



A Laboratory Study of the McGurk Effect in 324 Monozygotic and Dizygotic Twins

Guo Feng^{1,2,3}, Bin Zhou^{1,2}, Wen Zhou^{1,2*}, Michael S. Beauchamp⁴ and John F. Magnotti^{4*}

¹ CAS Key Laboratory of Behavioral Science, Institute of Psychology, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Beijing, China, ² Department of Psychology, University of Chinese Academy of Sciences, Beijing, China, ³ Psychological Research and Counseling Center, Southwest Jiaotong University, Chengdu, China, ⁴ Department of Neurosurgery and Care for Advanced MRI, Baylor College of Medicine, Houston, TX, United States

OPEN ACCESS

Edited by:

Robert J. Zatorre,
McGill University, Canada

Reviewed by:

Antoine Shahin,
University of California, Merced,
United States
Julian Keil,
University of Kiel, Germany

*Correspondence:

Wen Zhou
zhouw@psych.ac.cn
John F. Magnotti
magnotti@bcm.edu

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 29 May 2019

Accepted: 10 September 2019

Published: 04 October 2019

Citation:

Feng G, Zhou B, Zhou W, Beauchamp MS and Magnotti JF (2019) A Laboratory Study of the McGurk Effect in 324 Monozygotic and Dizygotic Twins. *Front. Neurosci.* 13:1029.
doi: 10.3389/fnins.2019.01029

Multisensory integration of information from the talker's voice and the talker's mouth facilitates human speech perception. A popular assay of audiovisual integration is the McGurk effect, an illusion in which incongruent visual speech information categorically changes the percept of auditory speech. There is substantial interindividual variability in susceptibility to the McGurk effect. To better understand possible sources of this variability, we examined the McGurk effect in 324 native Mandarin speakers, consisting of 73 monozygotic (MZ) and 89 dizygotic (DZ) twin pairs. When tested with 9 different McGurk stimuli, some participants never perceived the illusion and others always perceived it. Within participants, perception was similar across time ($r = 0.55$ at a 2-year retest in 150 participants) suggesting that McGurk susceptibility reflects a stable trait rather than short-term perceptual fluctuations. To examine the effects of shared genetics and prenatal environment, we compared McGurk susceptibility between MZ and DZ twins. Both twin types had significantly greater correlation than unrelated pairs ($r = 0.28$ for MZ twins and $r = 0.21$ for DZ twins) suggesting that the genes and environmental factors shared by twins contribute to individual differences in multisensory speech perception. Conversely, the existence of substantial differences within twin pairs (even MZ co-twins) and the overall low percentage of explained variance (5.5%) argues against a deterministic view of individual differences in multisensory integration.

Keywords: audiovisual fusion, multisensory integration, McGurk effect, twin studies, behavioral genetics

INTRODUCTION

Humans have the remarkable ability to combine information from the mouth of a conversation partner with information from their voice in order to facilitate communication (Campbell, 2008). The interaction of visual and auditory speech can be studied with stimuli containing mismatched visual and auditory components. When presented with certain combinations of incongruent speech, such as auditory *ba* paired with visual *ga*, individuals may perceive an illusory fused percept (*da*). This illusion, first described by McGurk and MacDonald (1976), has come to be known as the McGurk effect. The effect has been widely used to examine audiovisual speech processing across the lifespan and in clinical populations (Ross et al., 1998; de Gelder et al., 2003;

Burnham and Dodd, 2004; Rogers and Ozonoff, 2005; Smith and Bennetto, 2007; Tremblay et al., 2007; Pearl et al., 2009; Nath et al., 2011; Guiraud et al., 2012).

One surprising fact about the McGurk effect is that even though humans are experts in audiovisual speech perception, the illusion is not universal: even among healthy young adults, some participants never perceive the illusion, while others always do (Gentilucci and Cattaneo, 2005; Nath and Beauchamp, 2012; Stevenson et al., 2012; Strand et al., 2014; Basu Mallick et al., 2015; Magnotti and Beauchamp, 2015; Shahin, 2019). The origin of this individual variability is unknown. One obvious question is whether inherited differences in perceptual abilities influence McGurk perception, a question that can be explored with twin studies. Monozygotic twins share identical genetic makeup as well as their prenatal environment (since most MZ twins share the same placenta) while dizygotic twins share only 50% of their genetic material and never share a placenta. If MZ twins are more similar than DZ twins for a particular trait, it suggests that genetics and prenatal environment contribute to the development of the trait. The literature is mixed on the heritability of different cognitive and perceptual abilities. High heritability has been reported for some basic perceptual phenomena including binocular rivalry rate (Miller et al., 2010) and musical pitch recognition (Drayna et al., 2001); weaker heritability has been reported for olfactory sensitivity (Hubert et al., 1980), face perception (Zhu et al., 2010), and language skills such as vocabulary, syntax, and semantics (Hayiou-Thomas et al., 2012).

In order to determine the influence of genes and early environment on the McGurk effect, we examined perception of the McGurk effect in 324 healthy twins (mean age \pm SD = 16.8 \pm 2.1 years). 73 pairs of twins were monozygotic (MZ) and 89 pairs were same-gender dizygotic (DZ). To assess the stability of the McGurk effect, some of the participants ($n = 150$) were retested 2 years after the original test.

