, the volume of the original recording was set to 0% and the audio of the opposite recording at 400% volume was added. That is, in one manipulation the individual visually appears to be saying "la" but the audio input is "fa" and in the other case the individual visually appears to be saying "fa" but the audio input is "la". In the background of each of the four recordings (i.e. two congruent "fa-fa" and "la-la" trials and 2 incongruent "fa-la" and "la-fa" trials), a television screen appeared to be playing a series of audio-visual scenes. Each background clip was a four second, non-copyright stimulus obtained from YouTube and their accompanying sound was set to 60% volume. The purpose of such volume settings was to make it such that the syllables were more audible than the video, but that the background videos could still be heard.

These background distractors differed in saliency including positive, negative and neutral stimuli, and were chosen based on the results obtained from the categorization study. Out of 60 initial clips, 10 of the most agreed-upon clips in each of the three salience levels were used, yielding 30 clips total. In each salience group, three clips showed congruent audio and visual inputs (e.g., fa-fa), while seven clips were incongruent (e.g., fa-la).This imbalance in the number of incongruent and congruent trials maximized the number of trials that should elicit a McGurk effect (i.e. incongruent trials), while also maintaining a comparison group (i.e. congruent trials). This also allowed for the overall number of trials to be as low as possible while maintaining an adequate level of statistical power. These 30 recordings were then repeated, but the speaker's face was blurred using the app BlurEffect in the second round of clips. All 60 recordings were then combined into one single video with each clip playing one after another. To give time for participants to circle a response and act as a refresher, a four second black screen was inserted between each clip such that the entire experiment video was eight minutes long. A list of the finalized 30 audio-visual distractors can be found in Appendix C, along with a transcript that provides the contents of each of the 60 recordings within the final copy of the video

Informed Consent. All participants completed an informed consent process. A copy of the informed consent form for the main experiment can be found in Appendix D.

Testing Questionnaire. The participant was given a questionnaire with questions corresponding to the stimuli presentations, instructing them to indicate the sound they heard in each of the 60 clips within the video. A copy of this questionnaire can be found in Appendix E.

Demographics/Screening Questionnaire. A second questionnaire, administered following the experiment, included demographic related questions (i.e. age and gender identity)

as well as screening questions directly related to the experiment. A copy of this questionnaire can be found in Appendix F.

Debriefing

All participants completed a debriefing process in which participants were informed of the interest in both saliency and alteration of audio-visual distractors, and how that effects the McGurk response. The script utilized for debriefing purposes can be found in Appendix G.

Procedure

To validate the affective saliency classification of the stimuli in the current experiment (as either positive, negative, or neutral), 16 independent judges were asked to rate the stimuli as Positive, Negative, or Neutral. Any stimuli that were deemed to have an ambiguous valence were omitted and the 10 most agreed-upon audio-visual clips for each degree of saliency were used in the current experiment. Once the results of this categorization study were returned and analyzed, advertising of the current experiment began. Participants were recruited through word of mouth, announcements in Psychology 1001 classes at Grenfell Campus, and using posters displayed around said campus (i.e. on bulletin boards and walls). Individuals who displayed interest were sent an email with extra information regarding the study.

After setting up a suitable time slot, all recruited participants began with completing an informed consent form. Once signed, instructions were clearly given and testing for the current research project commenced. Participants were asked to sit in a chair facing a 13.0” MacBook Air computer screen on which the experiment video was presented. A blank wall was located directly behind the computer screen and the door of the experiment room was shut to prevent potential background noises. Participants were told they could adjust the computer screen to a position that was best seen from their perspective. The computer played the video at a constant

and comfortable volume for each participant. In addition, participants were specifically instructed to focus their attention on the speaker in the foreground of the video and directed to hold that attention until the speaker said the syllable for the third time. At this point, participants were able to look down and record a response during the four second transition, indicated by an empty black screen. This was done to ensure individuals did not hear two different sounds on each trial which could happen if attention is no longer on the face, yet the audio is still playing. For each of the 60 clips within the video, the participants were instructed to indicate (i.e. on the testing questionnaire) which syllabic sound they were hearing the speaker say (e.g. la, fa, tha, va, etc.).

Once the entire video (i.e. 60 stimuli) had been completed, the participants were asked to complete a brief demographics/screening questionnaire. These questions covered age, gender identity, whether they have normal vision and hearing, and questions reflecting on the experiment they had just completed. Upon completion of this questionnaire, participants went through a debriefing session before leaving the testing area which provided full disclosure of the research intentions as well as contact information to address any questions or concerns they might have.

Results

Categorization Study

Affective saliency means were calculated for each of the 60 audio-visual clips, indicating the percentage of participants who rated the stimuli as either positive, negative or neutral. All trials with less than 75% interrater agreement in any of the three conditions were omitted. Based on this criterion, questions 3, 5, 6, 9, 27, 29, 42, 53 and 56 were removed, as the highest percentage of agreeability in a single category ranged from 50%- 68.75%. These questions corresponded to the audio-visual clips which contained content of two people kissing (62.5% positive, 37.5% neutral, 0% negative), people riding an escalator (31.25% positive, 68.75% neutral, 0% negative), a graveyard (0% positive, 31.25% neutral, 68.75% negative), someone ordering coffee (43.75% positive, 50% neutral, 6.25% negative), birds chirping (68.75% positive, 31.25% neutral, 0% negative), an alarm going off (6.25% positive, 37.5% neutral, 56.25% negative), someone getting a haircut (43.75% positive, 56.25% neutral, 0% negative), trees swaying in the wind (37.5% positive, 56.25% neutral, 6.25% negative) and a kettle boiling (31.25% positive, 68.75% neutral, 0% negative), respectively. Out of the remaining 51 clips, highest percentages of agreement were used to determine which audio-visual clips would be used to represent the three degrees of affective salience.

Fifteen of the remaining 51 clips were said to have positive saliency by 100% of participants. These clips included someone getting a raise (question 7), people laughing (question 8), baby gender reveal (question 13), a vacation advertisement (question 14), a home makeover reveal (question 20), kids on Christmas morning (question 21), a child's birthday party (question 25), puppies (question 26), a wedding ceremony (question 30), a baby laughing (question 40), someone winning a race (question 43), someone winning a game show (question 49), someone

winning the lottery (question 54), a proposal (question 58) and a welcome home surprise (question 59). Because we wanted ten clips in each salience level, five clips were to be removed. By random choice, someone getting a raise (question 7), people laughing (question 8), a home makeover reveal (question 20), someone winning the lottery (question 54) and a welcome home surprise (question 59) were omitted. Because ten clips were already chosen based on completely unanimous results, other clips that showed highest agreeability percentages for positive saliency were also removed. These clips included a fireplace crackling (75% positive), a parade (93.75% positive), kittens (93.75% positive), people dancing (93.75% positive), a fountain in a park (75% positive) and a welcome home surprise (93.75% positive) which corresponded to questions 31, 33, 34, 39, 48 and 50, respectively.

