Research article

Neural evidence accounting for interindividual variability of the McGurk illusion

Antoine J. Shahin

Department of Cognitive and Information Sciences, University of California, Merced, CA 95343, United States and Center for Mind and Brain, University of California, Davis, CA 95618, United States

ARTICLE INFO

ABSTRACT

The McGurk illusion is experienced to various degrees among the general population. Previous studies have implicated the left superior temporal sulcus (STS) and auditory cortex (AC) as regions associated with this interindividual variability. We sought to further investigate the neurophysiology underlying this variability using a variant of the McGurk illusion design. Electroencephalography (EEG) was recorded while human subjects were presented with videos of a speaker uttering the consonant-vowels (CVs) /ba/ and /fa/, which were mixed and matched with audio of /ba/ and /fa/ to produce congruent and incongruent conditions. Subjects were also presented with unimodal stimuli of silent videos and audios of the CVs. They responded to whether they heard (or saw in the silent condition) /ba/ or /fa/. An illusion during the incongruent conditions was deemed successful when individuals heard the syllable conveyed by mouth movements. We hypothesized that individuals who experience the illusion more strongly should exhibit more robust desynchronization of alpha (7–12 Hz) at fronto-central and temporal sites, emphasizing more engagement of neural generators at the AC and STS. We found, however, that compared to weaker illusion perceivers, stronger illusion perceivers exhibited greater alpha synchronization at fronto-central and posterior temporal sites, which is consistent with inhibition of auditory representations. These findings suggest that stronger McGurk illusion perceivers possess more robust cross-modal sensory gating mechanisms whereby phonetic representations not conveyed by the visual system are inhibited, and in turn reinforcing perception of the visually targeted phonemes.

1. Introduction

The purpose of this study was to determine the neurophysiological factors that underlie interindividual variability of the McGurk illusion [1]. Several studies have examined the basis of this variability [2–6]. In a large sample study, Basu Mallick et al. [3] found that the robustness of the McGurk illusion not only varied across subjects but also across recordings (talkers) and tasks. They concluded that the McGurk illusion cannot be explained by a normal distribution, and as a result, assessing AV integration among populations using the McGurk illusion design with small samples can be misleading [7,8].

The neurophysiological underpinning of interindividual variability of the McGurk illusion has been addressed by Nath and Beauchamp [2]. They found that individuals who experienced the illusion more robustly exhibited greater fMRI activation in the left superior temporal sulcus (STS) for the illusory percepts than the weaker illusion perceivers. Their finding is not surprising given that the left STS is a multisensory integration site [9–11]. A recent EEG study [12] implicated the auditory cortex (AC) as a site indexing interindividual variability. Shahin et al.

[12] employed a variant of the classic McGurk illusion design, whereby the visually conveyed syllable is heard as opposed to a fused percept. In Shahin et al. individuals watched and listened to videos and audios of the consonant vowel (CV) syllables /ba/ and /fa/, which were mixed and matched to produce congruent and incongruent AV stimulus-pairs. An illusion was deemed successful when individuals heard the syllable conveyed by mouth movements. They found that stronger illusion perceivers exhibited distinct N1 auditory evoked potential (AEP) amplitudes for illusory 'ba' and 'fa'. The weaker illusion perceivers exhibited similar N1 amplitudes evoked by each of these two percepts. Together, Nath and Beauchamp [2] and Shahin et al. [12] suggest that distinct neurophysiological attributes at the AC and STS contribute to individual differences in the McGurk illusion.

In this study, we further examined the neurophysiology underlying interindividual variability by probing the dynamics of alpha band activity (7–12 Hz). Alpha is a well-established neurophysiological marker of sensory gating. Alpha enhancement (synchronization) marks disengagement (inhibition) of task-specific neural processes, while alpha suppression (desynchronization) indicates engagement (excitation) of

E-mail address: ashahin@ucmerced.edu.

task-specific neural processes [13–15]. Mazaheri et al. [13] utilized AV tasks and showed that when the task was visual-related, suppression of alpha was observed over occipital sites, while enhancement of alpha was observed over temporal sites, indicating engagement of the visual cortex and disengagement of the auditory cortex, respectively. The reverse was observed with an auditory-related task. Mazaheri et al. posited that the observed alpha dynamics heighten sensory efficiency of one modality in part by inhibiting activity in the irrelevant modality.

