



OPEN ACCESS

EDITED BY

Robert J. Zatorre,
McGill University, Canada

REVIEWED BY

Jean Vroomen,
Tilburg University, Netherlands
Virginia Penhune,
Concordia University, Canada

*CORRESPONDENCE

John F. Magnotti

✉ John.Magnotti@Penmedicine.upenn.edu

Michael S. Beauchamp

✉ Michael.Beauchamp@Penmedicine.upenn.edu

RECEIVED 20 November 2024

ACCEPTED 24 February 2025

PUBLISHED 28 March 2025

CITATION

Magnotti JF, Basu Mallick D, Feng G, Zhou B,
Zhou W and Beauchamp MS (2025) The
McGurk effect is similar in native Mandarin
Chinese and American English speakers.
Front. Psychol. 16:1531566.
doi: 10.3389/fpsyg.2025.1531566

COPYRIGHT

© 2025 Magnotti, Basu Mallick, Feng, Zhou,
Zhou and Beauchamp. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

The McGurk effect is similar in native Mandarin Chinese and American English speakers

John F. Magnotti^{1*}, Debshila Basu Mallick², Guo Feng³,
Bin Zhou⁴, Wen Zhou⁴ and Michael S. Beauchamp^{1*}

¹Department of Neurosurgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, United States, ²OpenStax, Rice University, Houston, TX, United States, ³Psychological Research and Counseling Center, Southwest Jiaotong University, Chengdu, Sichuan, China, ⁴Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

Humans combine the visual information from mouth movements with auditory information from the voice to recognize speech. A common method for assessing audiovisual speech perception is the McGurk effect: when presented with some incongruent pairings of auditory and visual speech syllables (e.g., the auditory speech sound “ba” dubbed onto the visual mouth movements for “ga”) individuals perceive a third syllable, distinct from the auditory and visual components. The many differences between Chinese and American culture and language suggest the possibility of group differences in the McGurk effect. Published studies have reported less McGurk effect in native Mandarin Chinese speakers than in English speakers, but these studies sampled small numbers of participants tested with a small number of stimuli. Therefore, we conducted in-person tests of the McGurk effect in large samples of Mandarin-speaking individuals from China and English-speaking individuals from the USA (total $N = 307$) viewing nine different stimuli. Averaged across participants and stimuli, we found similar frequencies of the McGurk effect between Chinese and American participants (48% vs. 44%). In both groups, there was high variability both across participants (range from 0% to 100%) and stimuli (14%–83%) with the main effect of culture and language accounting for only 0.2% of the variance in the data. The high variability inherent to the McGurk effect necessitates the use of large sample sizes to accurately estimate group differences and requires testing with a variety of McGurk stimuli, especially stimuli potent enough to evoke the illusion in the majority of participants.

KEYWORDS

McGurk effect, cultural differences, audiovisual speech, multisensory integration, individual differences

Introduction

Humans around the world communicate by speaking and listening face-to-face. During these interactions, we integrate the heard speech sounds with the seen mouth movements to increase both the speed and accuracy of speech perception (Peelle and Sommers, 2015; Ross et al., 2007; Sumby and Pollack, 1954; van Wassenhove et al., 2005). A common way to assess multisensory integration during speech perception is an illusion known as the McGurk effect (McGurk and MacDonald, 1976) in which individuals presented with incongruent auditory and visual syllables report hearing an entirely different syllable. The McGurk effect has become a popular assay of multisensory speech perception because it is easy to administer: both the stimulus and the response

consist of only a single syllable. However, some individuals do not experience the effect and instead perceive the auditory or visual components of the stimulus (Nath and Beauchamp, 2012; Stevenson et al., 2012). These individual differences are consistent across test-retest intervals of 12 months or longer, suggesting that they reflect stable differences in the propensity to integrate auditory and visual speech information (Basu Mallick et al., 2015).

