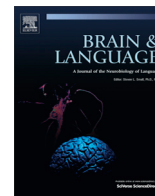


Contents lists available at [SciVerse ScienceDirect](http://SciVerse.Sciencedirect.com)

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Motor excitability during visual perception of known and unknown spoken languages



Swathi Swaminathan^{a,b}, Mairéad MacSweeney^c, Rowan Boyles^a, Dafydd Waters^c, Kate E. Watkins^a, Riikka Möttönen^{a,*}

^a Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, United Kingdom

^b Department of Psychology, University of Toronto, Mississauga, Ontario L5L 1C6, Canada

^c Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, United Kingdom

ARTICLE INFO

Article history:

Accepted 19 March 2013

Available online 3 May 2013

Keywords:

Bilingualism

Lipreading

Motor cortex

Action observation

Motor evoked potentials

Social cognition

Speech

Speechreading

Transcranial magnetic stimulation

ABSTRACT

It is possible to comprehend speech and discriminate languages by viewing a speaker's articulatory movements. Transcranial magnetic stimulation studies have shown that viewing speech enhances excitability in the articulatory motor cortex. Here, we investigated the specificity of this enhanced motor excitability in native and non-native speakers of English. Both groups were able to discriminate between speech movements related to a known (i.e., English) and unknown (i.e., Hebrew) language. The motor excitability was higher during observation of a known language than an unknown language or non-speech mouth movements, suggesting that motor resonance is enhanced specifically during observation of mouth movements that convey linguistic information. Surprisingly, however, the excitability was equally high during observation of a static face. Moreover, the motor excitability did not differ between native and non-native speakers. These findings suggest that the articulatory motor cortex processes several kinds of visual cues during speech communication.

Crown Copyright © 2013 Published by Elsevier Inc. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

1. Introduction

Viewing a speaker's articulatory movements influences speech perception. The well-known McGurk effect demonstrates that seeing an incongruent visual speech signal can modify perception of a clear auditory speech signal. For example, an auditory /ba/ and visual /ga/ are often heard as /da/ (McGurk & MacDonald, 1976). In everyday life, viewing a speaker's articulatory movements improves speech comprehension under challenging auditory circumstances, such as in noisy environments (Sumbly & Pollack, 1954). However, visual speech signals are not as intelligible as auditory speech signals. For example, consonants that share the same place of articulation are hard to discriminate from each other visually (e.g., bilabials /b/, /p/, and /m/) (see, e.g., MacDonald & McGurk, 1978). Speechreading (or "lipreading") is, therefore, a demanding skill and the ability to understand visual speech varies greatly between people (MacLeod & Summerfield, 1990).

Observation of the speaker's articulatory movements facilitates learning of non-native phonetic contrasts (Hardison, 2005; Hazan, Sennema, Iba, & Faulkner, 2005; Hirata & Kelly, 2010). Studies have

shown that non-native speakers use visual cues during audiovisual speech perception of their L2 and that they can learn to use these cues as efficiently as native speakers (Hardison, 1999, 2003; Hazan et al., 2006; Wang, Behne, & Jiang, 2008). There is also evidence that visual cues can enable non-native speakers to make phonetic distinctions that they are not sensitive to when only auditory cues are available (Navarra & Soto-Faroco, 2007).

It is possible to discriminate languages by viewing a speaker's articulatory movements. It has been shown that adults can visually discriminate spoken languages as long as at least one of the languages is their first (L1) or second (L2) language (Ronquest, Levi, & Pisoni, 2010; Soto-Faroco et al., 2007). Most likely, visual identification of languages is based on extracting both lexical and rhythmic information from visual speech signals (Ronquest et al., 2010). Interestingly, 4-month old monolingual and bilingual infants are able to discriminate spoken languages visually (Weikum et al., 2007). This skill, however, disappears in monolingual, but not bilingual, infants by the age of 8 months, suggesting that bilingualism enhances visual discrimination abilities in infancy (Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012; Weikum et al., 2007). Nevertheless, both monolingual and bilingual adults are typically able to extract visual cues from their native or non-native languages and discriminate languages visually (Ronquest et al., 2010; Soto-Faroco et al., 2007).

* Corresponding author.

E-mail address: riikka.mottonen@psy.ox.ac.uk (R. Möttönen).

Numerous neuroimaging studies have shown that viewing a speaker's articulatory gestures activates the superior temporal cortex and the inferior frontal and premotor regions (Calvert & Campbell, 2003; Hall, Fussell, & Summerfield, 2005; Paulesu et al., 2003; Pekkola et al., 2006; Turner et al., 2009). Importantly, viewing linguistically meaningless mouth movements, i.e., gurns, does not activate these fronto-temporal regions as strongly as viewing speech movements (Campbell et al., 2001; Hall et al., 2005; Turner et al., 2009). The frontal activity has been suggested to reflect involvement of the articulatory motor system in speechreading (e.g., Campbell et al., 2001; Paulesu et al., 2003).