Accordingly, we hypothesized that greater alpha desynchronization induced by the illusory percepts should be exhibited at fronto-central and temporal sites in stronger versus weaker illusion perceivers emphasizing the more robust engagement of the AC and STS in the stronger illusion perceivers.

2. Materials and methods

The current study used the publicly available data of Shahin et al. [12]. A detailed account of the methods can be found in Shahin et al. [12] and is discussed briefly below.

2.1. Subjects

Nineteen subjects participated in this study (mean age of 20.9 years with 1.8 SD). All subjects reported normal/corrected vision, normal hearing, and no histories of neurological or language disorders. Fourteen subjects were native English speakers, and 5 were fluent English speakers (Native languages, Chinese and Spanish). Fluent English speakers spoke English for a minimum of 10 years prior to participation. All subjects signed a written informed consent consistent with the guidelines of the University of California, Davis Institutional Review Board. Subjects were monetarily compensated for their participation.

2.2. Stimuli

Six acoustic CV syllables (3 /ba/, 3 /fa/) were combined with two silent videos of a speaker uttering /ba/ and /fa/ in congruent and incongruent manners. The experimental design also included unimodal tokens of the stimuli (auditory-only and visual-only). The /ba/ CVs were approximately 425 ms in duration, and the /fa/ CVs were always 50 ms longer due to the initial phone /f/. All acoustic CVs were matched in root mean square power. To circumvent confounding acoustic contributions to the EEG results, EEG data were averaged across the three acoustic tokens of each CV type.

2.3. Procedure

The behavioral responses and EEG data (64-channel BioSemi Active Two system, 10–20 Ag-AgCl electrode system, with Common Mode Sense and Driven Right Leg passive electrodes serving as grounds, A/D rate 1024 Hz) were acquired while participants sat 85 cm in front of a 24-in. Dell monitor. In an event-related mixed design, subjects were presented with combinations of the audio and videos of the CVs as well as unimodal auditory-only and visual-only stimuli. Following each trial presentation, participants made judgments about the syllable they heard, or in the visual-only case, about what they saw. The experiment consisted of six ˜10-min blocks, with 204 trials per block. Trials had a duration of 2700 ms plus a variable jitter of up to 500 ms. The stimulus combinations were as follows: In the auditory-only condition, subjects listened to audios of the CVs while watching a still image of the speaker. In the visual-only condition, subjects watched a silent video of the speaker uttering either CV. In the congruent condition, subjects watched and listened to congruent videos and audios of the speaker uttering /ba/ or /fa/. In the incongruent condition, subjects watched and listened to incongruent videos and audios of the CVs. Subjects indicated whether they heard /ba/ or /fa/ (or saw /ba/ or /fa/ in the visual-only
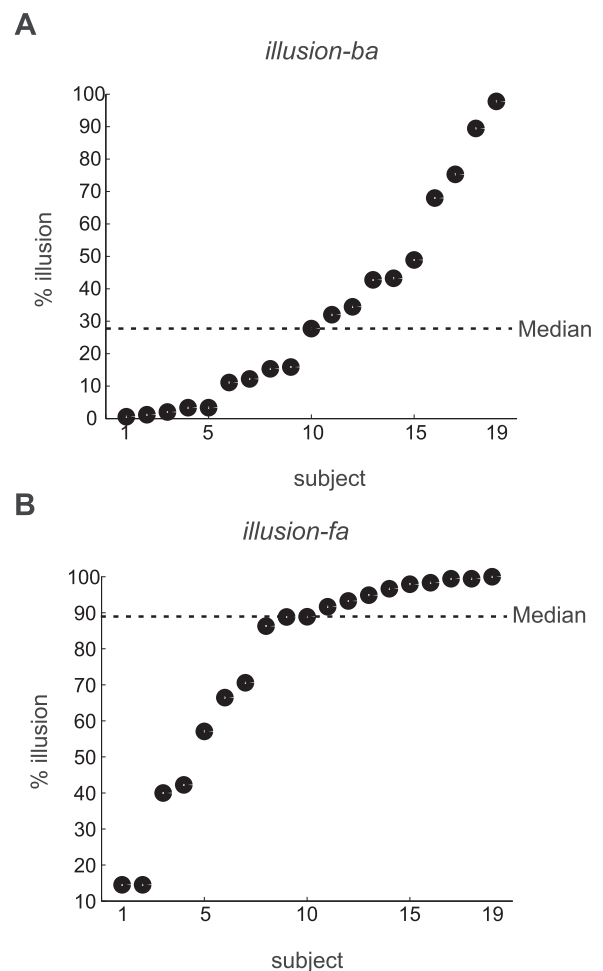

