Senior Projects Spring 2020   Bard Undergraduate Senior Projects

Spring 2020

The Way You Hear It, the Way You Judge It: Moral Decision-making and Moral Reasoning in Accented Speech

Yifan Gu
Bard College, yg1617@bard.edu

Recommended Citation

Bard

The way you hear it, the way you judge it:

Moral Decision-making and Moral Reasoning in Accented Speech

Senior Project submitted to

The Division of Science, Mathematics, and Computing

of Bard College

by

Yifan Gu

Annandale-on-Hudson, New York

May 2020

To Professor Stuart Stritzler-Levine

Once on the Shinkansen train to Sendai,
you heard it as the train approached the platform.

"Obento, obento!"

Having no idea of the meaning,
your heart thumped in excitement.

"The emotion in it!"

On fourth floor Stevenson Library,
begins our journey.

Acknowledgment

Table of Contents

Abstract

The previous studies have shown that people make different decisions not only after reading and also listening to moral dilemmas in a foreign language (L2) than in a native language (L1). This effect is named *Moral Foreign Language Effect* (MFLE). Emotion, which is considered to play a pivotal role in moral judgments, is also found to have a close interaction with sounds. The current research aims to (1) investigate whether the sound of different languages (i.e. accents) can also trigger the MFLE in listeners' moral decision-making and (2) examine the foreign accent effect on listeners' moral reasoning pattern. Chinese ESL college students were recruited as listeners of Chinese-accented and English-accented speech of moral dilemmas in Mandarin and English. However, although the study revealed a potential foreign accent effect on moral reasoning patterns in native-accented Chinese and foreign-accented Chinese, contradicting our predictions, a foreign accent effect on moral decisions and moral reasoning patterns was not detected. Neither the higher proficiency in L2 was found associated with moral reasoning patterns employed in L2-sounding speech. With potential explanations of the results, and future improvements and research directions in moral psychology are also discussed.

Keywords: moral decision-making; moral reasoning; native and foreign accents; foreign language effects; emotion; spoken word processing

Introduction

Under the rapid globalization, it has become a commonplace activity for people to make decisions and judgments based on communications in foreign languages. The change in our lifestyles elicits the question about whether processing information in a foreign language (L2), as opposed to a native language (L1), influences the way we process morally salient stimuli. Contradicting to the widely held belief that our moral standards remain constant and act as the major component of our identity (Strohminger & Nichols, 2013), researchers found that people make different decisions not only after reading but also listening to highly-conflict moral scenarios in a native language (L1) than in a foreign Language (Costa et al., 2014; Brouwer, 2019). This effect is named *Moral Foreign Language Effect* (MFEL) (Keysar, Hayakawa, & An, 2013). Although the cause of MFEL remains a debating topic, emotion is often considered playing a pivotal role in modulating moral judgments (Haidt, 2003; Moll et al, 2002). In combination with studies showing the close relationship between sound and emotions (Juslin & Sloboda, 2001; Koeda et al, 2013), the research gap between sound-emotion interaction and auditory MFLE appears.

This present study seeks to fill this research gap and examine the association between moral judgments, moral reasoning process and the sounds used to present moral scenarios. Specifically, native- and foreign accents are adopted as sound cues. In the following sections, I am going to first go over the previous research findings. Hypotheses derived from these studies are presented along with the method and experimenting materials. In the "Result" chapter, data analysis and result interpretation are included. Finally, potential explanations of the results and future research directions in moral psychology are also discussed.

Research Background

Mechanisms for Moral Judgments

Moral judgments, or moral decisions, has been defined as "evaluations (good vs.bad) of

the actions or character of a person that are made with respect to a set of virtues held to be

obligatory by a culture or subculture" (Hadit, 2001). Psychologists have spent decades

investigating the rationale behind moral judgments. For a long time, people take the idea that

moral judgments are driven by one certain type of mental processing. Early rationale claims that

moral judgments are made mainly by reasoning and reflecting (Kohlberg, 1984). While the

reasoning is considered by rationalists as the main process for one to generate a moral judgment,

they do still account for emotions. However, emotions such as sympathy are not considered as

the direct causes of moral judgments for rationalists although they may intervene in the

reasoning process in this theory, as shown in figure 1.
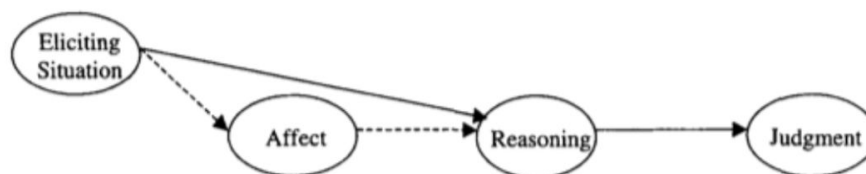
Figure 1. The rationalist model of moral judgement. Moral affects may be involved, but considered less
relavant (Hadit, 2001)

On the other hand, sentimentalists raise the argument that it is the emotional response that

has the primary influence on moral judgments when people are facing difficult moral situations

(Gibbard, 1990). A new social intuitionist approach was constructed, claiming that moral

judgments are made with quick intuitive responses, gut feelings, and emotions. It is after the

decision is made under instinct that people use reasoning to legitimize their conclusions (Haidt, 2001). The social intuitionist model, as shown in figure 2, suggests that the automatic evaluations and moral effects directly drive moral judgments. In addition, Haidt powered his model by suggesting the moral reasoning is rather interpersonal than private. In interpersonal discussions, the reasons given Person A may influence the intuition and judgments of Person B. Specifically, if Person B might want to form an alliance or to have emotional resonance with Person A, his/her intuition will follow Person A's reasoning, and make judgments and develop his/her own reasoning to support Person A's judgments.
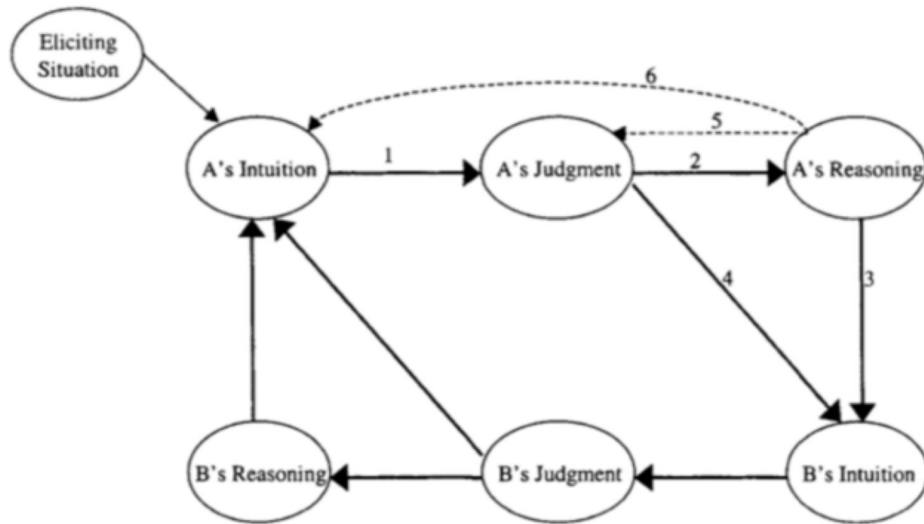
Figure 2. The social intuitionist model of moral judgment. The numbered links, drawn for Person A only, are (1) the intuitive judgment link, (2) the post hoc reasoning link, (3) the reasoned persuasion link, and (4) the social persuasion link. Two additional links are hypothesized to occur less frequently: (5) the reasoned judgment link and (6) the private reflection link. (Hadit, 2001)

Following this model, a growing number of researchers today incline to the hypothesis that emotion plays a significant role in the moral decision-making process. The theory was further developed with detailed emotion classifications, including "moral emotions", defined as "emotions that are linked into the interest or we fare either of society as a whole or at least of

persons other than the judge or agent" (Haidt, 2003) and "self-conscious" emotions, which

provide immediate punishment or reinforcement of behavior and feed back on one's social and

moral acceptability (Tagney et al, 2007, Clavien, 2009). Aside from the theoretical arguments,

neurobiological data have also been offered in support of the statement that moral judgments

were made from a complicated interaction between the brain parts that control cognition and

emotions. For instance, it is found that when evaluating morally salient visual stimuli (e.g., poor

children abandoned in the streets), our prefrontal cortex increased activity and presented an

equivalent degree of arousal compared with conditions where pleasant (e.g., beautiful

landscapes) and unpleasant visual stimuli (e.g., body lesions) were presented. This arousal was

not found when evaluating neutral visual stimuli (including objects, people and landscapes)

(Moll et al, 2002). In addition, according to the study results of Moll and de Oliveira-Souza,

comparing to the controls, patients with prefrontal cortex damage experience decreased guilt

answering a moral dilemma and gave more utilitarian answers, which were considered to be

"less emotional" and focusing on maximizing the well-being for the most individuals (Moll & de

Oliveira-Souza, 2007).

Today, many mixed models were developed based on these foundational models of

rationalists and sentimentalists. The two major hybrid models discussed today are shown in

Figure 3. Again, in both models, emotional stimuli are considered as a strong factor that affects

moral judgments. While the question of where is the exact point that the effect occurs was still

under debate, a number of experiments have demonstrated that emotion has a strong effect on

responses to moral scenarios and suggested that the effect of emotion could occur at multiple

stages including the stage interpreting scenarios, the stage interpreting the moral question, the

stage generating a moral judgment or the stage of reporting the judgment as a response. Given

these research backgrounds, we may speculate that emotional content is a crucial factor in how

people's decision-making and reasoning patterns operate within moral scenarios.
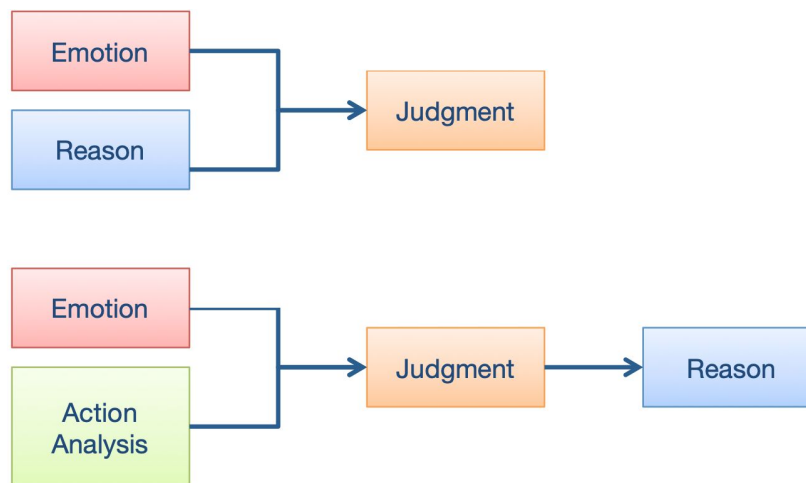
Figure 3. Mixed models of the moral processing. In the lower model, "Action Analysis" refers
to possible morally relevant actions in high-conflict moral dilemmas. Emotion is recruited to
resolve the outputs from these moral actions to make a judgment (Huebner, Dwyer, & Hauser,
2009).

Assessing Moral Judgments and Moral Reasoning

Based on different rationales, various methods have been created to assess our moral

judgments. The most frequently found measurement of moral judgments in MFLE research is

using binary moral dilemmas and categorizing the moral decisions into utilitarian and

non-utilitarian divisions. On the other hand, the Defining Issue Test II (DIT2), derived from

Lawrence Kohlberg's (1969) semi-structured interview provides not only a chance to assess

one's moral decision but also includes a set of five moral dilemmas to measure the moral

reasoning stage into three schemas: the Personal Interest schema, The Maintaining Norms

Schema and the Post-conventional Schema (Kohlberg, 1969; Rest et al, 1999). Specifically, Kohlberg argues that people progress from the Personal Interest schema to Maintaining Norms schema, and then to Post-conventional schema from childhood to adulthood. The Personal Interest schema focuses on the gains and losses each individual may personally experience within a moral dilemma. The Maintaining Norms Schema represents a society-wide moral perspective and emphasizes the importance to maintain social norms and established social orders. Finally, the Post-conventional schema is considered the most advanced moral reasoning stage. Thinking in this schema suggests all moral obligations should be examined based on criteria that emphasize shared ideals and open to scrutiny. That is to say, people who reach the post-conventional reasoning stage are capable of recognizing that rules and laws may be broken to protect individual rights depending on situations. It is distinguished from the other two schemas by challenging the established norms. Given this difference, Rest's measurement may be used to reveal the shift from the Maintaining Norms schema to the Post-conventional schema (Rest et al, 1999).

DIT2 involves relatively novel dilemmas that are less likely for participants to encounter beforehand in their daily life. These are not like other classic dilemmas, for instance, the trolly and footbridge dilemma: dilemmas that involve the idea of sacrificing one's life to saving more people. Hence, it is less likely that people have an established answer to the questions. For each moral dilemma in DIT2, participants are asked to decide what the protagonist ought to act. Then participants rate 12 items on a 5-point scale based on how they perceive the items' importance. Finally, they rank four items that best describe their understanding of the best solution to the dilemma. By examining the rate and ranks, the complexity of the cognitive process can be

presented by a concrete index. Therefore, DIT2 can be considered as a measurement that aims to move a step further from recording immediate responses and assess the thinking pattern that may also be altered by the use of different languages.