MATERIALS AND METHODS

Participants

All subjects gave written informed consent to participate in an experimental protocol which was carried out in accordance with relevant guidelines and regulations and was approved by the Institutional Review Board at Institute of Psychology, Chinese Academy of Sciences. Parental informed consent was obtained for subjects under 18 years of age. The subject pool consisted of twins enrolled in the BeTwiSt twin study (Chen et al., 2013). We tested all twins who were available for testing and who met our inclusion criteria, i.e., native Mandarin speakers with normal or corrected-to-normal vision and no history of speech, language, or hearing difficulties, resulting in a sample size of 162 pairs of twins (49% male; mean age \pm SD = 16.8 \pm 2.1 years) comparable to that of other twin studies (Zhu et al., 2010). Seventy three pairs were monozygotic and 89 pairs were same-gender dizygotic as determined by DNA analysis of short tandem-repeat polymorphisms crosschecked with phenotypic information (Chen et al., 2010, 2013). One

hundred and fifty participants underwent an additional testing session approximately 2 years after the initial session.

Stimuli and Procedure

Stimulus presentation and response collection were controlled by a PC running E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, United States). The McGurk stimuli consisted of nine digital audiovisual recordings, each \sim 2 s long (Basu Mallick et al., 2015; Magnotti and Beauchamp, 2015; Magnotti et al., 2015). Each stimulus was presented eight times in random order (Six subjects received 10 presentations of each stimulus; we analyze just the first 8 presentations for consistency with the rest of the subjects). Each stimulus contained an auditory recording of a syllable and a digital video of the face of the speaker enunciating a different, incongruent syllable (Table 1). There were five male speakers and three female speakers (the same female speaker appeared in two stimuli). Subjects were instructed to pay attention to each movie clip and report their percept by typing it via a standard keyboard into a computer. An open-choice design was used to minimize demand characteristics; no feedback was given. The auditory volume (\sim 75 dB) and the visual angle (\sim 5.2 \times 6.0°) of the stimulus were fixed across subjects. Co-twins were tested separately, but always on the same day and in the same testing room, so that variability in testing environment, if any, was comparable for MZ pairs and DZ pairs. In order to minimize testing time, there was no auditory-only or congruent audiovisual condition: an earlier study (Basu Mallick et al., 2015) using the same McGurk stimuli showed near-perfect accuracy for identifying auditory-only and congruent audiovisual syllables (accuracy of 98 and 99%, respectively).

Analyses

Subject responses were classified into four categories: auditory (responses corresponding to the auditory syllables), visual (responses corresponding to the visual syllables), McGurk (specific responses containing an element not contained in either the auditory or visual syllable, described in the original paper as “fused” responses, Table 1), and other (responses different from the previous three categories, e.g., “a”). For stimuli with double syllables (S3, S4, S5 in Table 1) each syllable in the response was coded independently (e.g., the response “dada” was coded as 1.0 McGurk; the response “daba” was coded as 0.5 McGurk and 0.5 auditory). The mean proportion of McGurk percepts across all presentations of all movie clips indexed one’s susceptibility to the McGurk effect.

TABLE 1 | McGurk stimuli.

Stimulus#	S1	S2	S3	S4	S5	S6	S7	S8	S9
Auditory component	Ba	Ba	Baba	Baba	Baba	Ba	Ba	Pa	Pa
Visual component	Ga	Ga	Gaga	Gaga	Gaga	Ga	Ga	Ka	Ka
McGurk percept	Da	Da	Dada	Dada	Dada	Da	Da	Ta	Ta

The stimuli consisted of 9 movie clips showing 8 talkers (5 males, 3 females; 1 female appeared in two movie clips) speaking syllables in which the voice (auditory component) and mouth movements (visual component) were incongruent.

To ensure our stimuli elicited a reliable McGurk effect in native Mandarin speakers, we compared a subset of the participants in the current study (the older twin of each twin pair) with a sample of native English speakers from the United States (Magnotti et al., 2015). Using the same stimuli and a similar open-choice procedure, we found no statistical difference between mean McGurk susceptibility for the Mandarin speakers (0.48) and the English speakers (0.44). Individuals from both groups spanned the entire range of McGurk perception (from 0 to 1) demonstrating a similar degree of interindividual variability across cultures.

To assess the consistency of individual differences, we tested both within-test reliability (odd-even split-half reliability) and test-retest reliability (Pearson's correlation) for the proportions of McGurk responses. For assessing the statistical significance of individual Pearson correlations, we report the associated 95% confidence interval and the results of the *t*-test assessing the null hypothesis of no linear relationship between the two variables being correlated.

To assess the similarity of McGurk illusion perception between cotwins, we separately calculated intraclass correlations for MZ and DZ pairs using the package "irr" in R and report the corresponding analysis of variance (ANOVA). To compare these correlations, we used Fisher's r-to-z transform to test the null hypothesis that the correlations are equal.

As a secondary measure of similarity, we determined how accurately we could predict each subject's McGurk illusion perception (root mean squared error across stimuli) using the co-twin's illusion perception. As a comparison, we also used the mean value across all subjects (except the subject being predicted) to predict each subject's illusion perception. To compare error rates between these methods, we used a non-parametric Wilcoxon signed-rank test.

We conducted additional control analyses to assess the robustness of our presented results. First, we used bootstrapped simulations to calculate the expected intraclass correlation between unrelated pairs of individuals for both illusion percepts and auditory percepts (visual and "other" responses were of lower frequency and were not examined). Next we used a linear mixed-effect model to assess the relationship between illusion perception and the collected demographic variables (fixed factors of age, gender, birth order, and all their interactions), with random effects of subject and stimulus. Finally, we looked for mean differences in the percentages of response types (auditory, illusion, visual, or other) between MZ and DZ twins.

