

Highlights

- Excitability of the lip area of M1 increases when listening to distorted speech
- M1 lip excitability increases regardless of the source of speech distortion
- Excitability of M1 lip area is also related to the hearing ability of the listener

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

3
4 4 Helen E. Nuttall^{a,b}, Daniel Kennedy-Higgins^b, Joseph T. Devlin^c, Patti Adank^b

5
6
7 5
8
9 6 a. Department of Psychology, Lancaster University, Lancaster, LA1 4YW

10
11
12 7 b. Department of Speech, Hearing and Phonetic Sciences, University College London,
13 Chandler House, 2 Wakefield Street, London, UK, WC1N 1PF

14
15
16 9 c. Department of Experimental Psychology, University College London, 26 Bedford
17 Way, London, UK, WC1H 0AP

18
19 11
20 12
21
22 13 Correspondence should be addressed to Helen Nuttall, Department of Psychology,
23 Lancaster University, Lancaster, LA1 4YW. E-mail: h.nuttall1@lancaster.ac.uk. Tel:
24 14
25 15 +441524592842

26
27 16
28 17
29 18
30 19
31 20
32 21
33 22
34 23
35 24
36 25
37 26
38 27
39 28
40 29
41 30
42 31
43 32
44 33
45 34
46 35
47 36
48 37
49 38
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 **Abstract**

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

29
30 Key words: Speech perception, motor cortex, transcranial magnetic stimulation,
31 motor evoked potentials.

32
33
34
35
36
37 **1. Introduction**

38
39 Successful speech perception is central to everyday communication and
40 quality of life. It is therefore surprising that understanding of the neural bases
41 underpinning speech perception remains limited. Although auditory-related areas

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

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

38 The precise function of observed motor activity during speech perception,
39 however, remains under active debate (Hickok et al., 2011; Lotto et al., 2009; Scott
40 et al., 2009). Recent theories suggest that motor activation may form the basis for
41 the mental simulation of perceived action, which may aid listeners when predicting

1 upcoming speech signals (Gambi & Pickering, 2013; Pickering & Garrod, 2013; Wilson
2 & Knoblich, 2005). Simulation theories of action perception argue that observing
3 actions results in the automatic generation of motor plans required to perform the
4 actions. Simulated motor plans are then used to inform forward models about the
5 co-ordination of one's own muscles to generate a simulated course of movement in
6 parallel with, or even in anticipation of, the movement being perceived (Grush,
7 2004). This type of forward model serves to anticipate others' actions as if they were
8 produced by the observer (Locatelli et al., 2012), and may be used to disambiguate
9 noisy, obscured, or ambiguous actions (Wilson and Knoblich, 2005). With regards to
10 speech perception, these types of conditions may involve listening to speech in the
11 presence of background noise, or listening to someone speaking in an unfamiliar
12 accent (Adank et al., 2012; Adank and Janse, 2009) or manner of speaking (Borrie et
13 al., 2013; Borrie and Schäfer, 2015).

14 Although it is well-established that perceiving speech draws upon
15 hierarchically organized temporo-frontal processing pathways (Davis and Johnsruide,
16 2007, 2003), it is not clear what role premotor and primary motor regions play
17 within this speech processing hierarchy. Knowledge of the nature of articulatory
18 motor representations and their sensitivity to speech is incomplete. For example, it
19 is unknown if, and how, increased motor excitability during perception of
20 challenging speech is modulated by the nature and extent of the speech distortion.
21 Accordingly, two possibilities currently exist for how the motor system responds to
22 distortion in the speech signal. The first is that articulatory motor regions may
23 activate whenever distortion is present in the speech signal, independent of the
24 form or type of speech distortion. The second possibility is that articulatory motor
25 regions may respond differently depending on the type of distortion in the speech
26 signal. If the former is true, it would suggest that the motor system acts as a self-
27 adjusting resource to provide additional information whenever auditory information
28 is found to be insufficient. Support for this prediction comes from demonstrations of
29 heightened motor excitability for both speech-internal distortion (Nuttall, Kennedy-
30 Higgins, Hogan, Devlin, & Adank, 2016) and speech-external distortion (Murakami et
31 al., 2011), yet these two sources of distortion have never been directly compared.