Moral Foreign-Language Effect

The moral foreign-language effect (MFLE) is detected robustly within this decade. It represents the situation that people make systematically different decisions to a moral dilemma (Cipolletti et al, 2016). One of the pioneer studies in this field shows that processing messages in L2 may increase psychological distance and attenuate emotions, resulting in relatively utilitarianism moral decisions (Costa et al., 2014). Costa et al. (2014) adapted the classic Footbridge moral dilemma:five people are tied to the train track and are about to be killed by an approaching train. To save their lives, the only way is to push a large bystander from the footbridge onto the tracks to stop the train from running over those tied to the track. In return, the life of the bystander will be sacrificed to do so. When the question is presented in the participants' L1, only 18% of the participants made a utilitarian decision to sacrifice the large man. However, when the dilemma is presented in the participants' L2, a surprising 44% of all answers chose to sacrifice the bystander(Geipel, Hadjichristidis & Surian, L, 2016). However, this pattern is not observed in a Switch dilemma. This dilemma asks participants to choose either to see a coming train crash into five workmen on the track or to switch the train using a lever and have the result of killing only one workman on the other track. Additionally, the use of a foreign language is found influencing the relative weight of intentions versus outcomes when judging the moral goodness in a given scenario (Geipel, Hadjichristidis & Surian, L, 2016).

So far, psychologists have successfully detected MFLE with Korean-English, English-Spanish, English-Hebrew, Spanish-Hebrew and English-French bilinguals (Costa et al, 2014; Geipel et al, 2015a, 2015b). Although the MFLE has been replicated and detected across a variety of languages, it remains in the center of debating why foreign-language use can affect our moral judgments. One widely discussed hypothesis comes along with the emotional rational for moral decision-making process. It suggests that thinking in an L2 elicits emotional distancing compared to thinking in one's L1, allowing a deliberative thinking process to take place instead (Keysar, Hyakawa & An, 2012; Costa et al, 2014; Cipolletti et al. 2016). Additionally, in Costa et al (2014), it was found that proficiency in L2 influences the size of the MFLE:high proficiency participants made less utilitarian decisions in an L2 than low proficiency participants. Costa and colleagues further explained this phenomenon speculating that high proficiency participants have developed more emotionality in L2 (Costa et al, 2014).

While all these discussions were based on moral judgment tests in the written form, it is observed that there is an auditory MFLE in both moral dilemmas recently in highly proficient Dutch-English bilinguals (Brouwer, 2019). In the first experiment, the Dutch-English bilinguals were asked to read nine moral dilemmas presented either in Dutch or English. In the second experiment, the participants from the same population were asked to listen to the same nine moral dilemmas either in Dutch or English. The results revealed that the MFLE, an increase in the rate of utilitarian decisions in an L2 compared to an L1.Remarkably, this significant MFLE was only detected in the listening experiment. Brouwer's result complied with was interpreted as revealing an important connection between the amount of emotionality in the language and moral decision making: that certain types of words and phrases are more emotionally intense in

the participants' L1 when heard than when read (Caldwell-Harris and Ayçiçeği-Dinn, 2009). For

instance, childhood reprimands such as "Go to your room" in L1 may elicit an emotion-memory

effect in which the sound of the words reminds native speakers of the emotion intensity they

received when they were little. This distinctive difference between "reading" and "listening"

leads to the following research question of how the sound of a language may convey different

emotions, alter people's perceptions and ultimately influence their moral judgments.

Psychological Accounts of Auditory MFLE

Sound and Emotion Perception

According to the auditory neuroscientist Seth Horowitz, "sound is a major emotional

driver for humans" (NPR, 2012). Not only can meaningful written words and languages convey

emotions, but also, previous research has found that people are sensitive to non-semantic sounds

such as vowel segments in terms of perceiving the emotions behind them (Juslin & Sloboda,

2001; Airas & Alku, 2004). Specifically, in one study that focused on the prosody-based

sound-emotion associations in Poetry, non-German-speaking participants were recruited to listen

to German poetries presented with joyful and sad prosodic cues (the expressive condition) and

German poetries presented with a neutral variant (the neutral condition). The results showed that

the non-German-speaking participants were able to identify the emotions accurately in

expressive conditions but not in the neutral condition (Kraxenberger, Menninghaus, Roth &

Scharinger, 2018), demonstrating the idea that emotions could be conveyed through speech

without understanding the language.

Although emotions are often universally recognized across different cultures, other sets

of studies have shown that native speakers of a language or intracultural group members

typically perform better in recognizing emotions contained in the language than those outside of

the language or culture (Graham, Hamblin, & Feldstein, 2001; Paulmann & Uskul, 2013).

Furthermore, it was found that a L1 speaker can perceive the emotions in non-verbal

vocalizations presented by other L1 speakers better than those who learned it as a second

language. In research conducted between Japanese and Canadian listeners, eight emotional

vocalizations and non-emotional vocalizations using "ah" portrayed by Canadian actors were

presented to all listeners. Canadian listeners presented higher sensitivity for negative emotions

including anger, disgust, fear, and pain, compared to Japanese listeners. Further, the results

revealed that while Japanese listeners can distinguish positive emotions, such as happiness and

pleasure, they presented difficulties distinguishing negative emotions such as disgust and fear

(Koeda et al, 2013). These findings could be related back to the Sentimentalist moral processing

mechanism that emphasizes emotion's role in the moral decision-making process. Given the

concept that anger and disgust are main "moral emotions" that can drive moral judgments (Haidt,

2003), the difficulties with detecting and distinguishing negative emotions in L2 speeches that

L2 learners have may influence their moral decisions. Subsequently, they may make more

utilitarian decisions and process the moral dilemmas within moderate moral reasoning stages that

focus on personal interests and social norms.

To return to the relationship between sound and emotion, one may ask why this

emotion-gap would appear between different language speakers. The source-filter model

developed in the field of linguistics may be considered an explanation for this phenomenon. This

model was proposed to understand the ways in which speech acoustics might provide

information about emotional state through the combination of sound source energy (i.e. the

vibration of vocal cords) and the filtering of this energy by the resonant properties of the vocal

tract (Kent, 1997). While the measures associated with frequency of speech (F_0) are most

commonly used in emotions research, the filter-related cues are examined less often. However,

these cues can be crucial for conveying emotions, since facial expressions can also influence

filtering effects. Moreover, the size and shape of the vocal tract can also mediate which

resonances are reinforced, which in turn can affect emotional response. In other words, our vocal

habits (e.g., muscle tension, pitch range, etc.) may be received as source- or filter-related cues for

emotions (Kent, 1997). Henceforth, we may infer that those, who are extremely familiar with one

language, including its tone, prosody and nuances in vowel segments, are much more capable of

detecting the subtle cues in a speech in this language. As a result, when listening to speeches in

less familiar sounds, people may make more utilitarian decisions and less likely to process moral

dilemmas with higher reasoning stages with an idea that rules may be broken for individual

morals.

From Sound to Accent

Vocal habits can be formed based on the phonation pattern of one's native language.

Henceforth, the differences in the frequency and tone of the speech due to our vocal habits forms

the concept of "accents" with which we are all familiar. Given the concept that sound

characteristics of different languages are recognized as salient signals of social group

membership and provide clues to the speaker's identity (Creel, 2016), as a type of sound, accent

can also accentuate the concept of "in-group" and "out-group". For instance, the Social Identity

Theory (Tajfel, 1959) suggested that it is likely that people form a sense of in-group and

out-group based on the sound of the language spoken. As a result of in-group bias, which is the

preference of people who share the same linguistic sound in this case, the out-groups are

considered more socially distanced and might be evaluated negatively to defend the self-esteem

for the in-group. Prior work has demonstrated that adults are very sensitive to novel accents, and

they often make negative social evaluations of non-standard speakers. For instance, speakers

with southeastern US accents were reported to sound less educated (Campbell-Kibler, 2007), and

customers rated employees with non-standard accents lower than those with a standard accent

(Wang, 2013). Even preschoolers are found to be sensitive to accent distance. After listening to

speakers with different accents, the five-year-olds were instructed to place each speaker on a

map with a reference point that represented their current location. The five-year-olds ended up

placing speakers with stronger accents at more distant locations than speakers with weaker

accents. Moreover, they were able to recognize foreign accents as being more "alien" compared

with regional accent speeches (Weatherhead, Friedman, & White, 2019). In another set of

studies, it is found that 5-6 year-olds prefer to be friends with fictional children who speak their

native language or accent (Kinzeler et al, 2012). These ideas are consistent with finding

indicating foreign-accented speech can negatively impact message reliability, even though these

foreign-accented speakers were merely reciting statements provided by native speakers (Lev-Ari

and Keysar, 2010).

Based on this, Researchers have further proposed a hypothesis associating the linguistic

distance to the social distance. It is also observed that people perceived low social attractiveness

on the speakers who speak a different accent from them, even though they were communicating

based on the use of the same language. It is only when they imitate the speaker's accent by

themselves, they report increased social attractiveness to the same speaker. This explained

reduced social distance through imitation as well (Adank et al., 2013). An increased social

distance from the different linguistic sounds of speech may result in inefficiency in conveying

emotions through the speeches. In other words, familiar sounds may eventually better convey the

emotions and eventually lead people to make less utilitarian decisions to moral questions,

understanding the situation comprehensively and ultimately adopting a more developed moral

reasoning stage.

While research found an accent-effect on social cognition, evidence of complicated

interactions between accent and our information-processing cognitive circuits were also offered.

Prior research has revealed that people process the language of a non-native speaker with fewer

details and suggested the adjustment to non-native speakers depends on working memory
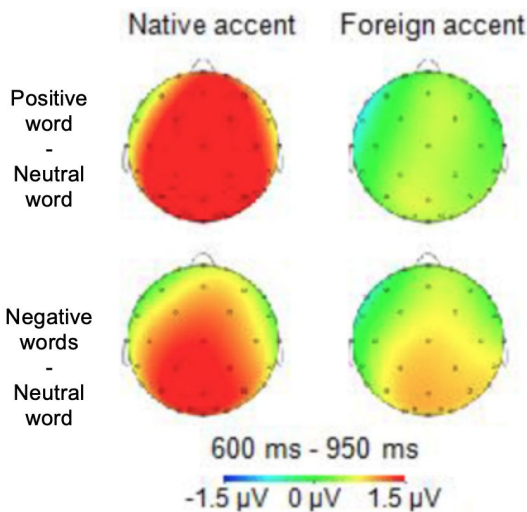
Figure 4. Brain topographic map of cerebral activity of Spanish native listeners listening to positive and negative words in native- or foreign- accented Spanish. To make accurate comparison, top panel shows the wave differences of positive word amplitudes minus neutral word amplitudes; and the bottom panel shows the result of negative word amplitudes minus neutral word amplitudes (Hatzidaki, Baus, & Costa, 2015).

(Lev-Ari and Keysar, 2013). Further, in Hatzidaki, Baus, and Costa's (2013) study, recordings of positive, negative and neutral words spoken in Standard Spanish and Spanish with four different foreign-accents were played for participants. As shown Figure 4, increased cerebral activity was detected when native speakers were processing emotional words in native-accented speech than in foreign-accented speech. This effect of emotional word processing

starts earlier in native than in foreign accent. Remarkably, when processing negative words in foreign accents, listeners were found having increased cerebral activity than processing positive words. This may suggest a heightened attention to negative words when processing content in foreign accents (Hatzidaki, Baus, & Costa, 2015). This alienation from emotions in foreign accents may lead to utilitarian decisions in moral judgments and a less deliberated thinking pattern.

Additionally, adding an extra tone or sound to the standardized language would result in a delayed onset of the semantic process no matter the familiarity of the sound, or to say, the familiarity of the accent, suggesting cognitive resources are required to comprehend foreign accents(Grey et al, 2018). Research was done looking at the differences processing speech in native-accented and L1-congruent foreign-accented conditions. The results show that L2 speakers of a given L1 are typically more experienced hearing their own foreign accent in L2, proving that the experience with a foreign-accent improves the intelligibility of the speech. In other words, the general difficulty of speech comprehension in L2 can be attenuated when L2 is presented with an L1 accent. This phenomenon is termed as the *Interlanguage Speech Intelligibility* (Bent, T., & Bradlow, A. R, 2003). It is thereby considerable whether listening to a familiar sound of their native language allows people to better understand and process the speech while listening to a foreign-accented speech requires more cognitive processes on the semantics. In actual examinations, this should be reflected in the presence of a more deliberate thinking process when people listen to the moral dilemmas in L1 accented L2 compared to messages in L2.

Based on these findings, it can be hypothesized that aside from the language type (L1 or L2), accent, the special tone of speeches, may further influence the way we receive and process information. There is a gap between the information processing with unfamiliar sounds and the MFLE that leads us to practice empirical studies and discover the relationship between the sound of language, moral reasoning and moral judgments.

The Present Study: Hypotheses and Method

Hypothesis

The goal of the present research is to (1) replicate the auditory MFLE, and (2) examine whether the power of auditory MELF could be mediated by the familiarity to the sound of the spoken language, specifically the familiarity to an accent.

