# Title: The role of hearing ability and speech distortion in the facilitation of articulatory motor cortex

Helen E. Nuttall[a,b], Daniel Kennedy-Higgins[b], Joseph T. Devlin[c], Patti Adank[b]

a. Department of Psychology, Lancaster University, Lancaster, LA1 4YW

b. Department of Speech, Hearing and Phonetic Sciences, University College London, Chandler House, 2 Wakefield Street, London, UK, WC1N 1PF

c. Department of Experimental Psychology, University College London, 26 Bedford Way, London, UK, WC1H 0AP

Correspondence should be addressed to Helen Nuttall, Department of Psychology, Lancaster University, Lancaster, LA1 4YW. E-mail: h.nuttall1@lancaster.ac.uk. Tel: +441524592842

**Abstract**

Excitability of articulatory motor cortex is facilitated when listening to speech in challenging conditions. Beyond this, however, we have little knowledge of what listener-specific and speech-specific factors engage articulatory facilitation during speech perception. For example, it is unknown whether speech motor activity is independent or dependent on the form of distortion in the speech signal. It is also unknown if speech motor facilitation is moderated by hearing ability. We investigated these questions in two experiments. We applied transcranial magnetic stimulation (TMS) to the lip area of primary motor cortex (M1) in young, normally hearing participants to test if lip M1 is sensitive to the quality (Experiment 1) or quantity (Experiment 2) of distortion in the speech signal, and if lip M1 facilitation relates to the hearing ability of the listener. Experiment 1 found that lip motor evoked potentials (MEPs) were larger during perception of motor-distorted speech that had been produced using a tongue depressor, and during perception of speech presented in background noise, relative to natural speech in quiet. Experiment 2 did not find evidence of motor system facilitation when speech was presented in noise at signal-to-noise ratios where speech intelligibility was at 50% or 75%, which were significantly less severe noise levels than used in Experiment 1. However, there was a significant interaction between noise condition and hearing ability, which indicated that when speech stimuli were correctly classified at 50%, speech motor facilitation was observed in individuals with better hearing, whereas individuals with relatively worse but still normal hearing showed more activation during perception of clear speech. These findings indicate that the motor system may be sensitive to the quantity, but not quality, of degradation in the speech signal. Data support the notion that motor cortex complements auditory cortex during speech perception, and point to a role for the motor cortex in compensating for differences in hearing ability.

Key words: Speech perception, motor cortex, transcranial magnetic stimulation, motor evoked potentials.

## 1. Introduction

Successful speech perception is central to everyday communication and quality of life. It is therefore surprising that understanding of the neural bases underpinning speech perception remains limited. Although auditory-related areas are thought to be at the heart of the neural architecture for understanding speech, there is accumulating evidence that areas extending beyond primary and association auditory cortices are important for successful speech perception. Cortical regions including, but not limited to, ventral premotor cortex, inferior frontal gyrus, and supplementary and primary motor areas have also been suggested to be involved in speech perception (Adank et al., 2012; Callan et al., 2010; D'Ausilio et al., 2009; Londei et al., 2010; Skipper et al., 2005; Tremblay et al., 2012). Indeed, it is now largely accepted that articulatory motor areas are active when we perceive speech (Bartoli et al., 2015; Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Möttönen & Watkins, 2009; Wilson, Saygin, Sereno, & Iacoboni, 2004). Furthermore, the motor system does not seem to activate in a binary fashion when listening to either speech or non-speech; instead, excitability of articulatory motor regions during speech perception appears to be graded depending on the clarity of speech (Murakami et al., 2011). Murakami et al. (2011) demonstrated that lip motor evoked potentials (MEPs), elicited by transcranial magnetic stimulation (TMS) to the lip area of primary motor cortex (M1), are enhanced when perceiving speech-in-noise relative to perceiving speech without noise. This finding has been interpreted to reflect increased excitability in the cortical motor representation of the lips when listening to degraded speech.

These MEP findings are in line with behavioural changes that have been observed after receiving online TMS to primary lip and tongue areas. Paired-pulse TMS to M1 lip was found to lead to faster (facilitated) reaction times to lip-articulated stimuli in noise, and similarly for tongue-articulated stimuli following tongue stimulation, but with no change to reaction time when listening to speech in quiet (D'Ausilio et al., 2012). Similar findings have been shown for premotor cortex by Meister and colleagues (2007), who used 1 Hz repetitive TMS, which has been shown to result in inhibitory effects (Fitzgerald et al., 2006), to demonstrate that ventral premotor cortex (PMv) contributes to the perception of speech-in-noise. Crucially, Sato et al (2009) also used inhibitory 1 Hz repetitive TMS and found that a contribution from PMv was absent when speech was presented without background noise, indicating that speech perception must be challenged before PMv contributes to listening to speech. These TMS findings resonate with fMRI observations of increased motor cortex recruitment during comprehension of degraded speech (Hervais-Adelman et al., 2012; Osnes et al., 2011). Taken together, data indicate that the motor cortex is preferentially engaged when listening to speech that is difficult to perceive, and that motor activation may be necessary for successful speech perception under challenging listening conditions.

The precise function of observed motor activity during speech perception, however, remains under active debate (Hickok et al., 2011; Lotto et al., 2009; Scott et al., 2009). Recent theories suggest that motor activation may form the basis for the mental simulation of perceived action, which may aid listeners when predicting upcoming speech signals (Gambi & Pickering, 2013; Pickering & Garrod, 2013; Wilson & Knoblich, 2005). Simulation theories of action perception argue that observing actions results in the automatic generation of motor plans required to perform the actions. Simulated motor plans are then used to inform forward models about the co-ordination of one’s own muscles to generate a simulated course of movement in parallel with, or even in anticipation of, the movement being perceived (Grush, 2004). This type of forward model serves to anticipate others' actions as if they were produced by the observer (Locatelli et al., 2012), and may be used to disambiguate noisy, obscured, or ambiguous actions (Wilson and Knoblich, 2005). With regards to speech perception, these types of conditions may involve listening to speech in the presence of background noise, or listening to someone speaking in an unfamiliar accent (Adank et al., 2012; Adank and Janse, 2009) or manner of speaking (Borrie et al., 2013; Borrie and Schäfer, 2015).

Although it is well-established that perceiving speech draws upon hierarchically organized temporo-frontal processing pathways (Davis and Johnsrude, 2007, 2003), it is not clear what role premotor and primary motor regions play within this speech processing hierarchy. Knowledge of the nature of articulatory motor representations and their sensitivity to speech is incomplete. For example, it is unknown if, and how, increased motor excitability during perception of challenging speech is modulated by the nature and extent of the speech distortion. Accordingly, two possibilities currently exist for how the motor system responds to distortion in the speech signal. The first is that articulatory motor regions may activate whenever distortion is present in the speech signal, independent of the form or type of speech distortion. The second possibility is that articulatory motor regions may respond differently depending on the type of distortion in the speech signal. If the former is true, it would suggest that the motor system acts as a self-adjusting resource to provide additional information whenever auditory information is found to be insufficient. Support for this prediction comes from demonstrations of heightened motor excitability for both speech-internal distortion (Nuttall, Kennedy-Higgins, Hogan, Devlin, & Adank, 2016) and speech-external distortion (Murakami et al., 2011), yet these two sources of distortion have never been directly compared.

Conversely, if the second possibility is true, and activity in motor regions is differentially modulated depending on the type of speech distortion, this indicates that the motor system operates in a form dependent manner during speech perception. Indeed, this is in line with the hypothesis that prediction signals generated by forward models during perception are ideally suited to disambiguate biological sources of variation (Sebanz et al., 2006); for example, when perceiving speech signals that are difficult to understand due to an unfamiliar manner of speech production. This possibility resonates with common-coding accounts of action perception, whereby the motor system is most responsive to observed actions that the observer has experience producing themselves (Calvo-Merino et al., 2005). In this case, the motor system's prediction signal would be less well-suited to assist action understanding when the difficulty arises from a non-motor source, such as speech-shaped background noise, for example, which does not constitute an imitable action. To date, whether, and how, motor facilitation is affected by speech distortion type, or extent of speech distortion, is unknown.

Moreover, the quality of the speech signal received by auditory cortex is at the heart of motor simulation. However, previous considerations of speech signal quality have been limited to the properties of the speech stimulus. This is not the only means by which speech signal quality can be degraded. The first point in the speech processing chain begins with the ear, where differences in mechanical and electrical function at the level of the cochlea and auditory nerve can contribute to discrepancies in how the speech signal is processed, even when individuals have clinically normal hearing (Bharadwaj et al., 2015; Harris et al., 2009; Ruggles et al., 2012). Accordingly, it is possible that differences in auditory processing at the ear modulate motor activity in a manner that is qualitatively similar to the effect of a degraded speech stimulus. Indeed, our previous study demonstrated a significant correlation between peripheral hearing acuity and the extent of lip MEP facilitation during distorted relative to clear speech perception, which was not present for hand MEPs (Nuttall, Kennedy-Higgins, Hogan, Devlin, & Adank, 2016). Precisely how hearing abilities relate to the extent of motor activation in optimal and suboptimal listening conditions has not been studied. Relatedly, Peelle and colleagues (2011) found that moderate declines in peripheral auditory processing led to a systematic down-regulation of neural activity in auditory regions during speech processing, and may also contribute to loss of gray matter volume in primary auditory cortex. If motor system activation is interlinked with speech signal quality, as motor simulation accounts would propose, it may be that hearing ability plays a role in engagement of the motor cortex during speech perception. A second aim of this study, therefore, was to explicitly test the relationship between hearing ability and speech motor excitability under different challenging listening conditions.