RESULTS

Different McGurk stimuli vary in efficacy, so we presented nine different McGurk stimuli in order to obtain a more generalizable estimate of each participant's susceptibility to the illusion. Seven of the stimuli consisted of pairings of auditory "ba" and visual "ga" and two of the stimuli consisted of pairings of auditory "pa" and visual "ka" (**Table 1**). Participants used an open-choice response with no feedback to reduce demand characteristics. Responses were classified into four categories: auditory, visual,

McGurk illusory fusion, and other (see **Supplementary Data Sheet S1**). Consistent with previous studies, there was substantial variability across different stimuli within the same participant. For instance, for stimulus 1 in the participant shown in **Figure 1A**, all of the responses corresponded to the auditory component of the stimulus while for stimulus 9 in the same participant, nearly all responses corresponded to the illusory fusion percept. Across participants, the most frequent report was the illusory fusion percept ($48 \pm 24\%$, mean percentage \pm SD) followed by reports of the auditory component of the stimulus ($25 \pm 22\%$). There were fewer visual ($12 \pm 12\%$) and other responses ($16 \pm 14\%$).

As suggested by the large standard deviations in the response percentages, there was substantial variability across participants. We defined the McGurk susceptibility for each subject as the mean proportion of McGurk fusion reports across all presentations of all stimuli. A plot of the susceptibility values showed a broad range, with some subjects never reporting the effect and others always reporting it (**Figure 1B**). This broad distribution could not be explained by random response selection, as subjects responded consistently: across even and odd trials within the first testing session, the Pearson correlation for the proportion of McGurk responses was 0.97 [95% CI: 0.96 to 0.97 , 95% CI; $t(322) = 68.1$, $p < 10^{-16}$; **Figure 1C**]. Across the first and second testing session, separated by 2 years (see **Supplementary Data Sheet S2**), the Pearson correlation was 0.55 [95% CI: 0.42 to 0.65 ; $t(148) = 7.9$, $p = 10^{-13}$; **Figure 1D**].

Given that individual differences in McGurk susceptibility were reliable, we next examined the effect of twin status. For some pairs of twins, there were substantial differences between twins (such as MZ twin pair #153, **Figures 2A,B**). For other twin pairs, responses were more similar between co-twins (MZ twin pair #87, **Figures 2C,D**).

To visualize these relationships, we calculated the mean susceptibility to the McGurk effect of each twin pair and plotted the distance from the mean for each co-twin for both MZ (**Figure 2E**) and DZ twins (**Figure 2F**). For quantification, we calculated intraclass correlations (Griffin and Gonzalez, 1995) for each twin type. The mean rate of McGurk percepts was significantly correlated for both MZ twins ($r = 0.28$, 95% CI: 0.06 to 0.48 , $F(72,73) = 1.79$, $p = 0.007$) and DZ twins ($r = 0.21$, 95% CI: -0.01 to 0.40 $p = 0.026$). The intraclass correlation was not significantly different between MZ and DZ twins ($z = 0.52$, $p = 0.60$). The rate of auditory percepts also showed a positive intraclass correlation for both MZ twins [$r = 0.25$, 95% CI: 0.02 to 0.45 , $F(72,73) = 1.65$, $p = 0.02$] and DZ twins [$r = 0.14$, 95% CI: -0.07 to 0.34 , $F(88,89) = 1.32$, $p = 0.10$] with no significant difference between MZ and DZ twins ($z = 0.7$, $p = 0.48$).

While the twin correlations were significantly different than zero, they were low in absolute terms. An $r = 0.23$ (intraclass correlation calculated across MZ and DZ twins) corresponds to only 5.5% of variance explained. As another way to understand this relatively low correlation, we estimated how well one twin pair's response to a particular stimulus could predict their co-twin's response to the same stimulus. Across subjects, prediction error (root-mean squared error) was $44 \pm 17\%$ for a given stimulus. Strikingly, if we instead use the mean value for a given

