Behavioral/Cognitive

Inhibitory TMS over Visual Area V5/MT Disrupts Visual Speech Recognition

Lisa Jeschke,[1] Brian Mathias,[2] and Katharina von Kriegstein[1]

[1]Chair of Cognitive and Clinical Neuroscience, Faculty of Psychology, Technische Universität Dresden, 01069 Dresden, Germany and [2]School of Psychology, University of Aberdeen, Aberdeen AB243FX, United Kingdom

During face-to-face communication, the perception and recognition of facial movements can facilitate individuals' understanding of what is said. Facial movements are a form of complex biological motion. Separate neural pathways are thought to processing (1) simple, nonbiological motion with an obligatory waypoint in the motion-sensitive visual middle temporal area (V5/MT); and (2) complex biological motion. Here, we present findings that challenge this dichotomy. Neuronavigated offline transcranial magnetic stimulation (TMS) over V5/MT on 24 participants (17 females and 7 males) led to increased response times in the recognition of simple, nonbiological motion as well as visual speech recognition compared with TMS over the vertex, an active control region. TMS of area V5/MT also reduced practice effects on response times, that are typically observed in both visual speech and motion recognition tasks over time. Our findings provide the first indication that area V5/MT causally influences the recognition of visual speech.

Key words: motion recognition; TMS; V5/MT; visual speech

Significance Statement

In everyday face-to-face communication, speech comprehension is often facilitated by viewing a speaker's facial movements. Several brain areas contribute to the recognition of visual speech. One area of interest is the motion-sensitive visual medial temporal area (V5/MT), which has been associated with the perception of simple, nonbiological motion such as moving dots, as well as more complex, biological motion such as visual speech. Here, we demonstrate using noninvasive brain stimulation that area V5/MT is causally relevant in recognizing visual speech. This finding provides new insights into the neural mechanisms that support the perception of human communication signals, which will help guide future research in typically developed individuals and populations with communication difficulties.

Introduction

Humans gain information about what a speaker is saying from their visible articulator movements. In everyday life, visual speech recognition, also called speech-reading or lip-reading (Campbell, 2008), is especially important in face-to-face conversations with high levels of auditory background noise, as it can facilitate comprehension in noisy environments or even enable it when auditory signals are absent (Sumby and Pollack, 1954; Grant and Seitz, 2000; Chandrasekaran and Ghazanfar, 2011). Visual speech recognition is also beneficial

for speech comprehension in hearing impairment (Auer and Bernstein, 2007; Rosemann et al., 2020) and contributes to neurodevelopmental disorders (Schelinski et al., 2014; van Laarhoven et al., 2016).

There has been a great deal of debate surrounding how and where in the brain visual speech recognition is conducted. Previous research has demonstrated that visual speech recognition relies on the perception of dynamic and configural facial features linked to dorsal and ventral visual neural pathways, respectively (Campbell et al., 2001; O'Toole et al., 2002; Bernstein and Liebenthal, 2014). One cortical region consistently implicated in visual speech processing is the posterior superior temporal sulcus (pSTS) and posterior superior temporal gyrus (pSTG), including a most sensitive subregion coined the temporal visual speech area (TVSA; Schultz et al., 2005; Bernstein et al., 2011; Riedel et al., 2015). Further, functional MRI studies showed that the motion-sensitive visual middle temporal area (V5/MT) responds to biological motion, including movements pertaining to visual speech (Calvert and Campbell, 2003; Paulesu et al., 2003; Peuskens et al., 2005; Borowiak et al., 2018).

In addition to complex, biological motion, there is strong evidence that V5/MT is involved in perceiving nonbiological simple

motion such as linearly or circularly moving shapes (Grèzes et al., 2001; Antal et al., 2004; Silvanto et al., 2006; Koivisto et al., 2010; Cai et al., 2014). While many studies using transcranial magnetic stimulation (TMS) have demonstrated that an inhibition of V5/MT leads to a reduction in individuals' recognition of nonbiological simple motion (Beckers and Zeki, 1995; Laycock et al., 2007; McKeefry et al., 2008; Grasso et al., 2018), its causal contribution to visual speech recognition remains unclear. Separate mechanisms for biological versus nonbiological motion have been suggested, and it remains unresolved whether V5/MT contributes to both (Miller et al., 2018). One approach argues that V5/MT might be an obligatory waypoint for information pertaining only to nonbiological, but not biological, motion such as visual speech. For instance, Mather et al. (2016) demonstrated that inhibitory TMS over V5/MT resulted in decreased nonbiological motion discrimination accuracy compared with the stimulation of a control site, but had no effects on the identification of point-light walkers. However, no study so far has investigated the effects of TMS over V5/MT on visual speech recognition in particular, despite fMRI findings showing a V5/MT response to visual speech, but not to walking stimuli (Santi et al., 2003). There may in fact be separate processing streams that subserve different types or characteristics of biological motion.

In contrast to the previous approach, O'Toole et al. (2002) argue that facial motion information transits through V5/MT and is then passed to more specialized regions such as the TVSA. V5/MT is also functionally and structurally connected to both the TVSA and visual ventral stream regions (Ethofer et al., 2011; Furl, 2015; Bernstein et al., 2018). Moreover, a case report on a patient with V5/MT lesions showed impaired visual speech recognition but intact biological motion recognition (Campbell et al., 1997). Note, however, that the lesion extended also into the left STS and left STG; hence, it is unclear whether this impairment was caused by the V5/MT and/or the STS/STG lesion. In summary, it is an open question whether V5/MT is causally involved in processing visual speech information.

The objective of the present study was to investigate the functional relevance of area V5/MT for visual speech recognition. We applied offline inhibitory TMS over V5/MT and an active control brain region and examined participants' subsequent visual speech recognition performance, as well as their recognition of simple nonbiological motion.

Materials and Methods

The Ethics Committee at Technische Universität Dresden approved the study (Aktenzeichen EK1701042019). We conducted the experiments at the Neuroimaging Center of Technische Universität Dresden. All participants gave their written informed consent and were reimbursed for their participation.

Participants

Twenty-four right-handed, native German speakers between the ages of 19 and 36 years (mean = 25.2, SD = 5.49) were included in the data analysis (7 males, 17 females). Of the 27 participants who were recruited for the study, 3 participants were excluded as they experienced mild head or jaw pain during stimulation. None of the participants reported a history of neurologic or psychiatric disorders, or any contraindications for MRI or TMS. During the recruitment process, all participants underwent an online screening for clinical conditions that are associated with an altered visual speech recognition, including developmental dyslexia and autism spectrum disorder (Schelinski et al., 2014; van Laarhoven et al., 2016). The screening contained an inquiry about an existing diagnosis of a developmental dyslexia, personally or in relatives, and the German version of the autism-spectrum quotient (AQ; mean AQ = 15.03, SD = 5.87; Freitag et

al., 2007). Before the fMRI testing, we applied the Freiburg Visual Acuity Test (Bach, 1996). All participants passed this test (normal or corrected decimal visual acuity > 0.8).

Design

The study consisted of four separate sessions for each participant (Fig. 1a). In the first session, we acquired structural and functional MRI data. The MRI data were used to determine each participant's individual V5/MT coordinates in both hemispheres. The second and third sessions consisted of an offline TMS protocol, which included the stimulation of bilateral V5/MT and two sites in the vertex region as an active control. We used a 2×2 repeated-measures design with the within-participant factors Task (visual speech task vs motion direction task) and Stimulation (V5/MT vs vertex). Additionally, we acquired baseline performance measurements in the behavioral tasks before the administration of stimulation in Sessions 2 and 3. The behavioral tasks were completed again in a fourth and final session during which no stimulation occurred. The behavioral data collected in Session 4 were used for exploratory purposes and to assess the effects of TMS on practice effects across more sessions. Sessions 2 and 3 were separated by 8 d on average (SD = 3.82 d; range, 3–14 d), and Session 4 followed 1–3 d later (mean = 1.91, SD = 0.81). The order of V5/MT and vertex stimulation across Sessions 2 and 3 was counterbalanced between participants.