Fig. 1. Individual percentages of illusory perception for A) *illusion-ba* (video /ba/, audio /fa/, heard 'ba') and B) *illusion-fa* (video /fa/, audio /ba/, heard 'fa').

condition) by pressing a keyboard button using their left middle or index finger, respectively. In the incongruent condition, an illusion was deemed successful when subjects heard the syllable conveyed by the speaker's mouth, not the audio. This led to two possible illusory percepts, *illusion-fa* (video /fa/, audio /ba/, heard 'fa') and *illusion-ba* (video /ba/, audio /fa/, heard 'ba').

2.4. Data analysis

Preprocessing was similar to the AEP analysis in Shahin et al. [12], except that the current data was epoched from -1250 to 1250 ms around sound onset (as opposed to -100 to 500 ms). Additionally, the current data were not filtered prior to conducting time-frequency analysis. Following preprocessing, EEG files were epoched according to each percept type (auditory-only, visual-only, and congruent 'ba' and 'fa', and *illusion-ba* and *illusion-fa*) and subjected to time-frequency analysis.

Conducting time-frequency analysis on stronger versus weaker illusion perceivers for the illusory percepts was not feasible since stronger illusion perceivers had more illusory trials than weaker illusion perceivers (Fig. 1). To circumvent this issue, time-frequency analysis was conducted on the incongruent trials of 'ba' and 'fa' collapsing across the *illusion* trials and trials when the illusion failed (*illusion-failure*). Subsequently, the two groups had similar overall number of trials per incongruent condition, with the exception that stronger illusion perceivers had more *illusion* trials than *illusion-failure* trials within the

condition. Thus, probing the illusory effect remained intact while maintaining similar EEG signal-to-noise ratios.

Time-frequency analysis was conducted by generating event-related spectral perturbation (ERSP) spectrograms for each channel, individual and percept type using the timef.m function of the EEGLAB toolbox. The analysis was based on a sinusoidal wavelet-based discrete Fourier transform (DFT) of the time-domain signal (500-ms Hanning window, 10 ms and 1 Hz steps) by employing two cycles at the lowest frequency and increasing number of cycles linearly with frequency. The baseline for adjusting ERSP activity was the pre-acoustic activity from -1250 to -500 ms.

2.5. Statistical analysis

Differences between groups were assessed using nonparametric permutation tests to compare spectral power differences at each time and frequency bin of the channel spectrograms for each percept type. All channels were included in the resampling to account for multiple comparisons (see Chau et al. [16] for more details of this method). Significance was determined based on null distributions derived from all samples of a 500-ms pre-acoustic stimulus period (-1000 to -400 ms, 60 samples per channel, 3840 samples across all channels) of the maximum values obtained in repeated resampling of the data (2000 permutations, p = 0.01). Post hoc analysis were conducted using Analysis of Variance (ANOVA) and Tukey's Honest Significant Difference Test.

3. Results

3.1. Behavior

Accuracy for the auditory-only, visual-only and congruent percept types were at ceiling. However, perception of either illusion was highly variable. Fig. 1 demonstrates the percentage of *illusion-ba* and *illusion-fa* trials that were experienced across individuals. First, note the large interindividual variability, especially for *illusion-ba*. Second, *illusion-fa* was much more potent (median ˜ 88%) compared to *illusion-ba* (median ˜ 28%). This variability is also seen in the classical McGurk design (e.g., video /ga/, audio /ba/, heard as 'da') [3]. One critical observation is that within the same population (19 subjects), not all subjects who experienced one illusion also experienced the other illusion with the same potency. Five subjects who were included in the stronger *illusion-ba* group were also included in the stronger *illusion-fa* group. While the data show a strong overlap among individuals for the two illusion types, they also suggest that individual variability for visually mediated illusions, such as the McGurk illusion, is stimulus specific.

To further confirm stimulus specificity of the McGurk illusion, Spearman rank correlation was conducted on individual illusion percentages across the two illusion (stimulus) types. The correlation yielded a Spearman's rank correlation coefficient of 0.34 at p = 0.14 (2-tailed). This non-significant result supports the conclusion that the strength of the McGurk illusion is not preserved across different stimuli within individuals.