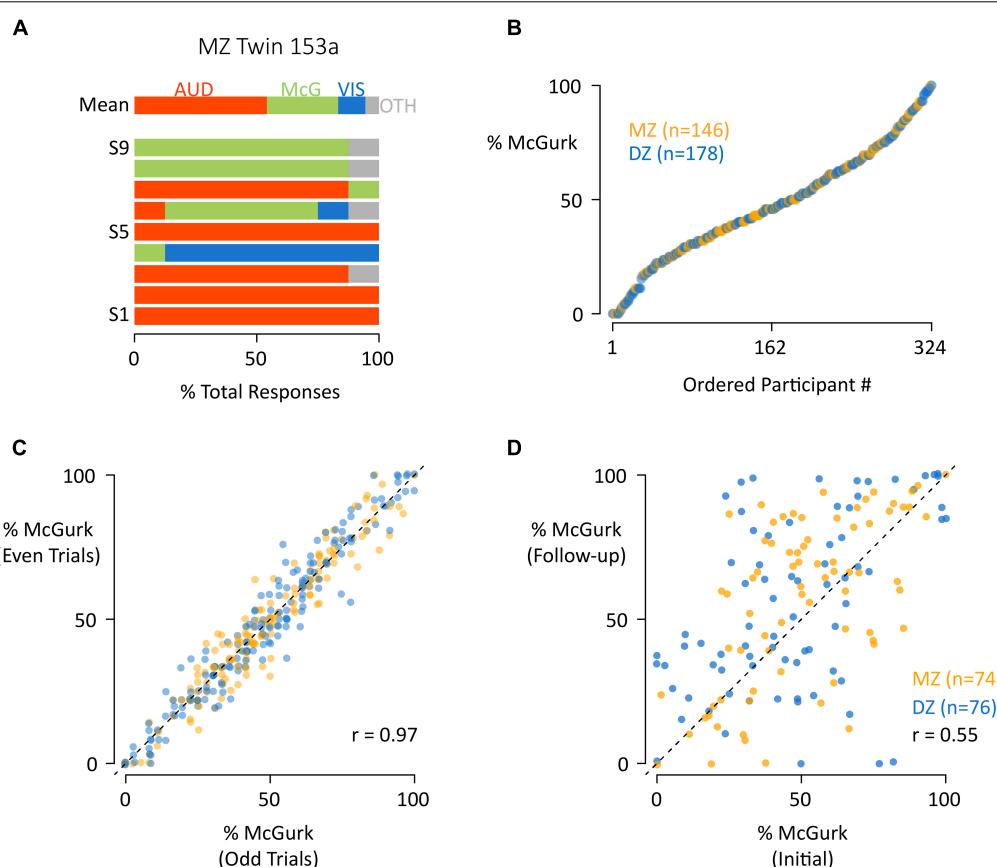


FIGURE 1 | The distribution and reliability of susceptibility to the McGurk effect. **(A)** For a single participant (MZ Twin 153a) the mean response across stimuli (row “Mean”) and for each individual stimulus (successive rows) are shown. For each stimulus presentation, the participant reported a percept, which was classified as auditory (AUD, red), McGurk (McG, green), visual (VIS, blue), or other (OTH, gray). The colored portions of each bar illustrate the percentage of each response type across repeated presentations of the same stimulus. For instance, the red color of the bar labeled “S1.” indicates that the participant reported an auditory percept for every presentation of stimulus “S1.” **(B)** For each participant, the mean percentage of McGurk reports across all stimuli (“% McGurk”) was calculated across stimuli, equivalent to the width of the green bar in the first row of **(A)**. Participants were ordered by % McGurk and plotted, one symbol per participant, orange for MZ twins, blue for DZ. **(C)** To assess reliability, for each participant the mean percentage of McGurk reports to even and odd trials were plotted, one symbol per participant, orange for MZ twins, blue for DZ. **(D)** In a follow-up experimental session approximately 2 years after the initial testing session, participants responded to the same stimuli. To assess reliability, for each participant the mean percentage of McGurk reports at the initial and follow-up sessions were plotted, one symbol per participant, orange for MZ twins, blue for DZ.

stimulus across all subjects (except the subject being predicted), prediction error improved to $35 \pm 10\%$ for a given stimulus ($p < 10^{-16}$, Wilcoxon signed-rank test).

Control Analyses

First, we verified that the intraclass correlations between actual twins (MZ and DZ twins combined, $r = 0.25$ and 0.17 for McGurk and auditory percepts, respectively) were higher than would be expected for randomly selected non-twin pairs for both McGurk illusion percepts (mean simulated $|r| < 0.01$, 95% bootstrap interval CI: -0.14 to 0.13) and auditory percepts (mean $|r| < 0.01$, 95% CI: -0.14 to 0.12).

Second, we examined demographic explanations for individual differences in McGurk susceptibility. Age, gender, birth order, and their interactions failed to explain a significant portion of the variance in McGurk susceptibility (linear mixed effects model with gender, age, birth order, and all two- and

three-way interactions as fixed factors; subject and stimulus were random effects; all parameter tests yielded $|t| < 1.15$, $p > 0.25$).

Finally, we verified that there were no mean differences between MZ and DZ groups in any percept category: McGurk, MZ vs. DZ, $48 \pm 23\%$ vs. $48 \pm 25\%$, $t(322) = -0.06$, $p = 0.95$; auditory: $23 \pm 22\%$ vs. $26 \pm 22\%$, $t(322) = 0.96$, $p = 0.34$; visual: $13 \pm 13\%$ vs. $12 \pm 11\%$, $t(322) = -0.97$, $p = 0.33$; other: $16 \pm 15\%$ vs. $15 \pm 13\%$, $t(322) = -0.56$, $p = 0.58$.

DISCUSSION

Twin studies have a long and troubled history of being used to support unjustified conclusions that human behavior is strongly genetically determined, reviewed in Joseph (2015). Strong claims about the heritability of behavior rest on many assumptions, including the conjecture that MZ and DZ twins experience equal