In the present study, we first aimed to disambiguate between form dependent and form independent accounts of how speech distortion modulates motor activation during speech perception, and second, we investigated whether normal variation in hearing ability impacts speech motor facilitation. To this end, in a first experiment, MEPs were elicited during perception of three different types of auditory stimuli: 1) clear speech stimuli presented without background noise, 2) speech stimuli distorted via a motor perturbation introduced during prior stimulus creation (motor distortion), and 3) speech-in-noise (noise distortion), where intelligibility was matched to the motor-distorted speech based on equating percent correct identification between the two degraded stimuli types. For speech-in-noise stimuli, clear speech stimuli were presented in a steady background of speech-shaped noise. In a second experiment, we recorded MEPs during perception of 1) clear speech, 2) speech-in-noise that was 75% intelligible, and 3) speech-in-noise that was 50% intelligible. The same clear speech and noise type from Experiment 1 were used in Experiment 2, but signal-to-noise ratio was varied. All speech stimuli were disyllabic vowel-consonant-vowel sounds containing a mixture of consonants that were either lip- (/aba/, /apa/) or tongue-articulated (/ada/, /ata/). Chance performance was therefore always 25%. Stimulation was thus used to determine whether motor facilitation in lip M1 is sensitive to the nature and extent of speech signal degradation. In both experiments, we also measured hearing sensitivity to examine whether hearing ability is related to motor facilitation when perceiving different types of speech distortion.

## 2. Methods

### 2.1 Subjects

In Experiment 1, eighteen subjects took part (eight males; average age: 22 years 8 months (± SD 3 months); age range: 18–28 years). Eighteen subjects also took part in Experiment 2 (seven males; average age: 22 years 6 months (± SD 3.8 months); age range: 18–30 years), two of whom had also participated in Experiment 1. All subjects in Experiments 1 and 2 were right-handed, monolingual, native speakers of British English, with normal language function and hearing thresholds. Handedness was established via self-report. Pure-tone audiometric hearing thresholds were established using a diagnostic audiometer (AD229b, Interacoustic A/S, Denmark) in accordance with The British Society of Audiology Recommended Procedure (The British Society of Audiology, 2011), across 0.5, 1, 2, and 4 kHz bilaterally. All subjects had clinically normal thresholds (≤20 dB HL). Subjects presented no TMS contraindications as assessed by the University College London TMS safety screening form, and did not report any neurologic/psychiatric disease, or that they were under the effect of neuroactive drugs. All subjects had a minimum high school-level education, with the majority currently studying at University level. Experiments were undertaken with the understanding and written consent of each subject, according to Research Ethics Board of University College London.

### 2.2 Speech stimuli

Stimuli used in Experiments 1 and 2 consisted of twenty vowel-consonant-vowel (VCV) syllables containing an equal distribution of lip- (/apa/, /aba/) or tongue-articulated (/ata/, /ada/) syllables. Two sets of the same twenty syllables were created: a set based on natural, normal articulation, and a set based on motor distortion, where the speaker produced the syllables whilst speaking with a tongue depressor. The tongue depressor was a flat wooden spatula with rounded ends, and was five inches long and one inch wide. Notably, the use of a tongue depressor was intended to impair perception of both tongue- and lip-articulated sounds due to the tongue depressor restricting both the tongue and lower lip movement. A tongue depressor was specifically chosen so as to introduce a motor-based distortion into the speech signal, to relate the speech perception challenge to a speech production difficulty (for further information about these stimuli, and how they are perceived, please see Nuttall et al., 2016). This enabled us to contrast clear speech against distorted speech produced by the same speaker, in contrast to imposing synthetic manipulations upon the spectral characteristics of the original clear speech.

All natural and motor-distorted speech stimuli were recorded in a sound-attenuated room and produced by a twenty-seven year old female British English speaker with a vocal pitch of 215 Hz. The same natural stimuli were used in Experiments 1 and 2. All stimuli were naturally produced to be of approximately the same duration (mean 975.25 ms) but were not synthetically manipulated to be precisely the same length. Stimuli varied by a standard deviation 60.77 ms. All stimuli were produced with natural falling intonation, with stress placed on the initial syllable. Audio digitizing was performed at 44.1 kHz with 16 bits. All syllables were amplitude root-mean-square normalized offline using Praat (Boersma and Weenink, 2016), and then presented using Matlab through ultra-shielded insert earphones (ER-3A; Etymotic Research, Inc., IL), at a comfortable level of around 65 dB SPL. For each subject, a stimulus list containing five occurrences of /apa/, /aba/, /ata/ and /ada/ stimuli was randomly permuted, and stimuli were presented according to this order for six blocks without cessation for all TMS conditions (120 stimuli in total). For the behavioural pre-tests, the same stimulus presentation procedure was also used, apart from in the speech-in-noise pre-test, where stimuli were presented for eight blocks (160 stimuli in total) to enable the presentation of several signal-to-noise ratios (SNRs) and improve validity of intelligibility quantification. The speech-shaped noise used in the speech-in-noise pre-test was created in Matlab (R2013a; The Mathworks Inc., Natick, MA) using a custom-written script, and contained the same long-term average spectrum as speech, but without amplitude modulation.

### 2.3 Design

Two separate experiments were undertaken to assess how activity in the motor system is modulated when listening to distorted speech. In Experiment 1, two different types of speech distortions, speech-in-noise and motor-distorted speech, were used to investigate if the speech motor cortex is sensitive to the quality of distortion in the speech signal. In Experiment 2, one type of speech distortion, speech-in-noise presented at different SNRs, was used to assess if the motor cortex is sensitive to the quantity of the speech distortion. Accordingly, in Experiment 2, two SNRs were used at which subjects were able to correctly identify 50% of the speech stimuli (moderate SNR, condition 1), or 75% of the speech (mild SNR, condition 2).

### 2.4 Pre-tests

In both Experiments 1 and 2, subjects took part in behavioural pre-tests to examine their ability to perceive the distorted and clear speech stimuli used in the TMS part of the experiments, prior to receiving any TMS (see Figure 1). Chance performance was always 25% in all pre-tests. Pre-tests were performed separately in a counter-balanced order, to test speech perception in the following conditions:

1. Natural Speech: normally articulated vowel-consonant-vowel (VCV) syllables, which contained an equal distribution of lip- (/apa/, /aba/) and tongue-articulated (/ata/, /ada/) syllables.
2. Speech-in-noise: listening to normally articulated VCV syllables from Natural Speech condition presented at different levels in speech-shaped unmodulated background noise to yield eight different SNRs. Stimuli contained an equal distribution of lip- (/apa/, /aba/) and tongue-articulated (/ata/, /ada/) syllables. Speech stimuli were presented at SNRs ranging from -3 to -24 dB SPL in steps of 3 dB SPL. This selection of SNRs was based on pilot data that indicated subjects were able to achieve a range of performance levels from near ceiling to chance performance across this range of SNRs for the Natural Speech stimuli.
3. Distorted Speech: listening to motor-distorted VCV syllables that had been produced whilst using a tongue depressor. Stimuli contained an equal distribution of lip- (/apa/, /aba/) and tongue-articulated (/ata/, /ada/) syllables.

All three behavioural pre-tests were tested in Experiment 1, whereas subjects completed only the first two pre-tests in Experiment 2, as no motor-distorted stimuli were used in the second experiment. Subjects were not given feedback on their performance during the pre-tests, or after the pre-tests.