The role of the articulatory motor system in speech perception is under active investigation and its importance is under debate (Hickok, 2010; Hickok, Houde, & Rong, 2011; Pulvermüller & Fadiga, 2010; Scott, McGettigan, & Eisner, 2009). One of the central claims of the motor theory of speech perception is that speech perception and production are tightly linked (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Lieberman & Mattingly, 1985). According to this theory the speaker's intended articulatory gestures (or "neuro-motor commands") are detected from the acoustic speech signal. This inverse (from sensory to articulatory-motor) modeling has been hypothesized to activate the motor brain areas that are important for controlling articulatory movements. Indeed, such activations have been found in some neuroimaging studies during listening to speech (Pulvermüller et al., 2005; Wilson, Saygin, Sereno, & Iacoboni, 2004). Interestingly, listening to non-native speech sounds activates the motor brain regions more strongly than listening to native speech sounds, whereas native speech sounds elicit stronger activity in the superior temporal regions (Callan, Jones, Callan, & Akahane-Yamada, 2004; Wilson & Iacoboni, 2006). This suggests that enhanced articulatory-motor processing complements auditory processing of non-native speech sounds that are more ambiguous than native speech sounds. Given that access to articulatory information is more direct during visual than auditory speech perception, it is plausible that the articulatory motor cortex is engaged in visual speech perception, perhaps even more strongly than in auditory speech perception. No previous studies have investigated differences in articulatory-motor processing of native and non-native visual speech.

Transcranial Magnetic Stimulation (TMS) provides a powerful tool to investigate the excitability of the articulatory motor system during speech perception (for a review, see Möttönen and Watkins 2012). Previous studies using TMS have shown that viewing and listening to speech enhances excitability of the lip representation in the left primary motor cortex (M1) (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Murakami, Restle, & Ziemann, 2011; Sundara, Namasivayam, & Chen, 2001; Watkins, Strafella, & Paus, 2003). These studies found that the excitability of the articulatory motor system is higher during observation of visual speech than during observation of visual noise (Murakami et al., 2011; Watkins et al., 2003), a fixation dot (Murakami et al., 2011), eye and brow movements (Watkins et al., 2003), or lateral jaw movements (Murakami et al., 2011). These control conditions either involved viewing non-biological stimuli or biological movements that are not performed by using the articulators. Therefore, the specificity of the enhanced excitability in the articulatory motor cortex during visual speech perception is unclear. If this enhanced excitability is speech-specific, viewing speech-related lip movements should excite the articulatory motor cortex more strongly than non-speech lip movements (i.e., gurns). Furthermore, if the articulatory motor cortex is involved in linking perceived articulatory gestures to internal motor codes and extracting linguistic cues from visual speech, excitability should be higher during observation of speech movements related to a known language than to a foreign language, which the observer is not experienced in producing and, which does not convey linguistic information to the observer.

In the present study, we used behavioral tasks and TMS to investigate processing of visual speech in native and non-native speakers of English. Using behavioral tasks, we measured visual language discrimination (English vs. Hebrew) and speechreading (English) skills. We aimed to determine whether our participants are able to discriminate languages visually and how language background (i.e., English as L1 or L2) influences visual speech perception skills. We also used TMS to examine the excitability of the articulatory motor cortex during observation of a known language (i.e., English), an unknown language (i.e., Hebrew), gurns and a still face. As a control, we examined the excitability of the hand motor cortex. The main aim of the TMS experiments was to examine the specificity of the excitability changes in the articulatory motor cortex. We hypothesized that the excitability in the articulatory motor cortex, but not in the hand motor cortex, is higher during observation of a known spoken language than an unknown spoken language or non-speech mouth movements. Furthermore, we aimed to determine whether the modulations of the motor excitability differ between native and non-native English speakers.

2. Materials and methods

2.1. Participants

Forty six healthy, right-handed adults participated in the study. Data from four additional participants were excluded from analysis due to unreliable MEP or artefacts in the EMG recordings. Handedness was tested using the Edinburgh Handedness Inventory (Oldfield, 1971).

Twenty four participants were native British English speakers (nine men); their mean age was 22 years (18–33 years). Native speakers were defined in this study according to the following criteria: (a) both parents spoke English at home and (b) the primary language in which they received their education up to the age of 18 years was English. Twelve native speakers were assigned to the Lip experiment and 12 were assigned to the Hand experiment.

Twenty two participants were non-native English speakers (nine men); their mean age was 26 years (20–35 years). They spoke 15 different languages as their native language. The non-native group was defined as (a) both parents spoke a language other than English at home, (b) at least until the age of 18 years, the participant had lived outside the United Kingdom in a country where English is not the dominant language, (c) the primary language in which they received education until the age of 18 years was not English. However, non-native English speakers were recruited from the UK and were, therefore, proficient enough in English to engage in employment or full time study in the UK. Average time spent as resident in the UK was 44 months. On average the non-native speakers had started learning English at the age of 11 years and estimated having become fluent at the age of 19 years. Nine non-native speakers were assigned to the Lip experiment and 13 were assigned to the Hand experiment.

All participants reported being unfamiliar with Hebrew and languages related to Hebrew such as Arabic. All participants had normal hearing and normal or corrected to normal eyesight and were screened prior to participation for contraindications to TMS. The study was performed under permission from the National Research Ethics Service.