Localization of V5/MT

Design and stimuli. A well established fMRI paradigm was used to localize V5/MT in each individual participant (Bridge et al., 2008). We chose to run a conventional "moving versus static" paradigm to aim for typical V5/MT coordinates. The stimuli were videos of random-dot kinematograms (RDKs) that consisted either of a single frame of static dots (static condition) or dots moving inward or outward in coherent motion (moving condition). We used 0.1° large white dots and a 0.2° large gray fixation point in the center of the screen in front of a black background. In the moving condition, 250 dots were generated overall and shifted with a speed of 4.7°/s. We created the stimuli using MATLAB R2019b (version 9.6.0.1047502) and Psychtoolbox version 3.0.15 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). Participants viewed a 187 cm remote mirrored screen (screen diagonal = 65 cm). The procedure consisted of 32 blocks, randomly alternating between the static and moving conditions (8× inward moving, 8× outward moving, 16× static).

Data acquisition parameters. Structural and functional images were acquired on a 3 T Tim Trio machine with a 32-receiver channels head coil (Siemens Healthcare).

Structural T1-weighted images were acquired (TR = 1.9 s; TE = 2.26 ms; flip angle = 9°; bandwidth = 200 Hz/pixel) with a voxel size of 1 mm^3 (176 sagittal slices).

Functional images were acquired with a gradient echoplanar imaging sequence (TR = 2.36 s; TE = 25 ms; flip angle = 80°; bandwidth = 2232 Hz/pixel; whole-brain coverage) with a voxel size of 3 mm^3 (slice thickness = 2.5 mm; spacing between slices = 0.5 mm).

Data preprocessing and analysis. The fMRI data were analyzed with SPM12 (Wellcome Trust Center for Neuroimaging, UCL, London, UK; www.fil.ion.ucl.ac.uk/spm) in a MATLAB environment (R2019b; version 9.6.0.1047502). Preprocessing of the images included calculating voxel deformation maps to correct for motion artifacts, realigning, unwarping, normalizing to Montreal Neurologic Institute (MNI) standard stereotactic space, and smoothing with a Gaussian filter of 8 mm full-width at half-maximum. We generated statistical parametric maps by modeling the evoked hemodynamic response for the static and moving conditions using the general linear model approach. The target V5/MT coordinates were retrieved with the contrast "moving" versus "static," computed for each participant at the first level by means of a t-contrast, with a threshold of $p < 0.001$, uncorrected. Each participant's individual V5/MT coordinates were determined within a cluster of a group contrast of the same localizer from a previous experiment (https://osf.io/9b4fn/). Within this area, we located each subject's coordinates with the highest t-value. The MNI coordinates of left V5/MT (mean coordinates: $x = -45.75 \pm 2.78$, $y = -72.86 \pm 5.06$, $z = 5.71 \pm 3.58$) and right V5/MT (mean coordinates:
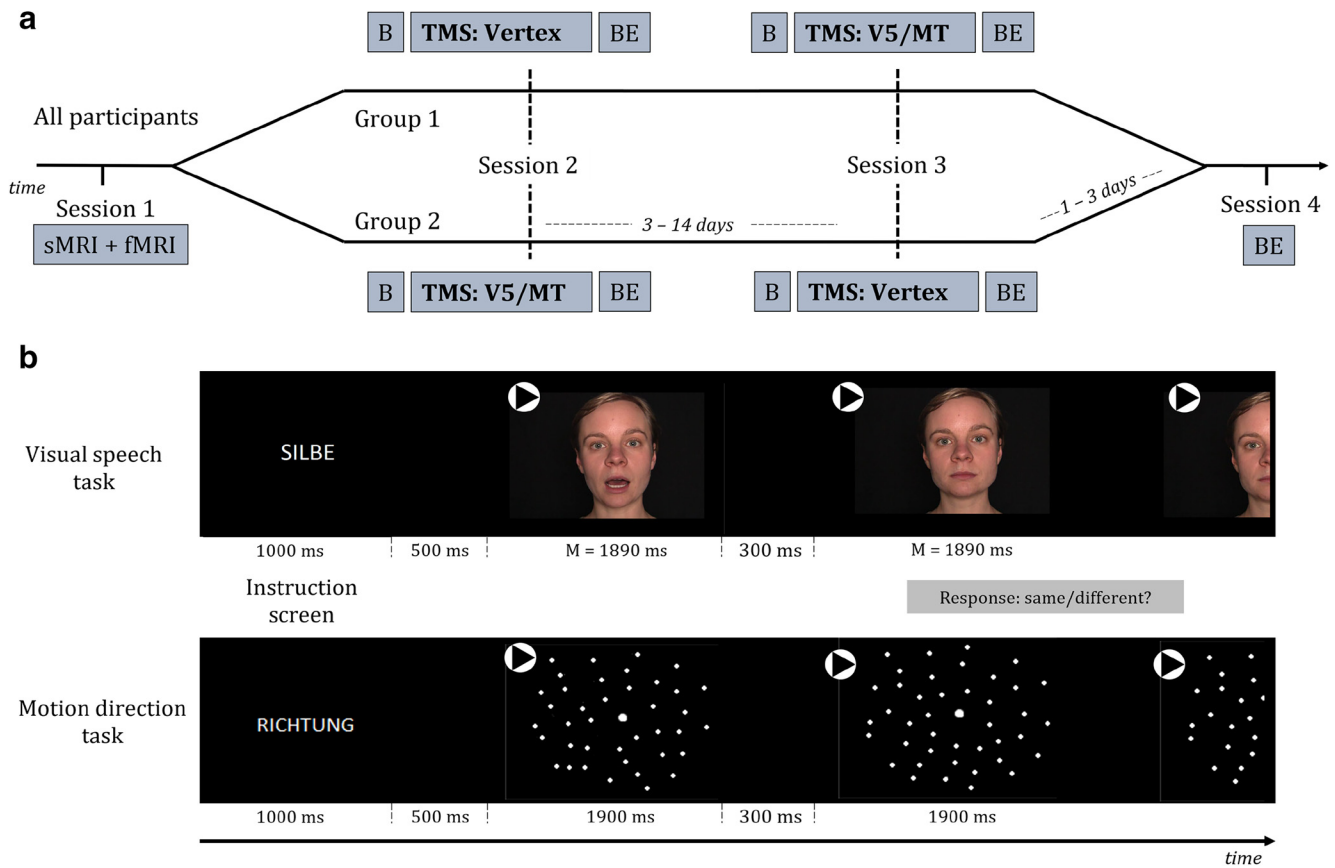
a

b

Figure 1. Experimental design. ***a***, Summary of 4 experimental sessions and counterbalancing of stimulation conditions across sessions. All participants underwent structural and functional MRI in their first session. In Sessions 2 and 3, we first measured participants' baseline performance (B) in the behavioral experiment. This was followed by the application of offline TMS over bilateral area V5/MT or two sites in the vertex region, and then followed by the main behavioral experiment (BE). ***b***, Example trials for the behavioral experiment. Participants performed two 1-back tasks: in the visual speech task ("SILBE"; English: "syllable"), participants responded whether the uttered syllable in the muted video was the same or a different one than in the video before. In the motion direction task ("RICHTUNG"; English: "direction"), participants indicated whether the dots moved in the same or a different direction as the immediately preceding stimulus.
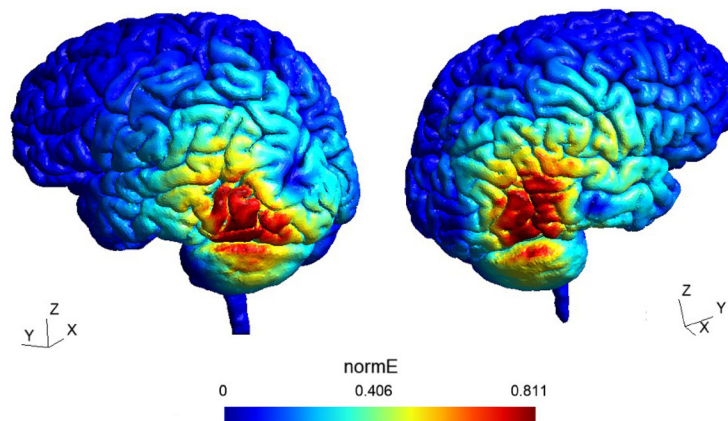

Figure 2. The average strength of the induced electrical field across participants displayed on the normalized cortical surface of an example subject ("Ernie" head model) for left and right V5/MT, computed with SimNibs software (Thielscher et al., 2015).