Of the remaining 30 audio-visual clips left to analyze, five were said to have negative saliency by 100% of participants. These clips included a homeless person asking for money (question 46), a person stubbing their toe (question 47), a storm destruction report (question 51), someone failing a test (question 52) and a person crying (question 60). In order to obtain ten stimuli in this condition, five more clips were chosen based on the second highest percentage of participants indicating negative saliency with agreeability of 93.75%. These clips included someone with the flu sneezing (question 10), a child in a dentist office (question 11), someone getting fired (question 16), someone sick in a hospital bed (question 19), a police car with lights flashing (question 28), a couple breaking up (question 35) and two people arguing (question 44). Since only five more were needed and not seven, clips including a kid in a dentist office (question 11) and someone sick in a hospital bed (question 19) were removed by random decision. Other clips that showed highest agreeability percentages for negative saliency that were greater than 75% but less than 93.75% were also removed. These clips included a baby crying

(75% negative), someone getting a needle (81.25% negative), a funeral (75% negative), an ambulance with sirens on (81.25% negative), army tanks (81.25% negative) and a haunted house scare scene (81.25% negative) which corresponded to questions 1, 2, 15, 22, 37 and 55, respectively.

Of the now 12 remaining clips, three of these were said to have neutral saliency by 100% of participants. These clips included someone reading the newspaper (question 18), people getting on a bus (question 23) and a detergent television advertisement (question 36). In order to obtain 10 neutral stimuli, we also looked at audio-visual clips that were said to have neutral saliency by 93.75% of participants. These clips included the weather network (question 4), someone typing on a keyboard (question 24), someone walking/footsteps (question 32), locking a car door (question 38), someone taking an elevator (question 41) and a person knocking on a door (question 57). To obtain the final stimulus for the neutral salience level, we looked at audio-visual clips that were said to have neutral saliency by 87.5% of participants. These clips included someone getting into a taxi (question 12), person pushing a grocery cart (question 17) and someone ringing a doorbell (question 45). Since only one more was needed and not three, clips including someone getting into a taxi (question 12) and a person pushing a grocery cart (question 17) were removed. This left us with ten stimuli in each salience category.

Audio-Visual Study

All Trials Combined

A 2 (congruency: congruent & incongruent) x 3 (saliency: positive, negative, & neutral) x 2 (distortion: not blurred & blurred) repeated measures analysis of variance (ANOVA) was conducted (see Table 1 for descriptive statistics). Results indicated that there was a significant main effect of congruency on the participants accuracy of reporting the correct auditory syllable,

$F(1.00, 33.00) = 13.74$, $p = .001$, $\eta p^2 = .29$, 95% CI [0.05, 0.18] (mean difference = 0.12). Participants reported the correct audio at a significantly higher accuracy for the congruent clips ($M = 0.95$, $SD = 0.13$) compared to the incongruent ($M = 0.84$, $SD = 0.14$).

Results also indicated a significant main effect of saliency on the participants accuracy of reporting the correct auditory syllable $F(1.87, 61.77) = 12.43$, $p < .001$, $\eta p^2 = .27$. Pairwise post hoc tests revealed that participants reported the correct audio at a significantly lower accuracy when the clips had neutral saliency ($M = 0.86$, $SD = 0.12$) in comparison to both positive saliency ($M = 0.91$, $SD = 0.10$) (mean difference = -0.04, $p = .001$, 95% CI [-0.06, -0.02]) and negative saliency ($M = 0.92$, $SD = 0.09$) (mean difference = -0.05, $p < .001$, 95% CI [-0.08, -0.03]. There was no significant difference in the participants accuracy of reporting the correct auditory syllable between clips with positive saliency and negative saliency (mean difference = -0.01, $p = .241$, 95% CI [-0.03, 0.01]).

Results also indicated that there was a significant main effect of distortion on the participants accuracy of reporting the correct auditory syllable, $F(1.00, 33.00) = 6.08$, $p = .019$, $\eta_p^2 = .16$, 95% CI [0.01, 0.07] (mean difference = 0.04). Participants reported the correct audio at a significantly higher accuracy for clips that were not blurred ($M = 0.91$, $SD = 0.09$) compared to those that were blurred ($M = 0.88$, $SD = 0.12$).

Table 1. *Descriptive Statistics of Accuracy Across Congruency, Distortion and Saliency Type*

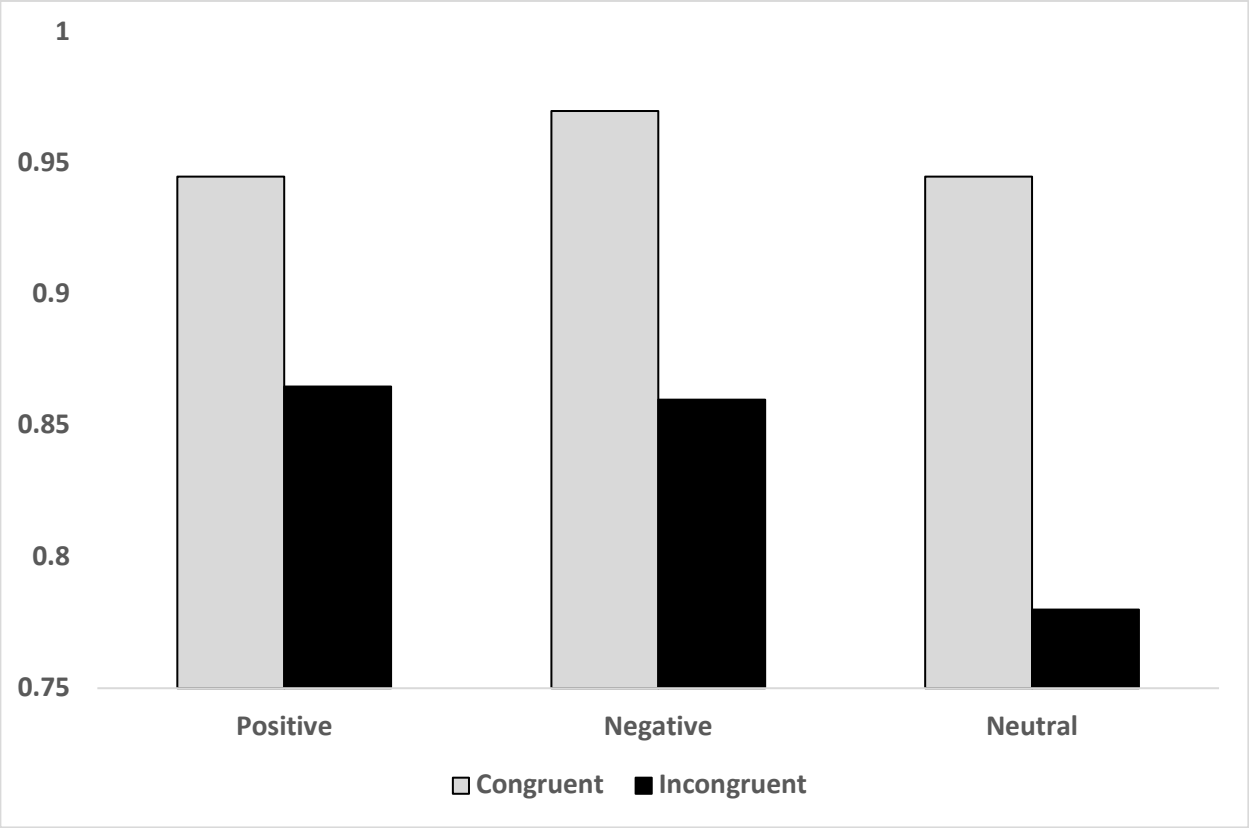
	Positive		Negative		Neutral	
	No Blur	Blur	No Blur	Blur	No Blur	Blur
Congruent						
M	0.96	0.93	0.99	0.95	0.95	0.94
SD	0.14	0.20	0.06	0.15	0.15	0.19
n	34	34	34	34	34	34
Incongruent						
M	0.89	0.84	0.89	0.83	0.80	0.76
SD	0.13	0.20	0.14	0.18	0.22	0.22
n	34	34	34	34	34	34