32 Conversely, if the second possibility is true, and activity in motor regions is
33 differentially modulated depending on the type of speech distortion, this indicates
34 that the motor system operates in a form dependent manner during speech
35 perception. Indeed, this is in line with the hypothesis that prediction signals
36 generated by forward models during perception are ideally suited to disambiguate
37 biological sources of variation (Sebanz et al., 2006); for example, when perceiving
38 speech signals that are difficult to understand due to an unfamiliar manner of
39 speech production. This possibility resonates with common-coding accounts of
40 action perception, whereby the motor system is most responsive to observed
41 actions that the observer has experience producing themselves (Calvo-Merino et al.,

1 2005). In this case, the motor system's prediction signal would be less well-suited to
2 assist action understanding when the difficulty arises from a non-motor source, such
3 as speech-shaped background noise, for example, which does not constitute an
4 imitable action. To date, whether, and how, motor facilitation is affected by speech
5 distortion type, or extent of speech distortion, is unknown.

6 Moreover, the quality of the speech signal received by auditory cortex is at
7 the heart of motor simulation. However, previous considerations of speech signal
8 quality have been limited to the properties of the speech stimulus. This is not the
9 only means by which speech signal quality can be degraded. The first point in the
10 speech processing chain begins with the ear, where differences in mechanical and
11 electrical function at the level of the cochlea and auditory nerve can contribute to
12 discrepancies in how the speech signal is processed, even when individuals have
13 clinically normal hearing (Bharadwaj et al., 2015; Harris et al., 2009; Ruggles et al.,
14 2012). Accordingly, it is possible that differences in auditory processing at the ear
15 modulate motor activity in a manner that is qualitatively similar to the effect of a
16 degraded speech stimulus. Indeed, our previous study demonstrated a significant
17 correlation between peripheral hearing acuity and the extent of lip MEP facilitation
18 during distorted relative to clear speech perception, which was not present for hand
19 MEPs (Nuttall, Kennedy-Higgins, Hogan, Devlin, & Adank, 2016). Precisely how
20 hearing abilities relate to the extent of motor activation in optimal and suboptimal
21 listening conditions has not been studied. Relatedly, Peelle and colleagues (2011)
22 found that moderate declines in peripheral auditory processing led to a systematic
23 down-regulation of neural activity in auditory regions during speech processing, and
24 may also contribute to loss of gray matter volume in primary auditory cortex. If
25 motor system activation is interlinked with speech signal quality, as motor
26 simulation accounts would propose, it may be that hearing ability plays a role in
27 engagement of the motor cortex during speech perception. A second aim of this
28 study, therefore, was to explicitly test the relationship between hearing ability and
29 speech motor excitability under different challenging listening conditions.

30 In the present study, we first aimed to disambiguate between form
31 dependent and form independent accounts of how speech distortion modulates
32 motor activation during speech perception, and second, we investigated whether
33 normal variation in hearing ability impacts speech motor facilitation. To this end, in a
34 first experiment, MEPs were elicited during perception of three different types of
35 auditory stimuli: 1) clear speech stimuli presented without background noise, 2)
36 speech stimuli distorted via a motor perturbation introduced during prior stimulus
37 creation (motor distortion), and 3) speech-in-noise (noise distortion), where
38 intelligibility was matched to the motor-distorted speech based on equating percent
39 correct identification between the two degraded stimuli types. For speech-in-noise
40 stimuli, clear speech stimuli were presented in a steady background of speech-
41 shaped noise. In a second experiment, we recorded MEPs during perception of 1)

1 clear speech, 2) speech-in-noise that was 75% intelligible, and 3) speech-in-noise
2 that was 50% intelligible. The same clear speech and noise type from Experiment 1
3 were used in Experiment 2, but signal-to-noise ratio was varied. All speech stimuli
4 were disyllabic vowel-consonant-vowel sounds containing a mixture of consonants
5 that were either lip- (/aba/, /apa/) or tongue-articulated (/ada/, /ata/). Chance
6 performance was therefore always 25%. Stimulation was thus used to determine
7 whether motor facilitation in lip M1 is sensitive to the nature and extent of speech
8 signal degradation. In both experiments, we also measured hearing sensitivity to
9 examine whether hearing ability is related to motor facilitation when perceiving
10 different types of speech distortion.