$x = 47.25 \pm 3.62$, $y = -67.32 \pm 4.35$, $z = 4.75 \pm 3.95$) of each individual participant were retransformed to individual space for the neuronavigated TMS. The average induced electric field on the mean coordinates is visualized in Figure 2.

Transcranial magnetic stimulation

TMS procedure and neuronavigation. Offline TMS was administered in each participant's second and third experimental sessions. We decided to carry out an offline stimulation because of several reasons. First,

online TMS protocols were ill suited for a visual speech experiment: we planned to disrupt motion processing for the whole stimulus length. While single pulses would not have lasted over the whole stimulus duration, a sufficient pulse train on the other hand needed a long intertrain interval that would not have been compatible with an *n*-back and similar tasks. Second, we planned to inhibit area V5/MT bilaterally. Together with presenting the stimulus on both visual hemifields, online TMS was not feasible since two coils on both scalp sites would have physically interfered with each other. Third, in a future planned study, we aimed to

apply the design to an fMRI experiment following TMS, for which off-line stimulation of area V5/MT would be needed. Each participant's active motor threshold (AMT) was measured at the beginning of the first TMS session to determine their individual stimulation intensity. Single TMS pulses starting with an intensity of 50% of stimulator output were applied to the left primary motor cortex (M1) while the participant was instructed to lightly tense their right index finger muscle (abductor pollicis brevis) by pressing it against their thumb; and the stimulation intensity was decreased until 5 of 10 pulses resulted in a motor evoked potential above a predetermined threshold of $50\,\mu V$. Motor thresholds are a stable parameter within individuals across several days (Kimiskidis et al., 2004; Rossi et al., 2009; Perellón-Alfonso et al., 2018). The respective target coordinates were converted from a standard site ($x = 37$; $y = -21$, $z = 58$) from a meta-analysis (Mayka et al., 2006).

To stimulate the vertex area, two sites 1.5 cm anterior and posterior of the computed vertex coordinate were determined. Based on the individual structural MR images, the vertex coordinates were transformed into individual space from MNI coordinates corresponding to the location of the Cz electrode in the 10–20 EEG system ($x = 0.8$, $y = -14.7$, $z = 73.9$) from a study by Okamoto et al. (2004) who projected 10–20 standard cranial positions over the MNI cortical surface.

We used stereotactic neuronavigation (BrainSight, Rogue Research) to precisely position the TMS coil over each individual's V5/MT, vertex, and M1 coordinates. The coil was set over the BrainSight-indicated entry points of the respective coordinates. The entry points were those sites on the participant's scalp that had the shortest distance to the target coordinates. We counterbalanced whether the participants received stimulation over the anterior or posterior vertex area first and if the left or right V5/MT would be stimulated first, respectively. Immediately after stimulating both hemispheres, the participant started with the behavioral experiment.

TMS parameters. Participants received continuous theta-burst stimulation (cTBS) with a stimulator (model MagPro X100, MagVenture) and a figure-of-eight MCF-B65 coil (double-circle diameter = 75 mm). For each site, 200 frames of a 60 ms duration were applied. Each frame contained three pulses that were separated by a 15 ms interval (600 pulses in total). An intertrain interval with a duration of 140 ms followed each frame, resulting in an overall stimulation duration of 40 s for each site. The stimulation was applied with an intensity of 100% of the AMT.

We used this stimulation protocol and design as a pilot study with a similar design was not effective or interpretable: 40% of pilot participants reported the effective stimulation as feeling more intense than sham stimulation. Because of potentially confounding effects of differences in stimulation intensity between effective and sham TMS conditions, we opted in the current study to use the active stimulation of the vertex rather than sham stimulation as an experimental control condition. Moreover, a potential main effect of stimulation would not allow drawing any conclusion about whether the stimulation effect would be specific to the stimulated area because of the presence versus absence of electromagnetic pulses alone. The current study did not include any of the pilot experiment participants. The pilot experiment methods and results are summarized at https://osf.io/9b4fn/.

Behavioral experiment

Stimuli. RDKs were used as stimuli in the motion direction task. The RDKs consisted of a fixation point in the center of the screen surrounded by white ~0.05° large dots with a density of 0.35 dots/°², moving with a speed of 3.15°/s. We created videos for 48 directions that were allocated equally over a 360° space. Since articulator movements in the visual speech videos started 10 frames (200 ms) after video onsets on average (Díaz et al., 2018), we constructed the RDKs so that the dots also remained static for the first 200 ms in each video. To ensure that the pattern of the static dots would not predict the subsequent motion direction, we generated seven videos per direction (i.e., 336 videos in total). Fifty percent of the dots moved coherently in a respective direction, and the other dots moved randomly. All videos had a duration of 1900 ms. Videos were generated using MATLAB R2019b (version 9.6.0.1047502) and Psychtoolbox (version 3.0.15; Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) using the "limited lifetime" algorithm (Pilly and Seitz, 2009).

For the visual speech task, stimuli were taken from an in-house stimulus database of video recordings of German native speakers uttering short syllables (Borowiak et al., 2018; Díaz et al., 2018). We selected 48 muted syllable videos of one female speaker in front of a black background. The syllables were composed in a vowel-consonant-vowel structure that was a combination of the vowels /a/, /e/, and /u/, and the consonants /f/, /p/, /n/, /r/, /s/, and /t/. The vowels and consonants corresponded to discriminable viseme classes in German (see Aschenberner and Weiss, "1 Phoneme-viseme mapping for German video-realistic audio-visual-speech-synthesis IKP-working paper NF 11"; available at: https://www.semanticscholar.org/paper/1-Phoneme-Viseme-Mapping-for-German-Video-Realistic-Aschenberner/c1535ec43b8d0c261aaa9ea4a85915a9f7fe9850). Videos were on average 1.89 ± 0.19 s long, and started and ended with the speakers' mouths closed.

Task procedure. Each block of the behavioral experiment started with a 1000-ms-long instruction screen presenting the word "direction" (German: "RICHTUNG") or "syllable" (German: "SILBE"), which instructed participants as to which task to perform, followed by a 500-ms-long black screen (Fig. 1b). During the motion direction task, participants watched the RDKs and responded by button press with the right hand to indicate whether the dots moved in the same or a different direction than those shown in the previous video. Similarly, in every "syllable" block, the visual speech videos were displayed, and participants responded to indicate whether the uttered syllable in the muted video was the same as or different from the one shown in the preceding video. Each block contained 28 trials separated by a 300-ms-long intertrial interval. The match/mismatch rate (i.e., same/different syllable/direction) was 50% overall, jittered from 40% to 60% across blocks, corresponding to a chance accuracy level of 50%. Another aim of our randomization was to keep both tasks similar in design and to ensure an equivalent level of difficulty between tasks as well as between blocks. Therefore, in every block of the motion direction task, all presented directions belonged to a single common quadrant within the field of view (e.g., between 0° and 90°) and had to be at least 15° apart. For each block of the visual speech task, four different syllables were chosen that all started with the same viseme, respectively. Thus, each block of the two tasks contained either four different syllables or four different directions.