For the interactions, results indicated that there was a significant interaction between congruency and saliency on the participants accuracy of reporting the correct auditory syllable, $F(1.61, 53.18) = 5.55$, $p = .010$, $\eta p^2 = .14$. (see Figure 1). Comparing across congruency levels, a paired-samples t-test was conducted on trials with positively salient clips and found a significantly higher accuracy of reporting the correct auditory syllable when the clip was congruent ($M = 0.95$, $SD = 0.15$) compared to incongruent ($M = 0.86$, $SD = 0.13$), $t(33) = 2.57$, $p = .015$. A second paired-samples t-test was conducted on trials with negatively salient clips and found a significantly higher accuracy of reporting the correct auditory syllable when the clip was congruent ($M = 0.96$, $SD = 0.12$) compared to incongruent ($M = 0.86$, $SD = 0.13$), $t(34) = 3.04$, $p = .005$. A third paired-samples t-test was conducted on trials with neutrally salient clips and

found a significantly higher accuracy of reporting the correct auditory syllable when the clip was congruent (M = 0.94, SD = 0.17) compared to incongruent (M = 0.78, SD = 0.18), $t(35) = 3.39$, p = .002. This was expected because the audio and video match in the congruent trials preventing the possibility of a McGurk response, while incongruent trials were designed to establish the McGurk illusion.

A repeated measures ANOVA examining accuracy on congruent trials across levels of saliency was conducted. Results indicated that there was no significant difference in the accuracy of reporting the correct auditory syllable according to the type of saliency, $F(1.37, 46.73) = 1.84$, p = .180, η_p^2 = .05. A second repeated measures ANOVA was conducted, this time examining the accuracy on incongruent trials across levels of saliency. Results indicated that there was a significant difference in the accuracy of reporting the correct auditory syllable of incongruent trials according to the type of saliency, $F(1.63, 55.42) = 12.09$, $p < .001$, $\eta_p^2 = .26$. Pairwise post hoc tests revealed that within incongruent trials, participants reported the correct audio at a significantly lower accuracy when the clips had neutral saliency (M = 0.79, SD = 0.18) in comparison to both positive saliency (M = 0.86, SD = 0.13) (mean difference = -0.08, p = .001, 95% CI [-0.12, -0.04]) and negative saliency (M = 0.87, SD = 0.13) (mean difference = -0.08, $p <$ .001, 95% CI [-0.12, 0.04]). There was no significant difference in the participants accuracy of reporting the correct auditory syllable in incongruent trials between clips with positive saliency and negative saliency (mean difference = -0.002, p = .890, 95% CI [-0.03, 0.03]).

Figure 1. *Interaction Between Congruency and Saliency Type*



There were no significant interactions between congruency and distortion, saliency and distortion, or congruency, saliency and distortion on the participants accuracy of reporting the correct auditory syllable, all *F*s < 1.

Additional Analyses

To test if the videos were distracting and if distraction depended on the level of saliency, a one-way repeated measures analysis of variance (ANOVA) was conducted. Results indicated that there was a significant difference between number of videos remembered and saliency type, $F(1, 70) = 9.64$, $p < .001$, $\eta_p^2 = .22$. Post hoc tests indicated that participants remembered significantly more positively salient audio-visual clips ($M = 4.22$, $SD = 2.06$) than audio-visual clips of negative ($M = 3.22$, $SD = 2.10$) (mean difference = 1.00, $p = .004$, 95% CI [0.33, 1.67]), and neutral ($M = 3.00$, $SD = 1.71$) (mean difference = 1.22, $p < .001$, 95% CI [0.67, 1.78])

saliency. There was no significant difference between the number of negative and the number of neutral audio-visual clips remembered (mean difference = 0.22, p = .441, 95% CI [-0.36, 0.80]).

Discussion

First and importantly, our results showed that we were successfully able to recreate the McGurk Effect. As participants reported the correct audio at a significantly higher accuracy for the congruent clips compared to the incongruent, across all trials combined, it became evident that a McGurk response was established. Using the McGurk illusion, we then investigated the impact of saliency type and visual distortion on such perceptions. Findings in the present study are mostly consistent with pre-existing research, although some minor inconsistencies did become apparent.