3.2. EEG

Subjects were divided into top and bottom 9 illusion perceivers for each illusory type (*illusion-ba, illusion-fa see Introduction*). The median subject was not included in either group.

Fig. 2AB presents the spectrotemporal activity distinguishing stronger versus weaker *illusion-ba* perceivers when they perceived *illusion-ba* and *illusion-fa* (undifferenced spectrograms are shown in Supplementary Fig. 1A). The two types of spectrograms at channel Fz represent the raw group difference as well as the permuted group difference (p = 0.01). In the permuted spectrograms, only activity that reached significance is displayed, while all other activity is zeroed

(masked by green color). The Figures also show the topographies of the group difference significant activity for alpha (7–12 Hz). The differences reached significance at other channels as well, but for simplicity only activity at Fz is shown. Note from the Fz spectrograms that larger and significant early (< 500 ms) alpha activity distinguished stronger versus weaker *illusion-ba* perceivers, and this activity was stronger for *illusion-ba* than *illusion-fa*. This result suggests that stronger *illusion-ba* perceivers exhibited specificity to the visually conveyed CV /ba/.

To further confirm this specificity, we conducted a post hoc analysis on the individual alpha activity. Individual alpha activity was isolated as follows: 1) For each channel spectrogram pixels that did not survive the permutation test during the *illusion-ba* and *illusion-fa* group contrasts (e.g., Fig. 2A and B masked, green color) were summed to create a mask; 2). The mask was then applied to the data of each individual to isolate the alpha power at each channel for *illusion-ba* and *illusion-fa* percept types; 3) The isolated alpha activity was averaged across frequency (7–12 Hz), time points and channels to yield one value for each individual and percept type. AN ANOVA was then conducted using these values with the independent variable being group (stronger versus weaker *illusion-ba* perceivers) and the dependent variable being percept type (*illusion-ba, illusion-fa*). The purpose of the ANOVA was to show an interaction with greater group differences for *illusion-ba* than *illusion-fa* (Fig. 2C), which turned out to be the case ($F_{(1, 16)}$ = 8.6, p = 0.01; η_p^2 = 0.35). A follow up Tukey test showed that the interaction was due to a larger alpha occurring for *illusion-ba* in the stronger versus weaker *illusion-ba* perceivers (p = 0.037). The groups did not exhibit differences for *illusion-fa* (p = 0.54).

Similarly, we examined the illusion efficacy in stronger versus weaker *illusion-fa* perceivers. The subjects were rearranged within groups to segregate the top and bottom 9 *illusion-fa* perceivers. Fig. 2DE presents the spectrotemporal activity distinguishing stronger versus weaker *illusion-fa* perceivers when they perceived *illusion-ba* and *illusion-fa* (undifferenced spectrograms are shown in Supplementary Fig. B). Note that from the Fz spectrograms, larger and significant early (-100 to 500 ms) alpha activity distinguished stronger versus weaker *illusion-fa* perceivers, and this activity was stronger for *illusion-fa* than *illusion-ba*. Indeed, there were no differences in activity between the groups for *illusion-ba* at Fz. Again, this means that for the stronger *illusion-fa* perceivers, the illusion was CV-specific. Indeed, an ANOVA (Fig. 2F) showed an interaction between group and illusory type ($F_{(1, 16)}$ = 6.3, p = 0.023; η_p^2 = 0.28). A follow up Tukey test showed that the interaction was due to a larger alpha occurring for *illusion-fa* in the stronger versus weaker *illusion-fa* perceivers (p = 0.04). The groups did not exhibit differences for *illusion-ba* (p = 0.4).

Qualitatively similar group alpha patterns were found for the visual-only (silent video) and congruent conditions (not shown). For these conditions however post hoc Tukey tests did not reach, but approached, significance (p = 0.06 to 0.1). No group alpha differences were found for the auditory-only condition (not shown). Accordingly, we may conclude that the group alpha pattern is not specific to the illusion, albeit it was stronger in the illusory conditions. Furthermore, the absence of sound in the visual-only condition suggest that interindividual variability is driven by the specificity to the visually conveyed stimulus.