Although many laboratory studies of psychological phenomena focus exclusively on native English speakers, the McGurk effect is an important exception. It has been studied across native speakers of Mandarin Chinese, Cantonese, Thai, and Japanese (Burnham and Lau, 1998; Chen and Hazan, 2007; Sekiyama, 1997; Sekiyama and Tohkura, 1991), Spanish, German, Hungarian (Fuster-Duran, 1996; Grassegger, 1995), Italian (Bovo et al., 2009), Finnish (Sams et al., 1998; Traunmüller and Öhrström, 2007), and Hebrew (Aloufy et al., 1996). The groups in these studies are defined both by cultural differences and by differences in their native language; in this paper, we group them and refer to them together as “intercultural.”

The strongest claim in the literature for intercultural differences in the McGurk effect involves comparisons between Asian and non-Asian cultures. Sekiyama and Tohkura (1991, 1993) reported a lower frequency of McGurk perception in native Japanese speakers than in native English speakers, and equal or lower frequency in Mandarin Chinese speakers than in Japanese speakers (Hayashi and Sekiyama, 1998; Sekiyama, 1997). In agreement with these results, Burnham and Lau (1998) found a lower frequency of McGurk perception in Cantonese speakers than English speakers although other studies comparing English and Chinese speakers did not find differences in McGurk frequency (Chen and Hazan, 2007, 2009).

Two major groups of hypotheses have emerged to explain intercultural differences in the McGurk effect. The linguistic hypothesis explains them via the properties of Asian languages. Tonal languages (such as Mandarin) and semi-tonal languages (such as pitch accents in Japanese) may increase reliance on auditory speech cues, decreasing the relevance of visual speech information (Sekiyama, 1997). Phonemes of Mandarin and Japanese may be easier to discriminate without visual cues than those of English, reducing the need for visual speech information to disambiguate speech sounds (Sekiyama and Burnham, 2008). The face-avoidance hypothesis explains them via the cultural milieu of the listener. In Japanese and Chinese cultures, direct viewing of the face can be considered impolite and hence may discourage people in these cultures from developing a strong reliance on the visual speech information required for perception of the McGurk effect (Sekiyama, 1997). There is some evidence that English-speaking children are better at visual-only identification of speech than Japanese children (Sekiyama and Burnham, 2008).

One potential problem with these findings of intercultural differences is that they were conducted before recent advances in our understanding of individual differences in the McGurk effect. Some native English speakers never perceive the illusion and others always perceive it (Magnotti and Beauchamp, 2015; Basu Mallick et al., 2015; Nath and Beauchamp, 2012; Stevenson et al., 2012; Strand et al., 2014). High variability means that large sample sizes are necessary for accurate statistical inference, but many studies of cultural differences in the McGurk effect have used small sample sizes (e.g., 10–14 participants, Bovo et al., 2009; Sekiyama,

1994, 1997), possibly resulting in inferential errors (Magnotti and Beauchamp, 2018).

Another difficulty in interpreting the literature is that stimuli from different talkers (or even different stimuli from the same talker) vary greatly in their ability to evoke the McGurk effect (Jiang and Bernstein, 2011; Magnotti and Beauchamp, 2015; Basu Mallick et al., 2015). This variability is problematic when cross-cultural studies use stimuli created from only two talkers (Bovo et al., 2009; Burnham and Lau, 1998; Hayashi and Sekiyama, 1998; Sekiyama, 1994, 1997; Sekiyama and Tohkura, 1993). Just as testing a small group of participants from a highly variable population is problematic, testing only a few McGurk stimuli can also lead to errors in inference due to the idiosyncrasies of individual stimuli, especially if only weak stimuli that rarely evoke the McGurk effect are tested.

To overcome these difficulties, we compared McGurk perception between a large sample of Mandarin-speaking individuals from China ($n = 162$) and a large sample of English-speaking individuals from the USA ($n = 145$) using a battery of nine McGurk stimuli from eight different talkers. In-person testing with a large sample of participants and stimuli allowed for the accurate estimation of intercultural differences in the McGurk effect.