11 12 13 **2. Methods**

14 15 **2.1 Subjects**

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

33 34 **2.2 Speech stimuli**

35 Stimuli used in Experiments 1 and 2 consisted of twenty vowel-consonant-
36 vowel (VCV) syllables containing an equal distribution of lip- (/apa/, /aba/) or
37 tongue-articulated (/ata/, /ada/) syllables. Two sets of the same twenty syllables
38 were created: a set based on natural, normal articulation, and a set based on motor
39 distortion, where the speaker produced the syllables whilst speaking with a tongue
40 depressor. The tongue depressor was a flat wooden spatula with rounded ends, and
41 was five inches long and one inch wide. Notably, the use of a tongue depressor was

1 intended to impair perception of both tongue- and lip-articulated sounds due to the
2 tongue depressor restricting both the tongue and lower lip movement. A tongue
3 depressor was specifically chosen so as to introduce a motor-based distortion into
4 the speech signal, to relate the speech perception challenge to a speech production
5 difficulty (for further information about these stimuli, and how they are perceived,
6 please see Nuttall et al., 2016). This enabled us to contrast clear speech against
7 distorted speech produced by the same speaker, in contrast to imposing synthetic
8 manipulations upon the spectral characteristics of the original clear speech.

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

31 32 33 **2.3 Design**

34 Two separate experiments were undertaken to assess how activity in the
35 motor system is modulated when listening to distorted speech. In Experiment 1, two
36 different types of speech distortions, speech-in-noise and motor-distorted speech,
37 were used to investigate if the speech motor cortex is sensitive to the quality of
38 distortion in the speech signal. In Experiment 2, one type of speech distortion,
39 speech-in-noise presented at different SNRs, was used to assess if the motor cortex
40 is sensitive to the quantity of the speech distortion. Accordingly, in Experiment 2,
41 two SNRs were used at which subjects were able to correctly identify 50% of the

1 speech stimuli (moderate SNR, condition 1), or 75% of the speech (mild SNR,
2 condition 2).

6 **2.4 Pre-tests**

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

- 12 1. Natural Speech: normally articulated vowel-consonant-vowel (VCV) syllables,
13 which contained an equal distribution of lip- (/apa/, /aba/) and tongue-
14 articulated (/ata/, /ada/) syllables.
- 15 2. Speech-in-noise: listening to normally articulated VCV syllables from Natural
16 Speech condition presented at different levels in speech-shaped
17 unmodulated background noise to yield eight different SNRs. Stimuli
18 contained an equal distribution of lip- (/apa/, /aba/) and tongue-articulated
19 (/ata/, /ada/) syllables. Speech stimuli were presented at SNRs ranging from -
20 3 to -24 dB SPL in steps of 3 dB SPL. This selection of SNRs was based on pilot
21 data that indicated subjects were able to achieve a range of performance
22 levels from near ceiling to chance performance across this range of SNRs for
23 the Natural Speech stimuli.
- 24 3. Distorted Speech: listening to motor-distorted VCV syllables that had been
25 produced whilst using a tongue depressor. Stimuli contained an equal
26 distribution of lip- (/apa/, /aba/) and tongue-articulated (/ata/, /ada/)
27 syllables.

28
29
30 All three behavioural pre-tests were tested in Experiment 1, whereas subjects
31 completed only the first two pre-tests in Experiment 2, as no motor-distorted stimuli
32 were used in the second experiment. Subjects were not given feedback on their
33 performance during the pre-tests, or after the pre-tests.