3.3. Alternative regression analysis

As a supplementary analysis (see Supplementary Fig. 2), regression analysis was conducted between the individual alpha power (mean across all channels) and the individual illusory percentages using all 19 subjects. This analysis is not ideal given that in several individuals illusory perception reached ceiling values. Nonetheless, using the entire mean of the post stimulus alpha power did not yield significant correlation. However, using the group masked values, as in previous section, produced significant correlation between alpha power and illusory perception for *illusion-ba* (R = 0.56, p = 0.01) and *illusion-fa* (R = 0.45, p = 0.06).
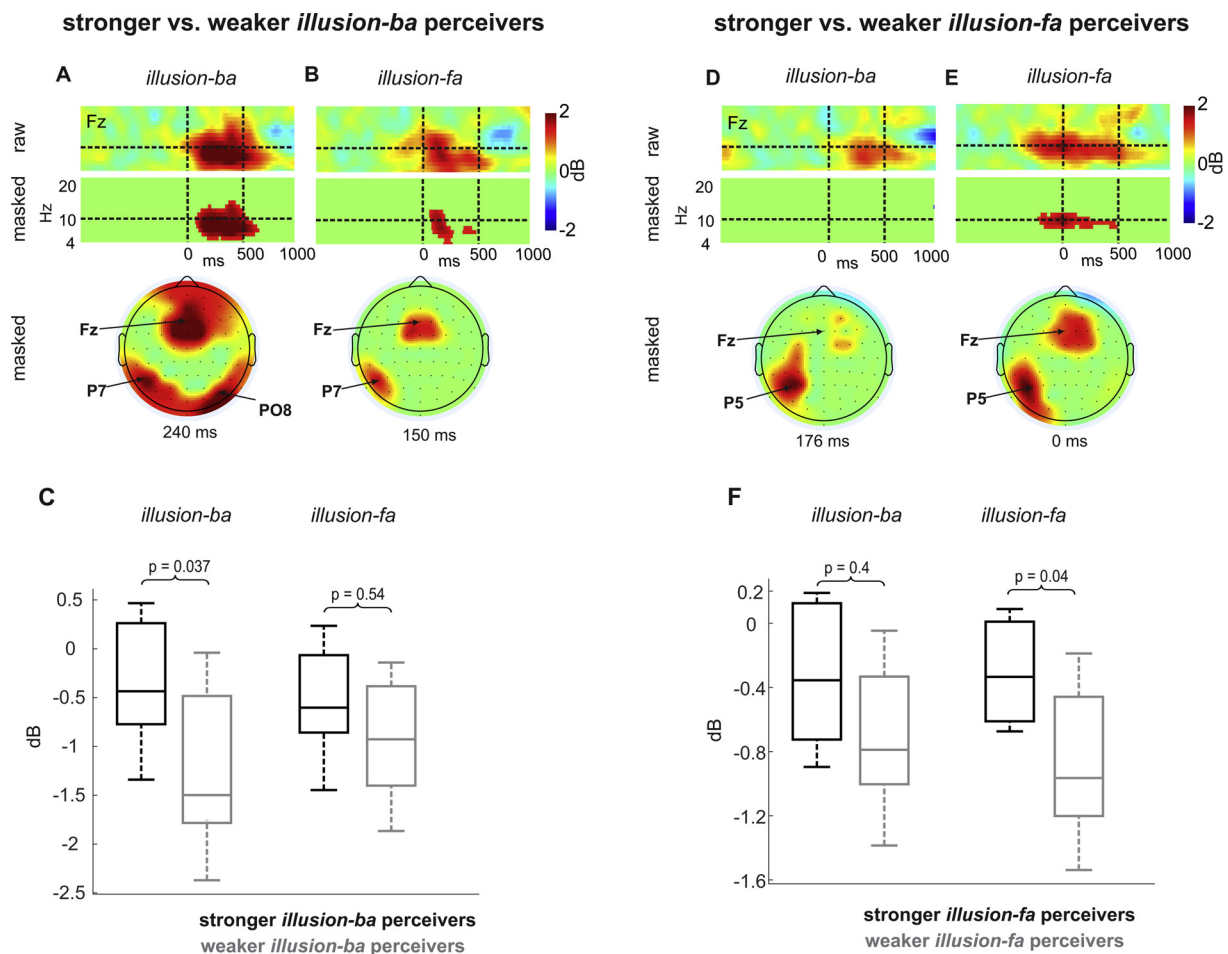
Fig. 2. Time-frequency spectrograms at channel Fz revealing oscillatory activity group differences (stronger vs. weaker *illusion-ba* perceivers) for *illusion-ba* (A) and *illusion-fa* (B). The top spectrograms represent the group raw differences. The bottom spectrograms represent the permuted differences (nonsignificant activity is masked by green color). The figures also show the topographies at latencies in which the groups exhibited maximum differences. C) Boxplots of data points that survived the permutation test for *illusion-ba* and *illusion-fa* of the stronger and weaker *illusion-ba* perceivers. D–F same as A–C, except the contrast is between stronger vs. weaker *illusion-fa* perceivers.