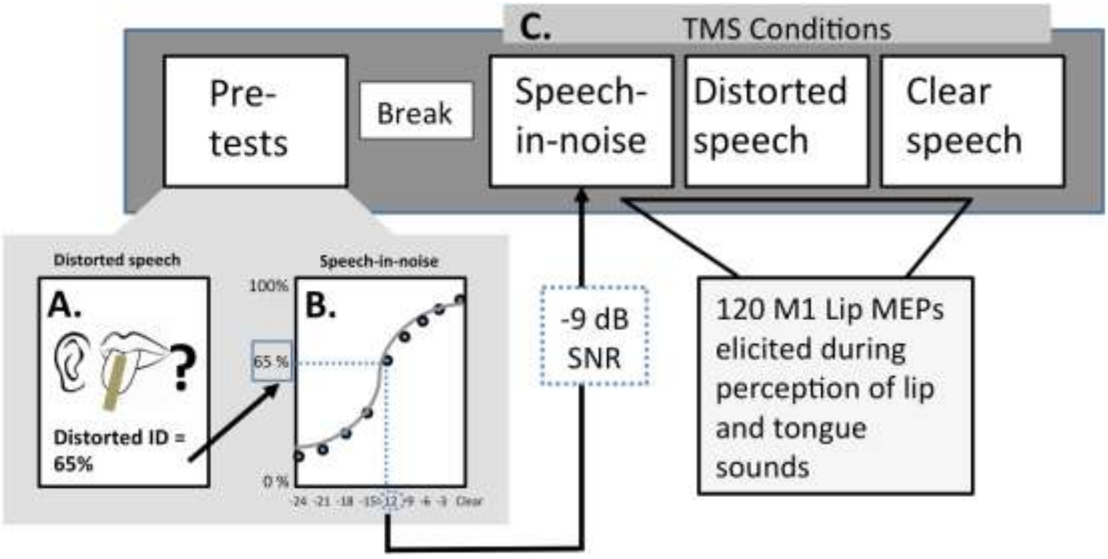

Figure 1. Experimental design of Experiment 1. Subjects initially took part in behavioural pre-tests, allowing intelligibility to be matched between the distorted speech pre-test (A) and the speech-in-noise pre-test (B). The SNR which led to equivalent performance in the speech-in-noise pre-test relative to performance on the distorted pre-test was then used for the speech-in-noise TMS condition (C). As Experiment 2 investigated quantity of speech distortion and not quality, the distorted speech pre-test and distorted speech TMS condition were not used in Experiment 2.

After completing the pre-tests in Experiment 1, performance accuracy (percent correct) for the motor-distorted stimuli was compared to the speech-in-noise test, to find the SNR at which subjects found the speech-in-noise stimuli to be as equally intelligible as the motor-distorted speech stimuli. The associated SNR was then used as the SNR in the TMS experiment (see Figure 1B). In Experiment 2, we established the SNRs to be used in the TMS experiment as those at which each subject was able to correctly identify 50% and 75% of the speech stimuli in the speech-in-noise pre-test.

Speech identification tasks were presented via Matlab using custom-written scripts. Subjects were asked to listen carefully to the speech sounds and to identify the consonant in the middle of the sound as either a ‘p’, ‘b’ ‘t’ or ‘d’ using a key press, as quickly as possible without compromising accuracy. A fixation cross was presented on the monitor until the end of stimulus presentation, after which subjects were prompted with a visual cue to enter their response selection. Subjects were given up to 2500 ms from the onset of the stimulus to make their response, after which the program would present the subsequent trial. Failure to respond during this time period would result in a null response for that particular trial.

### 2.5 TMS conditions

#### 2.5.1 Experiment 1

Experiment 1 was designed to test whether activity in lip M1 is differentially modulated when listening to different forms of degraded speech. Subjects were instructed to listen passively to the speech stimuli and to try to understand the speech where possible in both experiments. The order of experimental conditions in Experiment 1 was counter-balanced. The following three experimental conditions were tested:

1) Distorted: Listening to the motor-distorted speech produced using a tongue depressor.
2) Noise: Listening to clear speech presented in speech-shaped unmodulated noise at a fixed SNR. The SNR was individually determined for each subject based on their performance in the distorted speech pre-test, to equate intelligibility between the different forms of speech degradation (see Methods 2.4).
3) Natural: Listening to naturally articulated clear speech.

#### 2.5.2 Experiment 2

Experiment 2 was designed to test whether activity in the lip area of the primary motor cortex is differentially modulated by the extent of degradation compromising speech intelligibility. Here, we define intelligibility as the amount of speech information that subjects were successfully able to perceive in noise according to their performance on the speech-in-noise pre-test. The same speech and type of noise were used as in Experiment 1. The following three experimental conditions were tested in a counter-balanced order:

1) SNR 50%: Listening to clear speech presented in noise at an SNR at which the subject was able to identify 50% of the target speech.
2) SNR 75%: Listening to clear speech presented in noise at an SNR at which the subject was able to identify 75% of the target speech.
3) Natural: Listening to naturally articulated clear speech.

### 2.6 Transcranial magnetic stimulation

In both Experiments 1 and 2, monophasic single TMS pulses were generated by a Magstim 200$^2$ unit and delivered by a 70mm diameter figure-of-eight coil, connected through a BiStim$^2$ module (Magstim, Dyfed, UK) set to simultaneous discharge mode (inter-pulse spacing of 0 ms). The coil was placed tangential to the skull such that the induced current flowed from posterior to anterior under the junction of the two wings of the figure-of-eight coil. The lip area of M1 was found by using the functional ‘hot spot’ localization method, whereby application of TMS elicits an MEP from the contralateral muscle. Here, the coil position and orientation is adjusted in millimetre movements to ascertain the location on the scalp at which the most robust MEPs are elicited. This location was then marked on a cap and active motor threshold (aMT) determined, which constitutes the intensity at which TMS pulses elicited 5 out of 10 MEPs with an amplitude of at least 50 µV (Watkins, Strafella, & Paus, 2003; Watkins & Paus, 2004). In this way, we first located the hand area by asking subjects to perform a pinching action where the index finger was held against the thumb to activate first dorsal interosseous. Following this, the lip area ‘hot spot’ was identified by moving the coil ventrally and slightly anterior until an MEP was observed in the contralateral lip muscle, and the aMT identified (Möttönen et al., 2014). The intensity of the stimulator was then set to 120% of aMT for the stimulations applied during the experiment. The mean stimulator intensity (120% aMT ± SEM) used to elicit lip MEPs in Experiment 1 was 51.6% (±1.1), and in Experiment 2 was 48.3% (±0.8), of the maximum possible intensity.

After establishing TMS test intensity, all subjects then received three test blocks of single-pulse TMS to the lip area of M1 in the left hemisphere. During the TMS test blocks, subjects were presented with the speech stimuli (see Methods 2.2) and were asked to listen passively to the sounds. During the presentation of each speech stimulus, Matlab was used to externally trigger the TMS system, such that a TMS pulse was generated 100 ms after the onset of the consonant in each stimulus type. TMS timing was based on previous MEP studies where excitability of the left articulatory motor cortex has been found to occur from 100 ms during speech perception (Fadiga et al., 2002; Sato et al., 2010). All speech stimuli were accompanied by a TMS pulse; therefore, all trials were presented with TMS. There was a 2.5 s inter-stimulus delay between all auditory stimuli and an inter-stimulation delay of between 4.5-5 s. TMS test blocks lasted for approximately 9-10 minutes, allowing for the application of 120 TMS pulses. Participants were given short breaks in between TMS test blocks, during which time the coil was changed to prevent over-heating.

## 2.7 Electromyography

Electromyographic (EMG) activity was recorded from the lip muscle orbicularis oris using surface electrodes (Ag/AgCl; 10-mm diameter) in a non-Faraday caged, double-walled sound-attenuating booth. Electrodes were attached to the orbicularis oris on the right side of the mouth in a bipolar montage, with an electrode placed at the right temple serving as a common ground. To stabilize background EMG activity, subjects were trained for approximately five minutes to produce a constant level of contraction (approximately 20-30% of maximum voluntary contraction) of the lip muscles by pursing, which was verified via visual feedback of the ongoing EMG signal, in accordance with standard practice for lip aMT thresholding (Möttönen et al., 2014; Watkins et al., 2003). Contraction of the lip muscle also facilitates a lower motor threshold relative to when the muscle is at rest, enabling the use of lower levels of stimulation during the experiment. The raw EMG signal was amplified by a factor of 1000, band-pass filtered between 100–1000 Hz, and sampled at 5000 Hz online using a 1902 amplifier (Cambridge Electronic Design, Cambridge), and analog-to-digital converted using a Micro1401-3 unit (Cambridge Electronic Design, Cambridge). Continuous data were acquired and recorded using Spike2 software (version 8, Cambridge Electronic Design, Cambridge).

### 2.8 Data analysis

In Experiment 1, the proportion of correct responses for the distorted speech pre-test were first calculated for each individual. For the speech-in-noise pre-test, a logistic psychometric function based on a maximum likelihood criterion implemented via the Palamedes Matlab toolbox (Prins & Kingdom, 2009) was fitted to the proportion of correct responses at each SNR (-3 dB to -24 dB SPL) for each individual. To match intelligibility between the two pre-tests, the point on the psychometric function that was equivalent to the proportion of correct responses in the distorted speech task was found, and the associated SNR derived (see Figure 1A and B). This SNR was then used for the TMS speech-in-noise condition. In Experiment 2, the 50% and 75% performance points on the psychometric functions were obtained for each subject, and the SNRs associated with these levels of performance were used for the speech-in-noise TMS conditions.

For peripheral hearing sensitivity, pure-tone average (PTA) audiometric thresholds were computed across all octave frequencies measured in the audiometric hearing test (.5, 1, 2, and 4 kHz) for each ear. The overall PTA was computed by averaging the left and right ear PTAs, and was used in subsequent analyses (lower PTAs indicate more sensitive peripheral hearing). Raw hearing threshold data was unavailable for one subject in Experiment 1, so this data point was replaced with the mean.