Methods

Chinese participants

All participants gave written informed consent to participate in an experimental protocol approved by the Institutional Review Board of the Institute of Psychology of the Chinese Academy of Sciences. Parental informed consent was obtained for participants under 18 years of age. Participants were tested in-person and consisted of $n = 162$ Mandarin speakers native to China (82 female; mean age = 17 years, range = 14–23) recruited from the Beijing Twin Study project of the Institute of Psychology of Chinese Academy of Sciences (analysis was only conducted on the first-born of each twin pair). All participants reported normal or corrected-to-normal vision and no history of speech, language, or hearing difficulties.

American participants

All participants gave written informed consent to participate in an experimental protocol approved by the Institutional Review Board of Rice University. All participants were native to the USA and reported English as their primary language ($n = 145$; 97 female, mean age = 19 years, range = 18–26). All participants reported normal or corrected-to-normal vision and no history of speech, language, or hearing difficulties, and were tested in-person.

Data and code availability

Data and analysis code are available in the [Supplementary material](#).

Stimuli and procedure

The McGurk stimuli consisted of nine audiovisual recordings, lasting 2 s each. Each stimulus contained an auditory recording of a syllable and a video recording of the face of the same talker enunciating a different syllable. Four stimuli consisted of auditory “ba” and visual “ga” (AbaVga). Three stimuli consisted of double syllables, auditory “baba” paired with visual “gaga” (AbabaVgaga). Two stimuli consisted of auditory “pa” and visual “ka” (ApaVka). There were five male speakers and three female speakers (the same female speaker appeared in two stimuli). Stimuli were viewed at a distance of 40 cm and filled a 15” LCD display.

During the experiment, the stimuli were presented in random order. Participants in the China group saw each McGurk stimulus eight times; participants in the USA group saw each McGurk stimulus 10 times, but we analyzed only the first eight presentations to match the China group (the results were unchanged when all 10 presentations were analyzed).

Participants reported their percepts by speaking aloud and no feedback was given. Responses were recorded by the stimulus computer and transcribed by a research assistant. The USA group also viewed control stimuli (10 times each) intermixed with the McGurk stimuli: six congruent audiovisual syllables (“ba,” “ga,” “pa,” “ka,” “da,” “ta”) and two non-McGurk incongruent stimuli, which are similar to McGurk stimuli, but with the auditory and visual constituents reversed (AgaVba and AkaVpa) all spoken by the same female speaker.

Scoring responses

Responses to McGurk stimuli were categorized as follows. The responses “da” or “tha” (to AbaVga) and “ta” or “tha” (to ApaVka) were categorized as McGurk fusion responses. The responses “ba” (to AbaVga) and “pa” (to ApaVka) were categorized as auditory responses. The responses “ga” (to AbaVga) and “ka” (to ApaVka) were categorized as visual responses. Any other response was categorized as “other.” For AbabaVgaga stimuli, each syllable was coded separately (e.g., the response “dada” was coded as 1.0 McGurk; the response “bada” was coded as 0.5 McGurk and 0.5 auditory).

Across all subjects and stimuli, the McGurk responses (46%) and auditory responses (37%) were the most common. Visual responses (7%) and “other” responses (10%) were comparatively rare across stimuli and individuals. Only two stimuli had visual responses more than 15% of time, and only two stimuli had “other” responses more than 15% of time. This pattern of responding led to complementary percentages between McGurk and auditory responses, and thus we analyzed only McGurk responses to each stimulus.

Results

We compared the frequency of the McGurk effect in native Mandarin-speaking individuals from China and native English-speaking individuals from the USA across nine stimuli (Figure 1A). The overall frequency of McGurk responses for the China group

(Mean = 48%, standard error of the mean, SEM = 2%) was slightly greater than the USA group (Mean = 44%, SEM = 2%). In both groups, there was high variability across participants, with some participants in both groups never perceiving the illusion (0%) and some from both groups always perceiving the illusion (100%). Stimuli varied dramatically in their effectiveness (Figure 1B). The weakest stimulus (#1 in the ranking) evoked the McGurk effect on only 14% of trials (averaged across all participants), while the strongest stimulus (#9) evoked the McGurk effect on 83% of trials (Figure 1B).