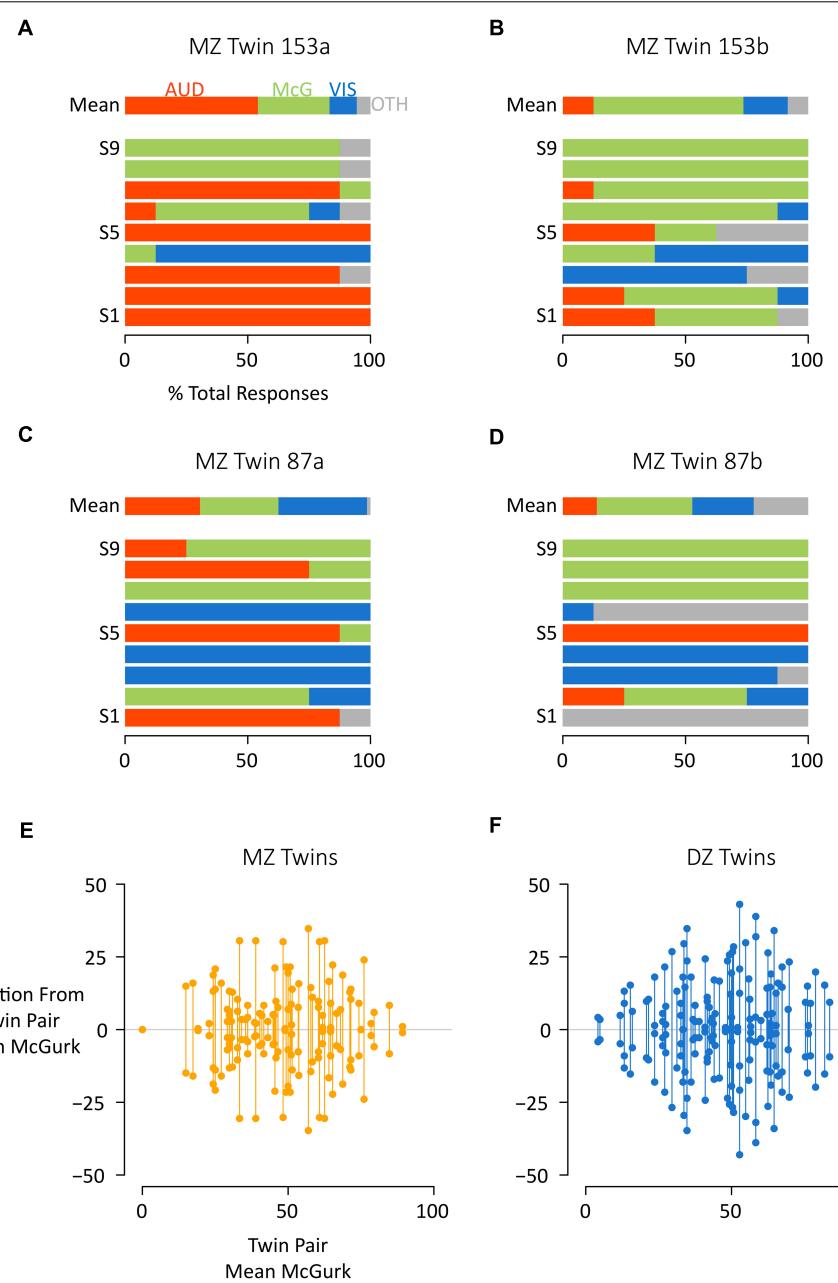


FIGURE 2 | Comparing McGurk susceptibility between and across co-twins. **(A)** For a single participant (MZ Twin 153a) the mean response across stimuli (row “Mean”) and for each individual stimulus (successive rows) are shown. For each stimulus presentation, the participant reported a percept, which was classified as auditory (AUD, red), McGurk (McG, green), visual (VIS, blue), or other (OTH, gray). The colored portions of each bar illustrate the percentage of each response type across repeated presentations of the same stimulus. For instance, the red color of the bar labeled “S1” indicates that the participant reported an auditory percept for every presentation of stimulus “S1.” **(B)** For the other co-twin in this pair (MZ Twin 153b) a different pattern of responses was noted. For instance, for stimulus “S1”, a majority of participant reports corresponded to the McGurk percept. **(C,D)** Co-twins MZ Twin 87A and 87B showed more similar responses across stimuli. **(E,F)** To examine individual twin pairs, the % McGurk reports within each twin pair was calculated, and the deviation from this mean plotted by showing two circles (one for each co-twin) connected by a vertical line. The average length of the vertical lines indicates the magnitude of the intraclass correlation. The left plot shows all MZ twins (orange circles), the right plot shows all DZ twins (blue circles). Within each plot, twin pairs were ordered by mean% McGurks, from least to most. For instance, the left-most orange symbol pair shows two co-twins both with 0% McGurk, while the right-most blue symbol shows two co-twins with 97 and 100% McGurk.

environments, so that greater behavioral similarity between MZ than DZ co-twins can only be explained by greater genetic similarity. However, *in utero*, approximately two-thirds of MZ

twins share placentas, while no DZ twins share placentas. Since most MZ co-twins are exposed to exactly the same concentration and temporal variation in developmentally active biochemical

factors in their shared circulation, MZ twins have much more similar prenatal environments than DZ twins, invalidating the equal environment assumption.

Keeping these caveats in mind, we found that the degree of susceptibility to the McGurk effect in MZ and DZ twins was significantly more similar than chance, demonstrating that shared genes and environment play some role in determining McGurk susceptibility. However, the percent variance explained (r^2) was only 5.5% (averaged across MZ and DZ twins) and there was no significant difference between MZ and DZ twins. Put another way, for both MZ and DZ twins, knowing how often one co-twin perceived the McGurk effect provided little information on the McGurk perception of the other co-twin.

This result argues against a deterministic view of individual differences in the McGurk effect. A growing body of literature shows that McGurk susceptibility is significantly, but weakly, correlated with a number of cognitive and personality factors, including the pattern of eye movements made when viewing a talking face (Gurler et al., 2015); lipreading skill (Strand et al., 2014; Brown et al., 2018); temporal binding window (Stevenson et al., 2012); and autistic traits (Van Laarhoven et al., 2019). Taken together, these findings suggest that audiovisual speech perception, as assayed with the McGurk effect, is similar to other complex human behaviors in that it is influenced by a host of factors, both genetic and environmental.

Stimulus Variability

The audiovisual recordings of *ba/ga* and *pa/ka* used by McGurk and MacDonald in their original study are lost (McGurk and MacDonald, 1976; Beauchamp, 2018; MacDonald, 2018) forcing each laboratory to create their own stimuli. These stimuli vary in efficacy, hindering across-study comparisons (Alsius et al., 2018). Other studies have reported high interstimulus variability across different stimulus exemplars of *ba/ga* (Jiang and Bernstein, 2011; Basu Mallick et al., 2015), *pa/ka* (Magnotti and Beauchamp, 2015) and *ba/fa* (Shahin et al., 2018; Shahin, 2019).