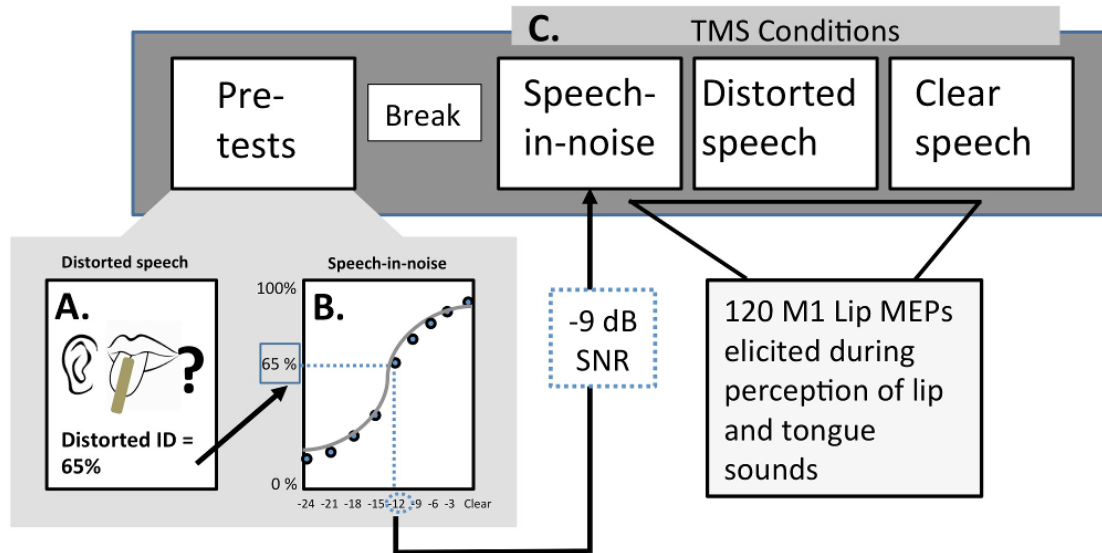


Figure 1. Experimental design of Experiment 1. Subjects initially took part in behavioural pre-tests, allowing intelligibility to be matched between the distorted speech pre-test (A) and the speech-in-noise pre-test (B). The SNR which led to equivalent performance in the speech-in-noise pre-test relative 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 200² unit and delivered by a 70mm diameter figure-of-eight coil, connected through a BiStim² module (Magstim, Dyfed, UK) set to simultaneous

1 discharge mode (inter-pulse spacing of 0 ms). The coil was placed tangential to the
2 skull such that the induced current flowed from posterior to anterior under the
3 junction of the two wings of the figure-of-eight coil. The lip area of M1 was found by
4 using the functional ‘hot spot’ localization method, whereby application of TMS
5 elicits an MEP from the contralateral muscle. Here, the coil position and orientation
6 is adjusted in millimetre movements to ascertain the location on the scalp at which
7 the most robust MEPs are elicited. This location was then marked on a cap and
8 active motor threshold (aMT) determined, which constitutes the intensity at which
9 TMS pulses elicited 5 out of 10 MEPs with an amplitude of at least 50 μ V (Watkins,
10 Strafella, & Paus, 2003; Watkins & Paus, 2004). In this way, we first located the hand
11 area by asking subjects to perform a pinching action where the index finger was held
12 against the thumb to activate first dorsal interosseous. Following this, the lip area
13 ‘hot spot’ was identified by moving the coil ventrally and slightly anterior until an
14 MEP was observed in the contralateral lip muscle, and the aMT identified (Möttönen
15 et al., 2014). The intensity of the stimulator was then set to 120% of aMT for the
16 stimulations applied during the experiment. The mean stimulator intensity (120%
17 aMT \pm SEM) used to elicit lip MEPs in Experiment 1 was 51.6% (\pm 1.1), and in
18 Experiment 2 was 48.3% (\pm 0.8), of the maximum possible intensity.

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

36 2.7 Electromyography

37 Electromyographic (EMG) activity was recorded from the lip muscle
38 orbicularis oris using surface electrodes (Ag/AgCl; 10-mm diameter) in a non-Faraday
39 caged, double-walled sound-attenuating booth. Electrodes were attached to the
40 orbicularis oris on the right side of the mouth in a bipolar montage, with an
41 electrode placed at the right temple serving as a common ground. To stabilize

1 background EMG activity, subjects were trained for approximately five minutes to
2 produce a constant level of contraction (approximately 20-30% of maximum
3 voluntary contraction) of the lip muscles by pursing, which was verified via visual
4 feedback of the ongoing EMG signal, in accordance with standard practice for lip
5 aMT thresholding (Möttönen et al., 2014; Watkins et al., 2003). Contraction of the lip
6 muscle also facilitates a lower motor threshold relative to when the muscle is at rest,
7 enabling the use of lower levels of stimulation during the experiment. The raw EMG
8 signal was amplified by a factor of 1000, band-pass filtered between 100–1000 Hz,
9 and sampled at 5000 Hz online using a 1902 amplifier (Cambridge Electronic Design,
10 Cambridge), and analog-to-digital converted using a Micro1401-3 unit (Cambridge
11 Electronic Design, Cambridge). Continuous data were acquired and recorded using
12 Spike2 software (version 8, Cambridge Electronic Design, Cambridge).