A linear mixed-effects model was fit with percent McGurk responses as the dependent measure, fixed factors of cultural group and stimulus, and random effect of participant. There was no main effect of cultural group [$X^2_{(1)} = 1.9, p = 0.17$]. There was a main effect of stimulus [$X^2_{(8)} = 1597.4, p < 10^{-16}$], driven by the wide range of effectiveness across stimuli, and a significant stimulus-by-cultural group interaction [$X^2_{(8)} = 124.4, p < 10^{-16}$] driven by greater McGurk for the USA group for stimulus 4 [*post-hoc* $t = -3.1, p = 0.002$], and greater McGurk for the China group for stimulus 7 [$t = 6.7, p = 10^{-11}$] and stimulus 8 [$t = 4.2, p = 10^{-5}$]; other stimuli showed no difference (all $ps > 0.1$).

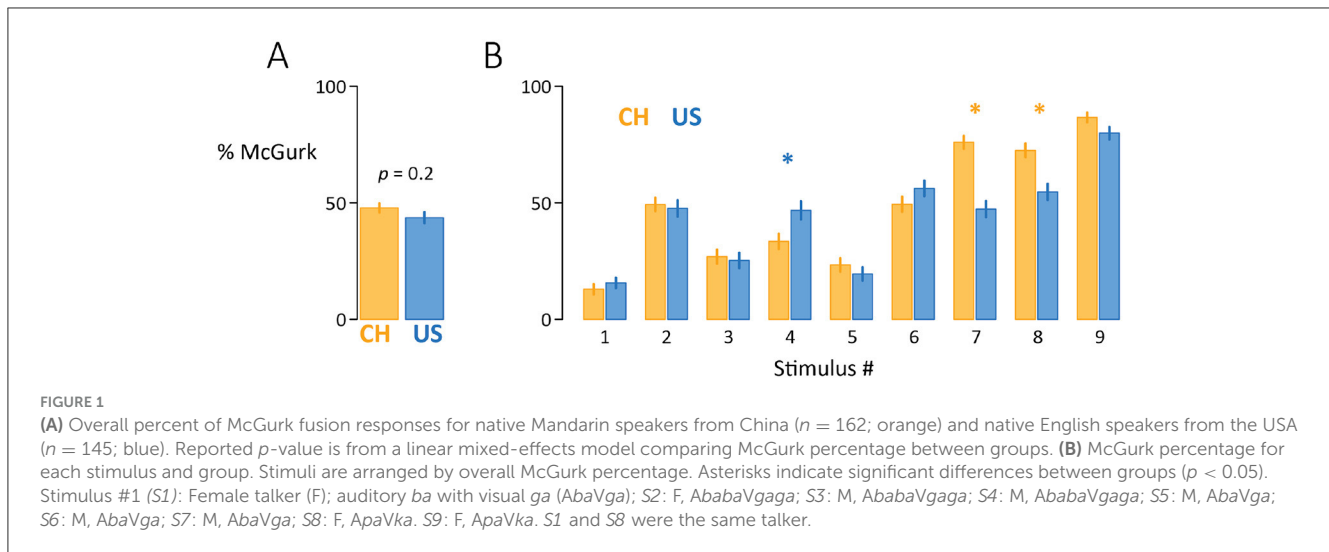
There were no obvious explanations for the differences between stimuli. Stimulus #4 (USA > CH) was a male talker saying *AbabaVgaga*, while stimulus #7 (USA < CH) was a male talker saying *AbaVga* and stimulus #8 (USA < CH) was a female talker saying *ApaVka*. Two stimuli recorded by the same female talker showed different response profiles, stimulus #1 (USA ~ CH) and stimulus #8 (USA < CH).