Our findings indicated a significant interaction between congruency and saliency type on accuracy of reporting the correct auditory syllable. A repeated measures ANOVA examining accuracy on congruent trials across levels of saliency was not significant. However, a similar ANOVA on incongruent trials, or rather the McGurk condition, revealed a significant main effect of saliency type. Pairwise post hoc tests revealed that both positively and negatively salient stimuli significantly increased the accuracy of reporting the correct auditory syllable when compared to stimuli with neutral saliency. Taken together these results suggested that the perception of the McGurk Effect was significantly reduced when stimuli contained either positively or negatively salient audio-visual distractors compared to clips with neutral stimuli. This reduction in the McGurk response can be attributed to the change in attentional state brought about by the salient audio-visual distractors. In particular, those with affective saliency were better at shifting attention such that the participant was no longer focused on the speaker's mouth as instructed. Knowing that displacement of focus from the speaker's mouth leads to less McGurk perceptions (Gurler et al., 2015), it is no surprise that this distraction would lead to a reduction in the McGurk response. These results would be consistent with general findings,

suggesting that pleasant and unpleasant emotional stimuli are more effective at drawing attention and altering fixations when compared to stimuli with neutral emotional associations (Calvo & Lang, 2004). Further, these results are consistent with the current study's saliency related hypotheses and the suggestion from previous research that affective saliency does indeed impact attention which in turn impacts multisensory integration (Coutrot et al., 2014).

Our first hypothesis stated that audio-visual distractors with both positive and negative affective salience would be capable of significantly weakening the overall perception of the McGurk Effect in comparison to neutral stimuli. This hypothesis was supported. Therefore, affective salience can be said to have a significant impact on our attention, leading to a reduction in the likelihood of experiencing the McGurk Effect. In contrast, although the current study supports the claim that affectively salient stimuli (i.e. both positive and negative) are best at capturing attentional fixation in comparison to neutral distractors, negatively salient stimuli did not result in more fixations than those with positive affect as suggested by previous research (Humphrey et al., 2012; Niu et al., 2012). Therefore, our second hypothesis, which was formulated on the basis of consistency with previous research, was not supported as no significant differences were found between negative and positive affect. It is possible that such findings can be attributed to the magnitude of saliency within the distractor stimuli. To maintain ethical standards and prevent the risk of potential harm to participants, we opted for stimuli that were not as polarized as they could have been. Avoiding extremely negative stimuli may have led to an insignificant result whereas, if used, a significant difference between the two might have been found. To determine how polarized the positive and negative stimuli were, a scale indicating saliency magnitude could be implemented into future research designs.

Our findings also indicated a main effect of distortion on participants accuracy of reporting the correct auditory syllable, under incongruent trials (i.e., McGurk conditions). Unlike reports that suggested the McGurk Effect remained equally as strong when the speaking stimulus was blurred or when using an unfiltered image (Buchan & Munhall, 2011), our results show that clips without distortion (i.e. no blurring) significantly increased the accuracy of reporting the correct auditory syllable than those lacking visual clarity (i.e. blurring). These results seemed to indicate that an absence of distortion in the visual stimulus was successfully capable in reducing perceptions of the McGurk Effect. Results are also inconsistent with studies indicating that blurring in one domain increases the influence of another un-blurred modality (Hazan et al., 2010). For the current study's results to be consistent with this notion, blurring should have increased the influence of the distractors leading to a reduction in the McGurk response as hypothesized. However, our hypothesis was not supported as the opposite effects occurred. In blurred conditions, the participants were significantly less accurate in reporting the correct auditory syllable in comparison to the not blurred condition, indicating that they were more susceptible to the McGurk response rather than less as predicted.

It is possible, that although affective saliency has been shown to be a useful manipulation for capturing attention and reducing the McGurk response, maybe blurring is more successful in altering participant fixations. Although our findings show that the McGurk Effect was reduced when blurring was removed, in contrast to Buchan and Munhall (2011), who proposed the McGurk Effect to remain equally as strong whether it was blurred or not, we must consider that the current study's additional manipulation, audio-visual distractors with differing levels of saliency, influenced the impact of stimulus blurring. If blurring was more distracting than the salient audio-visual clips and lead participants to fixate on the speaker's mouth in the foreground

instead of the background clips, then this difference in the strength of the distractor could serve as a possible explanation for the current results. Since our results suggest that no blurring significantly reduced the McGurk response, it seems possible that in the no blur conditions, participants were only faced with one distractor stimulus and were therefore inclined to look away from the face at the audio-visual clip irrespective of its saliency level. With two distractors (i.e. distortion and saliency), blurring may have been better at capturing attention, making participants more inclined to fixate on the face and more susceptible to a McGurk response.

In conclusion, the findings of the present study indicate that both affective saliency of cross-modal distractors (i.e. audio-visual) and distortion have a significant impact on susceptibility to the McGurk Effect. This research is important as it fulfills the suggestion of previous researchers to not only incorporate the significance of audiovisual stimuli, with context-sensitive information (Coutrot et al., 2014) but also the reliability of such affective salience on attentional tasks (Niu et al., 2012). The current study provides continued evidence of an important relationship between saliency and attention, while also offering new evidence for its impact on multisensory integration. While research supports a significant reduction in the McGurk response when stimuli with positive or negative saliency are present in comparison to neutral distractors, more research is necessary to confirm the possible interpretation of differences found due to distortion. One way to confirm the proposed explanation, would be to implement eye tracking technologies into future investigations. With the use of such equipment, exact gaze fixation locations could be pinpointed and the area that captured attention across distortion conditions could be determined.

References

Andersen, T., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21*, 301-308. doi: 10.1016/j.cogbrainres.2004.06.004

Andersen, T., Tiippana, K., Laarni, J., Kojo, I., & Sams, M. (2009). The role of visual spatial attention in audiovisual speech perception. *Speech Communication, 51*, 184-193. doi: 10.1016/j.specom.2008.07.004

Buchan, J., & Munhall, K. (2011). The influence of selective attention to auditory and visual speech on the integration of audiovisual speech information. *Perception, 40*, 1164-1182. doi: 10.1068/p6939

Calvo, M., & Lang, G. (2004). Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation and Emotion, 28*, 221-243. doi: 10.1023/B:MOEM.0000040153.26156.ed

Coutrot, A., Guyader, N., Ionescu, G., & Caplier, A. (2014). Video viewing: Do auditory salient events capture visual attention? *Annals of Telecommunications - Annales Des Télécommunications, 69*, 89-97. doi: 10.1007/s12243-012-0352-5

Fracasso, A., Targher, S., Zampini, M., & Melcher, D. (2013). Fooling the eyes: The influence of a sound-induced visual motion illusion on eye movements. *PLoS ONE, 8*, E62131. doi: 10.1371/journal.pone.0062131

Gibney, K., Aligbe, E., Eggleston, B., Nunes, S., Kerkhoff, W., Dean, C., & Kwakye, L. (2017). Visual distractors disrupt audiovisual integration regardless of stimulus complexity. *Frontiers in Integrative Neuroscience, 11*, 1. doi: 10.3389/fnint.2017.00001

Green, K., Kuhl, P., Meltzoff, A., & Stevens, E. (1991). Integrating speech information across talkers gender and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics, 50,* 524-536.