Many factors, including co-articulation (number of repetitions of the syllable, i.e., *ba* vs. *baba*) and speech rate (fast vs. slow) may contribute to the efficacy of a given McGurk exemplar (Magnotti et al., 2018a,b). In the absence of techniques for psychometrically manipulating McGurk stimuli as there are for simpler sensory stimuli (such as parametrically changing the volume of a tone to measure auditory sensitivity) participants in the present study viewed a battery of nine different *ba/ga* and *pa/ka* McGurk exemplars, with the average score across stimuli measuring the overall susceptibility of a given participant to the illusion. Because the same stimulus battery was presented to every individual, comparing scores within and across twin pairs allowed for an estimate of the magnitude of genetic and early environmental influences on McGurk susceptibility. Adding additional stimuli to the battery, such as *ba/fa* (Shahin, 2019), would increase generalizability at the cost of additional experimental time.

Participant Variability

In addition to interstimulus variability, our study confirms and extends previous reports demonstrating that different individuals vary greatly in their susceptibility to the McGurk

effect (Gentilucci and Cattaneo, 2005; Nath and Beauchamp, 2012; Strand et al., 2014; Basu Mallick et al., 2015; Magnotti and Beauchamp, 2015; Shahin, 2019). Our study used a sample size ($n = 324$) that is an order of magnitude larger than most published studies of the McGurk effect. This is important because in a phenomenon with high inter-subject variability, large sample sizes (> 100) are necessary to accurately estimate differences between groups (Magnotti and Beauchamp, 2018).

In order to obtain large sample sizes, online testing services such as Amazon Mechanical Turk make it possible to quickly and easily present stimuli to many participants. Although we have previously shown similar distributions of McGurk illusion perception between online and laboratory settings (Experiment 4 in Magnotti et al., 2018b), a concern with on-line testing of audiovisual speech is that different participants are likely to view the stimuli in different formats (e.g., a small tablet screen vs. a large monitor) and hear them under different listening conditions (e.g., headphones vs. loudspeakers; quiet vs. noisy environments) that are not under the control of the investigator. The present study used in-person testing in the laboratory, rather than on-line testing, allowing viewing and listening conditions to be carefully controlled and equated across participants. Therefore, the observed inter-subject variability cannot be attributed to the differences in viewing and listening conditions that are a possible confound in on-line testing. Given that there was only a small main effect of heritability/early environment on McGurk perception, we did not investigate the interaction between heritability/early environment and interstimulus or interparticipant variability.

Cultural Variability

All study participants were native speakers of Mandarin, while the McGurk stimuli were recorded by English speakers. However, the English syllables contained in the stimuli are common sounds in Mandarin, serving as acoustic subunits of Mandarin words with identical spellings in Pinyin, the official Romanization system for Standard Chinese. Although the McGurk effect was first discovered in native English speakers, previous studies have demonstrated that it is a cross-cultural phenomenon experienced by Mandarin speakers as well as native speakers of Cantonese, Finnish, German, Hebrew, Hungarian, Italian, Japanese, Spanish and Thai (Sekiyama and Tohkura, 1991; Grassegger, 1995; Aloufy et al., 1996; Fuster-Duran, 1996; Sekiyama, 1997; Burnham and Lau, 1998; Sams et al., 1998; Chen and Hazan, 2007; Traunmüller and Öhrström, 2007; Bovo et al., 2009).

In order to examine McGurk differences between cultures, in a previous publication we compared a subset of the participants in the current study (the older twin of each twin pair, $n = 162$) with a sample of native English speakers from the United States ($n = 145$) (Magnotti et al., 2015) using the same stimuli and a similar open-choice experimental procedure, we found no significant difference in mean McGurk susceptibility between native Mandarin and native English speakers ($48 \pm 2\%$ vs. $44 \pm 2\%$) and individuals from both groups spanned the entire range of McGurk susceptibility (from 0 to 100%). In the absence of strong intercultural differences in the illusion, there is no reason to think that the relatively small effect of genes and

environment on McGurk perception estimated for native Chinese speakers would be larger for native speakers of other languages, including English.

DATA AVAILABILITY STATEMENT

The datasets analyzed during the current study are included as **Supplementary Material** to this manuscript.

ETHICS STATEMENT

This study involving human participants were reviewed and approved by all subjects gave written informed consent to participate in an experimental protocol which was carried out in accordance with relevant guidelines and regulations and was approved by the Institutional Review Board at Institute of Psychology, Chinese Academy of Sciences. Parental informed consent was obtained for subjects under 18 years of age. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

MB conceived the study. MB and WZ designed the study. GF and BZ performed the testing and data collection. GF and JM performed the data analysis and interpretation under the

supervision of WZ and MB. All authors contributed to the writing of the manuscript and approved the final version of the manuscript for submission.

FUNDING

This work was supported by the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (QYZDB-SSW-SMC055), the National Natural Science Foundation of China (31671138), and the Beijing Twin Study project (BeTwiSt) of Institute of Psychology, Chinese Academy of Sciences. MB was supported by NIH R01NS065395. JM was supported by a training fellowship from the Gulf Coast Consortia, NLM Training Program in Biomedical Informatics (NLM Grant No. T15LM007093). Portions of this work were presented at the International Multisensory Research Forum in 2014 and 2016.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2019.01029/full#supplementary-material>

DATA SHEET S1 | Participants' responses to McGurk stimuli presented in the initial study of Experiment 1.

DATA SHEET S2 | Participants' responses to McGurk stimuli presented during follow-up study.