The behavioral experiments (i.e., poststimulation) consisted of 24 blocks overall, while the baseline task for the two TMS sessions (i.e., prestimulation) consisted of 12 blocks. For each participant, syllables and motion directions that had not been presented in the baseline task at the start of the first TMS session were used in the baseline task during the second TMS session. Before the baseline task, participants completed four additional practice blocks, each containing eight trials with feedback, to become familiar with the task. The assignment of the left and right buttons to the match and mismatch conditions was counterbalanced between participants. The experiment was run using Labvanced (Finger et al., 2017) software. Stimuli were presented on a 53 × 30 cm computer screen at a 60 cm distance from participants. At the end of the final session (Session 4), participants completed a questionnaire on potential differences between the TMS sessions, TMS side effects, and strategies used in the behavioral experiment.

Hypotheses and data analysis

Our main hypothesis (Hypothesis 1) was that the application of TMS over area V5/MT would lead to an increase in response times in both tasks compared with the application of TMS over the vertex area. In addition (Hypothesis 2), we expected that TMS over V5/MT would lead to a decreased practice effect (less of a decrease in response times over time) for both tasks in contrast to TMS over the vertex region. This prediction was based on prior findings that a repeated testing of visual speech leads to an improvement of performance over time (Lander and Davies, 2008; Riedel et al., 2015). This also applies to motion perception within RDKs (Zanker, 1999).

The dependent variable in our analyses was response time because TMS typically influences response times rather than accuracy and is therefore considered the standard measure for assessing effects of TMS (Pascual-Leone et al., 1996; Ashbridge et al., 1997; Sack et al., 2007;

Hartwigsen et al., 2017). We tested our hypotheses with linear mixed-effects models in R version 3.6.3 (see R Foundation for Statistical Computing, "R: a language and environment for statistical computing"; available at: https://www.r-project.org/). The mixed-effects model for both tests included fixed effects of Task (motion direction vs visual speech) and Stimulation (V5/MT vs vertex). To test Hypothesis 1, we modeled response times in the behavioral tasks completed during the two TMS sessions. Response times were computed as the time from video onset until the participant's response. To rule out falsely assigned responses, we excluded response times <200 ms. Incorrect responses were also excluded from the response time analysis. To test Hypothesis 2, we compared the performance of the baseline tasks and the behavioral experiment in the two TMS sessions. Specifically, we computed a ratio by dividing the response times of the behavioral experiment by the response times of the respective baseline for each participant, task, and stimulation condition. The random-effects structure of the mixed-effects models was determined using backward model selection, in which random-effects terms that accounted for the least variance were removed one by one until the fitted mixed model was no longer singular, that is, until variances of one or more linear combinations of random effects were no longer equal to zero. Both models included a random intercept by Subject and a random slope by Subject for the Task factor, and a random intercept by Stimulus. For exploratory purposes, we also analyzed accuracy (percentage correct) in the behavioral tasks using the same linear mixed-effects modeling approach. The accuracy model contained a random intercept by Subject and a random slope by Subject for the Stimulation factor, as well as a random intercept by Stimulus.

Results

Our primary hypothesis was that the application of inhibitory TMS over V5/MT would increase the times it took participants to recognize both visual speech and motion direction compared with inhibitory TMS over the vertex region. The linear mixed-effects model on response times revealed a significant main effect of Stimulation confirming our hypothesis: stimulation of V5/MT increased recognition response times compared with stimulation of the vertex region ($\beta = 16.42$; $t = 4.13$; $p < 0.001$; 95% CI = 8.64 24.20; $d = 0.05$; Fig. 3). Compared with vertex stimulation, V5/MT stimulation delayed the recognition of visual speech by 26 ms (SE = 4 ms) on average, and the recognition of motion direction by 16 ms (SE = 4 ms) on average (Fig. 3b).

In addition, the model revealed a significant main effect of Task (i.e., participants responded faster in the motion direction task than in the visual speech task; $\beta = 560.98$; $t = 18.64$; $p < 0.001$; 95% CI = 502.00, 619.96; $d = 4.06$). This was expected, as the critical viseme in the visual speech task occurred later than the motion onset in the motion task. The interaction of Stimulation and Task factors was not significant ($\beta = 9.41$; $t = 1.684$; $p = 0.092$; 95% CI = −1.55, 20.37; $d = 0.02$). The full mixed-model results are shown in Table 1.

Our second hypothesis was that the effect of V5/MT stimulation relative to vertex stimulation would diminish practice effects in both behavioral tasks (i.e., lessen the reduction in response times between baseline performance and the main behavioral performance). To account for such practice effects and intersubject variability in response times, we computed the ratio of response times measured during baseline (i.e., prestimulation) performance and response times measured during the behavioral experiment (i.e., poststimulation). The mean prestimulation and poststimulation response times for each task and stimulation condition are shown in Table 2. A more positive ratio indicated a greater TMS effect or a smaller practice effect over time. The mixed-effects model revealed a significant main effect of Stimulation: V5/MT stimulation resulted in a larger response time ratio than vertex stimulation ($\beta = 0.02$; $t = 5.658$;

$p < 0.001$; 95% CI = 0.01, 0.03; $d = 0.15$; Fig. 4). Thus, participants showed less improvement between the prestimulation baseline and the poststimulation experiment when area V5/MT was stimulated compared with when the vertex was stimulated. The full model results are shown in Table 3.

An exploratory mixed-effects model of accuracy scores indicated no significant main effect of Stimulation ($\beta = 0$; $t = -0.419$; $p = 0.676$; 95% CI = −0.03, 0.02; $d = 0.15$). In contrast, we found a significant main effect of Task on accuracy scores ($\beta = 0.03$; $t = 3.013$; $p = 0.003$; 95% CI = 0.01, 0.05; $d = 0.51$). Participants were more accurate in recognizing visual speech (mean = 86.1% correct, SE = 0.01%) than motion direction (mean = 84.5% correct, SE = 0.01%). Finally, the model showed a significant interaction of Task and Stimulation ($\beta = -0.03$; $t = -3.207$; $p = 0.001$; 95% CI = 0.04, −0.01; $d = 0.04$). The interaction was driven by a greater difference in accuracy between tasks during vertex stimulation ($\beta = 0.03$; $z = 3.01$; $p = 0.014$) compared with V5/MT stimulation ($\beta = 0.00$; $z = 0.35$; $p = 0.98$), as well as a significant effect of stimulation on visual speech recognition accuracy ($\beta = 0.03$; $z = 2.57$; $p = 0.0497$), but not motion direction accuracy ($\beta = 0.00$; $z = 0.42$; $p = 0.98$). The model results are shown in Table 4 and Figure 5.

We also calculated the ratio of prestimulation and poststimulation accuracy by dividing the accuracies following stimulation by the mean accuracies during baseline (i.e., prestimulation), analogous to the response time ratio. A more positive ratio indicated a smaller TMS effect or greater practice effect over time. We found a significant main effect of Stimulation ($\beta = 0.04$; $t = 5.223$; $p < 0.001$; 95% CI = 0.03, 0.06; $d = 0.06$). Participants showed a larger practice effect in the V5/MT stimulation condition (mean = 1.06, SE = 0.03) than in the vertex stimulation condition (mean = 1.01 SE = 0.03). Moreover, there was a significant main effect of Task ($\beta = 0.05$; $t = 2.227$; $p = 0.03$; 95% CI = 0.01, −0.10; $d = 0.73$), revealing a larger practice effect for the visual speech (mean = 1.07, SE = 0.04) than for the motion recognition task (mean = 1.01, SE = 0.02). The full model results can be found at https://osf.io/9b4fn/.

Based on the questionnaire completed following the TMS sessions, 67% of participants noticed no differences between the two TMS sessions, whereas ~13% correctly perceived the V5/MT stimulation as more intense or as increasing task difficulty. Reported side effects were mild headache (12.5% of participants), fatigue (8.33%), and neck tension (8.33%). Supplementary material can be found at https://osf.io/9b4fn/.