Gurler, D., Doyle, N., Walker, E., Magnotti, J., & Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye movements. *Attention Perception & Psychophysics, 77*, 1333-1341. doi: 10.3758/s13414-014-0821-1

Hazan, V., Kim, J., & Chen, Y. (2010). Audiovisual perception in adverse conditions: Language, speaker and listener effects. *Speech Communication, 52*, 996-1009. doi: 10.1016/j.specom.2010.05.003

Humphrey, K., Underwood, G., & Lambert, T. (2012). Salience of the lambs: A test of the saliency map hypothesis with pictures of emotive objects. *Journal of Vision, 12*, 1-15.

Jordan, T., & Thomas, S. (2011). When half a face is as good as a whole: Effects of simple substantial occlusion on visual and audiovisual speech perception. *Attention, Perception, & Psychophysics, 73*, 2270-2285. doi: 10.3758/s13414-011-0152-4

Jordan, T., McCotter, M., & Thomas, S. (2000). Visual and audiovisual speech perception with color and gray-scale facial images. *Perception & Psychophysics, 62*, 1394-1404. doi: 10.3758/BF03212141

Kawachi, Y., & Gyoba, J. (2006). Last but not least - Presentation of a visual nearby moving object alters stream/bounce event perception. *Perception, 35*, 1289-1294. doi: 10.1068/p5594

Mallick, D., Magnotti, J., & Beauchamp, M. (2015). Variability and stability in the McGurk effect: Contributions of participants, stimuli, time, and response type. *Psychonomic Bulletin & Review, 22*, 1299-1307. doi: 10.3758/s13423-015-0817-4

Marsja, E., Neely, G., & Ljungberg, J. (2018). Investigating deviance distraction and the impact of the modality of the to-be-ignored stimuli. *Experimental Psychology (Göttingen), 65*, 61-70. doi: 10.1027/1618-3169/a000390

McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748. doi: 10.1038/264746a0

Niu, Y., Todd, R., & Anderson, A. (2012). Affective salience can reverse the effects of stimulus-driven salience on eye movements in complex scenes. *Frontiers in Psychology, 3*, 336. doi: 10.3389/fpsyg.2012.00336

Paré, M., Richler, R., Hove, M., & Munhall, K. (2003). Gaze behavior in audiovisual speech perception: The influence of ocular fixations on the McGurk effect. *Perception & Psychophysics, 65*, 553-567. doi: 10.3758/BF03194582

Sekuler, R., Sekuler, A., & Lau, R. (1997). Sound alters visual motion perception. *Nature, 385*, 308. doi: 10.1038/385308a0

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*, 788. doi: 10.1038/35048669

Thomas, S., & Jordan, T. (2002). Determining the influence of Gaussian blurring on inversion effects with talking faces. *Perception & Psychophysics, 64*, 932-944. doi: 10.3758/BF03196797

Tiippana, K., Sams, M., & Andersen, T. (2001): Visual attention influences audiovisual speech perception. *AVSP-2001*, 167-171.

Walker, S., Bruce, V., & O'Malley, C. (1995). Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect. *Perception & Psychophysics, 57,* 1124-1133.

Appendix A

Participant Recruitment Script

Hi everyone, my name is Erin Ivany and I am a fourth-year psychology major working towards my BSc. As part of my honours thesis research, I am conducting a research study on the audio-visual interaction and I am looking for participants to volunteer. If you choose to participate you will simply watch a set number of audio-visual clips and record a response. The testing will take about 15 minutes of your time, and you will receive 1% course credit for participating. I am very flexible this semester with scheduling, so you can participate whenever you're free. I'm going to pass out a sign-up sheet for those of you who are interested. You just need to write down your name as well as email, and I will contact you with more information about the study, or you can email me at ebivany@grenfell.mun.ca. Thank you!

Information Email

Hi (name of potential participant),

Thanks so much for your interest in my honours project! Just sending you an email to provide a little more information about my research. The study is designed to investigate the mechanisms of cross-modality (i.e. how senses combine). Particularly, this study examines the auditory and visual stimulus integration.

The experiment will include watching a set number of audio-visual clips and having you record a response. The study will take about 15 minutes to complete and all responses will be analyzed on a group basis such that individual responses cannot be identified (i.e. completely anonymous).

Psychology (course number) is a participating course in these projects and participation in the study will earn you 1% course credit. If you are still interested, please let me know when you would be available for testing or to discuss the study a little more!

Cheers,

Erin

Research Participants Needed!

Erin Ivany supervised by Dr. Peter Stewart

- Can you trust your ears?
- Have you ever wondered if vision or hearing is your most powerful sense?

COME FIND OUT!

Volunteer 15-30 minutes of your time & join us in our research study on the audio-visual interaction.

This study has been approved by an ethics review process in the psychology program at Grenfell Campus, Memorial University of Newfoundland and has been found to be in compliance with Memorial University's ethics policy as well as Tri-council policy on Ethics.

ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca ebivany@grenfell.mun.ca

Appendix B

Degree of Saliency Informed Consent Form

The purpose of this Informed Consent Form is to ensure you understand the nature of this questionnaire and your involvement in it. This consent form will provide information about the study, giving you the opportunity to decide if you want to participate.

Researchers: This questionnaire is being conducted as part of the course requirements for Psychology 4959: Honours Project in Psychology II. The research is being conducted under the supervision of Dr. Peter Stewart.

Purpose: This questionnaire is designed to classify stimuli by their degree of emotional saliency. The goal of this study is to classify stimuli as positive, negative or neutral, based on participant responses.

Task Requirements: You will be asked to complete a brief questionnaire regarding the emotional saliency of 60 different stimuli. Your task is to indicate for each of these stimuli, whether you feel they are best described as positive, negative or neutral. It is asked that you indicate your response by circling only one of the three options. The last two questions will ask for demographics including age and gender identity.

Duration: The questionnaire will take approximately 10 minutes to complete.

Risks and Benefits: There are no obvious risks involved with your participation in this study.

Anonymity: All information will be analyzed and reported on a group basis. Thus, individual responses cannot be identified in any reporting of the results. All participant information will be kept on a password protected computer or in a locked cabinet.

Right to Withdraw: Your participation in this research is completely voluntary and you are free to stop participating at any time. However, once you complete this study your data cannot be removed. This is because we are not collecting any identifying information, and therefore we cannot link data to individual responses.

Contact Information: If you have any questions or concerns about the study, or are interested in the results of the study, please feel free to contact Erin Ivany at ebivany@grenfell.mun.ca or Dr. Peter Stewart at pstewart@grenfell.mun.ca.