In this study, an uninvestigated L1-L2 language pair, the Chinese(Mandarin)-English language pair, was adopted. Chinese participants who studied English as their second language were exposed to moral dilemmas presented in their L1 (Chinese), L2 (English), L1 - accented L2 (Chinese-accented English), and L2-accented L1 (English-accented Chinese). In all four conditions, participants' stages of moral development involved in analysing moral dilemmas were assessed. Based on the higher ability to convey emotions and providing detailed messages of L1demonstrated in prior studies, I hypothesized that (1)an auditory MFLE will appear: native Chinese speakers will make less utilitarian decisions when listening to moral dilemmas read in a native Chinese speech than in a native English speech. Also, I hypothesize that (2) participants will present a more advanced stage of moral thinking when exposed to Native Chinese audios than when exposed to Native English audios. I further hypothesized that (3) participants exposed

to questions read in native Chinese will perform advanced moral reasoning than when exposed to questions read in English-accented Chinese and (4) participants will report a higher stage of moral thinking when listening to moral dilemmas in Chinese-accented English compared to native English. Additionally, based on Costa et al (2014) and Brouwer's (2-10) study suggesting that proficiency in L2 may eliminate the MFLE (2019), I hypothesized that (5) higher proficiency in English will lead participants to adapt a higher stage of moral reasoning after encountered with moral dilemmas read in native English and in English-accented Chinese. Specifically, the number of years learning, exposed to English Environment and TOFEL scores will be positively related to the final scores in English-accented Chinese and native Native English conditions.

Method

Participants

Chinese students at Bard College were invited to participate through the school email system (see Appendix C for the invitation email). It was clearly stated that the experiment was open for native Chinese speakers only. A total of 35 Bard College students were determined eligible and completed the study. All participants' ages were equal or greater than 18 years old. None of the participants reported that they were not focused on the tasks. One participant was excluded before data analyses due to accidentally closing the survey webpage halfway through the completion. The final sample included 34 participants (*Mage* = 20.92, SD = 1.37), in which 13 are male and 21 are female.

Materials

Moral Decision-Making Assessment. In this study, a questionnaire with four moral

dilemmas (Famine, School Board, Cancer, Demonstration)was generated from the original DIT2

test. For each dilemma, participants were asked to decide what the protagonist ought to act.

Then, the participants were asked to rate the importance of 8 issues related to the dilemmas they

listened to on a 5-point likert scale (1 indicated not important at all and 5 indicated very

important). Specifically, participants were asked to consider to what extent, each issue was

important to be discussed and considered rather than answering the degree of agreement. Half of

these items fell into the Post-conventional Schema, while the other half fell into either the

Personal Interest schema or the Main Norms schema (Rest et al., 1999). Finally, participants

were asked to rank two issues that they believed were the most important and the second most

important issue to think about to achieve a best solution to the dilemma. The items belong to the

Personal Interest schema and Main Norms Schema were reverse scored. The two issues that the

participants ranked as with top two importance were scored based on how the issues are scored

from the rating task (e.g. An issue ranked as one of the top two that is scored/reverse-scored as 3

in the rating task was scored 3 in the ranking task). By examining the final total score of the

rating and ranking process, the moral reasoning state participated in each moral dilemma were

presented numerically. In this questionnaire, the highest score would be 50 and the lowest would

be 10. While the order of the four conditions was randomized, the dilemmas were presented in

the same order (Famine, School Board, Cancer, Demonstration). The questionnaire was referred

to as "Situation Questionnaire " in the experiment (see Appendix E for Situation Questionnaire).

Recordings of the Dilemmas. Four dilemmas chosen from DIT2 were translated into Chinese with assistance from the College Chinese Professor. The dilemma "Reporter" was excluded due to the context of democracy depicted in it. Most participants were from Mainland China and had never experienced choosing a government through voting in an election by themselves. As a result, it was used in the practice trail. Native Chinese and English speakers without prominent regional accents were responsible for recordings of native Chinese and native English. The audio recording of dilemmas read in English-accented Chinese was provided by a white American female undergraduate student who has learned Chinese as a second language for 6 years with one-year experience in Mainland China. The audio recording for the Chinese-accented English condition was provided by a female Chinese student from another American University on the west coast who has learned English for 12 years with 3-year experience in the United States. Both recorders were instructed to read the moral dilemmas in a neutral tone instead of making portrayals of specific emotions. All four dilemmas were recorded into four language versions. The audio for the practice trial was recorded by the Chinese experimenter in English. The pause between sentences was controlled within 1 second. Repetitions, corrections and nonverbal articulations (e.g. coughing) were also excluded from all the audios. The length of the recordings were measured from the first syllable to the last in the audio files based on the visualized sonic waves. There were length differences between recordings in native languages and speeches with a foreign accent ($Mdiff$ = 21s, SD = 17.15). Recordings of dilemmas in native Chinese and native English were controlled to be of the same length (F: 40s; SB: 59s; C: 33s; D: 50s), and recordings of accented Chinese and English were controlled to be of the same length (F: 55s; SB: 98s; C: 33s; D: 80s).

Manipulation Check. After listening to each moral dilemma and before completing the Situation Questionnaire, participants were asked to report their level of understanding of the dilemma by responding to seven statements on a 5-point Likert scale (1= Strongly disagree, 2 = Disagree, 3= Neutral, 4 = Agree, 5= Strongly Agree). One of these statements, "I have fully understood the situation described in the speech", aimed to check if the recording content, especially the accented ones, can be understood for most participants. The statement "The speaker presented a non-native accent speaking this language"was used to further assess whether participants perceive the differences between the audios used in a foreign-accented condition and a non-accented condition. The questionnaire was referred to as the "ClarityQuestionnaire " in the experiment(see Appendix F for Manipulation Check Questionnaire).

Demographic Questionnaire. After finishing all four Situation Questionnaires, participants were asked to provide additional demographic information, including age, gender, number of years learning English, number of years exposed to English Speaking Environment and their most recent TOEFL (Test of English as a Foreign Language) or IELTS (International English Language Testing System) score if applicable. Participants were also asked to report if they were paying attention to the tasks or not in this section (see Appendix G for Demographic Questionnaire).

Procedure

An email invitation was sent to the Chinese student group on campus inviting students whose first language is Mandarin. A google timesheet was also included for sign up. The time slots were provided from 10:00 am to 5:00 pm in a day (see Appendix C for the Recruitment

Email). Students were informed that the study was related to the field of linguistics and

evaluating the level of speaking for several speakers. Potential participants were also approached

by the experimenter at the college library. The eligibility (proficiency in both Mandarin and

English in listening and reading) was explained to all participants recruited online and offline.

All eligible participants were invited to a sound-proof room set up for the study within the

college library. All studies were conducted using English as the primary language by an Asian

female experimenter whose first language is Mandarin.

In the room, the experimenter greeted all participants and guided them to sit in front of a

computer with a speaker connected. Participants who provided informed consent (see Appendix

D for the Informed Consent Agreement) proceeded to the task on the computer screen. All

questionnaires were presented in one survey made with an online survey software named

SurveyGizmo.The experimenter then introduced a cover story and went through the tasks

included in the experiment. In the cover story, the aim was described as to investigate the

importance of clarity in speeches in terms of communication efficiency. The study was hence

introduced as a "Clarity Survey" , which was also used as a prominent title at the top of the

webpage throughout the entire task. Participants were told to pay attention listening to four voice

recordings depicting difficult situations in life in English and Mandarin and evaluate the clarity

of speeches right after each recording in the "Clarity Questionnaire". The "Situation

Questionnaire" was introduced as a further assessment and clarification of the participants'

understanding of the situations depicted in the recordings. A written instruction was also

provided as the experimenter explained the task on the screen. All content was identical with the

experimenter's explanation (see instructions in Appendix H for Experiment Script).

The participants were then guided into a practice trial to make sure they fully understand the task. In this practice trail, the "Reporter" dilemma was used with 8 issues selected randomly from DIT2. This dilemma was only presented in L1 accented L2 and the recording was made by the experimenter. Once the participants had no questions about each questionnaire, the experimenter told them that she will sit on a chair in the corner of the room (with participants' back to the experimenter) and work on other assignments. The chair was far enough that the experimenter was not able to see the contents on the computer. The participants were allowed to start once the experimenter sat down. The experimenter then started observing the participants' behavior during the task.

The last page of the online survey guided participants to notify the experimenter after submitting their answers. Participants were debriefed and presented with the original purpose of the study: investigating how people make systematically different decisions to a moral dilemma in a native and foreign language setting (see Appendix I for Debriefing Form). Participants were asked not to share the research question and experiment process with other students. Additionally, they were asked if they recognized any of the recordings as from someone they knew personally on campus. They were then thanked, handed the 5 dollar compensation, and asked to sign a payment receipt (see Appendix J for Payment Receipt). There were no time limits and the entire experiment took an average 30 minutes to complete.

Results

Data Preparation

Raw data were exported from Survey Gizmo, cleaned and arranged in Excel and analyzed with SPSS and R. All issues belonging to Personal Interest and Maintaining Norm Schema, including those chosen in the ranking questions, were reverse-scored. For instance, if an issue that could be applied to this situation was originally rated as 5 (great importance) by the listeners, this answer will be recorded in reverse, earning 1 score in the final scoring; if this issue was originally rated as 2 (little importance), a score of 4 was added to the final score. This was done to keep consistency, so that a higher score indicates a reasoning pattern closer to the Post-conventional Schema, and a lower score reflects one closer to the Personal Interest and Maintaining Norm Schema.The final total score in each language condition was calculated by adding up the eight scores received in the rating section and two scores received in the ranking section in the Situation Questionnaire. The issues that participants chose in ranking questions were recorded according to the schemas to which they belong, and this was recorded in a format of "the schema of the most important issue - schema of the second most import issue": P-P, P-IM, IM-P, and IM-IM (P = Post-conventional Schema, IM = Personal Interests & Maintaining Norms Schema). Two IELTS scores reported in demographic questionnaires were converted into TOEFL scores using the TOEFL iBT/ IELTS Total Score Comparison Tool constructed by Educational Testing Service (ETS).

Manipulation Check

Answers to the first question: "I fully understand the situation described in the speech"(Und) and the last question: "The speaker presented a non-native accent speaking this language"(Acc) in the clarity questionnaire were analyzed to check if the content was accessible

and if it was clear that a native or foreign accent was presented. As shown in table 1, in the

Native Chinese and Native English condition, participants reported that they had fully

understood the situation and considered the speaker as a native speaker of each language. In

Chinese-accented English condition and English-accented Chinese condition, participants

reported relatively high comprehension of the speech. They also reported that speakers had

strong non-native accents in these two conditions. The results revealed that these two variables

were precisely manipulated. Participants were able to detect the accent differences and also

understand the situations described in the speeches.

Table 1

Results of questions on understanding (Und) and accent presence (Acc) in four language conditions

	Und		Acc	
Condition	*M*	*SD*	*M*	*SD*
Native Chinese	4.79	.48	1.59	1.02
Native English	4.35	.98	2.08	1.504
Chinese-accented English	3.91	.87	4.76	.496
English-accented Chinese	4.09	.996	4.82	.387

*For question Und, a score of 5 indicates fully understood the content and 1 indicates did not understand at all. For question Acc, a score of 5 indicates that participants perceive a strong foreign accent, and 1 indicates they did not perceive a foreign accent.

Moral Assessment

Moral Decision-Making

The decisions made at the beginning of the situation questionnaire were categorized into

three categories: Cannot decide, Utilitarian decision, and Non-utilitarian decision as presented in

Figure 5. One-way chi-square tests were conducted to examine the relationship between

decisions and language conditions. The results demonstrated that there was a non-significant
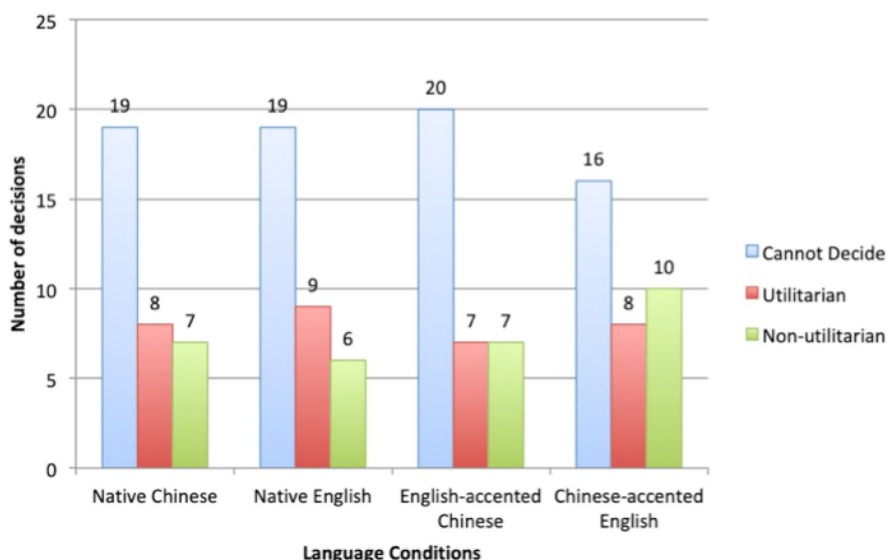
Figure 5. Types of moral decisions made in different language condition

relationship between language conditions and decision types, X^2 (6, N = 136) = 1.935, p = .925.

Neither a significant relationship between language conditions and decision types were detected

when decisions were classified into two categories: "Cannot decide" and "Decided" (both

utilitarian decisions and non-utilitarian decisions included), X^2 (3, N = 136) = 1.935, p = .785.

These results revealed that the differences in the sound of language were not found capable of

mediating people's tendency towards making a moral decision, or remain ambiguous in the four

language moral dilemmas adopted in this study.