2.2. Electromyography

In all participants electromyography (EMG) activity from the orbicularis oris (OO) muscle was recorded using two surface electrodes (22 × 30 mm ARBO neonatal electrocardiogram electrodes) attached on the right corners of the lower and upper lip. EMG

was also recorded from the first dorsal interosseus (FDI) muscle in the right hand using a pair of electrodes placed on the belly and tendon of the muscle. The raw EMG signal was amplified (gain: 1000), bandpass filtered (10–1000 Hz), and sampled (5000 Hz) using an amplifier, an analog-to-digital converter and a computer running Spike2 software (version 3, Cambridge Electronic Design).

2.3. Single pulse transcranial magnetic stimulation

Monophasic TMS pulses were generated by a Magstim 200 stimulator and applied using a 70-mm figure-eight coil (Magstim Co.). The lip area of left M1 was stimulated in the Lip experiment, and hand area of left M1 in the Hand experiment. The position of the coil was adjusted until a robust motor evoked potential (MEP) was observed in the contralateral target muscle. During the experiment the inter-pulse interval was 5–8 s. In this study, we applied single TMS pulses over the motor cortex simply to elicit MEPs in the target muscle. The MEPs provide a direct measure of the excitability of the motor pathways connecting the muscles with their cortical representations. The MEPs are large, when the motor excitability is high.

In the Hand experiment, stimulation intensity was determined as the lowest intensity that elicited robust MEPs (with mean amplitude of approximately 1 mV) in the resting FDI muscle in 10 consecutive trials. In the Lip experiment, stimulation intensity was set as the lowest intensity eliciting robust MEPs (with amplitude of at least 0.2 mV) in 10 consecutive trials in the resting OO muscle. The mean stimulator intensity (\pm SEM) used to elicit MEPs was 53.7% (\pm 1.57) in the Hand experiment and 65.3% (\pm 1.7) in the Lip experiment (of the maximum stimulator output).

2.4. Stimuli in TMS experiments

All videos were filmed with a female bilingual (English and Hebrew) speaker. In all clips the speaker's face from the base of her

neck to the top of her head were shown (see Fig. 1). Both Lip and Hand experiments included four different stimulus types: (1) *Known Speech* (i.e., English sentences), (2) *Unknown Speech* (i.e., Hebrew sentences), (3) *Gurns* (i.e., sequences of non-speech movements performed by the lips; the sequences ranged from having a single mouth movement repeated over to five different non-speech movements without repetition), (4) *Still mouth* (i.e., video clips of a speaker being silent; these clips included some natural eye blinks). The English sentences were taken from the Bamford–Kowal–Bench standard sentence lists (Bench, Kowal, & Bamford, 1979); these sentences were also translated to Hebrew. During the experiment 40 different stimuli belonging to Known speech, Unknown speech and Gurns conditions were presented. Also, five different Still mouth videos were presented eight times. The clips ranged between 2.5 and 3.5 s in duration with average clip durations matched across the four conditions. Although natural audio-visual speech was recorded, all videos were presented without the audio track.

The videos were presented in eight blocks of 20 clips on a 14-in. Dell monitor using Presentation software (NeuroBehavioral Systems). Each block included five videos from each of the four conditions presented in random order. One TMS pulse was delivered 1.5–2.5 s after the onset of each video. A fixation cross was presented for 3.5 s between the videos.

Participants were instructed to pay attention to the speaker's mouth and keep their lips and hand relaxed during the TMS experiment. They were also told that the model would be speaking English and another language.

2.5. Visual language discrimination task

Eighty videos of the speaker speaking English and Hebrew sentences were presented on the Dell monitor using Presentation software. The videos were the same as used during the TMS experiment. Participants were instructed to click the left mouse