Discussion

The current study used inhibitory TMS to investigate the causal relevance of cortical area V5/MT for the recognition of visual speech. Participants completed visual speech and nonbiological motion recognition tasks after undergoing TMS over V5/MT and a control site. We found that V5/MT stimulation slowed the recognition of both visual speech and simple, nonbiological motion relative to control TMS. Moreover, it reduced practice effects on response times. These findings support our main hypothesis that area V5/MT causally influences the recognition of visual speech.

To our knowledge, no previous study has investigated causal influences of area V5/MT to visual speech recognition. Causal contributions of area V5/MT are consistent with neuroimaging studies that report V5/MT responses during visual speech and biological motion recognition (Calvert and Campbell, 2003; Paulesu et al., 2003; Peuskens et al., 2005; Borowiak et al., 2018). Visual
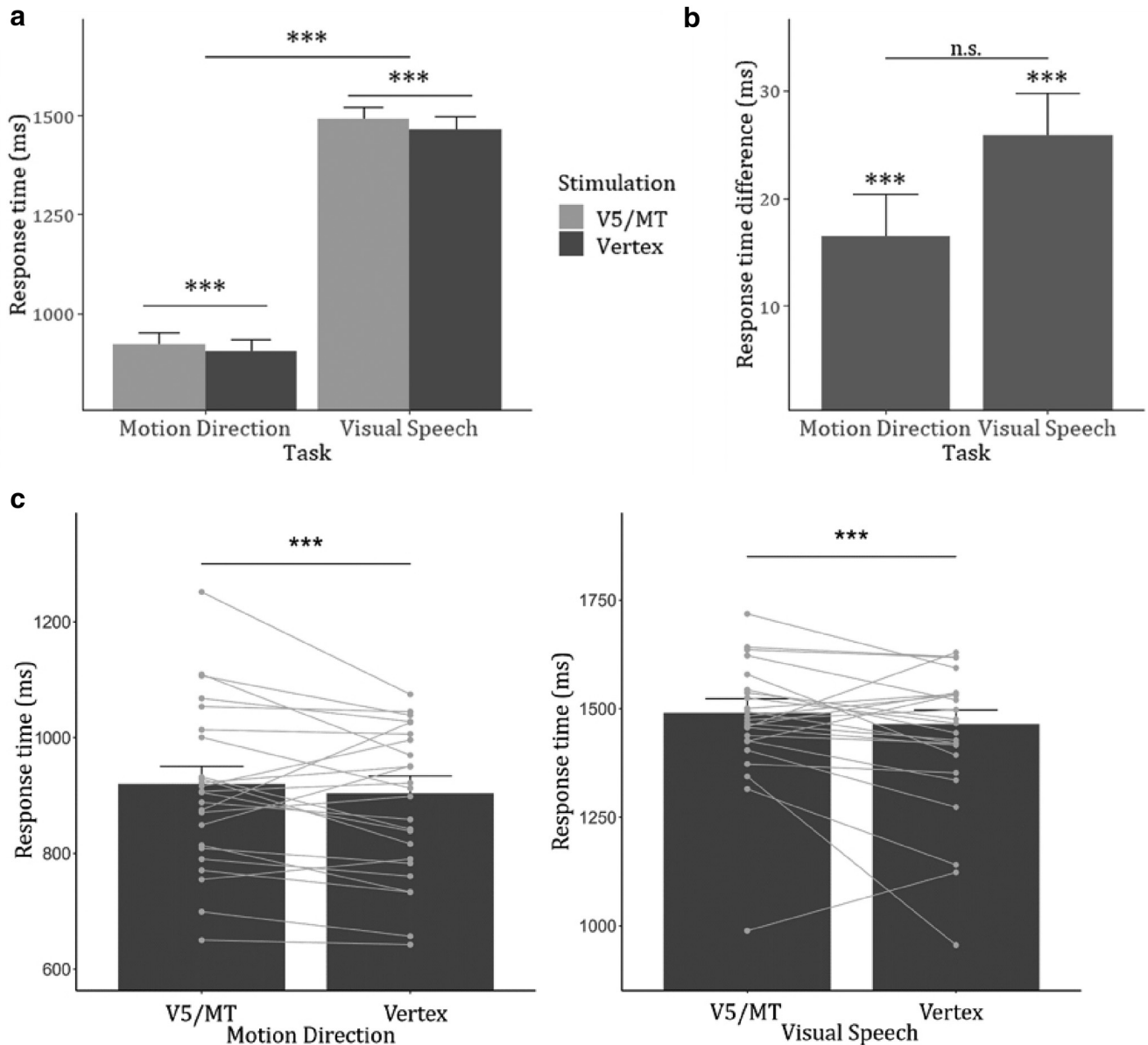
Figure 3. *a*, Effects of cTBS on response times in the motion direction and visual speech recognition tasks. Bilateral stimulation over V5/MT slowed response times in both tasks compared with stimulation over the vertex area. There was no significant interaction with Task. *b*, Response time difference (V5/MT stimulation − vertex stimulation) for each task. *c*, Individual data points for both stimulation conditions for each task. Error bars represent 1 SEM. ***$p < 0.001$, n.s. = not significant.

speech recognition performance typically increases over the course of a task (Lander and Davies, 2008; Sánchez-Panchuelo et al., 2012). Our findings show that V5/MT stimulation also interferes with such improvement.

The stimulation effect of V5/MT raises the question of how and to what extent it interacts with other cerebral regions and their respective functions during visual speech recognition. Findings in macaque brains have suggested a dual neural processing route for biological motion: one route for motion information via MT/V5, and a second route for visual form that bypasses MT/V5 (Bullier and Morel, 1990). Studies in humans corroborate this evidence (Miller et al., 2018). For instance, Mather et al. (2016) report a reduced recognition of motion direction for dots displaying a general drift, but not point-light walkers following V5/MT stimulation, indicating that certain types of biological motion (i.e., walking) may not necessarily depend on V5/MT.

In contrast, our experiment showed that V5/MT serves as a critical component for recognizing visual speech, which points toward an interconnected network rather than separate routes. Prior fMRI findings have corroborated this suggestion by revealing structures within both ventral and dorsal visual pathways that contribute to visual speech recognition (Campbell et al., 2001; Bernstein et al., 2011; Files et al., 2015). For instance, parts of the pSTS/pSTG respond specifically to multimodal speech processing (Beauchamp et al., 2002; Grossman et al., 2005; Hall et al., 2005; Schultz et al., 2005), and the functional integrity of pSTS/pSTG is relevant for visual speech recognition (Riedel et al., 2015). It remains uncertain whether other areas compensate for a V5/MT inhibition, leading visual speech recognition to remain intact. Moreover, research indicates a flexible use of certain areas and their functions when processing visual speech. Viewing natural speech uses a network that codes for visual form and movement (Campbell et al., 1997), and thus there might be an application of neuronal mechanisms dependent on the benefit for the specific task and stimulus (Giese and Poggio, 2003; Thirkettle et al., 2009), which is in line with the dorsal and

Table 1. Linear mixed-effects model effects of stimulation and task on response times

	Response times				
	β	CI	t	SE	p
Predictors					
Intercept	903.93	845.29–962.58	30.21	29.92	**<0.001**
Stimulation	16.42	8.64–24.20	4.13	3.97	**<0.001**
Task	560.98	1019.91–1052.06	18.64	30.09	**<0.001**
Stimulation * Task	9.41	502.00–619.96	1.68	5.59	0.092
Random effects		Variance	SD		
Subject					
Intercept		15,345	123.87		
Task		9454	97.23		
Stimulus					
Intercept		11,908	109.12		

Significant *p*-values are marked in bold.