If you have ethical concerns about the research, such as the way you have been treated or your rights as a participant, you may contact the Chairperson of the GC-REB at gcethics@grenfell.mun.ca. This study has been approved by an ethics review process in the psychology program at Grenfell Campus, Memorial University of Newfoundland and has been found to be in compliance with Memorial University's ethics policy.

By signing this form, I acknowledge that I am at least 19 years of age and/or a university student, and I have been informed of, and understand, the nature and purpose of the study, and I freely consent to participate.

_____ _____

Signature Date

Degree of Saliency Questionnaire

For each of the stimuli described below, please indicate whether you feel it is best described as a positive, negative or neutral stimulus. Simply circle the response that best fits your opinion. Remember, there are no right or wrong answers.

1. baby crying	NEGATIVE	NEUTRAL	POSITIVE
2. someone getting a needle	NEGATIVE	NEUTRAL	POSITIVE
3. two people kissing	NEGATIVE	NEUTRAL	POSITIVE
4. weather network	NEGATIVE	NEUTRAL	POSITIVE
5. riding an escalator	NEGATIVE	NEUTRAL	POSITIVE
6. graveyard	NEGATIVE	NEUTRAL	POSITIVE
7. someone getting a raise	NEGATIVE	NEUTRAL	POSITIVE
8. people laughing	NEGATIVE	NEUTRAL	POSITIVE
9. ordering coffee	NEGATIVE	NEUTRAL	POSITIVE
10. someone with the flu	NEGATIVE	NEUTRAL	POSITIVE
11. kid in dentist office	NEGATIVE	NEUTRAL	POSITIVE
12. getting in a taxi	NEGATIVE	NEUTRAL	POSITIVE
13. baby gender reveal	NEGATIVE	NEUTRAL	POSITIVE
14. vacation advertisement	NEGATIVE	NEUTRAL	POSITIVE
15. funeral	NEGATIVE	NEUTRAL	POSITIVE
16. someone getting fired	NEGATIVE	NEUTRAL	POSITIVE
17. pushing a grocery cart	NEGATIVE	NEUTRAL	POSITIVE
18. reading the newspaper	NEGATIVE	NEUTRAL	POSITIVE
19. someone sick in hospital bed	NEGATIVE	NEUTRAL	POSITIVE

20. home makeover reveal	NEGATIVE	NEUTRAL	POSITIVE
21. Christmas morning	NEGATIVE	NEUTRAL	POSITIVE
22. ambulance with sirens on	NEGATIVE	NEUTRAL	POSITIVE
23. getting on a bus	NEGATIVE	NEUTRAL	POSITIVE
24. typing on a keyboard	NEGATIVE	NEUTRAL	POSITIVE
25. child's birthday party	NEGATIVE	NEUTRAL	POSITIVE
26. puppies	NEGATIVE	NEUTRAL	POSITIVE
27. birds chirping	NEGATIVE	NEUTRAL	POSITIVE
28. cop car with lights flashing	NEGATIVE	NEUTRAL	POSITIVE
29. alarm going off	NEGATIVE	NEUTRAL	POSITIVE
30. wedding	NEGATIVE	NEUTRAL	POSITIVE
31. fireplace crackling	NEGATIVE	NEUTRAL	POSITIVE
32. walking down the road	NEGATIVE	NEUTRAL	POSITIVE
33. parade	NEGATIVE	NEUTRAL	POSITIVE
34. kittens	NEGATIVE	NEUTRAL	POSITIVE
35. breakup	NEGATIVE	NEUTRAL	POSITIVE
36. detergent television ad	NEGATIVE	NEUTRAL	POSITIVE
37. army tanks	NEGATIVE	NEUTRAL	POSITIVE
38. locking car door	NEGATIVE	NEUTRAL	POSITIVE
39. people dancing	NEGATIVE	NEUTRAL	POSITIVE
40. baby laughing	NEGATIVE	NEUTRAL	POSITIVE
41. taking an elevator	NEGATIVE	NEUTRAL	POSITIVE
42. getting a haircut	NEGATIVE	NEUTRAL	POSITIVE

43. winning a race	NEGATIVE	NEUTRAL	POSITIVE
44. two people arguing	NEGATIVE	NEUTRAL	POSITIVE
45. ringing a doorbell	NEGATIVE	NEUTRAL	POSITIVE
46. homeless person	NEGATIVE	NEUTRAL	POSITIVE
47. stubbing toe	NEGATIVE	NEUTRAL	POSITIVE
48. fountain in a park	NEGATIVE	NEUTRAL	POSITIVE
49. winning a game show	NEGATIVE	NEUTRAL	POSITIVE
50. surprise party	NEGATIVE	NEUTRAL	POSITIVE
51. storm aftermath/destruction	NEGATIVE	NEUTRAL	POSITIVE
52. someone failing test	NEGATIVE	NEUTRAL	POSITIVE
53. trees swaying in the wind	NEGATIVE	NEUTRAL	POSITIVE
54. winning the lottery	NEGATIVE	NEUTRAL	POSITIVE
55. haunted house scare scene	NEGATIVE	NEUTRAL	POSITIVE
56. kettle boiling	NEGATIVE	NEUTRAL	POSITIVE
57. knocking on a door	NEGATIVE	NEUTRAL	POSITIVE
58. proposal	NEGATIVE	NEUTRAL	POSITIVE
59. welcome home surprise	NEGATIVE	NEUTRAL	POSITIVE
60. adult crying	NEGATIVE	NEUTRAL	POSITIVE

What is your age? _____

What is your gender identity? (circle your response)

Male Female Other

If other, please specify: _____

Appendix C

Finalized Audio-Visual Distractors

Positive:

1) Winning a game show
2) Vacation advertisement
3) Birthday party
4) Puppies
5) Kids on Christmas morning
6) Person winning a race
7) Baby gender reveal
8) Proposal
9) Baby laughing
10) Wedding Ceremony

Negative:

1) Couple breaking up
2) Person stubbing toe
3) Person with the flu sneezing
4) Person failing a test
5) Cop car with sirens on and lights flashing
6) Homeless person asking for money
7) Person crying
8) Person getting fired
9) Two people arguing

10) Storm destruction report

Neutral:

1) Doorbell
2) Typing on a keyboard
3) People getting on a bus
4) Locking car door
5) Detergent television advertisement
6) Elevator
7) Walking/footsteps
8) Weather network
9) Knocking on a door
10) Person reading a newspaper