15 **2.8 Data analysis**

16 In Experiment 1, the proportion of correct responses for the distorted speech
17 pre-test were first calculated for each individual. For the speech-in-noise pre-test, a
18 logistic psychometric function based on a maximum likelihood criterion
19 implemented via the Palamedes Matlab toolbox (Prins & Kingdom, 2009) was fitted
20 to the proportion of correct responses at each SNR (-3 dB to -24 dB SPL) for each
21 individual. To match intelligibility between the two pre-tests, the point on the
22 psychometric function that was equivalent to the proportion of correct responses in
23 the distorted speech task was found, and the associated SNR derived (see Figure 1A
24 and B). This SNR was then used for the TMS speech-in-noise condition. In Experiment
25 2, the 50% and 75% performance points on the psychometric functions were
26 obtained for each subject, and the SNRs associated with these levels of performance
27 were used for the speech-in-noise TMS conditions.

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

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

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

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

23 24 25 **3. Results**

26 **3.1 Experiment 1**

27 28 **3.1.1 Behavioural data**

29 For the motor-distorted speech pre-test, performance ranged from 15% to
30 87.5% correct across subjects (chance performance = 25%; mean performance =
31 45.8%, S.D. 20%, Figure 2B), in line with previous observations (Nuttall et al., 2016).
32 Accordingly, this gave rise to a wide range of intelligibility-matched SNRs ranging
33 from -24 to -3.5 SNR dB SPL (Figure 2B) that yielded an equivalent level of
34 performance in the speech-in-noise task (mean -16 dB SPL, S.D. 7.5 dB SPL). Figure
35 2A shows the average proportion of correct responses across subjects for each SNR,
36 and the association between motor-distorted speech pre-test scores and matched
37 SNRs is shown in Figure 2B. Chance performance was also 25% in the speech-in-noise
38 pre-test.