To further examine my hypothesis that native Chinese speakers make less utilitarian

decisions after listening to the dilemma presented in native Chinese than in native English, I

conducted analyses to compare the prevalence of utilitarian decisions or non-utilitarian decisions

in native Chinese and native English Conditions for four dilemmas separately. In Figure 6,

Figure 6. Mean percentage of utilitarian decisions by dilemma and language. Only reports beside "Cannot decide" decisions in a moral dilemma were included into this analysis. "Total" represents the analysis that included all "Utilitarian" and "Non-utilitarian" responses in the four moral dilemmas.
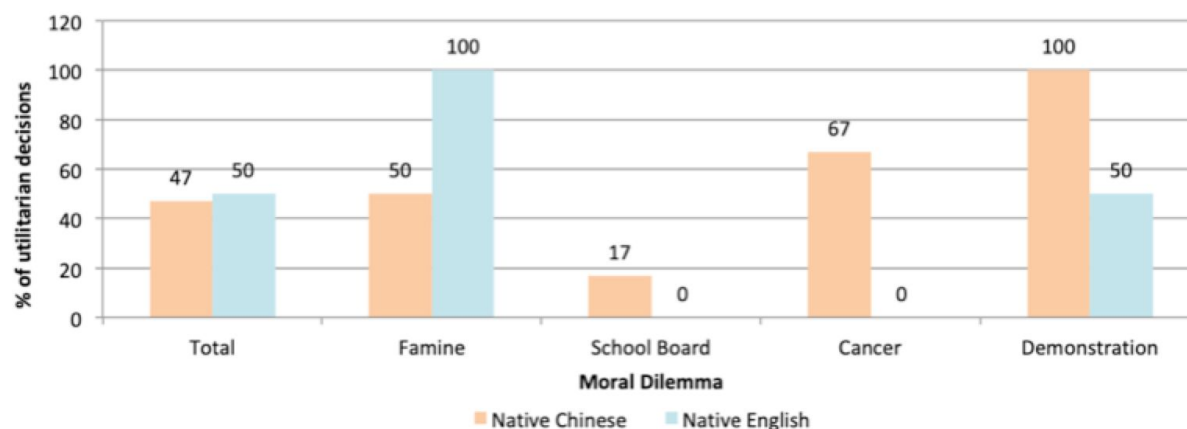
the percentage of utilitarian decisions made within the total number of "Decided"decisions

(excluding "Cannot decide" decisions) made in all conditions and in each dilemma were

presented. Since reports of "Cannot decide" were excluded from the analysis, this data filter

caused the small sample size of each dilemma. In the total analysis that included responses from

four moral scenarios, 47% of "Decided" reports collected in the Native Chinese condition were

utilitarian decisions ($N_{utilitarian} = 8$, $N_{non-utilitarian} = 7$). In Native English condition, 50 % of

"Decided" decisions were utilitarian decisions ($N_{utilitarian} = 9$, $N_{non-utilitarian} = 6$). A 2 (Language:

Native English or Native Chinese) x 2 (Decision type: Utilitarian or Non-Utilitarian Decision)

one-way Chi-square test result showed that after listening to moral scenarios read in native

English, native Chinese speakers did not make a different moral decision than after listening to

the scenarios in their native language, X^2 (1, $N = 30$) = .1357, $p = .713$. It was also revealed in

the result of the logistic regression model that there was no significant effect of Language,

estimate = - . 27, $se = .74$, z-value = -. 37, $p = .71$, on moral decision types. This data indicates

that for Chinese native speakers, the odds of making a utilitarian decision in Chinese (L1) was

equivalent to English (L2). In other words, inconsistent with previous research that suggested the existence of auditory MFLE (Brouwer, 2019), an auditory MFLE on moral decision-making was not found in this study.

Subsequently, I went a step further, analyzing if there was any auditory MFLE presented in a certain moral scenario. Since there were categories left with no responses after excluding the "Cannot decide" answers in all four dilemmas, Chi-square tests did not apply to the analysis between language conditions and decision types in each moral dilemma. Instead, Fisher's Exact Test was conducted to examine if there was a nonrandom association between moral decisions and language types in each moral dilemma. Again, the association was not found significant in any of the four moral scenarios, p_{famine} = .429, $p_{schoolBoard}$ = .99, p_{cancer} = .40, and $p_{demonstrationr}$ = .467, (see Appendix M for detailed Frequency table for Moral Decision x Language analysis), indicating that in all four moral dilemmas, the frequencies for native Chinese speakers to madea utilitarian decision or a non-utilitarian decision were equivalent when the dilemmas were read in native Chinese and in native English. With these data, we were not able to reject the null hypothesis that stated there are no differences in Chinese native speaker's moral decisions (utilitarian or non-utilitarian) when they listened to moral dilemmas in a native Chinese and in a native English. The auditory MFLE was absent in this data analysis.

Moral Reasoning Schema

A repeated measure ANOVA test was used to detect the differences between the total score for four different language conditions in the situation questionnaire. The result showed that there was no significant differences between each condition, $F(3, 31)$ = 1.474, p = .241, η2

=.125, and failed to reject the null hypothesis that participants present did not present different

stage of moral thinking when exposed to moral dilemmas recorded in different languages with

forign or native accents. However, the disposition of the total score from each language

condition matched my hypothesis. The mean total scores presented in Figure 7 showed
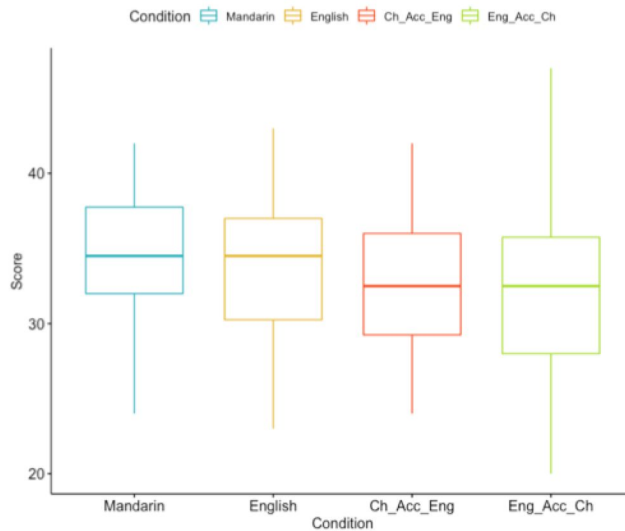
Figure 7. Box plot of the total score of situation questionnaire for four language conditions. The mean scores show a consistent declining trend of the total score in order of Mandarin > English > Chinese-accented English > English accented Chinese.

a consistent declining trend in order of Native Chinese ($M = 34.18$, $SD = 4.66$), Native

English($M = 33.561$, $SD = 5.1$) , Chinese-accented English ($M = 32.97$, $SD = 4.68$) , and

English-accented Chinese condition($M = 32.00$, $SD = 5.11$). That is to say, when the moral

dilemma was presented in native Chinese, participants (all of whom were native Chinese

speakers) had a tendency to adapt more advanced moral reasoning schema compared to the

situation in which moral dilemmas were presented in standard English. Although in Native

Chinese and English-accented Chinese conditions, the content in the audios were both Chinese,

when an English accent was added to the Chinese speech, participants presented a moral

reasoning schema that emphasizes rather more on personal interest and maintaining the social

norm. Still, the differences between the mean total scores remained subtle according to ANOVA.

We should treat this tendency carefully and interpret it with caution instead of suggesting it as

an affirmation.

Table 2

Paired t-test Results comparing total scores earned in different language

Paired language conditions	N	M_{diff}	SD_{diff}	t	p
Native Chinese – Native English	34	.62	5.90	.610	.546
Native Chinese – English-accented Chinese	34	2.18	5.95	2.13	.041[*]
Native English – Chinese-accented English	34	.588	6.68	1.221	.231

Due to the unexpected Covid-19 situation, I failed to gather a large enough sample to

adequately power this study. This deficiency was likely contributing to the null results obtained

in the repeated measure ANOVA (Power = .351). To further explore the interactions, I went

ahead and performed the focused comparisons I would have conducted were I to have obtained a

significant effect in the ANOVA. Specifically, to examine my prior hypothesis, I compared the

total scores in the four language conditions using paired-sample t-tests as shown in Table 2.

While this result should be interpreted with caution given the lack of a significant effect from the

ANOVA, it reflects a significant difference between the total score in the Native Chinese

condition and the English-accented Chinese condition, t (33)= 2.13, p = .041, d = .37. This

provided some evidence to the hypothesis that participants listening to moral dilemmas in

Chinese-accented English compared to native English present a stage of moral thinking incline

toward the Post-conventional schema.

Additionally, multiple chi-square tests were also conducted to examine the relationship

between the responses in ranking questions and language conditions. As described in the data

preparation section, all ranking responses were categorized into four conditions based on the three moral reasoning schemas. As can be seen by the frequencies cross tabulated in Table 3, the percentage of participants that gave ranking answers of P-P did not differ by language conditions, X^2 (9, $N = 136$) = 8.385, $p = .496$. Even in the chi-square test conducted between the two extreme language conditions, the Native Chinese condition and Native English condition reflected only a non-significant relationship between the language condition and ranking results, X^2 (3, $N = 34$) = 1.12954, $p = .7699$. Based on the results, we may infer that the differences in the sound of speech did not differentiate people's choice of two issues that they believe to be the most important ones. Specifically, in all four conditions, over 70% participants chose at least one issue that belonged to the Post-conventional schema as one of the two most important issues.

Table 3

Chi-square test results for subjects grouped by their answers to the ranking questions

	Ranking Condition*					
	IM-IM	IM-P	P-IM	P-P	X^2	p
Characteristics	n (%)	n (%)	n (%)	n (%)		
Language Condition						
Native Chinese	2 (5.9)	6 (17.6)	7 (20.6)	19 (55.9)	8.385	.496
Native English	2 (5.9)	8 (23.5)	4 (11.8)	20 (58.8)		
Chinese-accented English	3 (8.8)	6 (17.6)	11 (32.4)	14 (41.2)		
English-accented Chinese	3 (8.8)	7 (20.6)	12 (35.3)	12 (35.3)		

* IM = Personal Interest and Maintaining Norm Schema; P = Post-conventional Schema; Represented in order of [most important issue] – [Second most important issue]

Demographic Analysis

Familiarity with English

To examine my hypothesis that the familiarity with a second language will be positively

related to the total scores from Situational Questionnaire for L2-accented L1 and native L2

speeches, multiple correlation analyses were conducted to investigate the relationship between

variables that may reflect the familiarity to English, including (1) Age, (2)Number of Years

Learning English, (3) Number of Years Exposed to English Speaking Environments, (4) Most

Recent TOEFL / IELTS Score and the total scores in Native English and English-accented

Chinese conditions. Results in Table 4 indicated that while all other variable pairs were found to

have weak correlations, only the number of years learning English (*M*= 14.12*, SD*= 2.92) and

years exposed to English environments (*M*= 4.46*, SD*= 3.11) were strongly positively correlated,

r = .518, p <. 01. Additionally, the total scores in the Native English and English-accented

Chinese Conditions were significantly negatively correlated with age, r = - .347, p <. 05; r = -

.320, p <. 05 (see Appendix K for Correlation Scatterplot Matrix). This result showed that as the

age increases, the moral processing schema used when listening to moral dilemmas presented in

L2 sounds tended to incline to the Personal Interest and Maintaining Norms schema rather than

the Post-conventional schema.

Table 4
Means, standard deviations, and correlations between demographic variables and total scores

Variable	*M*	*SD*	1	2	3	4	5
1. Age	20.93	1.37					
2. Yrs_Learning	14.12	2.92	.180				
3. Yrs_Expose	4.46	3.11	-.029	.518**			
4. TOEFL	102.85	8.13	-.20	.071	.362*		
5. Eng_Total	33.56	5.11	-.347*	-.013	.091	.136	
6. EngAcc_Total	32.00	5.83	-.320*	-.075	-.025	-.017	.078

*Note. *Indicates p<. 05. ** Indicates p<. 01.*

To examine the mediating effect of age on the adaptation of moral reasoning schemas, a

linear regression was calculated to confirm that age is a significant predictor of the total scores in

Native English and English-accented Chinese conditions. While the equation for

English-accented Chinese condition was found not significant $F (1, 33) = 3.427$, $p = .073$, $R^2 =$

.069 a marginally significant regression equation was found in the Native English condition $F (1,$

$33) = 3.933$, $p = .056$, $R^2 = .082$. Either of these two models was sufficient enough to explain the

variability of the total scores received in Native English and English-accented Chinese

conditions.

A multiple linear regression analysis was further calculated to investigate if the

independent variables used as measures of the familiarity to English were significant predictors

of the dependent variable (total score in Situational Questionnaire) compared to age in Native

English and English-accented Chinese condition. Nine subjects were excluded from this

regression analysis due to lack of reporting TOEFL/IELTS scores. Again, only a marginally

significant regression equation was found, $F (4, 25) = 2.735$, $p = .056$, $R^2 = .217$, in Native

English condition. Still, the effect size of this model remained weak, $R^2 < .25$. That is to say, this

model that included measures of familiarity with English was not able to explain much in the

variation of the total score earned when a moral dilemma was presented in native English. The

previously marginally significant independent variable (Age) was the best and only significant

predictor in this equation ($p = .007$). Specifically, it was found that when age increased by 1, the

total score earned when listening to a moral dilemma in native English decreased by 0.5. The rest

of independent variables were not found as significant predictors of the total score. Meanwhile,

in the English-accented Chinese Condition, no significant regression equation was found, $F (4,$

$25) = 1.04$, $p = .410$, $R^2 = .006$. The familiarity with English was not found to be powerful

enough to modulate the moral reasoning schema one used to process moral dilemmas presented

in native English or English accented Chinese speeches.