Discussion

In a sample of 307 individuals and nine stimuli, similar frequencies of the McGurk effect were observed in native Mandarin speakers from China and native English speakers from the USA. In each group, there was high variability across participants (range from 0% to 100%) and stimuli (14%–83%). The large sample size allowed for the accurate estimation of effect sizes. The main effect of cultural group accounted for only 0.2% of variance in the frequency of McGurk perception; the interaction between cultural group and stimulus accounted for only 2% of the variance.

The finding of similar frequencies of McGurk perception in native Mandarin and native English speakers supports other evidence that the fundamentals of speech perception are similar between the two groups (Chen and Hazan, 2009; Hazan et al., 2010) but fail to replicate previous reports of significantly reduced susceptibility to the McGurk effect in native Chinese speakers (Burnham and Lau, 1998; Hayashi and Sekiyama, 1998; Sekiyama, 1997).

Studies that reported significant intercultural differences in the McGurk effect in native Chinese speakers tested many fewer participants than the present study, which did not find a significant intercultural effect. The pattern of a small study or studies showing significant effects followed by a null result from a well-powered study is not uncommon in the McGurk literature. In native Japanese speakers, studies with small sample sizes reported little



or no McGurk effect (Sekiyama and Tohkura, 1991, 1993) but two larger studies refuted this claim (Magnotti et al., 2024a; Tiippana et al., 2023). A small study claimed that skilled musician did not experience the McGurk effect (Proverbio et al., 2016), while larger studies reached a very different conclusion (Lee et al., 2024; Politzer-Ahles and Pan, 2019). Small studies reported differences in the McGurk effect between autistic and non-autistic adults but a highly-powered study ($n = 869$) found no difference (Jertberg et al., 2024).

A likely explanation for this pattern is that the large variability in individual susceptibility to the McGurk effect makes it impossible to precisely estimate McGurk frequency using the small sample sizes (<15 per group) typical of many McGurk group difference studies. A study with 15 participants per group would have only 18% power to detect a 10% difference in the frequency of the McGurk effect between groups (Magnotti and Beauchamp, 2018). In contrast, the current study (with an ~ 10 times larger n) had 83% power to find group differences as small as 10%. Finding a significant result with a small sample size is sometime taken as evidence that the effect must be robust, but this is a fallacy (Loken and Gelman, 2017). Instead, it usually reflects a “winner’s curse” where low power, publication bias, and a multiverse of data analysis choices combine to produce inflated estimates of effect size (Ferguson and Heene, 2012; Kuhberger et al., 2014; Lindstromberg, 2023; Magnotti and Beauchamp, 2018; Steegen et al., 2016).

Along with large sample sizes, an equally important ingredient for reproducibility is quantitative models that make specific predictions. As an example, in the Bayesian framework for multisensory integration, the contribution of a sensory modality to perception is inversely proportional to its reliability (Angelaki et al., 2009; Ernst and Banks, 2002). Applied to the McGurk effect, the Bayesian framework predicts that decreasing the reliability of auditory speech should increase the perceptual weighting of visual speech, resulting in increased McGurk effect (Magnotti and Beauchamp, 2017, 2015). Consistent with this hypothesis, in a clinical population with profound hearing loss (cochlear implant users), greater susceptibility to the McGurk

effect was observed than in controls (Stropahl et al., 2017). Age-related hearing loss is ubiquitous in older adults, and, as expected, the McGurk effect increased with age in a study with participant ages ranging from 18 to 75 years (Jertberg et al., 2024). The reliability of the auditory modality can be decreased experimentally by adding auditory noise to the stimulus, and, as predicted, this increased the McGurk effect within individual participants (Fixmer and Hawkins, 1998; Stacey et al., 2020).

In summary, individual variability in the McGurk effect necessitates large sample sizes (>50 participants per group) to accurately estimate group differences (Magnotti and Beauchamp, 2018). Variability across McGurk stimuli necessitates testing with a variety of McGurk stimuli, especially stimuli that are highly effective in evoking the effect (Magnotti et al., 2024b).