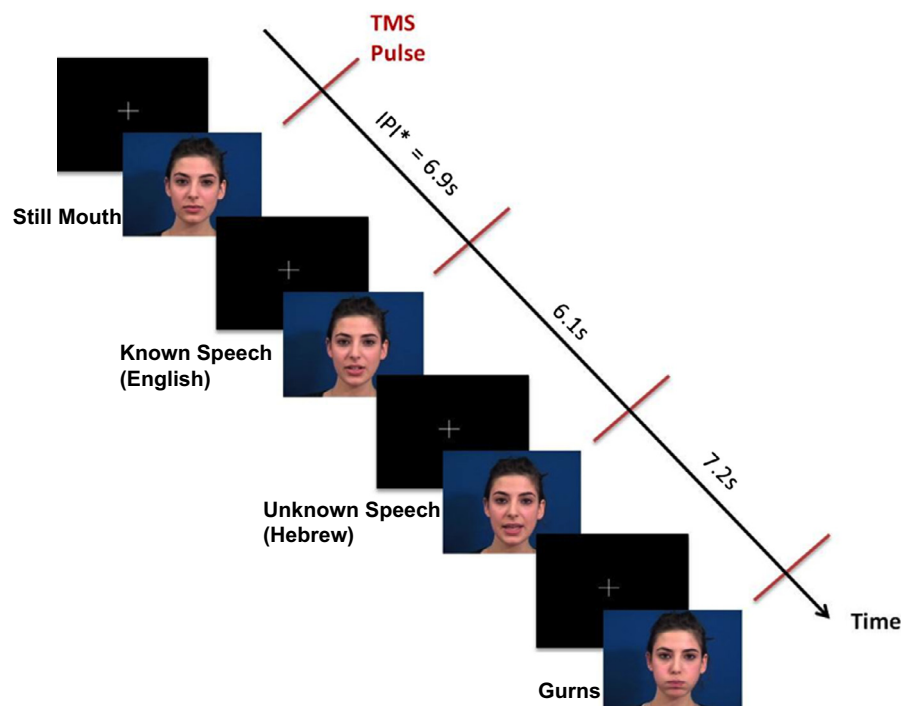


Fig. 1. Design of the TMS experiment. Four types of videos were presented in random order: still mouth, known speech, unknown speech and gurns. During each video, one TMS pulse delivered either over the lip or hand representation in the left motor cortex. Inter-pulse-interval (IPI) varied randomly between 5 and 8 s.

button if they thought the language spoken on a video was English and the right mouse button if they thought it was another language. There was an interval of 3.5 s between the videos.

2.6. Language Experience and Proficiency Questionnaire (LEAP-Q)

The LEAP-Q (Marian, Blumenfeld, & Kaushanskaya, 2007) is a questionnaire designed to acquire information about linguistic background and is suitable for use with monolinguals and bilinguals. For example, information regarding number of years spent in the UK and age of acquisition of L1 and L2 was obtained.

2.7. Reading test

The Vernon-Warden revised reading test (also known as the Kirklees reading test; Hedderly, 1996) was administered to all participants as a measure of English language proficiency. The test contains incomplete sentences that must be completed by selecting the correct word from a list of options. A time limit of 10 min is set.

2.8. Speechreading test

The Test of Adult Speechreading (TAS; Mohammed, Campbell, MacSweeney, Barry, & Coleman, 2006; Mohammed et al., 2005) is a standardized English speechreading test. During the test, participants are tested on speechreading accuracy for English words, sentences and stories. The maximum speechreading score is 45.

2.9. Procedure

After giving informed consent and filling the screening forms, participants first completed the Vernon Warden reading test and the LEAP-Q. Once these measures were completed, EMG electrodes were attached and the participants completed the TMS experiment. Then, participants completed the visual language discrimination task and the TAS.

2.10. Behavioral data analysis

Performance on the visual language discrimination task was analyzed using signal detection principles. The d' measure was calculated for each participant and one-sample t -tests was used to test if these values differed significantly from zero (i.e., chance performance). In addition, an independent-samples t -test was run to determine whether language discrimination ability differed between native and non-native English speakers.

Independent samples t -tests were run to determine differences between the native and non-native English speakers in speechreading ability measured using the TAS, and reading ability measured using the Vernon-Warden test. Finally, correlation analysis was run on the d' scores with the TAS and reading scores.

2.11. TMS data analysis

Peak-to-peak amplitude was used as the measure of MEP size. As expected, lip MEPs had shorter onset latency than the hand MEPs because the corticobulbar pathway to the lip muscle is shorter than corticospinal pathway to the hand muscle (Möttönen & Watkins, 2012). Amplitudes of lip MEP were calculated using the time window of 0.01–0.035 s after each TMS pulse. A time window of 0.02–0.04 s after the pulse was used for the hand MEPs.

For each participant, mean MEPs were calculated for each condition. Only MEPs that fell within two standard deviations of the condition mean were included in the analysis and mean MEPs were recalculated. The overall mean and standard deviations of MEPs across conditions was also calculated. MEPs in each condi-

tion were standardized to this overall mean, yielding a z -score for each condition.

Statistical analyses were carried out using SPSS (version 17.0, IBM). A three-way analysis of variance (ANOVA) was run on the z -scores with Stimulus type (Known speech, Unknown speech, Gurns, Still mouth) as a within-subjects factor, and Experiment (Lip, Hand) and Language group (Native, Non-native English speakers) as between subject factors. We also ran separate two-way ANOVAs for Lip and Hand experiments. Post hoc paired t -tests comparing different stimulus types were also conducted.

3. Results

3.1. Behavioral results

Reading, speechreading and language discrimination scores of native and non-native speakers of English are presented in Fig. 2. The native English speakers performed better than non-native speakers in the reading assessment ($t(44) = 6.29, p < .001$) and in the speechreading test ($t(44) = 2.63, p < .05$). There was no significant difference between the performance of native and non-native speakers in the language discrimination task ($t(44) = 1.70, p = .1$).