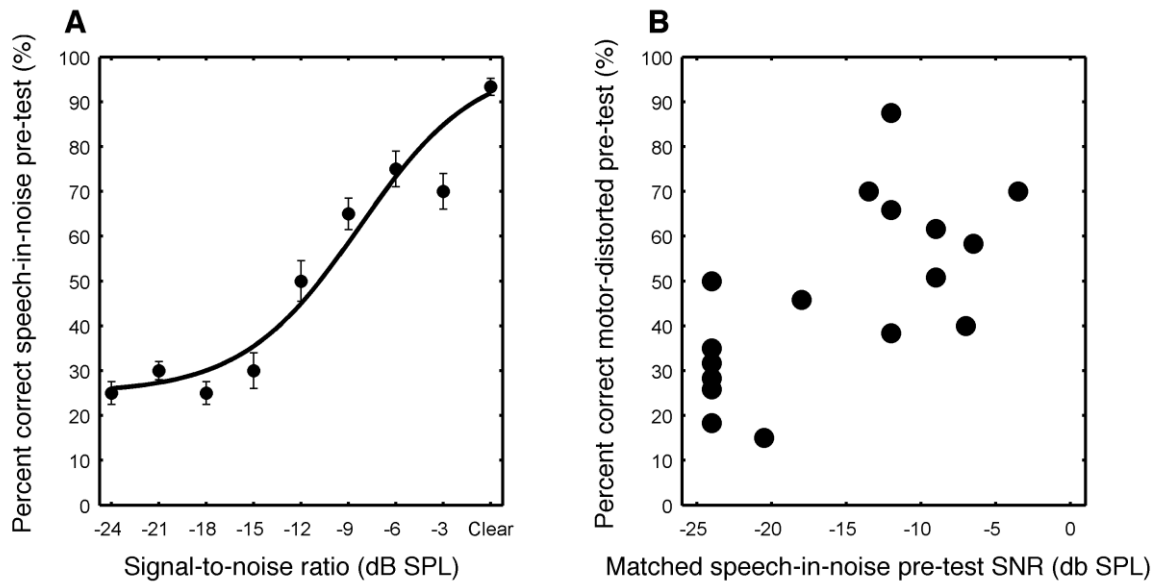


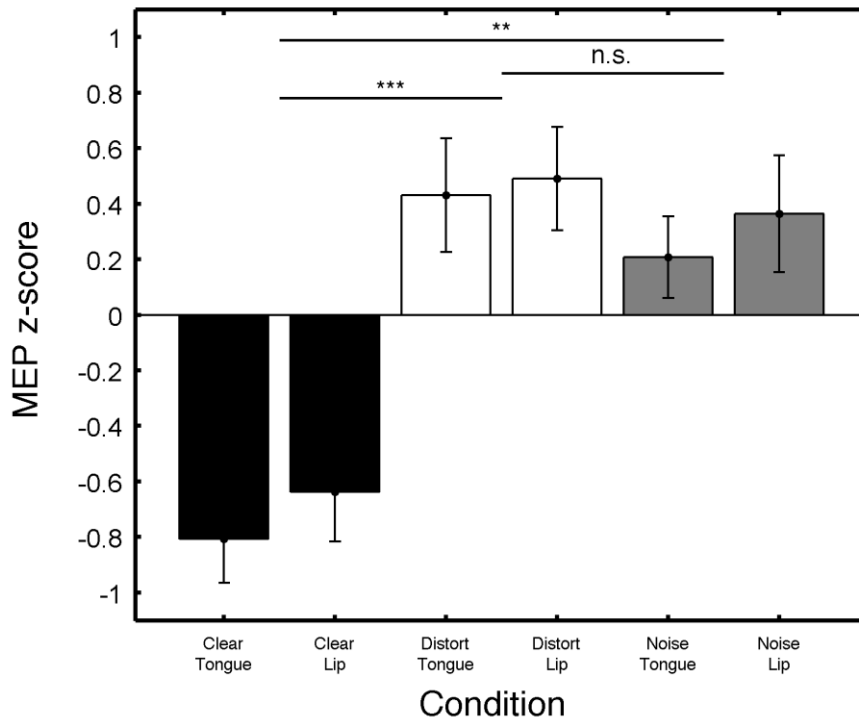
Figure 2. Psychometric function fitted to grand mean proportion of correct responses across SNRs and for clear speech \pm SEM (A). Association between motor-distorted pre-test scores and matched speech-in-noise pre-test SNRs (B).

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 \times articulator: $p = 0.9$; speech type \times hearing ability (pure tone average [PTA]) $p = 0.2$; articulator \times PTA: $p = 0.4$; speech type \times articulator \times 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

1 significant difference between MEPs recorded during perception of natural and
 2 distorted speech ($p = 0.001$), and between MEPs recorded during perception of
 3 natural speech and speech-in-noise ($p = 0.005$). However, the speech distortions did
 4 not differ significantly in their effect on MEP size ($p = 1.0$).



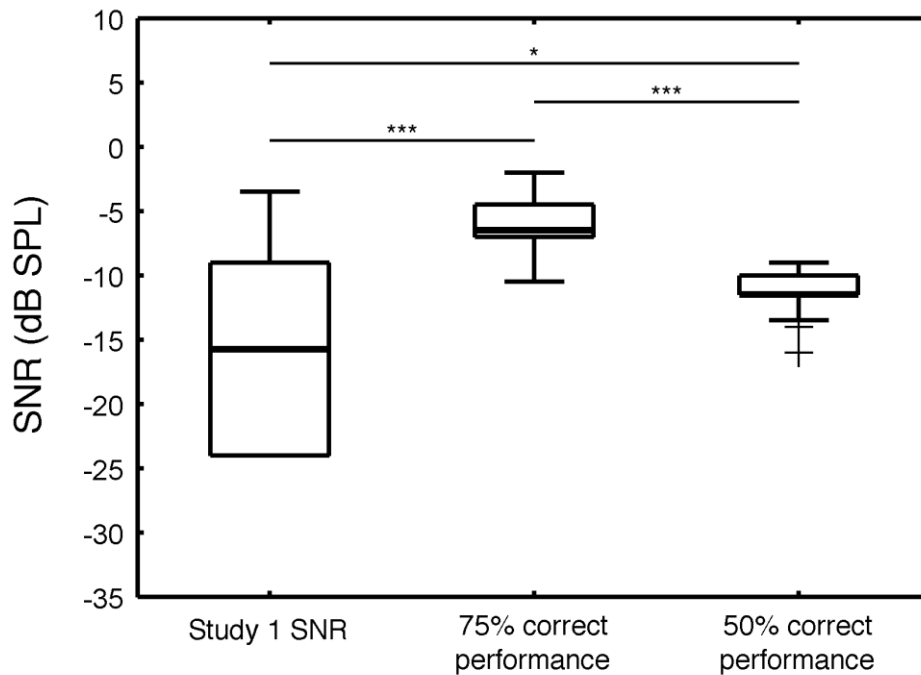
6 Figure 3. Mean standardized area under the curve (AUC) of MEPs elicited during
 7 perception of clear (black bars), motor-distorted (white bars), and speech-in-noise
 8 (grey bars) stimuli, articulated by the lips or tongue (\pm SEM). *** = $p \leq 0.001$; ** = $p \leq 0.01$;
 9 n.s. = not significant.