Table 2. Mean response times (in ms) for the baseline measurement (i.e., prestimulation) and the behavioral experiment (i.e., poststimulation)

Task	Stimulation		Mean	SD
Motion direction	V5/MT	Baseline	967	291
		Behavioral experiment	915	266
	Vertex	Baseline	985	265
		Behavioral experiment	899	253
Visual speech	V5/MT	Baseline	1516	311
		Behavioral experiment	1484	301
	Vertex	Baseline	1544	276
		Behavioral experiment	1463	304

Table 3. Linear mixed-effects model effects of stimulation and task on the pre-TMS and post-TMS response time ratios

		Behavioral experiment/baseline			
Response times (β)		CI	t	SE	p
Predictors					
Intercept	0.92	0.89–0.94	73.22	0.013	**<0.001**
Stimulation	0.02	0.01–0.03	5.66	0.004	**<0.001**
Task	0.03	0–0.06	1.91	0.017	0.056
Stimulation * Task	0	−0.01 to 0.01	0.93	0.005	0.352
		Variance	SD		
Random effects					
Subject					
Intercept		0.001	0.032		
Task		0.001	0.033		
Stimulus					
Intercept		0.005	0.071		

Significant *p*-values are marked in bold.

Table 4. Linear mixed-effects model effects of stimulation and task on mean accuracies

	Mean accuracy				
	β	CI	t	SE	p
Predictors					
Intercept	0.85	0.83–0.87	76.75	0.014	**<0.001**
Stimulation	−0.00	−0.03 to 0.02	−0.42	0.018	0.676
Task	0.03	0.01–0.05	3.01	0.010	**0.003**
Stimulation * Task	−0.03	−0.04 to −0.01	−3.21	0.007	**0.001**
		Variance	SD		
Random effects					
Subject					
Intercept		0.002	0.043		
Stimulation		0.003	0.051		
Stimulus					
Intercept		0.001	0.038		

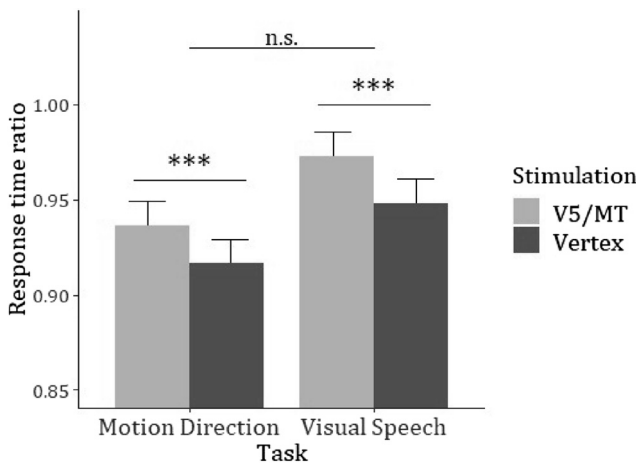Significant *p*-values are marked in bold.

Figure 4. Effects of cTBS on the ratio of prestimulation (baseline) and poststimulation (behavioral experiment) response times. A ratio <1.0 indicates faster responses following stimulation compared with baseline. Response times increased less when V5/MT was stimulated compared with when the vertex was stimulated in both behavioral tasks. This is reflected in larger ratios for the V5/MT stimulation condition relative to the vertex stimulation condition. There was no significant interaction between Task and Stimulation. Error bars represent 1 SEM. ***p < 0.001, n.s. = not significant.
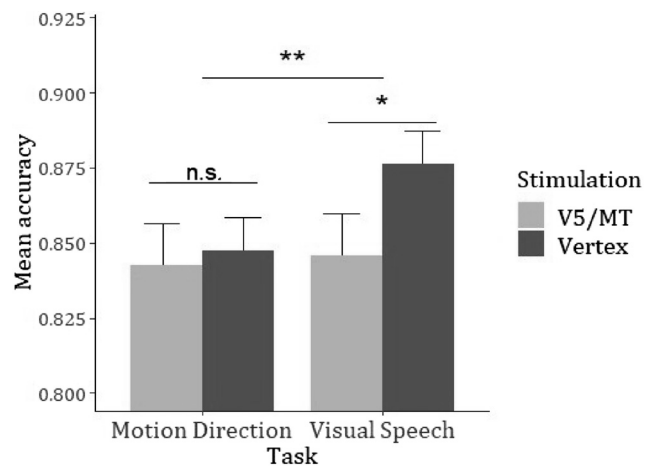

Figure 5. Effects of cTBS on accuracy in the motion direction and visual speech recognition tasks. Relative to vertex stimulation, V5/MT stimulation diminished accuracy in the visual speech task. The difference in mean accuracies between tasks was only significant during the vertex stimulation (p = 0.014), in contrast to the V5/MT stimulation (p > 0.05). Error bars represent 1 SEM. *p < 0.05, **p < 0.01, ***p < 0.001, n.s. = not significant.

ventral pathway contribution. Arnal et al. (2009) suggest an interaction of the STS and motion-sensitive cortex areas: signals from auditory and motion-sensitive areas synchronize when syllables are articulated, and the STS reflects a

feedback pathway for incongruent signals. Motion-relevant areas might be less used for the recognition of stimuli relying more on form-based strategies than in conditions retrieving dynamic visual cues. For instance, the TVSA could be recruited during the usage of the former, as opposed to V5/MT for the latter. This indicates the possibility of stimulation effects depending on the task design; TMS might differently affect visual speech recognition among multiple speaker identities than an intra-speaker speech recognition.

Our findings extend previous studies using inhibitory TMS over V5/MT on visual motion recognition tasks, such as speed and direction discrimination, motion detection, and implied motion perception (Beckers and Zeki, 1995; Kourtzi and Kanwisher, 2000; Schenk et al., 2005; Laycock et al., 2007; McKeefry et al., 2008; Silvanto and Cattaneo, 2010; Grasso et al., 2018). Similar to visual speech recognition, response times in the motion direction task decreased between prestimulation and poststimulation to a smaller degree following TMS over V5/MT compared with vertex stimulation. Given that simple visual tasks such as the recognition of random dot motion lead to rapid improvement in performance (Ahissar and Hochstein, 1997; Zanker, 1999), we have shown that V5/MT stimulation can interfere with such practice effects. However, our exploratory accuracy analyses indicate a speed–accuracy trade-off in which TMS over V5/MT led to slower, but more correct, responses in relation to the baseline performance.

Several prior studies have administered TMS to V5/MT using online stimulation protocols that are time locked to motion onsets, thus potentially aligning with sensitive periods for stimulation (Schenk et al., 2005; Sack et al., 2006; Laycock et al., 2007; McKeefry et al., 2008; Stevens et al., 2009; Alexander et al., 2018; Grasso et al., 2018). Early periods are thought to reflect direct communication between V5/MT and the visual sensory thalamus (Laycock et al., 2007), especially for high-frequency movement components such as those contained in visual speech (Beckers and Zeki, 1995; Grasso et al., 2018). Since recent research has demonstrated an important role of the visual thalamus for visual speech recognition (Díaz et al., 2018), we speculate that feedback between V5/MT and subcortical structures may also contribute to visual speech recognition. Our results also show that V5/MT can be inhibited by longer-term offline TMS protocols, complementing the few existing respective studies (Cai et al., 2014; Chakraborty et al., 2019). It is noteworthy, however, in contrast to the response times, TMS lowered accuracies only on the visual speech task and not the motion direction task. This is consistent with the view that TMS generally has stronger effects on response times (Ashbridge et al., 1997; Sack et al., 2007).

Active TMS conditions are often used preferentially to sham conditions as experimental controls, depending on the investigated brain region and somatosensory side effects of stimulation (Loo et al., 2000; Duecker and Sack, 2015). Nevertheless, TMS effects in active control site designs can be explained either by inhibition of the target region or by excitation of the control region. We argue that, in the current study, TMS had an inhibitory effect on V5/MT rather than an excitatory effect on the vertex for two reasons. First, to our knowledge, there are no findings in which the stimulation protocol for V5/MT (Cai et al., 2014; Chakraborty et al., 2019) has shown an excitatory rather than inhibitory effect on response times. Second, vertex stimulation does not induce increased BOLD responses at the stimulated site and, instead, even decreases activations of regions within the default mode network (Jung et al., 2016).