Table 5

Results of the Multiple Regression Analyses by Language Condition

Language Condition	t	p	β	F	p	adj. R^2
Native English						
Overall model				2.735	.056	.217
Age	-2.97	.007	-.572			
Years Learning English	-.013	.990	-.003			
Years Exposed to English	1.49	.151	.329			
TOEFL Score	-.850	.405	-.167			
English-accented Chinese						
Overall model				1.040	.410	.006
Age	-2.01	.057	-.437			
Years Learning English	.296	.770	.072			
Years Exposed to English	.277	.785	.069			
TOEFL Score	-.604	.552	-.124			

Comparison by Gender

The final sample included 14 respondents who self-identified as male and 20

respondents who self-identified as female. A 2 (Gender: Male or Female) x 4 (Language

Condition: Native Chinese, Native English, Chinese-accented English, or English-accented

Chinese) repeated measure ANOVA test and chi-square tests were conducted to examine

whether the moral decisions, the ranking results or the total scores in each language condition

were differed by the subjects' gender. Results of chi-square tests indicated that types of

decisions and ranking results reported in Situation Questionnaire were relatively equally

distributed in two gender populations, X^2 (3, N = *34*) = 8.385, p = .496; X^2 (3, N = *34*) = 4.587,

$p = .205$. In the repeated measure ANOVA, the main effect of gender was non-significant $F(1,$

$136) = 1.767$, $p = .193$, $\eta2 = .052$. Neither interaction effect was found significant, $F(1, 136) =$

$.261$, $p = .613$, $\eta2 = .008$, indicating that on the total scores in four different language conditions

were not significantly different in conditions that the subject is a male than it is a female (see

Appendix L for Gender Effect Analyses). These analyses were conducted to make sure gender

did not affect the participants' performances in each language conditions and to validate the

results of the previous regression analyses.

Observational Results and Other Feedbacks

After the practice trial ended and the task began, most participants did not ask questions

about the task during the experiment. While most participants leaned towards the computer

screen throughout the task, three participants leaned back in the chair when they heard the audio

recording in the Chinese-accented English condition. One of these three participants chuckled

while listening to the Chinese-accented English. These behaviors may be interpreted as signals of

participants in a relaxed mood listening to the Chinese-accented English audio. One participant

asked about the meaning of a word ("hotheads" in scenario School Board) during the task and

the experimenter responded in English. After the debriefing, all participants reported that the

practice trial was helpful in terms of explaining the tasks and the meaning of the questions.

Nearly a third of the participants (N = 8) reported that it was hard to keep in mind that they were

asked to rate the issues based on their importance instead of giving an answer of 'agree' or

'disagree' to them. In addition, one participant suspected whether the information provided in

each dilemma is leading to certain decisions. This participant further gave an example of the

Famine dilemma. Specifically, more background information about the Mustaq family was

provided than the rich man. The participant perceived this lack of information as guiding

listeners to choose "Should take the food" option.

Participants were also asked if they recognized any of the speeches coming from people

they knew personally. Four participants correctly recognized the native English speaker through

listening to the English-accented Chinese speech. One participant recognized the native English

speaker listening to the speech in native English. One participant was confident about who the

speaker of the Chinese-accented English audio was after completing the task, yet the guess was

confirmed to be wrong. None of the participants correctly recognized the speaker for native

Chinese and Chinese-accented English speeches.

Discussion

This study aimed to investigate the association between exposure to familiar sounds of

language and moral judgments. The analyses revealed that priming familiar and unfamiliar

sounds of language did not influence the moral decisions or the moral reasoning schema adopted.

The results in focused comparisons provided some support to the hypothesis (3) that when

participants were exposed to moral dilemmas read in native Chinese, they performed

Post-conventional moral reasoning more than when exposed to dilemmas read in

English-accented Chinese. Nonetheless, this could not be accepted as solid evidence to reject the

null hypothesis. The results neither supported the rest of my hypotheses that (1) there is an

auditory FLE on moral decision-making for Chinese-English bilinguals; (2)&(4) there is a main

effect of sounds of language on moral reasoning schema adopted to process moral dilemmas; and

(5) higher proficiency in English leads Chinese native speakers to employ a more advanced

moral reasoning schema after listening to moral scenarios spoken in native English and English

accents. Additionally, older age was found associated with a moral reasoning pattern inclining to

the Personal Interest Schema and Maintaining Norm Schema (Rest et al, 1999) when moral

dilemmas were presented in Native English.

The current study should be considered innovative in the sense of being one of the

leading studies that has examined the foreign language effect not merely on moral

decision-making but also on moral reasoning patterns. Although the results in this study have not

revealed a strong foreign language effect on the moral reasoning patterns that were developed

and renovated by James Rest (1999), they should have brought new aspects of thinking into the

discussion on MFLE: Will the FLE appear in moral reasoning models other than the one adopted

in DIT2? Does FLE have influences on specific information processing, or decision-making

strategies used to process morally salient stimuli? In other words, future research that analyzes

the relationship between foreign languages and moral judgments should not be limited to moral

decisions but should also include the moral reasoning process.

Additionally, while Brouwer (2019) has found an auditory MFLE, this study has further

provided insights into the role of sound in the auditory MFLE. This paper has scrutinized the

potential connections between sound and moral judgements with an exhaustive review on

previous studies. Further in the study design, we have adopted accented speech to explore

whether a foreign tone added on a native language has the same power as a foreign language to

modulate people's moral decisions. The influences of the vocal sounds that represent foreignness

(e.g. accents) would be a topic that needs more exploration in today's highly globalized world.

The novelty and creativity in this current study should be recognized and encouraged. Nonetheless, to understand the results critically and comprehensively, the inconsistencies with previous research, potential explanations of the results, and limitations of this study shall be also discussed in the following sections.

Inconsistencies with previous research

The results from this study were found inconsistent with the findings presented in the Brouwer (2019) study. In combination with the investigations on sound-emotion associations, it was inferred that the auditory MFLE stems from the better ability to convey emotions in familiar sounds compared to less familiar sounds. It was therefore predicted that (1) the auditory MFLE can be replicated in Chinese-English bilinguals; (2) Native Chinese speakers make less utilitarian decisions and present higher moral reasoning stages after listening to moral dilemmas in native Chinese than in native English, (3) in native Chinese than in English-accented Chinese (4) in Chinese-accented English than in native English. Since in the Brouwer (2019) study, it was also suggested that higher proficiency in L2 may eliminate the MFLE, it was further predicted that (5) there is a positive correlation between the proficiency in English and the moral reasoning schema adopted. Although some data support the third hypothesis and the general rules predicted in the second to the fourth hypothesis, the differences between different sound groups were not significant enough to back up any of the hypotheses. These inconsistencies could have arisen from the following aspects.

Emotional Content within Moral Dilemmas

In the Greene et al (2004) study, Greene and colleagues distinguished moral dilemmas into two categories: "personal" and "impersonal" dilemmas based on the emotion involved analyzing each dilemma. Personal dilemmas were characterized as a dilemma that involves (a) if a personal moral violation is likely to cause "serious bodily harm" in which (b) falls on particular individuals and (c) does not result from deflecting an existing threat on a different party(Greene et al, 2004). For instance, in the Footbridge dilemma, the choice of pushing the man off the bridge will cause serious physical harm to the man, who is vividly represented in the dilemma as the one who receives the harm. Finally, the action of pushing the man off captures "a notion of agency"(Greene et al, 2004). That is to say, the action is sprung directly from "you", the agent in the Footbridge dilemma, and from "your" will. The big man does not face an existing danger before the action of pushing is adopted. Hence, the physical harm is considered as "authored" rather than "edited" by this action. Therefore, the Footbridge matches all three characteristics and is considered as a Personal dilemma. On the other hand, any dilemma that lacks any of these three characteristics is considered as an impersonal dilemma. For example, the switch dilemma is considered as an impersonal dilemma in that it does not match the third characteristic: pulling the lever is considered as deflecting the existing threat between one person or five people.

Importantly, personal dilemmas are considered to contain more emotional content, specifically, a response of pushing one off the bridge is considered as more emotionally aversive than pulling the lever. In the subsequent study, Greene and colleagues (2008) argued that processing moral scenarios in an L2 may reduce the emotionality compared to an L1, leading to an increase in utilitarian decisions. Combining these studies on the association between emotion and sound, the MLFE can be predicted to be more prominent in personal dilemmas than

impersonal dilemmas. This pattern was further supported by data in the Cipolletti et al (2016) study: native English and native Spanish speakers did not show MFLE in their responses to the trolley dilemma, yet the MFLE was found in their responses to the footbridge dilemma.

However, in this study, the four dilemmas adopted were not classified with this personal-impersonal model beforehand. Only the Famine and Cancer dilemmas involved serious threat of bodily harm to the characters (the Mustaq Family and Mrs. Bennett) if certain actions were taken. Furthermore, it is remarkable that none of the dilemmas addressed the protagonist--the participants--as "you". The scenarios shared only perspectives from one side of the story in these highly-conflicting situations and implicitly encouraged the participants to put themselves in the shoes of the protagonists in order to make a decision. Given this fact, it is reasonable to speculate that all the dilemmas used in this study were considered less personal by listeners, as they might consider themselves as spectators and had no direct contact with characters within the scenarios. As a result, after listening to the moral dilemmas in native- or foreign-accented L1(Chinese) and L2(English), listeners perceived fewer emotions or attained less emotional resonance, leading to the absence of MFLE in this study.

Emotion Perception and L2 Proficiency

In this study, the hypothesis that with higher familiarity with L2, L1 speakers can perceive the emotional content in L2 accented speeches better was not supported by data. One potential reason is that the variables used in this study (the number of years learning English, the number of years exposed to English speaking environment, and scores of English Standard tests for ESL learners) were not the most appropriate ones to measure English proficiency. There

could be individual differences in the amount of effort one put into learning the language in a

given time span. In the real world, learning English for one year intently may achieve higher

proficiency in English than learning English for twelve years intermittently. This could be the

reason why no consistent association was detected between these variables and moral reasoning

stages.

In addition, English proficiency might not be considered equivalent to the ability to

identify the emotions expressed in sounds of English. In support of this explanation, Graham,

Arien and Stanley's (2001) study has found that Japanese and Spanish ESL learners at different

proficiency levels did not perform differently in identifying emotions expressed in oral English.

Graham and colleagues also suggested that for ESL learners to accurately judge emotions

portrayed through English accented speeches like native English speakers, they would have to

have extensive exposure in a native context or special attention to developing the skills

understanding emotions expressed in English (Graham, Arien & Stanley, 2001).

Meanwhile, a counter hypothesis to the prediction in this study is also worth mentioning:

second language proficiency might not be associated with our moral reasoning process. In the

Bhatara et al (2016) study, French native speakers with higher English ability were found

presenting less accurate identification of positive emotions in vocal expressions produced by

American actors but not found in identifications of negative emotions. That is to say that L2

skills may interfere with our recognition of some types of emotions in the L2 speech.

Particularly, positive emotions were likely to be affected by L2 skills. Yet as mentioned earlier,

the "moral emotions" that are powerful enough to drive moral judgments are usually negative

emotions such as anger and disgust (Haidt, 2003). Also, these negative emotions are more likely

to be triggered in many moral dilemmas. For instance, in the Footbridge dilemma, imagining

pushing the man off the bridge and seeing the train running over him might trigger disgust within

people, and hence cause them to make less utilitarian decisions. That means when listening to

moral dilemmas, individuals with higher L2 abilities and those with lower L2 abilities might

share the same ability to identify moral emotions in L2-sounding speech. Hence, it could be

hypothesized that second language proficiency is not associated with moral judgments, and the

results in this study would be found to comply with these rules.

Limited Role of Emotion in MFLE

Previous models developed to explain the process of making moral judgment suggested

that emotion might not be the only factor that drives our moral decision-making and moral

reasoning patterns. Vives, Aparici and Costa (2018) suggested this foreign language effect is

limited to moral decision-making in which emotions and intuitions play a role. It will not

modulate strategies used to evaluate moral dilemmas. Specifically, they found that the

outcome-bias processing strategy, comparing which outcome of the decisions is more desirable,

was not influenced in different languages. Although this finding did not provide enough

information about why the MFLE was not detected in moral decision-making in this study, this

could be the reason why the stage of moral reasoning used to process moral dilemmas was not

found modulated in native- and foreign- accents.

For instance, it is possible that contradicting the principle, that emotion is the key driven

factor of moral judgments, used to build hypotheses in this study, social and moral norms might

also interfere with moral reasonings. In Geipel et al (2015b), it was argued that the power of

social and moral norms could surpass the influence of the dilemma type (personal or impersonal) and emotions. For instance, the FLE was found in the impersonal Lost Wallet dilemma in which one was asked whether to keep the money in a wallet they found or to send it back to the owner. It was concluded that people are less likely to make a decision that violates everyday social and moral norms in a second language environment (Geipel et al, 2015b). Yet in our study, less obvious social norms were presented. It was only the Famine dilemma in which the social norm of "stealing is wrong" was relatively prominent. Hence, if this theory about social and moral norms has been affirmed to be true, the absence of MFLE in our study can be explained by the insufficient cues of social norms within the dilemma materials.