Transcript

1) POSITIVE-CONGRUENT
    - WINNING A GAME SHOW (100%)
    - NO BLUR
    - LA VISUAL – LA SOUND
    - POSSIBLE INTERPRETATION: “LA”
2) NEUTRAL-INCONGRUENT
    - DOORBELL (87.5%)
    - NO BLUR
    - LA VISUAL – FA SOUND
    - POSSIBLE INTERPRETATION: “THA”
3) NEGATIVE-INCONGRUENT
    - COUPLE BREAKING UP (93.75%)
    - NO BLUR
    - FA VISUAL – LA SOUND
    - POSSIBLE INTERPRETATION: “VA”
4) NEGATIVE-CONGRUENT
    - PERSON STUBBING TOE (100%)
    - NO BLUR
    - FA VISUAL – FA SOUND
    - POSSIBLE INTERPRETATION: “FA”

5) NEUTRAL-CONGRUENT
    - TYPING ON A KEYBOARD (93.75%)
    - NO BLUR
    - LA VISUAL – LA SOUND
    - POSSIBLE INTERPRETATION: “LA”
6) POSITIVE-INCONGRUENT
    - VACATION ADVERTISEMENT (100%)
    - NO BLUR
    - LA VISUAL – FA SOUND
    - POSSIBLE INTERPRETATION: “THA”
7) POSITIVE-INCONGRUENT
    - BIRTHDAY PARTY (100%)
    - NO BLUR
    - FA VISUAL – LA SOUND
    - POSSIBLE INTERPRETATION: “VA”
8) NEUTRAL-CONGRUENT
    - PEOPLE GETTING ON A BUS (100%)
    - NO BLUR
    - FA VISUAL – FA SOUND
    - POSSIBLE INTERPRETATION: “FA”

9) NEGATIVE-INCONGRUENT

- PERSON WITH THE FLU SNEEZING (93.75%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

10) NEUTRAL-INCONGRUENT

- LOCKING CAR DOOR (93.75%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

11) POSITIVE-INCONGRUENT

- PUPPIES (100%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

12) NEGATIVE-CONGRUENT

- PERSON FAILING A TEST (100%)
- NO BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

13) NEUTRAL-INCONGRUENT

- DETERGENT TELEVISION ADVERTISEMENT (100%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

14) NEGATIVE-INCONGRUENT

- COP CAR WITH SIRENS ON & LIGHTS FLASHING (93.75%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

15) POSITIVE-INCONGRUENT

- KIDS ON CHRISTMAS MORNING (100%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: ‘THA”

16) NEUTRAL-CONGRUENT

- ELEVATOR (93.75%)
- NO BLUR
- FA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “FA”

17) NEUTRAL-INCONGRUENT

- WALKING/FOOTSTEPS (93.75%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

18) NEGATIVE-CONGRUENT

- HOMELESS PERSON ASKING FOR MONEY (100%)
- NO BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

19) NEGATIVE-INCONGRUENT

- PERSON CRYING (100%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

20) POSITIVE-CONGRUENT

- PERSON WINNING A RACE (100%)
- NO BLUR
- FA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “FA”

21) POSITIVE-INCONGRUENT

- BABY GENDER REVEAL (100%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

22) NEUTRAL-INCONGRUENT

- WEATHER NETWORK (93.75%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

23) NEGATIVE-INCONGRUENT

- PERSON GETTING FIRED (93.75%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

24) NEGATIVE-INCONGRUENT

- TWO PEOPLE ARGUING (93.75%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

25) NEUTRAL-INCONGRUENT

- KNOCKING ON A DOOR (93.75%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

26) POSITIVE-CONGRUENT

- PROPOSAL (100%)
- NO BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

27) NEGATIVE-INCONGRUENT

- STORM DESTRUCTION REPORT (100%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

28) POSITIVE-INCONGRUENT

- BABY LAUGHING (100%)
- NO BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

29) NEUTRAL-INCONGRUENT

- PERSON READING A NEWSPAPER (100%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

30) POSITIVE-INCONGRUENT

- WEDDING CEREMONY (100%)
- NO BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

31) POSITIVE-CONGRUENT

- WINNING A GAME SHOW (100%)
- BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

32) NEUTRAL-INCONGRUENT

- DOORBELL (87.5%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

33) NEGATIVE-INCONGRUENT

- COUPLE BREAKING UP (93.75%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

34) NEGATIVE-CONGRUENT

- PERSON STUBBING TOE (100%)
- BLUR
- FA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “FA”

35) NEUTRAL-CONGRUENT

- TYPING ON A KEYBOARD (93.75%)
- BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

36) POSITIVE-INCONGRUENT

- VACATION ADVERTISEMENT (100%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

37) POSITIVE-INCONGRUENT

- BIRTHDAY PARTY (100%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

38) NEUTRAL-CONGRUENT

- PEOPLE GETTING ON A BUS (100%)
- BLUR
- FA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “FA”

39) NEGATIVE-INCONGRUENT

- PERSON WITH THE FLU SNEEZING (93.75%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

40) NEUTRAL-INCONGRUENT

- LOCKING CAR DOOR (93.75%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

41) POSITIVE-INCONGRUENT

- PUPPIES (100%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

42) NEGATIVE-CONGRUENT

- PERSON FAILING A TEST (100%)
- BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

43) NEUTRAL-INCONGRUENT

- DETERGENT TELEVISION ADVERTISEMENT (100%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

44) NEGATIVE-INCONGRUENT

- COP CAR WITH SIRENS ON & LIGHTS FLASHING (93.75%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

45) POSITIVE-INCONGRUENT

- KIDS ON CHRISTMAS MORNING (100%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: ‘THA”

46) NEUTRAL-CONGRUENT

- ELEVATOR (93.75%)
- BLUR
- FA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “FA”

47) NEUTRAL-INCONGRUENT

- WALKING/FOOTSTEPS (93.75%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

48) NEGATIVE-CONGRUENT

- HOMELESS PERSON ASKING FOR MONEY (100%)
- BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

49) NEGATIVE-INCONGRUENT

- PERSON CRYING (100%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

50) POSITIVE-CONGRUENT

- PERSON WINNING A RACE (100%)
- BLUR
- FA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “FA”

51) POSITIVE-INCONGRUENT

- BABY GENDER REVEAL (100%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

52) NEUTRAL-INCONGRUENT

- WEATHER NETWORK (93.75%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

53) NEGATIVE-INCONGRUENT

- PERSON GETTING FIRED (93.75%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

54) NEGATIVE-INCONGRUENT

- TWO PEOPLE ARGUING (93.75%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

55) NEUTRAL-INCONGRUENT

- KNOCKING ON A DOOR (93.75%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

56) POSITIVE-CONGRUENT

- PROPOSAL (100%)
- BLUR
- LA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “LA”

57) NEGATIVE-INCONGRUENT

- STORM DESTRUCTION REPORT (100%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

58) POSITIVE-INCONGRUENT

- BABY LAUGHING (100%)
- BLUR
- FA VISUAL – LA SOUND
- POSSIBLE INTERPRETATION: “VA”

59) NEUTRAL-INCONGRUENT

- PERSON READING A NEWSPAPER (100%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

60) POSITIVE-INCONGRUENT

- WEDDING CEREMONY (100%)
- BLUR
- LA VISUAL – FA SOUND
- POSSIBLE INTERPRETATION: “THA”

Appendix D

Informed Consent Form

The purpose of this Informed Consent Form is to ensure you understand the nature of this study and your involvement in it. This consent form will provide information about the study, giving you the opportunity to decide if you want to participate.