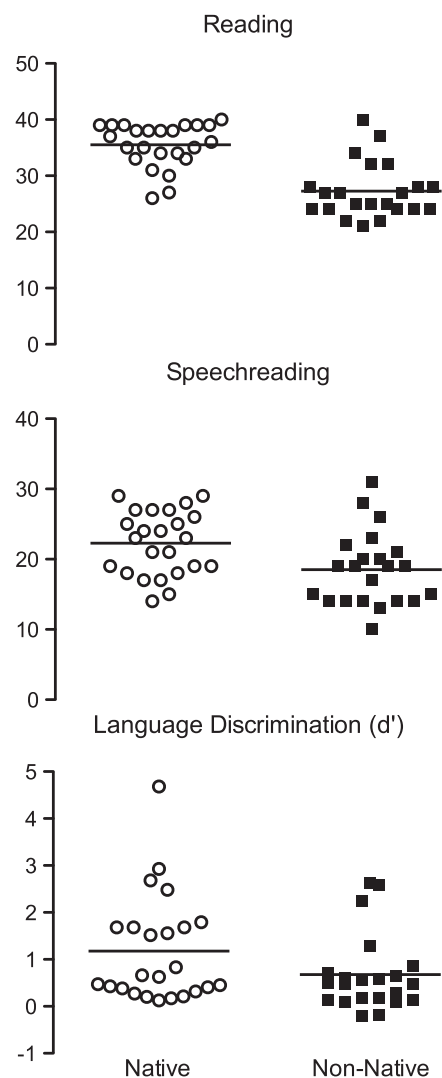


Fig. 2. Behavioral results. Performance of native ($n = 24$) and non-native ($n = 22$) English speakers in three behavioral tasks: reading, speechreading and language discrimination.

although there was a trend towards the native speakers being more accurate than non-native speakers. Both native ($t(23) = 5.07, p < 0.001$) and non-native ($t(21) = 3.90, p < 0.05$) speakers of English performed better than chance in this task, indicating that the participants were able to discriminate English from Hebrew visual speech. There were no differences in language discrimination, speechreading and reading abilities between participants in the Lip and Hand experiments.

There was a significant positive correlation between visual language discrimination ability and speechreading ability both in the native English group ($r(22) = 0.53, p = 0.008$) and in the non-native group ($r(20) = 0.642, p = .001$). Reading ability did not correlate with language discrimination or with speechreading ability in either group, or when the groups were combined.

3.2. TMS results

Fig. 3 presents the standardized MEP sizes, i.e., z-scores. A three-way ANOVA was conducted: Stimulus type, Experiment and Language group. This revealed a significant Stimulus type * Experiment (Lip vs. Hand) interaction ($F(3,126) = 7.98, p < .001$ (Greenhouse–Geisser corrected)). No significant main effects or interactions involving Language group (Native vs. Non-native) were found.

A two-way ANOVA for the Lip experiment showed a significant main effect of Stimulus type ($F(3,57) = 4.89, p < .01$). The main effect of Language group and the interaction between Stimulus type and Language group were non-significant. This indicates that the Stimulus type modulated the excitability of the motor lip representation in native and non-native speakers of English in a similar

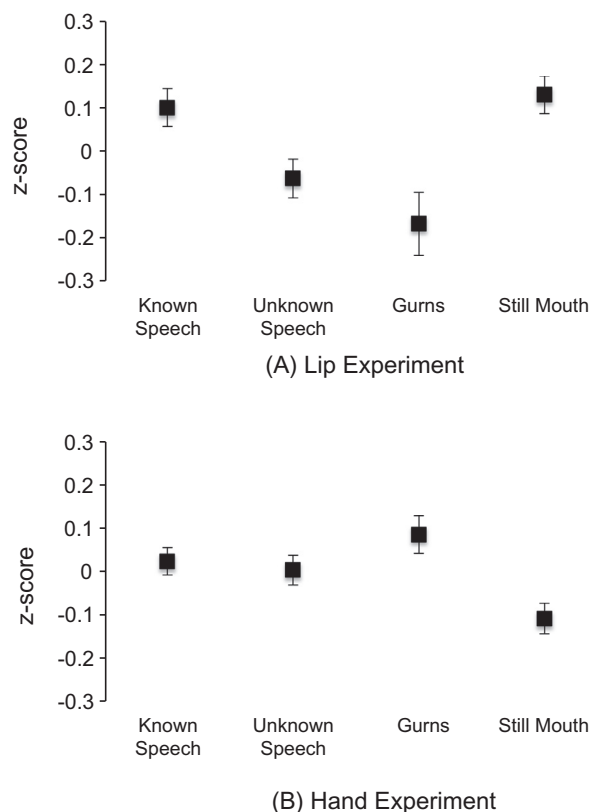


Fig. 3. TMS results. Standardized amplitudes of MEPs measured from the lip (A) and hand (B) muscles during observation of known speech (i.e., English), unknown speech (i.e., Hebrew), gurns and still mouth. Differences in MEP amplitudes between conditions reflect differences in the excitability of the lip and hand representations in the motor cortex.

manner. Therefore, we combined these groups in following analyses. Excitability of the motor lip area was higher when the participants observed Known speech (English) than when they observed Unknown speech (Hebrew) ($t(20) = 4.06, p = .001$) or Gurns ($t(20) = 2.34, p < .05$). The Unknown speech and Gurns did not differ from each other. The motor excitability was higher during observation of Still mouth than during observation of Unknown speech ($t(20) = 2.51, p < .05$) or Gurns ($t(20) = 3.17, p < .01$). Still mouth and Known speech did not differ from each other.