Finally, the effect due to inclusion of non-native English speakers was assessed by removing these subjects and rerunning the analyses. The results did not qualitatively change. In the original analysis, the 5 non-native subjects were divided across the two groups as follows: 1) Two were included in the stronger *illusion-fa* perceiver group and 3 in the weaker *illusion-fa* perceiver group; 2) One was included in the stronger *illusion-ba* perceiver group and 3 in the weaker *illusion-ba* perceiver group. The 5th subject was the median subject.

4. Discussion

This study provides evidence that interindividual variability observed with the McGurk illusion is reflected by stimulus-specific alpha activity. In contrast to our hypothesis, the results emphasize an inhibitory mechanism (greater alpha) as an attribute distinguishing stronger versus weaker illusion perceivers. Furthermore, the results showing that interindividual variability varies with stimulus type cautions against drawing conclusions about the efficacy of AV integration using one type of incongruent syllable-pairs.

It is not possible to assess the neural origins of the alpha activity distinguishing the two groups with certainty. The fronto-central and temporo-parietal topography distinguishing the two groups is consistent with auditory generators, because the poles of generators in the primary AC point in those directions [17]. A caveat of this deduction is that the posterior topography for auditory sources is usually more

widespread, as is the case in Fig. 2A topography, and more inferior (i.e., not reflected by a strong temporo-parietal focus in the P5/P7 channels). Topographies observed in Fig. 2BDE may be attributed to sources that are more closely beneath channels P5/P7 than what we expect from a source in the primary AC. This supports a case for sources in the posterior STS, where the sources have a similar orientation as the sources in the primary AC but are more posteriorly superficial (closer to the scalp). Alternatively, although less likely, the alpha topography may reflect visual sources. Typically, visual topography is represented by bifocal activity at the parieto-occipital channels (e.g., PO3/PO4) and frontal channels [18]. However, there are instances in which visual activity exhibits a similar topography to the current alpha topography. For example, evoked potentials of conscious minus unconscious face processing yield an N170 topography (if we disregard polarity) similar to the current one.

If indeed the current alpha pattern reflects auditory activity, we may posit that enhanced alpha synchronization in stronger illusion perceivers reflects cross-modal inhibition of irrelevant auditory representations. This sensory-gating process is consistent with the role of visual networks in constraining auditory processing [19,20]. The influence of vision on auditory neurophysiology has been consistently shown to be suppressive [21–23]. The consensus is that visual context renders some auditory information redundant, and therefore the auditory response (e.g., auditory evoked potentials or AEPs) is reduced. Some of these redundant auditory cues include sound onsets [23]. That

is, as the phonetic identity is identified by the visual percepts, acoustic cues like onsets become irrelevant and their representations are suppressed. The redundant cues may also include phonetic and lexical representations that are not conveyed by visual speech [19]. This deduction also follows from studies on memory retention, which show that as individuals retain more items in short-term memory, they exhibit greater alpha [14,24]. This alpha behavior is thought to reflect the inhibition of redundant sensory information that may impede the retention of additional items in memory. Alternatively, if the observed alpha in stronger illusion perceivers reflects visual sources (face processing), then the alpha enhancement may index inhibition of facial cues that are not relevant to the task (i.e., facial features other than the mouth).

Previous accounts suggest that a basis for interindividual variability of the McGurk illusion is the perceptual instability of the McGurk percept. For example, external factors, such as differences in tasks and recordings, can alter the robustness of the McGurk illusion [3]. This is noteworthy because visual influence on speech comprehension is mainly manifested in adverse listening situations – when the acoustic stimulus is degraded or unstable. This is consistent with the "inverse effectiveness" principle [25], which states that as the auditory signal becomes more degraded the benefit of visual cues to audition is increased. This principle has been substantiated in several studies [26–28], including neuroimaging data showing that in an AV task, the STS/G (G for gyrus) is more activated as the SNR decreases [29]. Results by Ross et al. [30], however, showed that visual benefit to speech comprehension is only inversely proportional to the acoustic SNR up to a point (intermediate SNR). One may speculate that individuals who have acquired stronger cross-modal connections and thus substantially benefit from visual cues in noisy situations, experience stronger cross-modal inhibition of redundant auditory representations even in the absence of noise. This automatic inhibitory mechanism lead to a more robust visual streamlining of auditory representations and a more robust McGurk illusion in these individuals.