Age in Moral Judgments

It was found that older age was associated with lower total scores when moral scenarios were presented in native English. A potential explanation for this discovery is that older listeners are less sensitive to vocal cues in foreign-accented speech compared to younger listeners. In a study where younger listeners (18 - 26 years old) and older listeners (65 - 80 years old) were gathered and listened to pronounced foreign accents, older listeners were less likely to catch the sound cues and distinguish the speech from background noises than younger listeners (Gordon-Salant et al, 2013). Hence the older listeners may also catch fewer emotional cues in a second language. Yet still, this might not be sufficient to explain the finding in this study, since the age span adopted in the Gordon-Salant et al(2013) study was much larger than that in this study. Again, because age was found incapable of accurately predicting the final scores used to measure moral reasoning stages ($R^2 < .25$), we should remain cautious about giving credits to this

negative correlation found between age and total score.

Limitations

Underpowered Analysis

Due to the COVID-19 epidemic, a limited number of participants were recruited in this study. The possibility of rejecting the null hypothesis when the alternative hypotheses are true remained low in the data analysis. Meanwhile, since all participants were international students from a College in the United States, they might have more experiences identifying emotional cues in English accents within their daily interactions with American students and professors. This population involved in this study may not perfectly represent all Chinese-English bilinguals. Subsequently, the result found in this study might neither be applied to all Chinese-English bilinguals.

Control of Audio Recordings

Length of accented speeches. As mentioned in the material section, there were length differences between recordings in native accents and in foreign accents ($Mdiff = 21s$, $SD = 17.15$). Foreign accented speeches were significantly longer than native accented speeches. The longer exposure to foreig- accented audio recordings may have increased fatigue in participants, leading to a reduced number of cues detected from the speech. On the other hand, it is also possible that the longer exposure to foreign accents allowed participants to collect more emotional cues from the less familiar accents. In addition, the audio recordings in this study were not precisely measured and controlled in terms of sound qualities (e.g. bit depth and sample rate).

In future studies, the speeches should be better controlled through (1) controlling the length in each accent to be the same; (2) controlling the sound qualities (e.g. using a common bit depth of 32 and a common sampling frequency of 44,100 Hertz) to reason out the potential influences of nuances in sound.

Personal connections with speakers. Since Bard College has a relatively small campus, the possibility of listeners recognizing speakers in audio recordings is fairly high. In participant feedback, four participants recognized the native English speaker correctly, and one participant made a wrong guess of who the native Chinese speaker was. If the participant recognized the speaker as someone whom they knew personally, emotional connections to the content might be built more quickly and easily, leading the listeners to make less utilitarian decisions and adapt to a more deliberate reasoning pattern. Furthermore, participants might also feel more relaxed listening to a sound from someone who they know and miss some sound cues instead.

Improvements to the Situational Questionnaire

Reliability of the measurement. The original scoring manual for DIT2 was preserved by the University of Alabama. However, after a number of attempts to contact people in charge, I failed to receive any responses from the school. As a result, the Situation Questionnaire was developed from the original DIT2 with help of two faculty from the psychology department. The moral reasoning stage was further divided into two general stages. However, while the reliability of results in DIT2 was relatively high (Cronbach's alpha = .81) (Rest et al, 1999), it should be highlighted that neither internal nor external reliability of this new measurement was tested beforehand. Hence, there remains potential that the measurements with this questionnaire may

vary from one to another individual, for each dilemma was not measuring the same moral

reasoning stage.

Dilemmas involve political aspects. Some dilemmas used in this questionnaire may be

challenged for being closely related to individuals' political views. The Demonstration dilemma

is an exemplar of this limitation. In the Demonstration, the country name "United States" was

explicitly presented. It was hence reasonable to point out the consideration that people's choices

could be altered based on their political positions rather than emotions. Additionally, a student

protest in Hong Kong that lasted for months caught people's attention recently. During the Hong

Kong protest, it was also widely debated whether students were protesting in an acceptable way

for Chinese citizens. Given the fact that all participants in this study were from Mainland China,

they might have an established answer to the Demonstration dilemma prior to completing this

questionnaire. Moreover, the results from these dilemmas might not be used to predict the

choices of listeners from other countries: native Spanish listeners, for example, might provide a

contradicting choice after listening to the dilemma, compared to Chinese listeners.

Conclusion

The main purpose of the present research was to examine the relationship between

language sounds, moral decision-making, and moral reasoning patterns. Specifically, this study

aims to replicate the recently-found auditory MFLE in a new language pair, the Chinese -

English pair, to have a more concrete understanding of the reason why auditory MFLE exists.

This study also investigated whether foreign-accented speech can also trigger an auditory MFLE

in comparison with native accented speech. Although some evidence supported that listeners

adopt higher moral reasoning stages in their native language than in a foreign-accented native

language, the evidence was not conclusive. Instead, the results overall showed that listeners

neither presented a difference in their moral decision-making nor in moral reasoning patterns

after listening to the highly-conflicted moral scenarios in native- or foreign-accented speech.

       To deepen our understanding of the relationship between sounds, moral decision-making,

and moral reasoning patterns, future research may dive into topics such as whether the

foreign-accented speech would cause MFLE in a certain type of moral dilemmas (e.g. personal

moral dilemmas). It is also encouraged to develop creative measurements and theories that are

capable of describing moral reasoning strategies or patterns in new systems aside from Kolberg's

theories. Discoveries on the relationship between the moral decision-making and moral

reasoning stages are also expected to see. Hopefully, the present work will provide help and

inspiration to future explorations in this field of moral psychology.

References

Adank, et al. (2013). Accent Imitation Positively Affects Language Attitudes. Frontiers

in Psychology. doi: 10.3389/fpsyg.2013.00280

Airas M., Alku P. (2004) Emotions in Short Vowel Segments: Effects of the Glottal Flow as

Reflected by the Normalized Amplitude Quotient. *Lecture Notes in Computer Science*,

3068.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. Journal of

the Acoustical Society of America, 114, 1600–1610.

Bhatara, A., Laukka, P., Boll-Avetisyan, N., Granjon, L., Elfenbein, H. A., & Bänziger, T.

(2016). Second Language Ability and Emotional Prosody Perception. Plos One, 11(6).

doi: 10.1371/journal.pone.0156855

Brouwer, Susanne (2019). The auditory foreign-language effect of moral decision making in

highly proficient bilinguals, *Journal of Multilingual and Multicultural Development*.

Romero-Rivas, Carlos. (2019) Foreign accents reduce false recognition rates in

the DRM paradigm. *Journal of Cognitive Psychology.*

Caldwell-Harris, C. L., and A. Ayçiçeği-Dinn. 2009. "Emotion and Lying in a Non-native

Language." *International Journal of Psychophysiology* 71: 193–204.

Campbell-Kibler, K. (2007). Accent, (ing), and the social logic of listener perceptions.

American Speech, 82, 32–64.

Cipolletti, H., McFarlane, S., & Weissglass, D. (2016). The moral foreign-language effect.

Philosophical Psychology, 29, 23-40.

Clavien, Christine. (2009) Gibbard's expressivism: An interdisciplinary critical analysis.

Philosophical Psychology. 22, 456-485

Costa A, Foucart A, Hayakawa S, Aparici M, Apesteguia J, et al. (2014). Your Morals

Depend on Language. *PLoS ONE* 9(4), e94842.

Geipel, J, Hadjichristidis, C and Surian, L (2015a). Foreign language affects the contribution

of intentions to moral judgment. *Cognition*, 154, 34-39.

Geipel, J, Hadjichristidis, C and Surian, L (2015b). The Foreign Language Effect on Moral

Judgment: The Role of Emotions and Norms.*PLoS ONE*, Vol 10(7), e0131529.

http://dx.doi.org/10.1371/journal.pone.0131529.

Geipel, J, Hadjichristidis, C and Surian, L (2016). Foreign Language Affects the Contribution

of Intentions and Outcomes to Moral Judgment. *Cognition*, 154. pp. 34-39.

https://doi.org/10.1016/j.cognition.2016.05.010

Gibbard, A. (1990). Wise choices, apt feelings: A theory of normative judgment. Cambridge,

MA: Harvard University Press.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., Cohen, J. I., & Waldroup, C.

(2013). Recognition of accented and unaccented speech in different maskers by younger

and older listeners. The Journal of the Acoustical Society of America, 134(1), 618–627.

doi: 10.1121/1.4807817

Graham, C. R., Hamblin, A., and Feldstein, S. (2001). Recognition of emotion in English

voices by speakers of Japanese, Spanish, and English. *Int. Rev. Appl. Linguist. Lang.*

Teach. 39, 19-37. Doi: 10.1515/iral/39.1.19.

Greene, J., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases

of cognitive conflict and control in moral judgment. *Neuron*, 44, 389–400.

Greene, J., S. Morelli, K. Lowenberg, L. Nystrom, and J. Cohen. 2008. "Cognitive Load

      Selectively Interferes with Utilitarian Moral Judgment." *Cognition* 107: 1144–1154.

Grey, S., Schubel, L. C., Mcqueen, J. M., Hell, J. G. Van. (2018). Processing foreign-

      accented speech in a second language: Evidence from ERPs during sentence

      comprehension in bilinguals.

Haidt, J. 2001. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to

      Moral Judgment. *Psychological Review,* 108, 814-834.

Haidt, J. (2003). The moral emotions. In R.J Davidson, K.R. Scherer, & H.H. Goldsmith

      (Eds.), *Handbook of affective sciences*. Oxford: Oxford University Press, 852-870.

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech

      intelligibility benefit for native speakers of Mandarin: Production and perception of

      English word-final voicing contrasts. *Journal of phonetics, 36*(*4*), 664–679.

Hatzidaki, A., Baus, C., & Costa, A. (2015). The way you say it, the way I feel it: emotional

      word processing in accented speech. *Frontiers in Psychology*, 6. doi:

      10.3389/fpsyg.2015.00351

Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral

      psychology. *Trends in Cognitive Sciences*, 1–6. doi: 10.1016/j.tics.2008.09.006

Juslin, P.N., Sloboda, J.A. (2001). *Music and emotion: theory and research*. New York:

      Oxford University Press, 504.

Kent, R.D. (1997). The speech sciences. San Diego: Singular Publishing.

Keysar, Boaz, Hayakawa, Sayuri L., & An, Sun Gyu (2012). The Foreign-Language Effect:

      Thinking a Foreign Tongue Reduces Decision Biases. *Psychological Science.* 23(6),

661-668.

Kinzler, K.D., Shutts, K., & Spelke, E.S. (2012). Language-based social preferences among

children in South Africa. *Language Learning and Development*, *8*, 215–232.

Koeda, M., Belin, P., Hama, T., Masuda, T., Matsuura, M., & Okubo, Y. (2013). Cross-

Cultural Differences in the Processing of Non-Verbal Affective Vocalizations by

Japanese and Canadian Listeners. *Frontiers in Psychology*, 4. doi:

10.3389/fpsyg.2013.00105

Kohlberg, L. (1984). Essays on moral development: The psychology of moral development.

San Francisco, CA: Harper & Row.

Kohlberg, L. (1969). Stage and sequence: The cognitive development approach to

socialization. In D. A. Goslin (Ed.). *Handbook of socialization theory,* 347-480.

Kraxenberger, M., Menninghaus, W., Roth, A., & Scharinger, M. (2018). Prosody-Based

Sound-Emotion Associations in Poetry. Frontiers in Psychology, 9. doi:

10.3389/fpsyg.2018.01284

Moll, J., de Oliveira-Souza, R., Eslinger, P. E., Bramati, I. E., Mourão-Miranda , J.,

Andreiuolo, P. A., & Pessoa, L. (2002). The neural correlates of moral sensitivity: a

functional magnetic resonance imaging investigation of basic and moral emotions. *J.

Neurosci*. 22, 2730–2736

Moll, J., & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian

brain. *Trends in cognitive sciences*, 11(8), 319-321.

NPR. (2012, September 7). Sound A Major Emotional Driver For Humans. All Things

Considered.

Paulmann, S., & Uskul, A. K. , Cross-cultural emotional prosody recognition: evidence from

      Chinese and British listeners. *Cogn Emot,* 28(2), 230-244.

Rest, J., Cooper, D., Coder, R., Masanz, J., & Anderson, D. (1974). Judging the important

      issues in moral dilemmas—an objective test of development *Developmental*

      Psychology, 10, 491-501.

Rest, et al (1999) DIT2: Devising and Testing a Revised Instrument of Moral Judgment.

      Journal of Educational Psychology, Vol 91, No.4, 644-659. ISSN 0022-0663

Strange, W. (1995). *Speech perception and linguistic experience: Issues in cross-language*

      research: *York Press.*

Strohminger, N. & Nichols. (2013). The essential moral self. *Cognition,* 131, 159-171.

Tangney, J., Stuewig, J., & Mashek, D. (2007). Moral emotions and moral behavior. *Annual*

      Review of Psychology, 58, 345-72

Vives, M.-L., Aparici, M., & Costa, A. (2018). The limits of the foreign language effect on

      decision-making: The case of the outcome bias and the representativeness heuristic. Plos

      One, 13(9). doi: 10.1371/journal.pone.0203528

Wang, Ze. Arndt, Aaron D., Singh, Surendra N., Biernat, M., Kansas, L., & Liu, Fan. (2013).