For the TMS data, individual EMG sweeps starting 40 ms before the TMS pulse and ending 40 ms post-stimulation were exported offline from the recording software into Matlab, where mean MEPs were calculated for each speech type, and lip or tongue sound combination, per subject. Lip MEPs exhibit substantial variability for reasons that are not fully understood, but may reflect biological variation such as the phase in the Mu rhythm at the time the MEP was elicited; neural recruitment; circadian rhythms (e.g., Keil et al., 2014; Kiers et al., 1993; Wassermann, 2002).

Individual averages were rectified and the integrated area under the curve (AUC) of the rectified EMG signal of each individual mean MEP was calculated from 8-35 ms post-stimulation, which captured the time at which lip MEPs are expected to occur (Devlin and Watkins, 2007). Deriving the AUC from averages in this manner was therefore useful to determine the systematic structure of the time-locked MEP, which removed subjective bias associated with identifying the MEP window for each individual trial. MEP AUCs were then converted into standard scores within subjects, to control for inter-subject variability. The standardized AUCs of MEPs were used in the statistical analyses. The average height of the pre-TMS baseline EMG activity was also computed, and paired t-tests confirmed that there were no significant differences between baseline activity levels between clear and distorted conditions, and clear and noise conditions, in Experiment 1 nor 2, indicating that baseline activity did not influence MEPs across conditions.

Statistical analyses were carried out using SPSS (version 22.0, IBM). A two-way repeated measures analysis of covariance (RM ANCOVA) was conducted on standardised MEP AUC as the dependent variable, with Speech Type (Natural, Distorted, Noise), and Articulator (Tongue, Lip), as within-subjects factors. The overall PTA was included in the model as a covariate based on a previously established association between hearing thresholds and the dependent variable (Nuttall et al., 2016), to examine how hearing ability may modulate the effect of Speech Type. Bonferroni-corrected comparisons and Pearson’s correlations were also conducted where appropriate.

## 3. Results

### 3.1 Experiment 1

#### 3.1.1 Behavioural data

For the motor-distorted speech pre-test, performance ranged from 15% to 87.5% correct across subjects (chance performance = 25%; mean performance = 45.8%, S.D. 20%, Figure 2B), in line with previous observations (Nuttall et al., 2016). Accordingly, this gave rise to a wide range of intelligibility-matched SNRs ranging from -24 to -3.5 SNR dB SPL (Figure 2B) that yielded an equivalent level of performance in the speech-in-noise task (mean -16 dB SPL, S.D. 7.5 dB SPL). Figure 2A shows the average proportion of correct responses across subjects for each SNR, and the association between motor-distorted speech pre-test scores and matched SNRs is shown in Figure 2B. Chance performance was also 25% in the speech-in-noise pre-test.

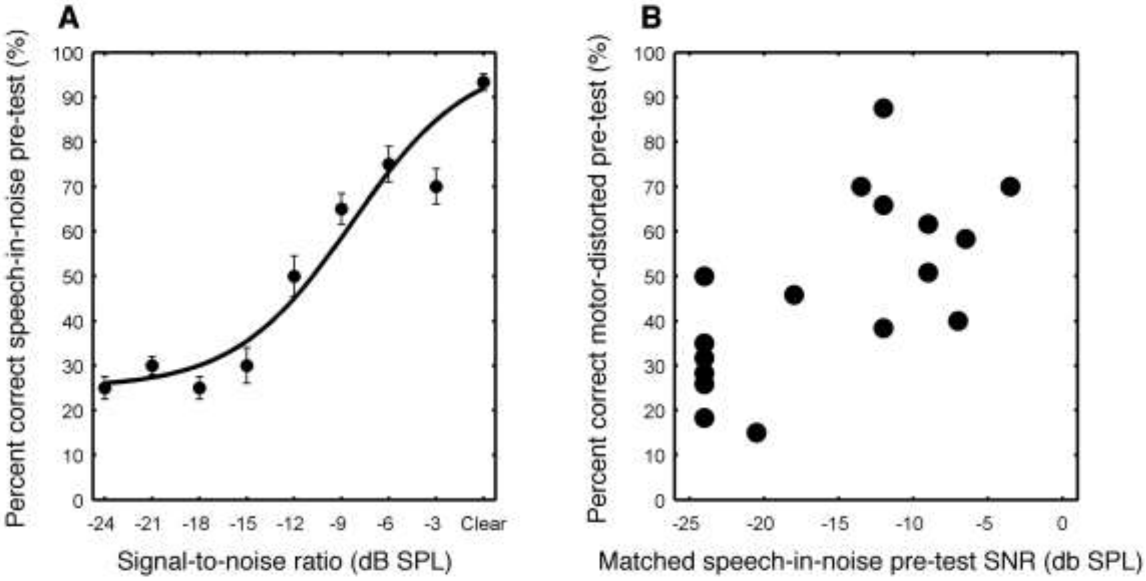

Figure 2. Psychometric function fitted to grand mean proportion of correct responses across SNRs and for clear speech ± SEM (A). Association between motor-distorted pre-test scores and matched speech-in-noise pre-test SNRs (B).

### 3.1.2 Motor evoked potentials

The TMS experiment (Figure 1C) tested how different types of degraded speech modulate the excitability of the lip motor area. The mean differences in lip MEP amplitude as a function of speech type and articulator is reflected in Figure 3. On average, lip MEPs in the distorted speech condition were greater than those in the natural condition, and same trend was also observed for lip MEPs in the speech-in-noise condition. There were no observable differences between MEPs elicited during perception of different types of degraded speech, or differences depending on place of articulation of the speech stimulus. These observations were confirmed by a two-way repeated measures ANCOVA which, after controlling for the effect of hearing threshold, identified a significant main effect of speech type ($F(2,30) = 9.38$; $p = 0.001$; partial $\eta_2 = 0.38$, indicating that MEPs were differentially modulated depending on the type of speech being perceived. All other main effects and interactions were non-significant. There were no significant interactions between MEPs and hearing ability (articulator: $p = 0.2$; speech type × articulator: $p = 0.9$; speech type × hearing ability (pure tone average [PTA]) $p = 0.2$; articulator × PTA: $p = 0.4$; speech type × articulator × PTA: $p = 0.7$).

To explore the effect of speech type on lip MEPs, Bonferroni-corrected pairwise comparisons were conducted between speech type pairs. This confirmed a significant difference between MEPs recorded during perception of natural and distorted speech (*p* = 0.001), and between MEPs recorded during perception of natural speech and speech-in-noise (*p* = 0.005). However, the speech distortions did not differ significantly in their effect on MEP size (*p* = 1.0).

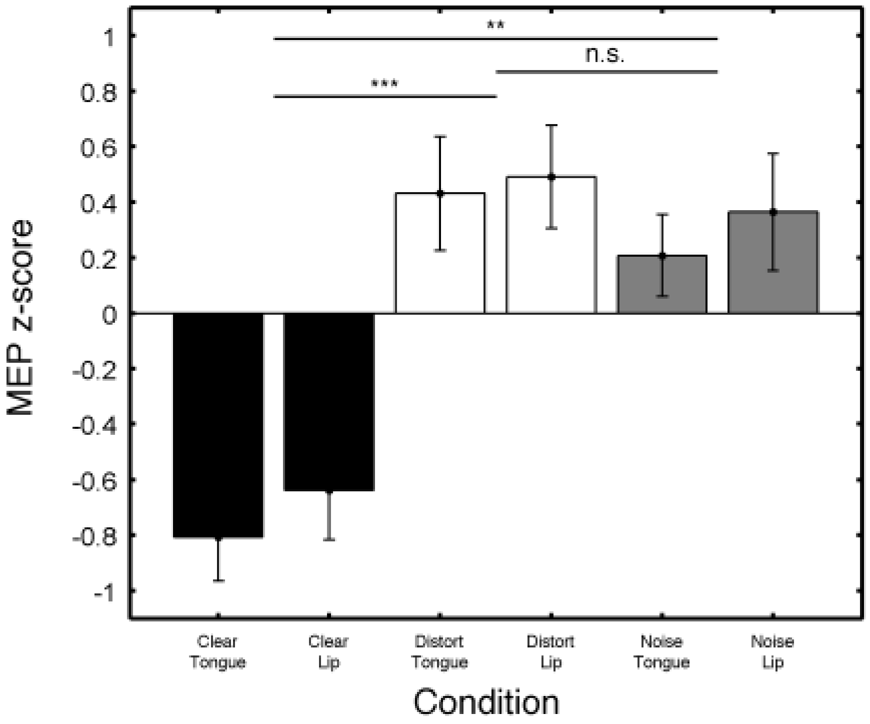

Figure 3. Mean standardized area under the curve (AUC) of MEPs elicited during perception of clear (black bars), motor-distorted (white bars), and speech-in-noise (grey bars) stimuli, articulated by the lips or tongue (± SEM). *** = $p \leq 0.001$; ** = $p \leq 0.01$; n.s. = not significant.