REFERENCES

- Aloufy, S., Lapidot, M., and Myslobodsky, M. (1996). Differences in susceptibility to the "blending illusion" among native hebrew and english speakers. *Brain Lang.* 53, 51–57. doi: 10.1006/brln.1996.0036
- Alsius, A., Paré, M., and Munhall, K. G. (2018). Forty years after hearing lips and seeing voices: the mcgurk effect revisited. *Multisens. Res.* 31, 111–144. doi: 10.1163/22134808-00002565
- Basu Mallick, D., Magnotti, J. F., and Beauchamp, M. S. (2015). Variability and stability in the McGurk effect: contributions of participants, stimuli, time, and response type. *Psychon. Bull. Rev.* 22, 1299–1307. doi: 10.3758/s13423-015-0817-4
- Beauchamp, M. S. (2018). Introduction to the special issue: forty years of the mcgurk effect. *Multisens. Res.* 31, 1–6. doi: 10.1163/22134808-00002598
- Bovo, R., Ciorda, A., Prosser, S., and Martini, A. (2009). The mcgurk phenomenon in italian listeners. *Acta Otorhinolaryngol. Ital.* 29, 203–208.
- Brown, V. A., Hedayati, M., Zanger, A., Mayn, S., Ray, L., Dillman-Hasso, N., et al. (2018). What accounts for individual differences in susceptibility to the mcgurk effect? *PloS one* 13:e0207160. doi: 10.1371/journal.pone.0207160
- Burnham, D., and Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: perception of an emergent consonant in the McGurk effect. *Dev. Psychobiol.* 45, 204–220. doi: 10.1002/dev.20032
- Burnham, D., and Lau, S. (1998). "The effect of tonal information on auditory reliance in the McGurk effect," in *Proceedings of the AVSP'98 International Conference on Auditory-Visual Speech Processing*, 1998, Sydney.
- Campbell, R. (2008). The processing of audio-visual speech: empirical and neural bases. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 1001–1010. doi: 10.1098/rstb.2007.2155
- Chen, J., Li, X., Chen, Z., Yang, X., Zhang, J., Duan, Q., et al. (2010). Optimization of zygosity determination by questionnaire and DNA genotyping in chinese adolescent twins. *Twin Res. Hum. Genet.* 13, 194–200. doi: 10.1375/twin.13.2.194
- Chen, J., Li, X., Zhang, J., Natsuki, M. N., Leve, L. D., Harold, G. T., et al. (2013). The beijing twin study (BeTwiSt): a longitudinal study of child and adolescent development. *Twin Res. Hum. Genet.* 16, 91–97. doi: 10.1017/thg.2012.115
- Chen, Y., and Hazan, V. (2007). "Developmental factor in auditory-visual speech perception-the mcgurk effect in mandarin-chinese and english speakers," in *Proceedings of the AVSP'07 International Conference on Auditory-Visual Speech Processing*, Hilvarenbeek.
- de Gelder, B., Vroomen, J., Annen, L., Masthof, E., and Hodiamont, P. (2003). Audio-visual integration in schizophrenia. *Schizophr. Res.* 59, 211–218. doi: 10.1016/s0920-9964(01)00344-9
- Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., and Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science* 291, 1969–1972. doi: 10.1126/science.291.5510.1969
- Fuster-Duran, A. (1996). "Perception of conflicting audio-visual speech: an examination across Spanish and German," in *Speechreading by Humans and Machines*, eds D. G. Stork, and M. E. Hennecke, (Berlin: Springer), 135–145.
- Gentilucci, M., and Cattaneo, L. (2005). Automatic audiovisual integration in speech perception. *Exp. Brain Res.* 167, 66–75. doi: 10.1007/s00221-005-0008-z
- Grassegger, H. (1995). "McGurk effect in german and hungarian listeners," in *Proceedings of the International Congress of Phonetic Sciences*, Stockholm.
- Griffin, D., and Gonzalez, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychol. Bull.* 118, 430–439. doi: 10.1037/0033-2909.118.3.430
- Guiraud, J. A., Tomalski, P., Kushnerenko, E., Ribeiro, H., Davies, K., Charman, T., et al. (2012). Atypical audiovisual speech integration in infants at risk for autism. *PLoS One* 7:e36428. doi: 10.1371/journal.pone.0036428
- Gurler, D., Doyle, N., Walker, E., Magnotti, J., and Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye

- movements. *Atten. Percep. Psychophys.* 77, 1333–1341. doi: 10.3758/s13414-014-0821-1
- Hayiou-Thomas, M. E., Dale, P. S., and Plomin, R. (2012). The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Dev. Sci.* 15, 233–249. doi: 10.1111/j.1467-7687.2011.01119.x
- Hubert, H. B., Fabsitz, R. R., Feinleib, M., and Brown, K. S. (1980). Olfactory sensitivity in humans: genetic versus environmental control. *Science* 208, 607–609. doi: 10.1126/science.7189296
- Jiang, J., and Bernstein, L. E. (2011). Psychophysics of the McGurk and other audiovisual speech integration effects. *J. Exp. Psychol.* 37, 1193–1209. doi: 10.1037/a0023100
- Joseph, J. (2015). *The Trouble with Twin Studies: A Reassessment of Twin Research in the Social and Behavioral Sciences*. New York, NY: Routledge.
- MacDonald, J. (2018). Hearing lips and seeing voices: the origins and development of the 'McGurk Effect' and reflections on audio-visual speech perception over the last 40 years. *Multisens. Res.* 31, 7–18. doi: 10.1163/22134808-00002548
- Magnotti, J. F., Basu Mallick, D., and Beauchamp, M. S. (2018a). Reducing playback rate of audiovisual speech leads to a surprising decrease in the McGurk effect. *Multisens. Res.* 31, 19–38. doi: 10.1163/22134808-00002586
- Magnotti, J. F., Smith, K. B., Salinas, M., Mays, J., Zhu, L. L., and Beauchamp, M. S. (2018b). A causal inference explanation for enhancement of multisensory integration by co-articulation. *Sci. Rep.* 8:18032. doi: 10.1038/s41598-018-36772-8
- Magnotti, J. F., Basu Mallick, D., Feng, G., Zhou, B., Zhou, W., and Beauchamp, M. S. (2015). Similar frequency of the McGurk effect in large samples of native mandarin chinese and american english speakers. *Exp. Brain Res.* 233, 2581–2586. doi: 10.1007/s00221-015-4324-7
- Magnotti, J. F., and Beauchamp, M. S. (2015). The noisy encoding of disparity model of the McGurk effect. *Psychon. Bull. Rev.* 22, 701–709. doi: 10.3758/s13423-014-0722-2
- Magnotti, J. F., and Beauchamp, M. S. (2018). Published estimates of group differences in multisensory integration are inflated. *PloS one* 13:e0202908. doi: 10.1371/journal.pone.0202908
- McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0
- Miller, S. M., Hansell, N. K., Ngo, T. T., Liu, G. B., Pettigrew, J. D., Martin, N. G., et al. (2010). Genetic contribution to individual variation in binocular rivalry rate. *Proc. Natl. Acad. Sci. U S A.* 107, 2664–2668. doi: 10.1073/pnas.0912149107
- Nath, A. R., and Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage* 59, 781–787. doi: 10.1016/j.neuroimage.2011.07.024
- Nath, A. R., Fava, E. E., and Beauchamp, M. S. (2011). Neural correlates of interindividual differences in children's audiovisual speech perception. *J. Neurosci.* 31, 13963–13971. doi: 10.1523/jneurosci.2605-11.2011
- Pearl, D., Yodashkin-Porat, D., Katz, N., Valevski, A., Aizenberg, D., Sigler, M., et al. (2009). Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Compr. Psychiatry* 50, 186–192. doi: 10.1016/j.comppsych.2008.06.004
- Rogers, S. J., and Ozonoff, S. (2005). Annotation: what do we know about sensory dysfunction in autism? a critical review of the empirical evidence. *J. Child Psychol. Psychiatry* 46, 1255–1268. doi: 10.1111/j.1469-7610.2005.01431.x
- Ross, D. E., Buchanan, R. W., Medoff, D., Lahti, A. C., and Thaker, G. K. (1998). Association between eye tracking disorder in schizophrenia and poor sensory integration. *Am. J. Psychiatry* 155, 1352–1357. doi: 10.1176/ajp.155.10.1352
- Sams, M., Manninen, P., Surakka, V., Helin, P., and Kättö, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences: effects of word meaning and sentence context. *Speech Commun.* 26, 75–87. doi: 10.1016/s0167-6393(98)00051-x
- Sekiyama, K. (1997). Cultural and linguistic factors in audiovisual speech processing: the McGurk effect in Chinese subjects. *Percept. Psychophys.* 59, 73–80. doi: 10.3758/bf03206849
- Sekiyama, K., and Tohkura, Y. (1991). McGurk effect in non-english listeners: few visual effects for japanese subjects hearing japanese syllables of high auditory intelligibility. *J. Acoust. Soc. Am.* 90(4 Pt 1), 1797–1805. doi: 10.1121/1.401660
- Shahin, A. J. (2019). Neural evidence accounting for interindividual variability of the McGurk illusion. *Neurosci. Lett.* 707:134322. doi: 10.1016/j.neulet.2019.134322
- Shahin, A. J., Backer, K. C., Rosenblum, L. D., and Kerlin, J. R. (2018). Neural mechanisms underlying cross-modal phonetic encoding. *J. Neurosci.* 38, 1835–1849. doi: 10.1523/JNEUROSCI.1566-17.2017
- Smith, E. G., and Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *J. Child Psychol. Psychiatry* 48, 813–821. doi: 10.1111/j.1469-7610.2007.01766.x
- Stevenson, R. A., Zemtsov, R. K., and Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *J. Exp. Psychol.* 38, 1517–1529. doi: 10.1037/a0027339
- Strand, J., Cooperman, A., Rowe, J., and Simenstad, A. (2014). Individual differences in susceptibility to the McGurk effect: links with lipreading and detecting audiovisual incongruity. *J. Speech Lang. Hear. Res.* 57, 2322–2331. doi: 10.1044/2014_JSLHR-H-14-0059
- Traunmüller, H., and Öhrström, N. (2007). Audiovisual perception of openness and lip rounding in front vowels. *J. Phon.* 35, 244–258. doi: 10.1016/j.wocn.2006.03.002
- Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lepore, F., and Théoret, H. (2007). Speech and non-speech audio-visual illusions: a developmental study. *PLoS One* 2:e742. doi: 10.1371/journal.pone.0000742
- Van Laarhoven, T., Stekelenburg, J. J., and Vroomen, J. (2019). Increased sub-clinical levels of autistic traits are associated with reduced multisensory integration of audiovisual speech. *Sci. Rep.* 9:9535. doi: 10.1038/s41598-019-46084-0
- Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., et al. (2010). Heritability of the specific cognitive ability of face perception. *Curr. Biol.* 20, 137–142. doi: 10.1016/j.cub.2009.11.067

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Feng, Zhou, Zhou, Beauchamp and Magnotti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.