Relatedly, although the precise site localization in the current study was achieved using fMRI-guided neuronavigation, one cannot rule out a stimulation of immediately adjacent areas. Kolster et al. (2010) defined three regions within the V5/MT cluster that adjoin the core V5/MT region: the ventral part of the medial superior temporal area, the fundus of the superior temporal area, and the V4 transitional zone. Located deeper within the sulci than V5/MT and, although considered motion sensitive, they are thought to respond more preferentially to shape relative to V5/MT. It remains unknown whether, dependent on interindividual differences in V5/MT depth and location, neighboring areas were costimulated in some participants. Regardless, potential stimulation of shape-related regions in the V5/MT vicinity cannot explain the motion-related TMS effects observed in the current study.

In summary, we consider our results as a promising first indication that V5/MT causally affects visual speech recognition. While neuroimaging studies have already revealed multiple sites associated with visual speech recognition processes, including V5/MT (Campbell et al., 2001; Calvert and Campbell, 2003; Paulesu et al., 2003; Hall et al., 2005; Blank and von Kriegstein, 2013; Borowiak et al., 2018), here we extend those findings by showing that V5/MT is causally involved, which is not possible using fMRI. Concerning biological motion, this may also argue against a dichotomy of involvement versus noninvolvement of V5/MT and suggests instead that the contribution of V5/MT depends on subtypes of biological motion and their respective parameters, such as velocity. This has already been implicated by findings on low-level simple motion, where the recognition of rapid motion was found to be more susceptible to a V5/MT inhibition compared with slower motion, especially early after stimulus onsets (Grasso et al., 2018). Since visual speech rarely occurs in isolation, but rather within complex audiovisual environments, our results also highlight the importance of V5/MT for speech recognition in natural settings. Visual signals are a crucial component in audiovisual speech perception: seeing a talker improves the comprehension of their auditory speech (Sumby and Pollack, 1954), and, incongruently, visual speech signals can even interfere with speech recognition (McGurk and Macdonald, 1976). Moreover, individuals on the autism spectrum, characterized by decreased visual speech recognition abilities, have displayed reduced responses in visual movement areas including V5/MT, but not in other speech regions (Borowiak et al., 2018). The finding that V5/MT can causally impact visual speech recognition is a step toward better understanding of the neural mechanisms supporting the perception of human communication signals both in typically developed individuals as well as in populations with communication difficulties.

References

Ahissar M, Hochstein S (1997) Task difficulty and the specificity of perceptual learning. Nature 387:401–406.

Alexander B, Laycock R, Crewther DP, Crewther SG (2018) An fMRI-neuronavigated chronometric TMS investigation of V5 and intraparietal cortex in motion driven attention. Front Hum Neurosci 11:638–639.

Antal A, Nitsche MA, Kruse W, Kincses TZ, Hoffmann KP, Paulus W (2004) Direct current stimulation over V5 enhances visuomotor coordination by improving motion perception in humans. J Cogn Neurosci 16:521–527.

Arnal LH, Morillon B, Kell CA, Giraud AL (2009) Dual neural routing of visual facilitation in speech processing. J Neurosci 29:13445–13453.

Ashbridge E, Walsh V, Cowey A (1997) Temporal aspects of visual search studied by transcranial magnetic stimulation. Neuropsychologia 35:1121–1131.

Auer ET, Bernstein LE (2007) Enhanced visual speech perception in individuals with early-onset hearing impairment. J Speech Lang Hear Res 50: 1157–1165.

Bach M (1996) The Freiburg Visual Acuity test–automatic measurement of visual acuity. Optom Vis Sci 73:49–53.

Beauchamp MS, Lee KE, Haxby JV, Martin A (2002) Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34:149–159.

Beckers G, Zeki S (1995) The consequence of inactivating areas v1 and v5 on visual motion perception. Brain 118:49–60.

Bernstein LE, Liebenthal E (2014) Neural pathways for visual speech perception. Front Neurosci 8:386.

Bernstein LE, Jiang J, Pantazis D, Lu ZL, Joshi A (2011) Visual phonetic processing localized using speech and nonspeech face gestures in video and point-light displays. Hum Brain Mapp 32:1660–1676.

Bernstein M, Erez Y, Blank I, Yovel G (2018) An integrated neural framework for dynamic and static face processing. Sci Rep 8:7036.

Blank H, von Kriegstein K (2013) Mechanisms of enhancing visual-speech recognition by prior auditory information. Neuroimage 65:109–118.

Borowiak K, Schelinski S, von Kriegstein K (2018) Recognizing visual speech: reduced responses in visual-movement regions, but not other speech regions in autism. Neuroimage Clin 20:1078–1091.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

Bridge H, Thomas O, Jbabdi S, Cowey A (2008) Changes in connectivity after visual cortical brain damage underlie altered visual function. Brain 131:1433–1444.

Bullier J, Morel A (1990) Anatomical segregation of two cortical visual pathways in the macaque monkey. Vis Neurosci 4:555–578.

Cai P, Chen N, Zhou T, Thompson B, Fang F (2014) Global versus local: double dissociation between MT+ and V3A in motion processing revealed using continuous theta burst transcranial magnetic stimulation. Exp Brain Res 232:4035–4041.

Calvert GA, Campbell R (2003) Reading speech from still and moving faces: the neural substrates of visible speech. J Cogn Neurosci 15:57–70.

Campbell R (2008) The processing of audio-visual speech: empirical and neural bases. Philos Trans R Soc Lond B Biol Sci 363:1001–1010.

Campbell R, Zihl J, Massaro D, Munhall K, Cohen MM (1997) Speechreading in the akinetopsic patient, L.M. Brain 120:1793–1803.

Campbell R, MacSweeney M, Surguladze S, Calvert G, McGuire P, Suckling J, Brammer MJ, David AS (2001) Cortical substrates for the perception of face actions: an fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). Brain Res Cogn Brain Res 12:233–243.

Chakraborty A, Tran TT, Silva AE, Giaschi D, Thompson B (2019) Continuous theta burst TMS of area MT+ impairs attentive motion tracking. Eur J Neurosci 5:7289–7300.

Chandrasekaran C, Ghazanfar AA (2011) When what you see is not what you hear. Nat Neurosci 14:675–676.

Díaz B, Blank H, von Kriegstein K (2018) Task-dependent modulation of the visual sensory thalamus assists visual-speech recognition. Neuroimage 178:721–734.

Duecker F, Sack AT (2015) Rethinking the role of sham TMS. Front Psychol 6:210.

Ethofer T, Gschwind M, Vuilleumier P (2011) Processing social aspects of human gaze: a combined fMRI-DTI study. Neuroimage 55:411–419.

Files BT, Tjan BS, Jiang J, Bernstein LE (2015) Visual speech discrimination and identification of natural and synthetic consonant stimuli. Front Psychol 6:878.

Finger H, Goeke C, Diekamp D, Standvoß K, König P (2017) LabVanced: a unified JavaScript framework for online studies. In: International conference on computational social science (Cologne): IC2S2 2017. Cologne, Germany: Deutsche Forschungsgemeinschaft.

Freitag CM, Retz-Junginger P, Retz W, Seitz C, Palmason H, Meyer J, Rösler M, Von Gontard A (2007) Evaluation der deutschen version des autismus-spektrum-quotienten (AQ): die Kurzversion AQ-k. Z Klin Psychol Psychother 36:280–289.

Furl N (2015) Structural and effective connectivity reveals potential network-based influences on category-sensitive visual areas. Front Hum Neurosci 9:253.

Giese MA, Poggio T (2003) Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4:179–192.

Grant KW, Seitz P-F (2000) The use of visible speech cues for improving auditory detection of spoken sentences. J Acoust Soc Am 108:1197–1208.