      'You lost me at hello': How and when accent-based biases are expressed and

      suppressed. *International Journal of Research in Marketing.* 30, 185 - 196.

Weatherhead, D., Friedman, O. & White, Katherine S. (2019). Preschoolers are sensitive to

      accent distance. *Journal of Child Language.* 46, 1058-1072

Appendices

Appendix A: IRB Approval Letter

Appendix B: Human Subjects Protection Training Certificate

Appendix C: Study Invitation Email

Appendix D: Informed Consent Agreement

Appendix E: Situation Questionnaire

Appendix F: Manipulation Check Questionnaire (Clarity Questionnaire)

Appendix G: Demographic Questionnaire

Appendix H: Experiment Script

Appendix I: Debriefing Form

Appendix J: Payment Receipt

Appendix K: Correlational Analysis Plot

Appendix L: Gender effect analyses

Appendix M: Frequency Table: Moral Decisions x Language Conditions

Appendix A:
IRB Approval Letter

Bard College   Institutional Review Board

Date: October 29, 2019
To: Yifan Gu
Cc: Betsy Hoffman, Stuart Levine
From: Laura Kunreuther, IRB Chair
Re: 'The auditory foreign-language accent effect of moral decision-making'

DECISION: APPROVED

Dear Yifan,

The Bard Institutional Review Board reviewed your proposal, and require no further changes. We only ask that you clarify in your debriefing statement that your study aims to measure the effect of accents on moral decision making, a fact that you did not disclose before they participated so as to not affect their responses. Your proposal is approved through October 29, 2020. Your case number is 2019OCT29-GU.

Please notify the IRB if your methodology changes or unexpected events arise.

We wish you the best of luck with your research.

Laura Kunreuther
kunreuth@bard.edu
IRB Chair

Appendix B:
Human Subjects Protection Training Certificate

Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that
Yifan Gu successfully completed the NIH Web-based training course
"Protecting Human Research Participants".

Date of completion: 03/05/2018.

Certification Number: 2671662.

Appendix C:
Recruitment Email

Hello all,

I am Yifan Gu, a senior psychology student at Bard College. I am currently conducting a psychology study for my senior project.

The study should concern students who are interested in social psychology and linguistics. In the study, I would like to ask for your help to evaluate the level of clarity for several English and Chinese speakers. I am looking for participants who are native Mandarin speakers and have English as their second language. They are expected to have an advanced English listening and reading skill to fully understand the materials. Each participant will get 5 dollars for participating.

If you are interested and would like to have the chance to find out more about the study before coming to any decision, please do not hesitate to contact me at yg1617@bard.edu or 845-853-5088. You can also contact the project advisor Stuart Levine at levine@bard.edu for more information. You would be under no obligation to take part.

The use of email to recruit participants for this study has been approved by the Institutional Review Board at Bard College.

Sincerely,

Yifan Gu

Appendix D:
Informed Consent Agreement

Informed Consent Agreement

Introduction. You are invited to participate in a research study conducted by member of the Psychology Department at Bard College. The purpose of this this study is to examine the importance of clarity in speeches in terms of communication efficiency. Please read the following information carefully.

What you will do in the study. You will listen to recordings of four difficult life situations by different people and rate the clarity of the speech in recordings. You will also rate and rank 8 issues based on how you perceive the issues' importance to show your understanding of the situation. The study will take about 20-30 minutes to complete.

Risks and Benefits. The tasks included in this study are unlikely to elicit feelings of discomfort. Each participant will receive a $5 payment for participating. You may also receive indirect benefits from learning about the research process and the previous research that motivates the present work.

Your rights as a participant. Your participation in this experiment is completely voluntary. You may withdraw from the experiment at any time without penalty, just inform the experimenter. The experimenter will tell you more about the study at the end of the session.

Confidentiality. The records of this study will be stored securely. There will be no way to connect you to your answers. I will not include any information that will make it possible to connect or identify you in any sort of report I might publish.

Contact. If you have any other questions or concerns about this research, please contact Yifan Gu at yg1617@bard.edu or the project advisor Stuart Levine at levine@bard.edu. If you have questions about your rights as a research participant, please contact the Bard College Institutional Review Board at irb@bard.edu. If you experience any emotional distress during the study, you are encouraged to contact the National Alliance on Mental Illness's (NAMI's) HelpLine at 1-800-950-6264.

STATEMENT OF CONSENT:
"The nature and purpose of this research have been sufficiently explained to me. I have been given an opportunity to ask questions, and my questions have been answered to my satisfaction. I have been told whom to contact if I have additional questions. I have read this consent form and agree to be in this study, with the understanding that I may withdraw at any time."

By signing the line below, I am indicating that I am in agreement with the above statement of consent and further certify that I am at least 18 years of age.

_____   _____
Participant signature   Date

_____
Participant name (printed)

_____
Experimenter signature

Appendix E:
Situation Questionnaire

Famine
The small village in northern India has experienced shortages of food before, but this year's famine is worse than ever. Some families are even trying to feed themselves by making soup from tree bark. Mustaq Singh's family is near starvation. He has heard that a rich man in his village has supplies of food stored away and is hoarding food while its price goes higher so that he can sell the food later at a huge profit. Mustaq is desperate and thinks about stealing some food from the rich man's warehouse. The small amount of food that he needs for his family probably wouldn't even be missed.

饥荒
这座位于印度北部的村庄以前就常常遭遇食物紧缺。但今年的饥荒比往年都要严重。一些家庭甚至开始用树皮煮汤充饥。穆斯塔克一家人已经濒临饿死的边缘。他听说村庄里有个富人储存了大量的粮食，准备囤到价格高涨时出售，以获得巨大的利润。 穆斯塔克不顾一切，打算从富人的粮仓里偷取一些粮食。他的一家人所需要的这少量的粮食可能根本不会引起富人的注意。

What should Mustaq Singh do? Do you favor the action of taking food? *

○ Should take the food

○ Cannot decide

○ Should not take the food

Rate the following issues in terms of importance. *

	No	Little	Some	Much	Great
Shouldn't the community's laws be upheld?	○	○	○	○	○
Does the rich man have any legal right to store food when other people are starving?	○	○	○	○	○
Does the rich man deserve to be robbed for being so greedy?	○	○	○	○	○
Is Mustaq Singh courageous enough to risk getting caught for stealing?	○	○	○	○	○
Isn't it only natural for a loving father to care so much for his family that he would steal?	○	○	○	○	○
What values are going to be the basis for social cooperation?	○	○	○	○	○
Would stealing bring about more total good for everybody concerned or wouldn't it?	○	○	○	○	○
Are laws getting in the way of the most basic claim of any member of a society?	○	○	○	○	○

Consider the 8 issues above and rank the most important issue. *
-- Please Select --

Please rank the second most important issue. *
-- Please Select --

Cancer

Mrs. Bennett is 62 years old, and in the last phases of colon cancer. She is in terrible pain and asks the doctor to give her more pain-killer medicine. The doctor has given her the maximum safe dose already and is reluctant to increase the dosage because it would probably hasten her death. In a clear and rational mental state, Mrs. Bennett says that she realizes this; but she wants to end her suffering even if it means ending her life. Should the doctor giver her an increased dosage?

癌症

62岁的本内特女士如今处在结肠癌晚期。她十分痛苦，希望医生给她开更多止疼药。医生已经给她开了安全用量范围内的最大剂量，并不愿意再加大剂量，因为过多的止疼药将会加速她的死亡。Bennett 女士神志清醒时对此表示理解。但她还是希望能够摆脱痛苦，虽然这样做可能意味着死亡。医生应该给她增加剂量吗？

Do you favor the action of giving more medicine? *

○ Should give her an increased dosage

○ Cannot decide

○ Should not give her an increased dosage

Rate the following issues in terms of importance. *

	No	Little	Some	Much	Great
If Mrs. Bennett dies, would the doctor be legally responsible for malpractice?	○	○	○	○	○
Isn't the doctor obligated by the same laws as everybody else if giving an overdose would be the same as killing her?	○	○	○	○	○
Wouldn't the society be better off without so many laws about what doctors can and cannot do?	○	○	○	○	○
Would the doctor show more sympathy for Mrs Bennett by giving the medicine or not?	○	○	○	○	○
Is helping to end another's life ever a responsible act of cooperation?	○	○	○	○	○
Shouldn't society protect everyone against being killed?	○	○	○	○	○
Where should society draw the line between protecting life and allowing someone to die if the person wants to?	○	○	○	○	○
Does the state have the right to force continued existence of those who don't want to live?	○	○	○	○	○

Consider the 8 issues above and rank the most important issue. *

-- Please Select --

Please rank the second most important issue. *

-- Please Select --

School Board

Mr. Grant has been elected to the School Board District 190 and was chosen to be Chairman. The district is bitterly divided over the closing of one of the high schools. One of the high schools has to be closed for financial reasons, but there is no agreement over which school to close. During his election to the School Board, Mr. Grant had proposed a series of "Open Meetings" in which members of the community could voice their opinions. He hoped that dialogue would make the community realize the necessity of closing one high school. Also he hoped that through open discussions, the difficulty of the decision would be appreciated, and that the community would ultimately support the school board decision. The first Open Meeting was a disaster. Passionate speeches dominated the microphones and threatened violence. The meeting barely closed without fist-fights. Later in the week, school board members received threatening phone calls. Mr. Grant wonders if he ought to call off the next Open Meeting.

学区委员会

格兰特先生被选为了190学区委员会的会长。该区在是否在关闭一所高校的问题上产生了分歧。因为财务原因，一所高校将被迫关闭，但人们还未在关闭哪一所高校的问题上达成一致。 在Grant 先生参选时，他曾保证举办"公开会议"，让社区成员发表意见。他希望通过会议上的对话可以使大家明白关闭高校是必要的。他也希望通过会议，大家可以理解委员会作出决定的不易并支持他们的决定。第一次公开会议的画面十分惨烈。人们情绪激动，并以暴力相威胁。会议几乎是在拳脚相向中结束。之后几天，委员会成员收到了威胁电话。格兰特先生不知道他是否应该取消下次的公开会议。

Do you favor calling off the next Open Meeting? *

○ Should call off the next open meeting

○ Cannot decide

○ Should have the next open meeting

Rate the following issues in terms of importance. *

	No	Little	Some	Much	Great
Is Mr. Grant required by law to have Open Meetings on major school board decisions?	○	○	○	○	○
Would the community be even angrier with Mr. Grant if he stopped the Open Meetings?	○	○	○	○	○
Does Mr. Grant have the authority to expel troublemakers from the meetings or prevent them from making long speeches?	○	○	○	○	○
If the school board is threatened, does the chairman have the legal authority to protect the Board by making decisions in closed meetings?	○	○	○	○	○
Does Mr. Grant have another procedure in mind for ensuring that divergent views are heard?	○	○	○	○	○
What effect would stopping the discussion have on the community's ability to handle controversial issues in the future?	○	○	○	○	○
Is the trouble coming from only a few hotheads, and is the community in general really fair-minded and democratic?	○	○	○	○	○
What is the likelihood that a good decision could be made without open discussion from the community?	○	○	○	○	○

Consider the 8 issues above and rank the most important issue. *

-- Please Select --

Please rank the second most important issue. *

-- Please Select --

Demonstration

Political and economic instability in a South American country prompted the President of the United States to send troops to "police" the area. Students at many campuses in the U.S.A. have protested that the United States is using its military might for economic advantage. There is widespread suspicion that big oil multinational companies are pressuring the President to safeguard a cheap oil supply even if it means loss of life. Students at one campus took to the streets in demonstrations, tying up traffic and stopping regular business in the town. The president of the university demanded that the students stop their illegal demonstrations. Students then took over the college's administration building, completely paralyzing the college. Are the students right to demonstrate in these ways?

示威

南美某国不稳定的政治经济形势使了美国总统派兵去"维护治安"。许多在美国的学生发出抗议，认为美国政府以军事力量来确保经济优势。人们之间广泛流传着一个猜测：大型跨国石油公司在向总统施压来保证便宜的石油供应渠道，即使代价是人命。一所校园的学生为了游行示威强占街道，造成堵车，并且使得镇上商家无法正常营业。学校校长要求学生们停止非法示威。学生们接着抢占了学校行政大楼，造成学校系统的完全瘫痪。学生们以这样的方式示威是对的吗？

Do you favor the action of demonstrating in this way? *

○ Should continue demonstrating in this way

○ Cannot decide

○ Should not continue demonstrating in this way

Rate the following issues in terms of importance. *

	No	Little	Some	Much	Great
Do the students realize that they might be arrested and fined, and even expelled from school?	○	○	○	○	○
Shouldn't the authorities be respected by students?	○	○	○	○	○
Isn't it everyone's duty to obey the law, whether one likes it or not?	○	○	○	○	○
Are the authorities to blame by giving in to the greed of the multinational oil companies?	○	○	○	○	○
Why should a few people like Presidents and business leaders have more power than ordinary people?	○	○	○	○	○
Does this student demonstration bring about more or less good in the long run to all people?	○	○	○	○	○
Is taking over a building consistent with principles of justice?	○	○	○	○	○
Will the public blame all students for the actions of a few student demonstrators?	○	○	○	○	○

Consider the 8 issues above and rank the most important issue. *
-- Please Select --

Please rank the second most important issue. *
-- Please Select --

The first four issues fall into Personal Interest and Maintaining Norms Schema, and the last four issues fall into the Post-conventional Schema.