A two-way ANOVA for the control Hand experiment also showed a significant main effect of Stimulus type ($F(3,69) = 3.66, p < .05$). The main effect of Language group and the interaction between Stimulus type and Language group were non-significant. This indicates that the stimuli modulated the excitability of the motor hand representation in native and non-native speakers of English in a similar manner. Therefore, we combined these groups in following analyses. The excitability of the motor hand representation did not differ between Known speech, Unknown speech and Gurns. However, the excitability was lower during observation of Still mouth than during observation of Gurns ($t(24) = -2.91, p < .01$) or Unknown speech ($t(24) = -2.73, p < .05$).

4. Discussion

The behavioral results indicate that both native and non-native English speakers were able to visually discriminate languages at a level significantly higher than chance. This replicates previous findings (Ronquest et al., 2010; Soto-Faroco et al., 2007) using a new language pair: English and Hebrew. Differences in speech rhythms are likely to be an important cue in visual language discrimination. Ronquest et al. (2010) showed that it is possible to discriminate between stress-timed (i.e., English) and syllable-timed (i.e., Spanish) languages using rhythmic information. English and Hebrew are both stress-timed languages, but the dominant stress patterns are different. English is a trochaic language (i.e., strong–weak pattern is most frequent), whereas Hebrew is an iambic language (i.e., weak–strong pattern is most frequent) (see e.g., Segal, Nir-Sagiv, Kishon-Rabin, & Ravid, 2009). It is possible that the participants used subtle differences in speech rhythms to discriminate these two languages.

Our findings also demonstrate that visual language discrimination ability was positively correlated with English speechreading ability. This suggests that both tasks measure the ability to extract linguistic cues from visual speech. Native speakers performed significantly better than the non-native speakers on the English speechreading test. However, the groups did not differ significantly on the visual language discrimination task, but nonetheless there was a trend in the same direction, i.e., native speakers performed better than non-native speakers. As expected, the native speakers performed better in the reading test that assessed English language proficiency. Interestingly, however, reading ability did not correlate with accuracy in either visual speech perception task. This suggests that some non-native speakers, who had relatively low language proficiency, were relatively good in extracting visual cues from English visual speech. In sum, the behavioral results demonstrated that both native and non-native English speakers were able to extract visual cues from English speech.

The main aim of the TMS experiments was to examine the specificity of the excitability changes in the articulatory motor cortex during observation of a speaker's face. We found that the lip MEPs were smaller during observation of non-speech mouth movements (i.e., gurns) than articulatory movements related to a known language (i.e., English). This finding supports the hypothesis that the enhanced excitability of the articulatory motor system is speech-specific. The result is in line with the results of Murakami et al.

(2011) who found enhanced lip MEPs during speech observation relative to observation of lateral jaw movements. This result is also in line with findings from imaging studies showing increased inferior frontal activations when viewing speech compared to gurns (Hall et al., 2005; Turner et al., 2009). Also, as hypothesized, the lip MEPs were significantly larger during visual perception of a known than an unknown language, providing further evidence for the specificity of the enhanced excitability in the articulatory motor cortex. There are several, not necessarily independent, explanations for this important finding. First, it is possible that viewing a familiar speech rhythm contributed to the enhanced excitability of the articulatory motor cortex during observation of English visual speech. Second, it has been suggested that motor activation during perception is established by association learning (Heyes, 2010). During speech production we can hear (and sometimes see) our own speech, thereby establishing an association between production and perception. Such an association can be established only if the appropriate experience is available. All participants in the present study had experience in producing and perceiving English, but not Hebrew. The difference in the excitability of the articulatory motor cortex may, therefore, reflect differences in the strength of sensorimotor associations for English and Hebrew speech. We propose that the enhanced sensorimotor resonance during visual perception of a known language facilitates linguistic processing. However, further studies, using for example repetitive TMS to disrupt functioning of the articulatory motor cortex, are needed to investigate the causal role of the sensorimotor processes in speechreading.

Importantly, the control experiment demonstrated that the excitability of the hand motor cortex did not differ during observation of known speech, unknown speech and gurns. This indicates that changes in motor excitability are somatotopic during visual speech perception and is in agreement with previous studies (Murakami et al., 2011; Watkins et al., 2003). Observation of communicative hand movements has been shown to enhance excitability in the hand motor cortex in the left hemisphere (Möttönen, Farmer, & Watkins, 2010).

Surprisingly, no significant differences in the excitability of the articulatory motor cortex were found between viewing the still mouth and English speech. The excitability of the articulatory motor cortex was higher when viewing the still mouth than gurns and unknown speech. Furthermore, the excitability of the hand motor cortex was lower when viewing a still mouth than speech and gurns. These results were unexpected in the light of previous imaging studies that found stronger activity in the inferior frontal cortex during observation of speech than a still face (Campbell et al., 2001; Paulesu et al., 2003;). This discrepancy may be partly due to the differences in sensitivity to changes in motor excitability between fMRI and TMS (see Möttönen & Watkins, 2012). There are several possible explanations for enhanced excitability during observation of a still face. First, in the current experiment it was difficult to predict whether the still face would start to speak, because the still face videos were randomized with the speech videos. During the initial 0.5 s of each speech video, the speaker remained still. Possibly, participants were anticipating speech when seeing the still face and, therefore, the excitability of the articulatory motor cortex was enhanced. Second, the still face is a communicative signal. In the monkey brain, there are neurons in the ventral prefrontal cortex that respond more strongly to faces than to non-faces and code gaze, expressions and identity information (e.g., Tsao, Schweers, Moeller, & Freiwald, 2008; Romanski & Diehl, 2011). These neurons have been suggested to be important for social communication (Romanski, 2012). In humans, another person's still face conveys information that is important for turn taking during face-to-face conversations. Stillness of a conversational partner usually indicates one's turn to speak. It is possible

that this association between seeing a still face and speech production underlies the enhancement of motor excitability. Thus, our results are in line with the idea that the sensorimotor processes are important for controlling conversations (Scott, McGettigan, & Eisner, 2009).