## 3.2 Experiment 2

### 3.2.1 Behavioural data

The results from the speech-in-noise pre-test in Experiment 2 were used to find the two SNRs at which subjects could correctly identify 50% and 75% of the speech stimuli in each subject. The SNRs associated with 50% correct performance ranged between -16 and -9 dB SPL across subjects (mean SNR -11.5 dB SPL, S.D. 1.7 dB SPL). The SNRs associated with 75% correct performance ranged between -10.5 and -2 dB SPL across subjects (mean SNR -6 dB SPL, S.D 2.1 dB SPL, Figure 4). A paired t-test confirmed that the SNRs associated with 50% performance and the SNRs associated with 75% performance were significantly different ($t(17) = 16.6$, *p* <

0.001). Notably, independent t-tests (equal variances not assumed) confirmed that both sets of SNRs were also significantly different from the SNRs used in Experiment 1 (Figure 4), which were matched to the intelligibility of the motor-distorted speech (SNR 50%: $t(18.76) = -2.58$, $p = 0.018$, SNR 75%: $t(19.67) = -5.56$, $p < 0.001$).

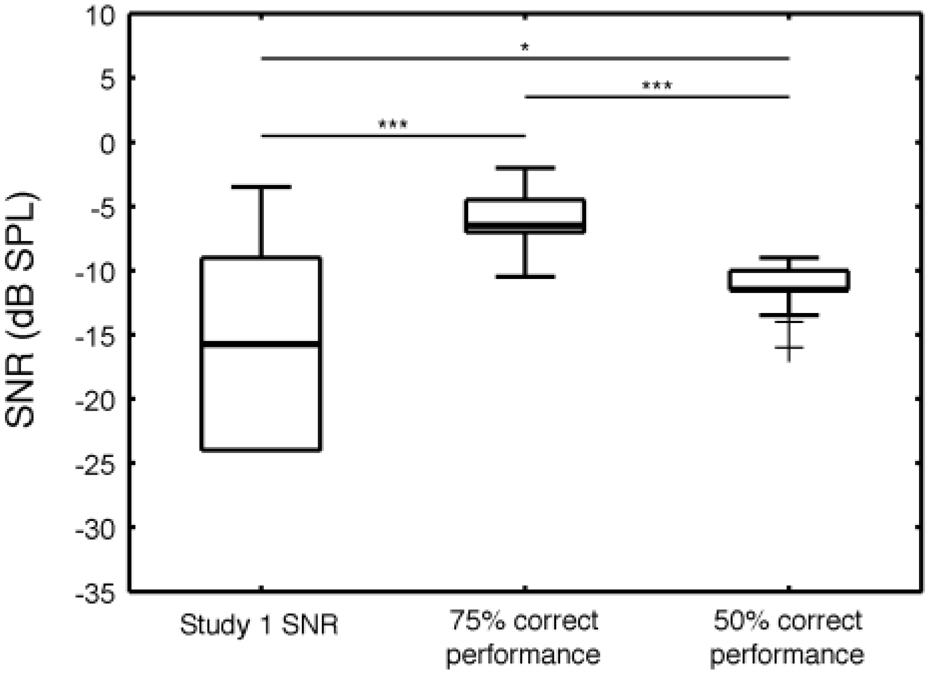

Figure 4. Box plots of the SNRs at which 75% and 50% correct performance was achieved on the speech-in-noise pre-test in Experiment 2, and of the SNRs from Experiment 1 for comparison. SNRs from Experiment 1 ranged widely due to being matched to performance on the motor-distorted speech pre-test. In Experiment 2, the two SNRs associated with 50% and 75% correct performance for each subject were subsequently used in the two speech-in-noise TMS conditions. *** = $p < 0.001$; * = $p < 0.05$.

### 3.2.2 Motor evoked potentials

Experiment 2 tested how the quantity of distortion in the speech signal modulated the excitability of the motor system. To this end, MEPs were elicited during perception of speech presented at mild and moderate SNRs.

The mean difference in lip MEP size as a function of speech type and articulator is reflected in Figure 5. On average, there were no observable differences between MEPs elicited during perception of speech-in-noise at varying levels of intelligibility, or differences depending on place of articulation of the speech stimulus. After controlling for the effect of hearing ability, a two-way repeated measures ANCOVA found that the main effect of speech type was non-significant ($F(2,32) = 0.5$; $p = 0.6$). However, the covariate, hearing ability, was significantly related to the effect of speech type ($F(2,32) = 3.6$, $p = 0.03$, partial $\eta_2 = .2$), suggesting that MEPs were differentially modulated by speech type in a manner dependent on the listener's peripheral hearing acuity. All other main effects and interactions were non-significant (articulator: $p = 0.8$; speech type × articulator: $p = 0.9$; articulator × PTA: $p = 0.8$; speech type × articulator × PTA: $p = 0.9$).

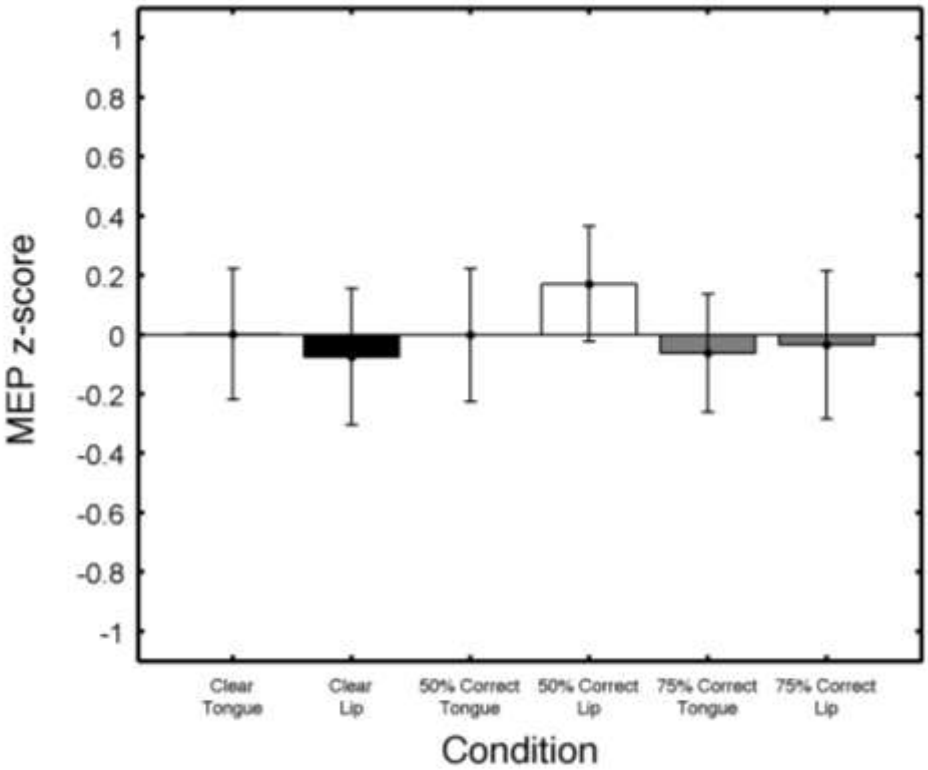

Figure 5. Mean standardized area under the curve (AUC) of MEPs elicited during perception of speech-in-noise when speech was 50% intelligible (white bars), 75% intelligible (grey bars), and when speech was presented without noise (black bars), for syllables articulated by either the lips or the tongue (± SEM).

A correlation analysis was conducted in order to characterize the significant interaction. To test the relationship between hearing ability and motor activity, we performed correlations between subjects' hearing ability and 1) mean MEP area during perception of clear speech, 2) mean MEP area during perception of speech that was 50% intelligible, and 3) mean MEP area during perception of speech that was 75% intelligible (Figure 6). Pearson's correlations established a significant positive relationship between MEPs recorded during perception of clear speech and hearing ability (*r*(18) = .6, *p* = 0.01), indicating that subjects with relatively worse (i.e. less sensitive) hearing showed greater MEPs when perceiving clear speech, compared to listeners with better hearing. However, when speech was 50% intelligible, there was a moderate negative correlation, such that individuals with better (more negative) hearing showed more motor facilitation when listening in the noise condition, (*r*(18) = -.5, *p* = 0.04). No correlation was observed between hearing and MEPs recorded when speech was 75% intelligible (*r*(18) = -.1, *p* = 0.75).