4.1. Caveats and considerations

There are few caveats of the current study. First, even though the two groups were drawn from a population with similar characteristics (healthy, young college students), the results may have been compromised by the low number of individuals per group (n = 9). Second, the conclusion that stronger illusion perceivers have greater alpha and hence more robust inhibitory mechanisms is not a unique. Because alpha power in stronger illusion perceivers was relative to the alpha power of the weaker illusion perceivers, the alpha group-differences may also be attributed to greater cross-modal excitatory mechanisms in the weaker illusion perceivers.

5. Conclusions

The current results suggest that interindividual variability of the McGurk illusion may be a byproduct of heightened cross-modal inhibition of auditory information that is not conveyed by the visual system in stronger illusion perceivers. The neurophysiological locus of the variability may be the auditory cortex, STS and/or even the visual system. Subsequent studies using more spatially superior methods (e.g., fMRI) are better suited to pinpoint the source.

Data accessibility

The original raw data are available at https://data.mendeley.com/datasets/yydw84284f/1.

Funding

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.neulet.2019.134322.

References

[1] H. McGurk, J. Macdonald, Hearing lips and seeing voices, Nature 264 (1976) 691–811, https://doi.org/10.1038/264746a0.

[2] A.R. Nath, M.S. Beauchamp, A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion, Neuroimage 59 (2012) 781–787, https://doi.org/10.1016/j.neuroimage.2011.07.024.

[3] D. Basu Mallick, J.F. Magnotti, M.S. Beauchamp, Variability and stability in the McGurk effect: contributions of participants, stimuli, time, and response type, Psychon. Bull. Rev. 22 (2015) 1299–1307, https://doi.org/10.3758/s13423-015-0817-4.

[4] V.A. Brown, M. Hedayati, A. Zanger, S. Mayn, L. Ray, N. Dillman-Hasso, J.F. Strand, What accounts for individual differences in susceptibility to the McGurk effect? PLoS One 13 (2018) e0207160, , https://doi.org/10.1371/journal.pone.0207160.

[5] N.T. Abbott, A.J. Shahin, Cross-modal phonetic encoding facilitates the McGurk illusion and phonemic restoration, J. Neurophysiol. 120 (2018) 2988–3000.

[6] J.-L. Schwartz, A reanalysis of McGurk data suggests that audiovisual fusion in speech perception is subject-dependent, J. Acoust. Soc. Am. 127 (2010) 1584–1594, https://doi.org/10.1121/1.3293001.

[7] J.F. Magnotti, M.S. Beauchamp, Published estimates of group differences in multisensory integration are inflated, PLoS One 13 (2018) 1–13, https://doi.org/10.1371/journal.pone.0202908.

[8] J. Magnotti, D. Basu Mallick, G. Feng, B. Zhou, W. Zhou, M. Beauchamp, Similar frequency of the McGurk effect in large samples of native Mandarin Chinese and American English speakers, Exp. Brain Res. 233 (2015) 2581–2586.

[9] G.A. Calvert, R. Campbell, M.J. Brammer, Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex, Curr. Biol. 10 (2000) 649–657, https://doi.org/10.1016/S0960-9822(00)00513-3.

[10] M.S. Beauchamp, A. Nath, S. Pasalar, fMRI-guided TMS reveals that the STS is a cortical locus of the McGurk effect, J. Neurosci. 30 (2010) 2414–2417, https://doi.org/10.1523/JNEUROSCI.4865-09.2010.fMRI-guided.

[11] L.M. Miller, M. D'Esposito, Perceptual fusion and stimulus coincidence in the cross-modal integration of speech, J. Neurosci. 25 (2005) 5884–5893, https://doi.org/10.1523/JNEUROSCI.0896-05.2005.

[12] A.J. Shahin, K.C. Backer, L.D. Rosenblum, J.R. Kerlin, Neural mechanisms underlying cross-modal phonetic encoding, J. Neurosci. 38 (2018) 1835–1849.