Grasso PA, Làdavas E, Bertini C, Caltabiano S, Thut G, Morand S (2018) Decoupling of early V5 motion processing from visual awareness: a

matter of velocity as revealed by transcranial magnetic stimulation. J Cogn Neurosci 30:1517–1531.

Grèzes J, Fonlupt P, Bertenthal B, Delon-Martin C, Segebarth C, Decety J (2001) Does perception of biological motion rely on specific brain regions? Neuroimage 13:775–785.

Grossman ED, Battelli L, Pascual-Leone A (2005) Repetitive TMS over posterior STS disrupts perception of biological motion. Vision Res 45:2847–2853.

Hall DA, Fussell C, Summerfield AQ (2005) Reading fluent speech from talking faces: typical brain networks and individual differences. J Cogn Neurosci 17:939–953.

Hartwigsen G, Bzdok D, Klein M, Wawrzyniak M, Stockert A, Wrede K, Classen J, Saur D (2017) Rapid short-term reorganization in the language network. Elife 6:e25964.

Jung J, Bungert A, Bowtell R, Jackson SR (2016) Vertex stimulation as a control site for transcranial magnetic stimulation: a concurrent TMS/fMRI study. Brain Stimul 9:58–64.

Kimiskidis VK, Papagiannopoulos S, Sotirakoglou K, Kazis DA, Dimopoulos G, Kazis A, Mills KR (2004) The repeatability of corticomotor threshold measurements. Neurophysiol Clin 34:259–266.

Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C, Cornelissen F (2007) What's new in Psychtoolbox-3. Perception 36:1–16.

Koivisto M, Mäntylä T, Silvanto J (2010) The role of early visual cortex (V1/V2) in conscious and unconscious visual perception. Neuroimage 51:828–834.

Kolster H, Peeters R, Orban GA (2010) The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. J Neurosci 30:9801–9820.

Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. J Cogn Neurosci 12:48–55.

Lander K, Davies R (2008) Does face familiarity influence speechreadability? Q J Exp Psychol (Hove) 61:961–967.

Laycock R, Crewther DP, Fitzgerald PB, Crewther SG (2007) Evidence for fast signals and later processing in human V1/V2 and V5/MT+: a TMS study of motion perception. J Neurophysiol 98:1253–1262.

Loo CK, Taylor JL, Gandevia SC, McDarmont BN, Mitchell PB, Sachdev PS (2000) Transcranial magnetic stimulation (TMS) in controlled treatment studies: are some "sham" forms active? Biol Psychiatry 47:325–331.

Mather G, Battaglini L, Campana G (2016) TMS reveals flexible use of form and motion cues in biological motion perception. Neuropsychologia 84:193–197.

Mayka MA, Corcos DM, Leurgans SE, Vaillancourt DE (2006) Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. Neuroimage 31:1453–1474.

McGurk H, Macdonald J (1976) Hearing lips and seeing voices. Nature 264:746–748.

McKeefry DJ, Burton MP, Vakrou C, Barrett BT, Morland AB (2008) Induced deficits in speed perception by transcranial magnetic stimulation of human cortical areas V5/MT+ and V3A. J Neurosci 28:6848–6857.

Miller L, Agnew HC, Pilz KS (2018) Behavioural evidence for distinct mechanisms related to global and biological motion perception. Vision Res 142:58–64.

O'Toole AJ, Roark DA, Abdi H (2002) Recognizing moving faces: a psychological and neural synthesis. Trends Cogn Sci 6:261–266.

Okamoto M, Dan H, Sakamoto K, Takeo K, Shimizu K, Kohno S, Oda I, Isobe S, Suzuki T, Kohyama K, Dan I (2004) Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. Neuroimage 21:99–111.

Pascual-Leone A, Wassermann EM, Grafman J, Hallett M (1996) The role of the dorsolateral prefrontal cortex in implicit procedural learning. Exp Brain Res 107:479–485.

Paulesu E, Perani D, Blasi V, Silani G, Borghese NA, De Giovanni U, Sensolo S, Fazio F (2003) A functional-anatomical model for lipreading. J Neurophysiol 90:2005–2013.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442.

Perellón-Alfonso R, Kralik M, Pileckyte I, Princic M, Bon J, Matzhold C, Fischer B, Šlahorová P, Pirtošek Z, Rothwell J, Kojovic M (2018) Similar effect of intermittent theta burst and sham stimulation on corticospinal

excitability: a 5-day repeated sessions study. Eur J Neurosci 48:1990–2000.

Peuskens H, Vanrie J, Verfaillie K, Orban GA (2005) Specificity of regions processing biological motion. Eur J Neurosci 21:2864–2875.

Pilly PK, Seitz AR (2009) What a difference a parameter makes: a psychophysical comparison of random dot motion algorithms. Vision Res 49:1599–1612.

Riedel P, Ragert P, Schelinski S, Kiebel SJ, von Kriegstein K (2015) Visual face-movement sensitive cortex is relevant for auditory-only speech recognition. Cortex 68:86–99.

Rosemann S, Smith D, Dewenter M, Thiel CM (2020) Age-related hearing loss influences functional connectivity of auditory cortex for the McGurk illusion. Cortex 129:266–280.

Rossi S, Hallett M, Rossini PM, Pascual-Leone A (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol 120:2008–2039.

Sack AT, Kohler A, Linden DEJ, Goebel R, Muckli L (2006) The temporal characteristics of motion processing in hMT/V5+: combining fMRI and neuronavigated TMS. Neuroimage 29:1326–1335.

Sack AT, Kohler A, Bestmann S, Linden DEJ, Dechent P, Goebel R, Baudewig J (2007) Imaging the brain activity changes underlying impaired visuospatial judgments: simultaneous fMRI, TMS, and behavioral studies. Cereb Cortex 17:2841–2852.

Sánchez-Panchuelo RM, Francis ST, Schluppeck D, Bowtell RW (2012) Correspondence of human visual areas identified using functional and anatomical MRI in vivo at 7 T. J Magn Reson Imaging 35:287–299.

Santi A, Servos P, Vatikiotis-Bateson E, Kuratate T, Munhall K (2003) Perceiving biological motion: dissociating visible speech from walking. J Cogn Neurosci 15:800–809.

Schelinski S, Riedel P, von Kriegstein K (2014) Visual abilities are important for auditory-only speech recognition: evidence from autism spectrum disorder. Neuropsychologia 65:1–11.

Schenk T, Ellison A, Rice N, Milner AD (2005) The role of V5/MT+ in the control of catching movements: an rTMS study. Neuropsychologia 43:189–198.

Schultz J, Friston KJ, O'Doherty J, Wolpert DM, Frith CD (2005) Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. Neuron 45:625–635.

Silvanto J, Cattaneo Z (2010) Transcranial magnetic stimulation reveals the content of visual short-term memory in the visual cortex. Neuroimage 50:1683–1689.

Silvanto J, Lavie N, Walsh V (2006) Stimulation of the human frontal eye fields modulates sensitivity of extrastriate visual cortex. J Neurophysiol 96:941–945.

Stevens LK, McGraw PV, Ledgeway T, Schluppeck D (2009) Temporal characteristics of global motion processing revealed by transcranial magnetic stimulation. Eur J Neurosci 30:2415–2426.

Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26:212–215.

Thielscher A, Antunes A, Saturnino GB (2015) Field modeling for transcranial magnetic stimulation: a useful tool to understand the physiological effects of TMS. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp 222–225. Piscataway, NJ: IEEE.

Thirkettle M, Benton CP, Scott-Samuel NE (2009) Contributions of form, motion and task to biological motion perception. J Vis 9(3):28, 1–11.

van Laarhoven T, Keetels M, Schakel L, Vroomen J (2016) Audio-visual speech in noise perception in dyslexia. Dev Sci 21:e12504.

Zanker JM (1999) Perceptual learning in primary and secondary motion vision. Vision Res 39:1293–1304.