Our TMS results revealed no differences in motor excitability between native and non-native English speakers during visual perception of spoken languages. The results suggest that the articulatory motor system is engaged in a similar manner during visual perception of spoken language in native and non-native speakers. The lack of differences in motor excitability may seem surprising in the light of behavioral results, showing a native-speakers advantage in the speechreading task and a similar (non-significant) trend in the language discrimination task. There are several possible explanations for the lack of difference in motor excitability between the groups. First, it should be noted that the groups did not differ significantly in their ability to discriminate English and Hebrew sentences and that this task was difficult for both groups. Although the performance was better than chance in both native and non-native groups, it was still rather poor in both groups (mean d' values of around 1). Thus, the stimuli used in the TMS experiment were equally difficult for native and non-native speakers. Second, in the current study the motor excitability was measured during passive observation of the stimuli, not during active visual speech perception tasks. Further studies are needed to examine whether motor excitability is higher during active speechreading than passive viewing of a speaker's articulatory movements and whether motor excitability would differ between native and non-native speakers during active speechreading. Third, a large fronto-temporal network of brain areas is involved in speechreading (Calvert & Campbell, 2003; Hall et al., 2005; Paulesu et al., 2003; Pekkola et al., 2006; Turner et al., 2009). The articulatory motor cortex is only one node in this network and it is possible that other areas are more sensitive to differences in speechreading ability. For example, activity in the left superior temporal gyrus and posterior cingulate cortex during visual perception of spoken sentences has been shown to positively correlate with speechreading scores (Hall et al., 2005). Fourth, a larger sample size may be needed to reveal subtle differences in motor excitability between native and non-native speakers. In sum, although the present study did not reveal any differences in motor excitability between native and non-native speakers during visual perception of spoken languages, further studies are needed to investigate the effect of language background on involvement of the articulatory motor system in speech perception. For example, studies investigating perception of auditory, visual and audiovisual speech signals in a large number of participants that differ in their age of acquisition and proficiency of the second language are needed.

In summary, our findings show that visual perception of a face speaking a known language or a face not speaking enhances excitability of the articulatory motor cortex relative to visual perception of a face that speaks an unknown language or performs non-speech movements with his/her mouth. Thus, visual signals that the perceiver is experienced in producing and perceiving during face-to-face conversations enhance excitability in the articulatory motor cortex. The enhanced sensorimotor resonance during visual perception of known speech (relative to unknown speech and gurns) is likely to support extraction of linguistic (e.g., phonetic and rhythmic) cues. On the other hand, the relatively high motor excitability during observation of a face not speaking suggests that the articulatory motor cortex also processes non-linguistic visual cues during speech communication. The speaker's silent face is an important communicative cue that can be used, for example, to control turn taking in conversations. Altogether, the findings suggest that sensorimotor processes may have multiple functions in

visual perception of a speaker's face during face-to-face communication. Future studies should aim to further characterize how the articulatory motor cortex contributes to extraction of both linguistic and non-linguistic cues from the speaker's face.

Acknowledgments

We thank Lialin Rodem-Stibbe for acting as a bilingual speaker and translating the English sentences to Hebrew. The study was funded by MRC Career Development Fellowship to R.M., by Wellcome Trust Career Development Fellowship to M.M. and by Wellcome Trust Project Grant to K.E.W and R.M.