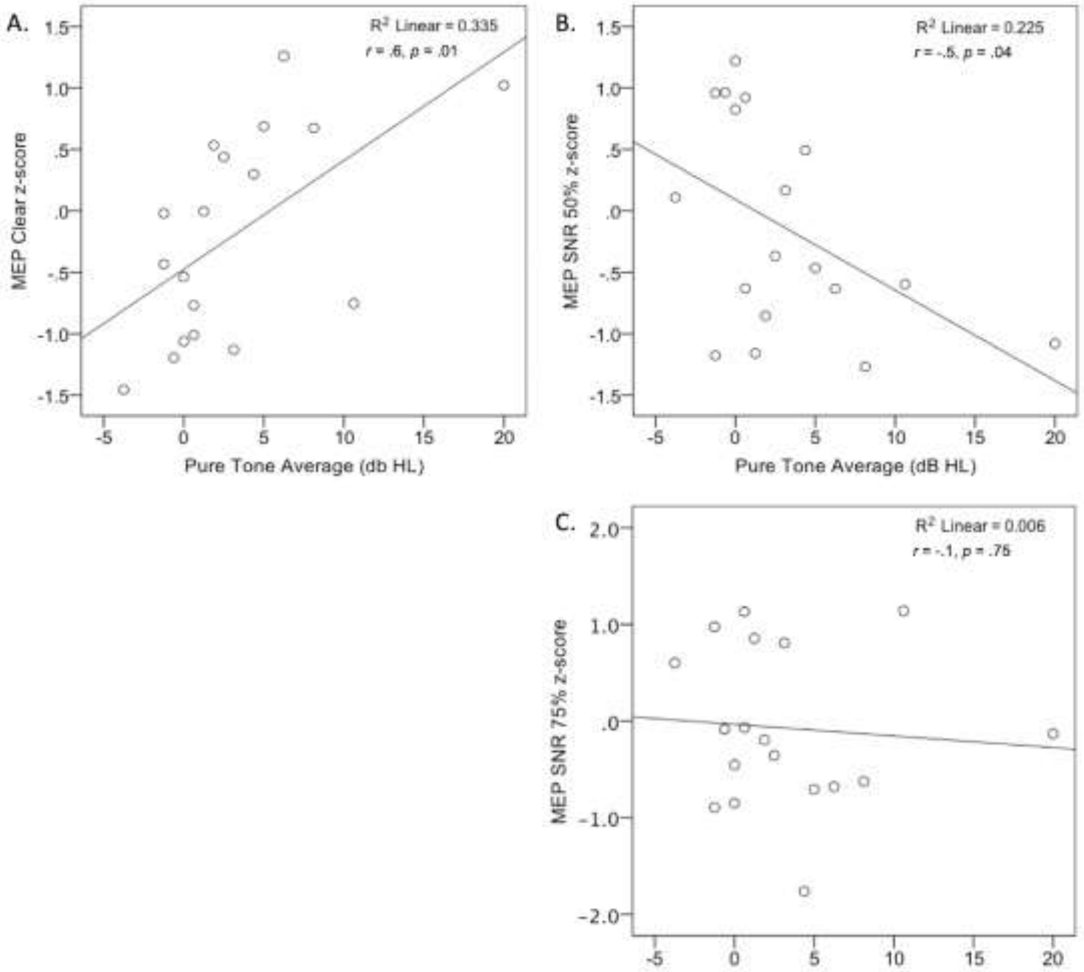

Figure 6. Correlations between pure tone average (PTA) hearing sensitivity, and MEP area during perception of clear speech (A), speech-in-noise when intelligibility was 50% (B), and speech-in-noise when intelligibility was 75% (C).

Taken together, these data suggest that listeners with less sensitive hearing exhibit more motor activity during perception of clear speech relative to speech-in-noise compared to listeners with better hearing, who tend to show more motor cortex activation when listening to speech-in-noise relative to clear speech. As such, this difference appears to obfuscate the effect of speech type on motor activity, which averages out when considered at the level of the group mean.

## 4. Discussion

The present study first aimed to establish whether motor facilitation resulting from degraded speech perception is modulated by the type and extent of degradation (form dependent), or is independent of the form of degradation (form independent). Second, we aimed to establish whether hearing ability is associated with articulatory motor activation during speech perception. Experiment 1 replicated earlier findings that showed that activity in the motor system is enhanced when listening to distorted speech (Murakami et al., 2011; Nuttall et al., 2016). Experiment 1 also demonstrated that both speech external distortions (noise), and speech internal distortions (speech produced with a motor perturbation), led to equivalent increases in MEP area. This finding suggests that the motor system does not differentiate between speech external and speech internal types of distortion. Experiment 2 used significantly less adverse SNRs compared to Experiment 1, and did not find evidence of motor system facilitation when speech could be identified with either 50% or 75% accuracy in background noise. However, there was a significant interaction between noise condition and hearing ability, demonstrating that motor facilitation during perception of speech-in-noise is modulated by hearing ability, but only when the SNR is moderate. Here, young adults with less sensitive hearing showed significantly more motor activation when listening to clear speech. Conversely, listeners with better hearing showed greater motor activation (facilitation) during perception of speech-in-noise.

Our observation in Experiment 1 that the speech motor system did not differentiate between different forms of speech distortion supports a form independent role of the primary motor system during speech perception. Similar studies on orofacial neurophysiology are limited, but our findings are in line with studies on observation of degraded visual hand stimuli, which highlight a sensitivity of the motor cortex when presenting more difficult to produce finger-tapping actions relative to simple finger-tapping actions (Roosink and Zijdewind, 2010). Similarly, the motor cortex has been shown to remain sensitive to hand action observation even when the hand grasping action cannot be fully observed (Valchev et al., 2015; Villiger et al., 2011). Taken together, these data may suggest that when perceiving action-related information such as speech, the motor system may operate as a reflexive gain control mechanism, whereby after a listener has received insufficient auditory information, regardless of the source of insufficiency, motor activity increases in order to provide supplementary information and improve the detection of the speech signal. Such functionality would not need to be sensitive to differences in the acoustic form of the distortion; it would simply attempt to compensate for the missing information. Form independent motor facilitation complements contemporary ideas on the neural hierarchy involved in spoken language processing, whereby temporal auditory areas at the origin of cortical speech processing are sensitive to surface acoustic form, such as speaker identity and intelligibility. Somatomotor regions are considered to be at the pinnacle of the processing hierarchy, and potentially most involved with higher-level linguistic abstraction, and not explicitly considered to be sensitive to variations in acoustic form (Evans and Davis, 2015). Indeed, Evans and Davis (2015) observed blood oxygen level dependent (BOLD) changes to indicate that the somatomotor region is sensitive to syllable identity but not to surface acoustic form, with adjacent motor areas demonstrating greater activity for degraded relative to clear speech. Whilst motor areas may be sensitive to the presence of distortion in the signal, it appears that they are insensitive to the nature or quality of the distortion.

Unlike Evans and Davis (2015), and other TMS studies (Bartoli et al., 2015; Fadiga et al., 2002; Roy et al., 2008), we did not find evidence that the motor system is sensitive to syllable identity (somatotopic for speech). This may be due to the fact that we recorded MEPs from one articulator only, orbicularis oris, which in combination with the sensitivity of our design did not emphasize somatotopic speech activity during passive speech perception. Contrasting activity in M1 lip with M1 tongue, for example, during perception of the same lip- and tongue-articulated stimuli, presents a more optimal test of somatotopy. Although Evans and Davis (2015) did not specifically demarcate different articulators in their fMRI analyses, it is possible that they observed identity-specific changes during perception of syllables due to the detection of BOLD activity associated with more than one articulator. Furthermore, engaging in an active task, as opposed to passive listening as in the present study, requires an explicit need to differentiate places of articulation during speech perception, and this may also increase the likelihood of detecting subtle, somatotopic responses. Lastly, new evidence has come to light that suggests that the sensorimotor cortex activates along an acoustic, and not somatotopic, dimension during speech perception. Cheung and colleagues (2016) found that sounds that have similar acoustic properties but different motor configurations, such as ‘b’ and ‘d’, activate the motor cortex in similar ways when perceiving speech. This may also relate to why we did not find somatotopic MEP differences.

In Experiment 2, we did not find a significant main effect of noise on MEPs elicited during perception of speech-in-noise, in contrast to Experiment 1. This may be in part related to the significantly less adverse SNRs used in Experiment 2 relative to Experiment 1. It should be noted that the results of Experiment 2 do not constitute a total replication of Experiment 1. Experiments 1 and 2 use significantly different noise levels (see Figure 4 and Results 3.2.1 for further details), with SNRs in Experiment 1 being significantly lower and thus more difficult than in Experiment 2. Therefore, we interpret the speech-in-noise data from Experiment 2 as complementary to Experiment 1, and not contradictory. However, the significant interaction between noise condition and hearing status in Experiment 2 revealed that the effect of noise on the articulatory motor system was dependent on hearing ability. Specifically, we observed in Experiment 2 that listeners with less sensitive hearing did not demonstrate motor facilitation when perceiving speech in background noise. Instead, they showed greater motor activity during clear speech perception. If these individuals engage motor resources less during challenging listening conditions then this may impact their ability to perceive speech-in-noise. Extracting a speaker's voice from background noise is essential for everyday communication, and is often challenging even for young adults with good hearing and normal cognitive abilities (Neff and Green 1987; Assmann & Summerfield, 2004). Successful detection of a target message is in part dependent on a listener's ability to extract the spectral information in the acoustic signal during the initial stages of speech processing at the ear. It is thought that small differences in hearing thresholds, as well as threshold differences across frequencies, impact upon the subsequent neural representation of the speech signal in the central auditory pathway (Nuttall, Moore, Barry, Krumbholz, & de Boer, 2015). Differences in central auditory processing could potentially regulate changes in motor engagement during speech perception, which in combination with additional auditory and cognitive factors, may be implicated in why normally hearing individuals vary in the extent to which they activate the speech motor system during perception.