Author’s note

This study was originally published with the title “Similar frequency of the McGurk effect in large samples of native Mandarin Chinese and American English speakers” (Magnotti et al., 2015). The original publication showed images (still frames) from the McGurk video stimuli. To protect the privacy of those shown in the images, in 2024 the original publication was retracted and removed from the publisher’s website, with the agreement of the authors. Because the retraction was unrelated to the study data, the study was republished in the present manuscript without the McGurk stimulus images to allow the findings to remain available to the scientific community. The present manuscript was given a new title to avoid confusion with the retracted publication.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of the Institute of Psychology of the Chinese Academy of Sciences and the Institutional Review Board of Rice University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Parental informed consent was obtained for participants under 18 years of age.

Author contributions

JM: Writing – original draft, Writing – review & editing. DB: Writing – original draft, Writing – review & editing. GF: Writing – original draft, Writing – review & editing. BZ: Writing – original draft, Writing – review & editing. WZ: Writing – original draft, Writing – review & editing. MB: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The research was funded by the National Institutes of Health (R01NS065395 to MB). The funder played no role in the design of the research.

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
- Angelaki, D. E., Gu, Y., and DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Curr. Opin. Neurobiol.* 19, 452–458. doi: 10.1016/j.conb.2009.06.008
- 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
- Bovo, R., Ciorba, A., Prosser, S., and Martini, A. (2009). The McGurk phenomenon in Italian listeners. *Acta Otorhinolaryngol. Ital.* 29, 203–208.
- Burnham, D., and Lau, S. (1998). “The effect of tonal information on auditory reliance in the McGurk effect,” in *Paper Presented at the AVSP’98 International Conference on Auditory-Visual Speech Processing*.
- Chen, Y., and Hazan, V. (2007). “Developmental factor in auditory-visual speech perception—the McGurk effect in Mandarin-Chinese and English speakers,” in *Paper Presented at the AVSP’07 International Conference on Auditory-Visual Speech Processing*.
- Chen, Y., and Hazan, V. (2009). Developmental factors and the non-native speaker effect in auditory-visual speech perception. *J. Acoust. Soc. Am.* 126, 858–865. doi: 10.1121/1.3158823
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Ferguson, C. J., and Heene, M. (2012). A vast graveyard of undead theories: publication bias and psychological science’s aversion to the null. *Perspect. Psychol. Sci.* 7, 555–561. doi: 10.1177/1745691612459059
- Fixmer, E., and Hawkins, S. (1998). “The influence of quality of information on the McGurk effect,” in *Presented at the AVSP*.
- 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, Heidelberg: Springer), 135–143. doi: 10.1007/978-3-662-13015-5_9
- Grassegger, H. (1995). “McGurk effect in German and Hungarian listeners,” in *Paper Presented at the Proceedings of the International Congress of Phonetic Sciences*, Stockholm.
- Hayashi, Y., and Sekiyama, K. (1998). “Native-foreign language effect in the McGurk effect: a test with Chinese and Japanese,” in *Paper Presented at the AVSP’98 International Conference on Auditory-Visual Speech Processing*.
- Hazan, V., Kim, J., and Chen, Y. (2010). Audiovisual perception in adverse conditions: language, speaker and listener effects. *Speech Commun.* 52, 996–1009. doi: 10.1016/j.specom.2010.05.003
- Jertberg, R. M., Begeer, S., Geurts, H. M., Chakrabarti, B., and Van der Burg, E. (2024). Age, not autism, influences multisensory integration of speech stimuli among adults in a McGurk/MacDonald paradigm. *Eur. J. Neurosci.* 59, 2979–2994. doi: 10.1111/ejn.16319
- Jiang, J., and Bernstein, L. E. (2011). Psychophysics of the McGurk and other audiovisual speech integration effects. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1193–1209. doi: 10.1037/a0023100
- Kuhberger, A., Fritz, A., and Scherndl, T. (2014). Publication bias in psychology: a diagnosis based on the correlation between effect size and sample size. *PLoS ONE* 9:e105825. doi: 10.1371/journal.pone.0105825
- Lee, H.-H., Groves, K., Ripollés, P., and Carrasco, M. (2024). Audiovisual integration in the McGurk effect is impervious to music training. *Sci. Rep.* 14:3262. doi: 10.1038/s41598-024-53593-0
- Lindstromberg, S. (2023). The winner’s curse and related perils of low statistical power – spelled out and illustrated. *Res. Methods Appl. Linguist.* 2:100059. doi: 10.1016/j.rmal.2023.100059
- Loken, E., and Gelman, A. (2017). Measurement error and the replication crisis. *Science* 355, 584–585. doi: 10.1126/science.aal3618
- 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