Researchers: This study is being conducted as part of the course requirements for Psychology 4959: Honours Project in Psychology II. This research is being conducted under the supervision of Dr. Peter Stewart.

Purpose: The study is designed to investigate the mechanisms of cross-modality (i.e., how senses combine). Particularly, the study will examine the auditory and visual stimulus integration.

Task Requirements: You will first be asked to listen to 60 brief audio-visual clips, each lasting only a few seconds, and indicate which sound you heard from the speaker in the video. Following the 60 trials, you will be asked to complete a brief demographic questionnaire and a few questions regarding your participation in the study.

Duration: The study will take approximately 30 minutes to complete.

Risks and Benefits: There are no obvious risks involved with your participation in this study. If you are currently taking a participating psychology course at Grenfell Campus you will receive 1% course credit for your participation in this study.

Anonymity: All information will be analyzed and reported on a group basis. Thus, individual responses cannot be identified in any reporting of the results. All participant information will be kept on a password protected computer or in a locked cabinet (AS 335).

Right to Withdraw: Your participation in this research is completely voluntary and you are free to stop participating at any time. However, once you complete this study and leave the testing area, your data cannot be removed. This is because we are not collecting any identifying information, and therefore we cannot link data to individual responses.

Contact Information: If you have any questions or concerns about the study, or are interested in the results of the study, please feel free to contact Erin Ivany at ebivany@grenfell.mun.ca or Dr. Peter Stewart at pstewart@grenfell.mun.ca.

If you have ethical concerns about the research, such as the way you have been treated or your rights as a participant, you may contact the Chairperson of the GC-REB at gcethics@grenfell.mun.ca. This study has been approved by an ethics review process in the psychology program at Grenfell Campus, Memorial University of Newfoundland and has been found to be in compliance with Memorial University's ethics policy.

By signing this form, I acknowledge that I am at least 19 years of age and/or a university student and I have been informed of, and understand, the nature and purpose of the study, and I freely consent to participate.

_____ _____

Signature Date

Appendix E

Testing Questionnaire

For each clip, indicate which sound you heard by circling one of the choices below. If the sound you heard is not included, please write the sound you heard in the blank space provided.

Clip 1

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 2

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 3

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 4

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 5

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 6

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 7

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 8

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 9

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 10

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 11

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 12

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 13

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 14

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 15

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 16

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 17

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 18

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 19

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 20

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 21

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 22

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 23

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 24

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 25

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 26

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 27

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 28

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 29

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 30

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 31

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 32

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 33

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 34

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 35

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 36

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 37

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 38

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 39

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 40

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 41

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 42

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 43

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 44

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 45

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 46

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 47

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 48

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 49

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 50

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 51

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 52

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 53

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 54

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 55

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 56

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 57

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 58

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 59

La Fa Tha Va

If none of the above, what did you hear? _____

Clip 60

La Fa Tha Va

If none of the above, what did you hear? _____

Appendix F

Demographics/Screening Questionnaire

1) Do you have normal, or corrected to normal (i.e. glasses, contacts) vision?

   YES NO

2) Do you have normal, or corrected to normal (i.e. hearing aid) hearing?

   YES NO

3) What is your age? _____

4) What is your gender identity?

   MALE FEMALE OTHER

If other, please specify: _____

5) Did you notice the television in the background?

   YES NO

6) If yes, what do you remember seeing and/or hearing on the screen? Check all that apply.

   _____ WINNING A GAME SHOW

   _____ DOORBELL

   _____ COUPLE BREAKING UP

   _____ PERSON STUBBING TOE

   _____ TYPING ON A KEYBOARD

_____ VACATION ADVERTISEMENT

_____ BIRTHDAY PARTY

_____ PEOPLE GETTING ON A BUS

_____ PERSON WITH THE FLU SNEEZING

_____ LOCKING CAR DOOR

_____ PUPPIES

_____ PERSON FAILING A TEST

_____ DETERGENT TELEVISION ADVERTISEMENT

_____ COP CAR WITH SIRENS ON & LIGHTS FLASHING

_____ KIDS ON CHRISTMAS MORNING

_____ ELEVATOR

_____ WALKING/FOOTSTEPS

_____ HOMELESS PERSON ASKING FOR MONEY

_____ PERSON CRYING

_____ PERSON WINNING A RACE

_____ BABY GENDER REVEAL

_____ WEATHER NETWORK

_____ PERSON GETTING FIRED

_____ TWO PEOPLE ARGUING

_____ KNOCKING ON A DOOR

_____ PROPOSAL

_____ STORM DESTRUCTION REPORT

_____ BABY LAUGHING

_____ PERSON READING A NEWSPAPER

_____ WEDDING CEREMONY

7) Would you say that the television distracted you during this task?

YES NO

8) What percentage of the time did you find the television to be distracting?

0-25% 26-50% 51-75% 76-100%

9) Would you say that the blurred clips distracted you during this task?

YES NO

10) What percentage of the time did you find the blurring to be distracting?

0-25% 26-50% 51-75% 76-100%

Appendix G

Debriefing Script

First of all, I would like to say thank-you for completing the study. Secondly, I would like to inform you of the specific interest of the study you have just completed. Before the study, you were informed about the interest in cross-modality between the audio and visual fields and individual perceptions of the McGurk Effect. However, we did not inform you of our interest in both saliency and alteration of audio-visual distractors, and how that effects the McGurk Effect response. As you might have noticed from the last few questions on the final questionnaire, we were interested in whether the distractors in the background and their degree of saliency (i.e. positive, negative, or neutral) had any effect on your perception of the McGurk Effect. We were also curious if blurring a portion of the visual stimulus had any effect on McGurk Effect perceptions. You now have the option to withdraw the data that you provided but after you leave the room, data can no longer be excluded as no identifying information was provided and your info will not be distinguishable from any other participant. If you have any questions or concerns, you can ask now or contact us through the emails provided on the informed consent form. Thank you so much for your participation.