Nonetheless, in listeners with better hearing, we did observe evidence of motor facilitation during perception of speech in the most adverse noise condition compared to clear speech. This does not preclude the possibility that other levels of intelligibility, SNR, or types of noise masking may differentially modulate auditory-motor activity. Interestingly, we have previously observed that individuals with better hearing show less facilitation when listening to speech which is internally-distorted relative to listeners with less sensitive hearing (Nuttall et al., 2016). These diverging relationships are most likely linked to the influence of different auditory processes during speech perception, as the effect of background noise has a very different impact on the auditory system compared to the effect of speech that is difficult to recognize due to the manner of speaking. For example, background noise causes neural adaption at the inner hair cell-auditory nerve junction, and as a result is thought to engage auditory efferent fibers to provide relief from background noise masking (Guinan, 2006). Therefore, different speech distortions do not affect the auditory system equally, and thus correlations between hearing and cortical motor facilitation during different types of speech distortion may be mediated by different auditory mechanisms. We also cannot establish if exposure to the numerous stimuli in the pre-test in Experiment 2, which always preceded the TMS experiment, may have affected listening strategy, motor activation, or adaptation during MEP recording, as subjects were not exposed to any speech stimuli prior to TMS in the previous study (Nuttall et al. 2016). Notably, hearing ability did not interact with motor activity in Experiment 1, which used significantly more adverse speech distortions than in Experiment 2. There was also no correlation between MEPs recorded during perception of clear speech in Experiment 1 and hearing ability (r = -0.2, p = 0.5). This is most likely due to the presence of significantly smaller and less variable MEPs in Experiment 1 than in Experiment 2 (t(34) = -2.7, p = 0.01; equal variances not assumed). Listeners' hearing abilities were not significantly different across the two experiments (t(34) = -0.4, p = 0.66), and therefore it is unlikely that differences In hearing ability led to the difference in correlations between Experiments 1 and 2.

In light of the current data, two outstanding questions remain: 1) how much distortion needs to be present in the speech signal before motor activity is engaged, and 2) how does the extent of motor activity interact with encoding of the auditory signal at the ear, and beyond. The findings of Experiment 1 and 2 give some indication of these relationships. In Experiment 1, the mean SNR used in the noise condition was -16 dB SPL, whereas in Experiment 2, the mean SNR for the mild noise condition (SNR 75%) was -6 dB SPL, and for the moderate noise condition (SNR 50%), -11.5 dB SPL. This difference in SNRs between experiments arises from the intelligibility matching in Experiment 1. Participants achieved a wide range of scores on the motor-distorted pre-test, ranging from 15% to 87.5%. If participants performed around or below chance (25%) on the motor-distorted pre-test, this would result in a considerably adverse SNR being used in the TMS speech-in-noise condition, to equate performance to around chance in the speech-in-noise pre-test. Conversely, in Experiment 2, the SNRs were chosen based solely on the 50% and 75% performance points in the speech-in-noise pre-test. Therefore, a greater range of more adverse SNRs were used in Experiment 1 compared to Experiment 2.

This difference in SNR usage could explain why we did not observe a main effect of noise in Experiment 2, as SNR may need to be considerably adverse (~ ≤12 dB SPL) for any effect of noise on motor facilitation to be borne out on a group level. However, it may be the case that for higher (less adverse) SNRs, inter-individual differences in auditory processing moderate whether motor cortices are engaged during speech perception, accounting for why we observed a relationship between hearing ability and motor activity in Experiment 2, but not Experiment 1.

Related work from Du and colleagues (2014) using fMRI has demonstrated that speech motor areas exhibit significant phonetic discrimination activation for SNRs above –6 dB SPL, whereas bilateral auditory cortices encode phoneme-specific information only when the noise is absent or extremely weak (SNR > 8 dB). Notably, the most adverse SNR tested by Du and colleagues was -8 dB SPL, and how to equate the BOLD signal and MEP is not straightforward. However, Du et al's findings, as well as our findings, support the possibility that the motor system assists the auditory system during speech perception under challenging listening conditions. This is in line with recent work from Murakami and colleagues (2015), who found that disrupting motor function caused subtle but significant phonological errors, the extent of which was dependent upon the level of background noise. Notably, however, the authors observed substantial neurophysiological changes during passive speech perception, which were linked only to limited effects on behavioural speech perception. Therefore, we concur with Murakami et al. (2015) that sensorimotor mapping via the dorsal stream may not be essential for speech perception, but that it contributes to subtle phonological analyses when sensory information is insufficient.

A limitation of our study is that there is no non-speech control condition or attentional control. However, we have previously established (Nuttall et al., 2016) that hand MEP control data are not modulated by speech distortion, unlike lip MEPs. This would suggest that lip MEP differences resulting from speech distortion are not the result of a non-specific global attention effect. If it were the case that distorted speech is more salient or attentionally engaging than clear speech, then hand MEPs should have also been facilitated during perception of distorted speech. This is in line with previous findings from related literature, where control site and control task effects differed from speech motor effects (Meister et al., 2007; Möttönen and Watkins, 2009; Murakami et al., 2011; Watkins et al., 2003). A non-speech condition, however, would have confirmed if distorted speech added further motor facilitation on top of the classical (clear) speech-related motor facilitation. Therefore, we cannot confirm if the MEP facilitation we have observed in this study is speech-specific, although previous research suggests that this is likely (Fadiga et al., 2002; Meister et al., 2007; Watkins et al., 2003). Future studies could benefit from exploring attention and listening effort manipulations and non-speech conditions to clarify these associations.

In conclusion, the present study tested whether activity in the motor system is modulated by the type and extent of distortion in the speech signal. Data indicated that the motor system assists speech perception when listening is difficult, both when listening to motor-distorted speech and speech-in-noise (Experiment 1). However, when background noise only minorly or moderately compromises speech intelligibility, the motor system does not show facilitation when considered at a group level. Instead, when hearing ability is taken into account, motor facilitation at moderate levels of noise can be observed in individuals with better hearing, and less activation in noise can be observed for individuals with relatively worse hearing (Experiment 2). These data underline the importance of individual differences and their effect on the role of the motor system in speech perception.

## Acknowledgements

This work was supported by a Project Grant from The Leverhulme Trust (RPG-2013-254). Our thanks go to The Leverhulme Trust, and the individuals who participated in this study.

## References

Adank, P., Davis, M.H., Hagoort, P., 2012. Neural dissociation in processing noise and accent in spoken language comprehension. Neuropsychologia 50, 77–84. doi:10.1016/j.neuropsychologia.2011.10.024

Adank, P., Janse, E., 2009. Perceptual learning of time-compressed and natural fast speech. J. Acoust. Soc. Am. 126, 2649–59. doi:10.1121/1.3216914

Bartoli, E., D'Ausilio, A., Berry, J., Badino, L., Bever, T., Fadiga, L., 2015. Listener-Speaker Perceived Distance Predicts the Degree of Motor Contribution to Speech Perception. Cereb. cortex 25, 281–8. doi:10.1093/cercor/bht257

Bharadwaj, H.M., Masud, S., Mehraei, G., Verhulst, S., Shinn-Cunningham, B.G., 2015. Individual differences reveal correlates of hidden hearing deficits. J. Neurosci. 35, 2161–72. doi:10.1523/JNEUROSCI.3915-14.2015

Boersma, P., Weenink, D. (2016). Praat: doing phonetics by computer [Computer 16 program]. Version 6.0.12, retrieved 26 January 2016 from 17 http://www.praat.org/

Borrie, S.A., McAuliffe, M.J., Liss, J.M., O'Beirne, G.A., Anderson, T.J., 2013. The role of linguistic and indexical information in improved recognition of dysarthric speech. J. Acoust. Soc. Am. 133, 474–82. doi:10.1121/1.4770239

Borrie, S.A., Schäfer, M.C.M., 2015. The Role of Somatosensory Information in Speech Perception: Imitation Improves Recognition of Disordered Speech. J. Speech. Lang. Hear. Res. 58, 1708–16. doi:10.1044/2015_JSLHR-S-15-0163

Callan, D., Callan, A., Gamez, M., Sato, M., Kawato, M., 2010. Premotor cortex mediates perceptual performance. Neuroimage 51, 844–58. doi:10.1016/j.neuroimage.2010.02.027

Calvo-Merino, B., Glaser, D.E., Grèzes, J., Passingham, R.E., Haggard, P., 2005. Action observation and acquired motor skills: an FMRI study with expert dancers. Cereb. Cortex 15, 1243–9. doi:10.1093/cercor/bhi007

Cheung, C., Hamiton, L.S., Johnson, K., Chang, E.F., 2016. The auditory representation of speech sounds in human motor cortex. Elife 5. doi:10.7554/eLife.12577

D'Ausilio, A., Bufalari, I., Salmas, P., Fadiga, L., 2012. The role of the motor system in discriminating normal and degraded speech sounds. Cortex 48, 882–887. doi:10.1016/j.cortex.2011.05.017

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., Fadiga, L., 2009. The Motor Somatotopy of Speech Perception. Curr. Biol. 19, 381–385. doi:10.1016/j.cub.2009.01.017

Davis, M.H., Johnsrude, I.S., 2007. Hearing speech sounds: top-down influences on the interface between audition and speech perception. Hear. Res. 229, 132–47. doi:10.1016/j.heares.2007.01.014

Davis, M.H., Johnsrude, I.S., 2003. Hierarchical processing in spoken language comprehension. J. Neurosci. 23, 3423–31.