References

- Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, *13*(3), 108–112.
- Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage*, *22*(3), 1182–1194.
- Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: The neural substrates of visual speech. *Journal of Cognitive Neuroscience*, *15*(1), 57–70.
- Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G. A., MacGuire, P., Suckling, J., et al. (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). *Cognitive Brain Research*, *12*, 233–243.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*(2), 399–402.
- Hall, D. A., Fussell, C., & Summerfield, A. Q. (2005). Reading fluent speech from talking faces: Typical brain networks and individual differences. *Journal of Cognitive Neuroscience*, *17*(6), 939–953.
- Hardison, D. M. (1999). Bimodal speech perception by native and nonnative speakers of English: Factors influencing the McGurk effect. *Language Learning*, *49*, 213–283.
- Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Journal of Applied Psycholinguistics*, *24*, 495–522.
- Hardison, D. M. (2005). Variability in bimodal spoken language processing by native and nonnative speakers of English: A closer look at effects of speech style. *Speech Communication*, *47*, 73–93.
- Hazan, V., Sennema, A., Faulkner, A., Ortega-Llebaria, M., Iba, M., & Chung, H. (2006). The use of visual cues in the perception of non-native consonant contrasts. *The Journal of the Acoustical Society of America*, *119*, 1740–1751.
- Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audio-visual perceptual training on the perception and production of consonants in Japanese learners of English. *Speech Communication*, *47*, 360–378.
- Hedderly, R. G. (1996). Vernon Warden reading test restandardised 1993 and 1994. *Dyslexia Review*, *7*(2).
- Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience & Biobehavioral Reviews*, *34*(4), 575–583.
- Hickok, G. (2010). The role of mirror neurons in speech and language processing. *Brain and Language*, *112*(1), 1–2.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*(3), 407–422.
- Hirata, Y., & Kelly, S. D. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. *Journal of Speech, Language, and Hearing Research*, *53*, 298–310.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461.
- Liberman, A., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.
- MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception. *Perception & Psychophysics*, *24*(3), 253–257.
- Macleod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology*, *24*(1), 29–43.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech and Hearing Research*, *50*(4), 940–967.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *263*, 746–748.
- Mohammed, T., Campbell, R., MacSweeney, M., Barry, F., & Coleman, M. (2006). Speechreading and its association with reading among deaf, hearing and dyslexic individuals. *Clinical Linguistics & Phonetics*, *20*(7–8), 621–630.
- Mohammed, T., MacSweeney, M., Campbell, R., Hansen, P., Coleman, M., & Milne, E. (2005). Visual motion coherence detection: A perceptual correlate of speechreading skill. *Perception*, *34*, 205–216.
- Möttönen, R., Farmer, H., & Watkins, K. E. (2010). Lateralization of motor excitability during observation of bimanual signs. *Neuropsychologia*, *48*(10), 3173–3177.
- Möttönen, R., & Watkins, K. E. (2012). Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology*, *26*(9), 1103–1118.
- Murakami, T., Kestle, J., & Ziemann, U. (2011). Observation-execution matching and action inhibition in human primary motor cortex during viewing of speech-related lip movements or listening to speech. *Neuropsychologia*, *49*(7), 2045–2054.
- Navarra, J., & Soto-Faroco (2007). Hearing lips in a second language: Visual articulatory information enables the perception of second language sounds. *Psychological Research Psychologische Forschung*, *71*, 4–12.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.
- Paulesu, E., Perani, D., Blasi, V., Silani, G., Borghese, N. A., De Giovanni, U., et al. (2003). A functional-anatomical model for lipreading. *Journal of Neurophysiology*, *70*, 2005–2013.
- Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., & Sams, M. (2006). Attention to visual speech gestures enhances hemodynamic activity in the left planum temporale. *Human Brain Mapping*, *27*(6), 471–477.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, *11*(5), 351–360.
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado, F., Hauk, O., & Shtyrov, Y. (2005). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, *103*(20), 7865–7870.
- Romanski, L. M. (2012). Integration of faces and vocalizations in ventral prefrontal cortex: Implications for the evolution of audiovisual speech. *Proceedings of National Academy of Sciences*, *109*, 10717–10724.
- Romanski, L. M., & Diehl, M. M. (2011). Neurons responsive to face-view in the primate ventrolateral prefrontal cortex. *Neuroscience*, *189*, 223–235.
- Ronquest, R. E., Levi, S. V., & Pisoni, D. B. (2010). Language identification from visual-only speech signals. *Attention, Perception & Psychophysics*, *72*(6), 1601–1613.
- Scott, S. K., McGettigan, C., & Eisner, F. (2009). A little more conversation, a little less action – Candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*, *10*, 295–302.
- Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, *23*(9), 994–999.
- Segal, O., Nir-Sagiv, B., Kishon-Rabin, L., & Ravid, D. (2009). Prosodic patterns in Hebrew child-directed speech. *Journal of Child Language*, *35*, 629–656.
- Soto-Faroco, S., Navarra, J., Weikum, W. M., Vouloumanos, A., Sebastian-Galles, N., & Werker, J. F. (2007). Discriminating languages by speech-reading. *Perception & Psychophysics*, *69*(2), 218–231.
- Sumbly, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *26*(2), 212–215.
- Sundara, M., Namasivayam, A. K., & Chen, R. (2001). Observation-execution matching system for speech: A magnetic stimulation study. *NeuroReport*, *12*(7), 1341–1344.
- Tsao, D. Y., Schweser, N., Moeller, S., & Freiwald, W. A. (2008). Patches of face-selective cortex in the macaque frontal lobe. *Nature Neuroscience*, *11*(8), 877–879.
- Turner, T. H., Fridriksson, J., Baker, J., Eoute, D., Jr., Bonilha, L., & Rorden, C. (2009). Obligatory Broca's area modulation associated with passive speech perception. *NeuroReport*, *20*(5), 492–496.
- Wang, Y., Behne, D. M., & Jiang, H. (2008). Linguistic experience and audio-visual perception of non-native fricatives. *The Journal of the Acoustical Society of America*, *124*(3), 1716–1726.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989–994.
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, *316*, 1159.
- Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying producibility: Evidence for the sensorimotor nature of speech perception. *NeuroImage*, *33*(1), 316–325.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), 701–702.