Appendix F:
Manipulation Check Questionnaire (Clarity Questionnaire)

Clarity Questionnaire

Please rate the following statements on a 5-point scale.

(1= Strongly disagree, 2 = Disagree, 3= Neutral, 4 = Agree, 5= Strongly Agree) *

	1	2	3	4	5
I fully understand the situation described in the speech.	○	○	○	○	○
The speaker demonstrated good speed control, speaking in a medium pace.	○	○	○	○	○
The speaker demonstrated good volume control, speaking with moderate volume.	○	○	○	○	○
In this speech, sounds, syllables, words and phrases were well jointed together.	○	○	○	○	○
The speaker stopped and paused a lot.	○	○	○	○	○
The speaker managed to accurately represent what happens in the situation.	○	○	○	○	○
The speaker presented a non-native accent speaking this language.	○	○	○	○	○

Appendix G:
Demographic Questionnaire

Demographic Questionnaire

Please provide the following information.

7. Age:

[   ]

8. Gender

○ Female

○ Male

○ Others

○ Prefer not to answer

9. Number of Years Learning English

[   ]

10. Number of Years Exposed to English Speaking Environment (e.g.Study in institutions' where English is the primary language of instruction is English; Live in an English-Speaking country; etc.)

[   ]

11. Most Recent TOEFL / IELTS Score (If applicable)

[   ]

12. Should we include your responses in our data?
Please select no if you were not paying attention to the tasks. *

○ Yes, I was paying attention to the audios and questions.

○ No, please exclude my data.

Appendix H:
Experiment Script

Lab Visit Experimenter Script

Invitation:

• EMAIL ON-CAMPUS CHINESE STUDENTS THE EXPERIMENT INVITATION.

• EMAIL INTERESTED PARTICIPANT WHO REPLIED:

    1. A SUMMARY OF THE STUDY AND ELIGIBILITY.

    2. A TIMESHEET TO SIGNUP FOR EXPERIMENT.

• APPROACH POTENTIAL PARTICIPANTS AT COLLEGE LIBRARY.

> *Hi! Sorry to bother you. I'm a senior in psychology and I'm running a study for my SPROJ on the third floor. To complete the study, I would like to ask for your help to evaluate the level of clarity for several English and Chinese speakers. You will get 5 dollars for participating. Would you like to participate?*

_____

Pre-session Reminders:

• ONE DAY BEFORE THE SCHEDULED SESSION, EMAIL PARTICIPANT A REMINDER (TIME/PLACE) AND A ELECTRONIC VERSION CONSENT FORM.

_____

Setup:

• TURN ON THE LIGHT AND COMPUTER IN THE SOUND-PROOF ROOM.

• OPEN THE CLARITY SURVEY AND FULL-SIZE THE WEBPAGE. CHECK THE SOUND DEVICE, SET THE GENERAL VOLUME TO 50.

• PLACE A PEN/PENCIL ON THE DESK.

_____

Introduction:

• INVITE THE PARTICIPANT TO VISIT THE SOUND-PROOF ROOM IN LIBRARY.

> *Thank you again for helping me with my SPROJ research! You can take a seat in front of the computer now.*

[Hand the consent form and a pen.]

Before we start, I'd like you to read over the consent form, which explains the study, its risks and benefits, your rights as a participant, and contact information of the researcher. Your consent indicates that you're participating in the study voluntarily. You may withdraw from the experiment at any time and still receive the payment for participating. Your data will be completely anonymous, which means I'll never be able to match your name to your responses.

If you wish to withdraw, please let me know and no questions will be asked.
Please let me know if you have any questions regarding the consent form.

[Participants sign the consent form.]

Thank you for participating! You may choose the option"Yes, I read through and signed the consent form" and then click on "Next" button. I will explain the tasks and you will first go through a practice trial.

This study aims to investigate the importance of clarity in speeches in terms of communication efficiency. In the following task, you will listen to 4 voice recordings of difficult life situations in either English or Chinese. After listening to each recording, you will answer two list of questions based on recording to show your understanding of the content.

In the first questionnaire, the Clarity Questionnaire, you will rate 7 statements about the clarity of the speech you listened to on a 5-point scale.
Then, you will complete the Situation Questionnaire. It is a further assessment and clarification of your understanding of the situations depicted in recordings. In this questionnaire, you will first make a decision on how should the protagonist act. You will

then rate 8 items in terms of their importance regarding to decision making process in the given situation. Lastly, you will rank the two items you found as the most important ones.

Please notice that there is no right or wrong answer to all the questionnaires.

Do you have any questions?

Good. Now, let's go through the task with a Practice Trial: Reporter. You may click on "Next" now.

[Participants click on "Next"]

Please click on the "Play/Arrow" button to play the audio.
Please listen carefully. *You will only be able to listen to the audio once.*

Now, Please complete the Clarity Questionnaire.
Now, Please complete the Situation Questionnaire.

[Go through practice trial]

Great! You completed the Practice Trial.
The following task will be in the same format. Do you have any questions?

Good.I will be sitting in the corner and working on other assignments. Feel free to let me know if you run into any problems. You may click on 'Next' when you feel ready to start the task.

[Experimenter sits in the corner]

<u>Debriefing</u>

Thank you so much for participating! How did everything go?

I would like to share you a little more about my study. I am interested in how people make systematically different decisions when faced with a moral dilemma in different language settings. Some studies show that people make different moral decisions when the moral dilemmas are presented in different languages, both in written forms and auditory forms. Instead of evaluating the speakers' speaking skill, the goal of this study is rather to investigating the effect of the sound of the language on the moral decision-making process. Hence, I used American accented Chinese and Chinese accented American to test. The second form, the "Situation Questionnaire", is used to assess your decision making process instead of merely understanding of the situation.

Here, I have a written debriefing form with more information about the study, as well as contact information in case there are any further questions. Would you like a copy of this form?

[If yes, hand the Debriefing Form.]

*Another important thing is, **I'd like to ask you to not share any information about the study with others as it might influence other people's performance on the task**. Is that okay?*

Thank you again and here's your payment for participating in the study. Please sign this receipt to confirm that you have received the payment.

[Hand the compensation and payment receipt to the participant.]

Appendix I:
Debriefing Form

Debriefing Form

Thank you for participating!

The primary goal of this study is to investigate whether the sound of language can influence one's moral decision-making process. Previous studies state that linguistic information can be more emotional in the native language when heard than when read to individuals. We reasoned that the sound of a native language (a native language accent) may be able to convey stronger emotion to native speakers and ultimately sway their moral judgments. Therefore in the tasks, we presented moral dilemmas with a range of accents and asked our participants to respond to the Situation Questionnaires to assess their moral judgments.

Again, there is no right or wrong answer to these questionnaires. Your data will be completely anonymous, which means we'll never be able to match your name to your responses. You may contact the experimenters if you are interested in the study result.

Please do not share any information about the study with others as it might influence other people's performance on the task.

Contact. If you have any other questions or concerns about this research, please contact Yifan Gu at yg1617@bard.edu or the project advisor Stuart Levine at levine@bard.edu. If you have questions about your rights as a research participant, please contact the Bard College Institutional Review Board at irb@bard.edu. If you experience any emotional distress during the study, you are encouraged to contact the National Alliance on Mental Illness's (NAMI's) HelpLine at 1-800-950-6264.

Thank you again for participating in this study!
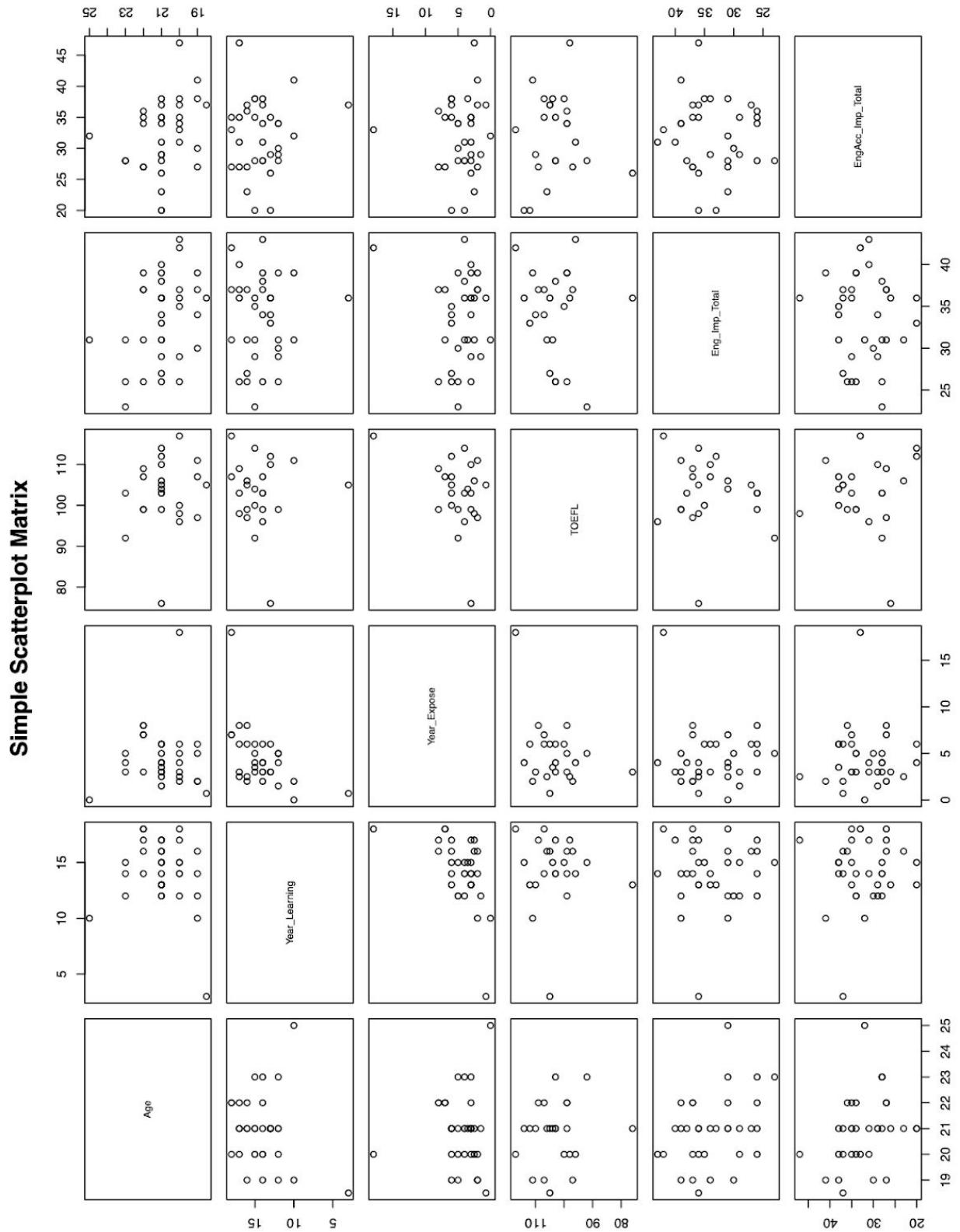
Appendix J:
Payment Receipt

Cash Receipt

I acknowledge the receipt of $5.00 for participation in the study
conducted by Yifan Gu from Bard College.

Name: _____

Signature: _____

Date: _____

Appendix K:
Correlational Analysis Plot

Appendix L:
Gender Effect Analyses

Table

Chi-square test and ANOVA results for subjects grouped by gender

	Gender		X^2 or F	p
	Male	Female		
Characteristics	N (%)	N (%)		
Decision				
Cannot Decided	10 (76.9)	9 (42.9)	8.385	.496
Utilitarian	2 (15.4)	6 (28.6)		
Non-utilitarian	1 (7.7)	6 (28.6)		
Ranking Results				
IM-IM	1 (7.7)	1 (4.8)	4.587	.205
IM-P	2 (15.4)	4 (19.0)		
P-IM	5 (38.5)	2 (9.5)		
P-P	5 (38.5)	14 (66.7)		
Score in Language Condition				
L1, Mean (SD)	32.62 (3.43)	35.14 (5.12)	1.767	.193
L2, Mean (SD)	33.15 (4.86)	33.80 (5.35)		
L1-L2, Mean (SD)	32.76 (4.49)	33.10 (4.91)		
L2-L1, Mean (SD)	30.38 (5.47)	33.00 (5.83)		

Appendix M:
Frequency Table: Moral Decisions x Language Conditions

Table

Frequency of moral decisions (Utilitarian or Non-utilitarian) by Language (Native Chinese or Native English) in each moral dilemma.

	Moral Decision		
Moral dilemma	Utilitarian	Non-utilitarian	*p**
Famine			
Native Chinese	2	2	.429
Native English	4	0	
School Board			
Native Chinese	1	5	.99
Native English	0	4	
Cancer			
Native Chinese	2	1	.40
Native English	0	3	
Demonstration			
Native Chinese	2	0	.467
Native English	2	2	
Total			
Native Chinese	7	8	. 713
Native English	6	9	

*p-values conducted to examine the association between moral decision and language for each moral dilemma were from results of Fisher's Exact test. The *p*-value conducted to examine for the same association for total results from all four moral dilemmas was from results of Chi-square analysis.