[13] A. Mazaheri, M.R. van Schouwenburg, A. Dimitrijevic, D. Denys, R. Cools, O. Jensen, Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities, Neuroimage 87 (2014) 356–362, https://doi.org/10.1016/j.neuroimage.2013.10.052.

[14] O. Jensen, A. Mazaheri, Shaping functional architecture by oscillatory alpha activity: gating by inhibition, Front. Hum. Neurosci. 4 (2010) 1–8, https://doi.org/10.3389/fnhum.2010.00186.

[15] N. Weisz, T. Hartmann, N. Müller, I. Lorenz, J. Obleser, Alpha rhythms in audition: cognitive and clinical perspectives, Front. Psychol. 2 (2011) 1–15, https://doi.org/10.3389/fpsyg.2011.00073.

[16] W. Chau, A.R. McIntosh, S.E. Robinson, M. Schulz, C. Pantev, Improving permutation test power for group analysis of spatially filtered MEG data, Neuroimage 23 (2004) 983–996, https://doi.org/10.1016/j.neuroimage.2004.07.007.

[17] M. Scherg, J. Vajsar, T.W. Picton, A source analysis of the late human auditory evoked potentials, J. Cogn. Neurosci. 1 (1989) 336–355, https://doi.org/10.1162/jocn.1989.1.4.336.

[18] N. Pratt, A. Willoughby, D. Swick, Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence, Front. Hum. Neurosci. 5 (2011) 1–9, https://doi.org/10.3389/fnhum.2011.00057.

[19] J.E. Peelle, M.S. Sommers, Prediction and constraint in audiovisual speech perception, Cortex 68 (2015) 169–181, https://doi.org/10.1016/j.cortex.2015.03.006.

[20] J. Bhat, L.M. Miller, M.A. Pitt, A.J. Shahin, Putative mechanisms mediating tolerance for audiovisual stimulus onset asynchrony, J. Neurophysiol. 113 (2015) 1437–1450, https://doi.org/10.1152/jn.00200.2014.

[21] V. Van Wassenhove, K.W. Grant, D. Poeppel, M. Halle, Visual speech speeds up the neural processing of auditory speech, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 1181–1186, https://doi.org/10.1073/pnas.0408949102.

[22] J. Besle, A. Fort, C. Delpuech, M.H. Giard, Bimodal speech: early suppressive visual effects in human auditory cortex, Eur. J. Neurosci. 20 (2004) 2225–2234, https://doi.org/10.1111/j.1460-9568.2004.03670.x.

[23] H. Shatzer, S. Shen, J.R. Kerlin, M.A. Pitt, A.J. Shahin, Neurophysiology underlying influence of stimulus reliability on audiovisual integration, Eur. J. Neurosci. 48 (2018) 2836–2848, https://doi.org/10.1111/ejn.13843.

[24] H. Van Dijk, I.L.C. Nieuwenhuis, O. Jensen, Left temporal alpha band activity increases during working memory retention of pitches, Eur. J. Neurosci. 31 (2010) 1701–1707, https://doi.org/10.1111/j.1460-9568.2010.07227.x.

[25] M.A. Meredith, B.E. Stein, Spatial factors determine the activity of muitisensory neurons in cat superior colliculus, Brain Res. 5 (1986) 350–354.

[26] N. Erber, Auditory-visual perception in speech, J. Speech Hear. Dis. 40 (1975) 481–492.

[27] J.J. O'Neill, Contributions of the visual components of oral symbols to speech comprehension, J. Speech. Hear. Dis. 19 (1954) 429.

[28] W.H. Sumby, I. Pollack, Visual contribution to speech intelligibility in noise, J. Acoust. Soc. Am. 26 (1954) 212–215, https://doi.org/10.1121/1.1907309.

[29] D.E. Callan, C.A. Je, A. Jones, K. Munhall, A.M. Callan, C. Kroos, E. Vatikiotis-bateson, Neural processes underlying perceptual enhancement by visual speech gestures, Neuroreport 14 (2003) 213–218, https://doi.org/10.1097/01.wnr.0000095492.38740.8f.

[30] L.A. Ross, D. Saint-amour, V.M. Leavitt, D.C. Javitt, J.J. Foxe, Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments, Cereb. Cortex 17 (2007) 1147–1153, https://doi.org/10.1093/cercor/bhl024.