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1531566/full#supplementary-material>

- 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. (2017). A causal inference model explains perception of the McGurk effect and other incongruent audiovisual speech. *PLoS Comput. Biol.* 13:e1005229. doi: 10.1371/journal.pcbi.1005229
- 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
- Magnotti, J. F., Lado, A., and Beauchamp, M. S. (2024a). The noisy encoding of disparity model predicts perception of the McGurk effect in native Japanese speakers. *Front. Neurosci.* 18:1421713. doi: 10.3389/fnins.2024.1421713
- Magnotti, J. F., Lado, A., Zhang, Y., Maasø, A., Nath, A., and Beauchamp, M. S. (2024b). Repeatedly experiencing the McGurk effect induces long-lasting changes in auditory speech perception. *Commun. Psychol.* 2, 1–10. doi: 10.1038/s44271-024-00073-w
- McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0
- 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
- Peelle, J. E., and Sommers, M. S. (2015). Prediction and constraint in audiovisual speech perception. *Cortex* 68, 169–181. doi: 10.1016/j.cortex.2015.03.006
- Politzer-Ahles, S., and Pan, L. (2019). Skilled musicians are indeed subject to the McGurk effect. *R. Soc. Open Sci.* 6:181868. doi: 10.1098/rsos.181868
- Proverbio, A. M., Massetti, G., Rizzi, E., and Zani, A. (2016). Skilled musicians are not subject to the McGurk effect. *Sci. Rep.* 6:30423. doi: 10.1038/srep30423
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., and Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cereb. Cortex* 17, 1147–1153. doi: 10.1093/cercor/bhl024
- 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. (1994). McGurk effect and incompatibility: a cross-language study on auditory-visual speech perception. *Stud. Essays Behav. Sci. Philos.* 14, 29–62.
- 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 Burnham, D. (2008). Impact of language on development of auditory-visual speech perception. *Dev. Sci.* 11, 306–320. doi: 10.1111/j.1467-7687.2008.00677.x
- 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
- Sekiyama, K., and Tohkura, Y. (1993). Inter-language differences in the influence of visual cues in speech perception. *J. Phon.* 21, 427–444. doi: 10.1016/S0095-4470(19)30229-3
- Stacey, J. E., Howard, C. J., Mitra, S., and Stacey, P. C. (2020). Audio-visual integration in noise: influence of auditory and visual stimulus degradation on eye movements and perception of the McGurk effect. *Atten. Percept. Psychophys.* 82, 3544–3557. doi: 10.3758/s13414-020-02042-x
- Steege, S., Tuerlinckx, F., Gelman, A., and Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* 11, 702–712. doi: 10.1177/1745691616658637
- 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. Hum. Percept. Perform.* 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
- Stropahl, M., Schellhardt, S., and Debener, S. (2017). McGurk stimuli for the investigation of multisensory integration in cochlear implant users: the Oldenburg audio visual speech stimuli (OLAVS). *Psychon. Bull. Rev.* 24, 863–872. doi: 10.3758/s13423-016-1148-9
- Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215. doi: 10.1121/1.1907309
- Tiippana, K., Ujii, Y., Peromaa, T., and Takahashi, K. (2023). Investigation of cross-language stimulus-dependent effects on the McGurk effect with Finnish and Japanese speakers and listeners. *Brain Sci.* 13:1198. doi: 10.3390/brainsci13081198
- 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
- van Wassenhove, V., Grant, K. W., and Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1181–1186. doi: 10.1073/pnas.0408949102