Devlin, J.T., Watkins, K.E., 2007. Stimulating language: insights from TMS. Brain 130, 610–22. doi:10.1093/brain/awl331

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc. Natl. Acad. Sci. U. S. A. 111, 7126–31. doi:10.1073/pnas.1318738111

Evans, S., Davis, M.H., 2015. Hierarchical Organization of Auditory and Motor Representations in Speech Perception: Evidence from Searchlight Similarity Analysis. Cereb. Cortex. doi:10.1093/cercor/bhv136

Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G., 2002. Speech listening specifically modulates the excitability of tongue muscles: A TMS study. Eur. J. Neurosci. 15, 399–402.

Gambi, C., Pickering, M.J., 2013. Prediction and imitation in speech. Front. Psychol. 4, 340. doi:10.3389/fpsyg.2013.00340

Grush, R., 2004. The emulation theory of representation: motor control, imagery, and perception. Behav. Brain Sci. 27, 377-96-442.

Guinan, J.J., 2006. Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear Hear. 27, 589–607. doi:10.1097/01.aud.0000240507.83072.e7

Harris, K.C., Dubno, J.R., Keren, N.I., Ahlstrom, J.B., Eckert, M. a, 2009. Speech recognition in younger and older adults: a dependency on low-level auditory cortex. J. Neurosci. 29, 6078–87. doi:10.1523/JNEUROSCI.0412-09.2009

Hervais-Adelman, A.G., Carlyon, R.P., Johnsrude, I.S., Davis, M.H., 2012. Brain regions recruited for the effortful comprehension of noise-vocoded words. Lang. Cogn. Process. 27, 1145–1166. doi:10.1080/01690965.2012.662280

Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor integration in speech processing: computational basis and neural organization. Neuron 69, 407–22. doi:10.1016/j.neuron.2011.01.019

Keil, J., Timm, J., Sanmiguel, I., Schulz, H., Obleser, J., Schönwiesner, M., 2014. Cortical brain states and corticospinal synchronization influence TMS-evoked motor potentials. J. Neurophysiol. 111, 513–9. doi:10.1152/jn.00387.2013

Kiers, L., Cros, D., Chiappa, K.H., Fang, J., 1993. Variability of motor potentials evoked by transcranial magnetic stimulation. Electroencephalogr. Clin. Neurophysiol. Potentials Sect. 89, 415–423. doi:10.1016/0168-5597(93)90115-6

Locatelli, M., Gatti, R., Tettamanti, M., 2012. Training of manual actions improves language understanding of semantically related action sentences. Front. Psychol. 3, 547. doi:10.3389/fpsyg.2012.00547

Londei, A., D'Ausilio, A., Basso, D., Sestieri, C., Gratta, C. Del, Romani, G.-L., Belardinelli, M.O., 2010. Sensory-motor brain network connectivity for speech comprehension. Hum. Brain Mapp. 31, 567–80. doi:10.1002/hbm.20888

Lotto, A.J., Hickok, G.S., Holt, L.L., 2009. Reflections on mirror neurons and speech perception. Trends Cogn. Sci. 13, 110–4. doi:10.1016/j.tics.2008.11.008

Meister, Wilson, Deblieck, Wu, 2007. Report The Essential Role of Premotor Cortex in Speech Perception. Curr. Biol. 17, 1692–1696. doi:10.1016/j.cub.2007.08.064

Möttönen, R., Rogers, J., Watkins, K.E., 2014. Stimulating the lip motor cortex with transcranial magnetic stimulation. J. Vis. Exp. doi:10.3791/51665

Möttönen, R., Watkins, K.E., 2009. Motor representations of articulators contribute to categorical perception of speech sounds. J. Neurosci. 29, 9819–9825. doi:10.1523/JNEUROSCI.6018-08.2009

Murakami, T., Kell, C.A., Restle, J., Ugawa, Y., Ziemann, U., 2015. Left Dorsal Speech Stream Components and Their Contribution to Phonological Processing. J. Neurosci. 35, 1411–1422. doi:10.1523/JNEUROSCI.0246-14.2015

Murakami, T., Restle, 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, 2045–2054. doi:10.1016/j.neuropsychologia.2011.03.034

Neff, D.L., Green, D.M., 1987. Masking produced by spectral uncertainty with multicomponent maskers. Percept. Psychophys. 41, 409–15.

Nuttall, H.E., Kennedy-Higgins, D., Hogan, J., Devlin, J.T., Adank, P., 2016. The effect of speech distortion on the excitability of articulatory motor cortex. Neuroimage 128, 218–226. doi:10.1016/j.neuroimage.2015.12.038

Nuttall, H.E., Moore, D.R., Barry, J.G., Krumbholz, K., de Boer, J., 2015. The influence of cochlear spectral processing on the timing and amplitude of the speech-evoked auditory brain stem response. J. Neurophysiol. 113, 3683–91. doi:10.1152/jn.00548.2014

Osnes, B., Hugdahl, K., Specht, K., 2011. Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. Neuroimage 54, 2437–45. doi:10.1016/j.neuroimage.2010.09.078

Peelle, J.E., Troiani, V., Grossman, M., Wingfield, A., 2011. Hearing loss in older adults affects neural systems supporting speech comprehension. J. Neurosci. 31, 12638–43. doi:10.1523/JNEUROSCI.2559-11.2011

Pickering, M.J., Garrod, S., 2013. An integrated theory of language production and comprehension. Behav. Brain Sci. 36, 329–47. doi:10.1017/S0140525X12001495

Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analysing 18 psychophysical data. Retrieved 18 February 2016 from 19 http://www.palamedestoolbox.org

Roosink, M., Zijdewind, I., 2010. Corticospinal excitability during observation and imagery of simple and complex hand tasks: implications for motor rehabilitation. Behav. Brain Res. 213, 35–41. doi:10.1016/j.bbr.2010.04.027

Roy, A.C., Craighero, L., Fabbri-Destro, M., Fadiga, L., 2008. Phonological and lexical motor facilitation during speech listening: A transcranial magnetic stimulation study. J. Physiol. Paris 102, 101–105. doi:10.1016/j.jphysparis.2008.03.006

Ruggles, D., Bharadwaj, H., Shinn-Cunningham, B.G., 2012. Why middle-aged listeners have trouble hearing in everyday settings. Curr. Biol. 22, 1417–22. doi:10.1016/j.cub.2012.05.025

Sato, M., Buccino, G., Gentilucci, M., Cattaneo, L., 2010. On the tip of the tongue: Modulation of the primary motor cortex during audiovisual speech perception. Speech Commun. 52, 533–541. doi:10.1016/j.specom.2009.12.004

Sato, M., Tremblay, P., Gracco, V.L., 2009. A mediating role of the premotor cortex in phoneme segmentation. Brain Lang. 111, 1–7. doi:10.1016/j.bandl.2009.03.002

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. Nat. Rev. Neurosci. 10, 295–302. doi:10.1038/nrn2603

Sebanz, N., Bekkering, H., Knoblich, G., 2006. Joint action: bodies and minds moving together. Trends Cogn. Sci. 10, 70–6. doi:10.1016/j.tics.2005.12.009

Skipper, J.I., Nusbaum, H.C., Small, S.L., 2005. Listening to talking faces: motor cortical activation during speech perception. Neuroimage 25, 76–89. doi:10.1016/j.neuroimage.2004.11.006

Tremblay, P., Sato, M., Small, S.L., 2012. TMS-induced modulation of action sentence priming in the ventral premotor cortex. Neuropsychologia 50, 319–26. doi:10.1016/j.neuropsychologia.2011.12.002

Valchev, N., Zijdewind, I., Keysers, C., Gazzola, V., Avenanti, A., Maurits, N.M., 2015. Weight dependent modulation of motor resonance induced by weight estimation during observation of partially occluded lifting actions. Neuropsychologia 66, 237–45. doi:10.1016/j.neuropsychologia.2014.11.030

Villiger, M., Chandrasekharan, S., Welsh, T.N., 2011. Activity of human motor system during action observation is modulated by object presence. Exp. brain Res. 209, 85–93. doi:10.1007/s00221-010-2522-x

Wassermann, E.M., 2002. Variation in the response to transcranial magnetic brain stimulation in the general population. Clin. Neurophysiol. 113, 1165–71.

Watkins, K., Paus, T., 2004. Modulation of Motor Excitability during Speech Perception: The Role of Broca’s Area. J. Cogn. Neurosci. 16, 978–987. doi:10.1162/0898929041502616

Watkins, K.E., Strafella, a. P., Paus, T., 2003. Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41, 989–994. doi:10.1016/S0028-3932(02)00316-0

Wilson, M., Knoblich, G., 2005. The case for motor involvement in perceiving conspecifics. Psychol. Bull. 131, 460–73. doi:10.1037/0033-2909.131.3.460

Wilson, S.M., Saygin, A.P., Sereno, M.I., Iacoboni, M., 2004. Listening to speech activates motor areas involved in speech production. Nat. Neurosci. 7, 701–2. doi:10.1038/nn1263