12 3.2 Experiment 2

14 3.2.1 Behavioural data

15 The results from the speech-in-noise pre-test in Experiment 2 were used to
 16 find the two SNRs at which subjects could correctly identify 50% and 75% of the
 17 speech stimuli in each subject. The SNRs associated with 50% correct performance
 18 ranged between -16 and -9 dB SPL across subjects (mean SNR -11.5 dB SPL, S.D. 1.7
 19 dB SPL). The SNRs associated with 75% correct performance ranged between
 20 -10.5 and -2 dB SPL across subjects (mean SNR -6 dB SPL, S.D. 2.1 dB SPL, Figure 4). A
 21 paired t-test confirmed that the SNRs associated with 50% performance and the
 22 SNRs associated with 75% performance were significantly different ($t(17) = 16.6, p <$

1 0.001). Notably, independent t-tests (equal variances not assumed) confirmed that
 2 both sets of SNRs were also significantly different from the SNRs used in Experiment
 3 1 (Figure 4), which were matched to the intelligibility of the motor-distorted speech
 4 (SNR 50%: $t(18.76) = -2.58, p = 0.018$, SNR 75%: $t(19.67) = -5.56, p < 0.001$).
 5



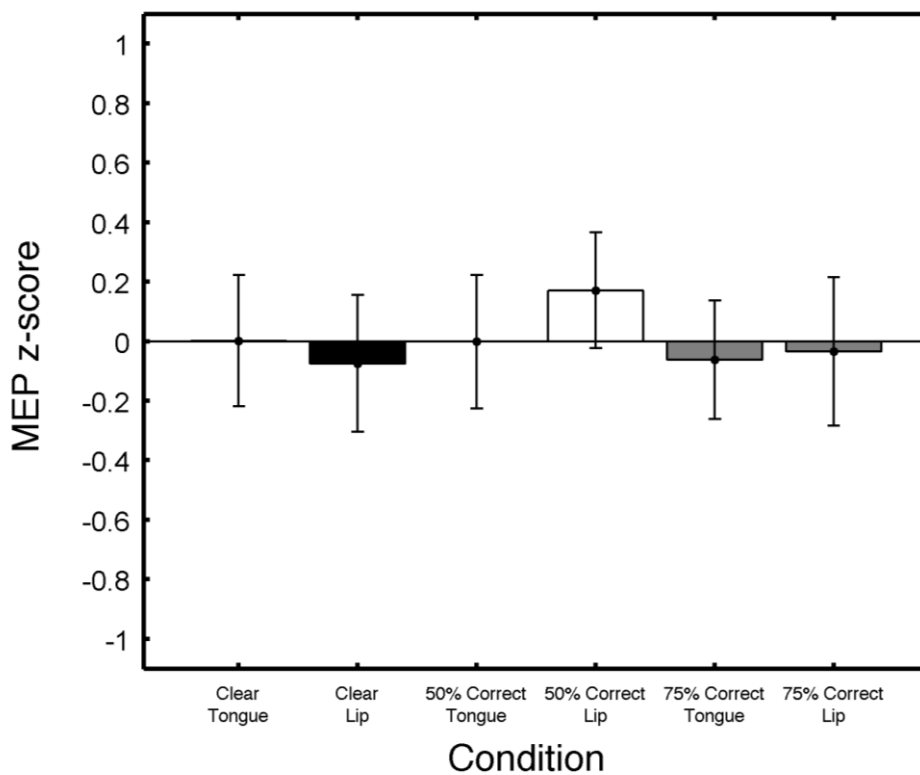
6
 7
 8
 9 Figure 4. Box plots of the SNRs at which 75% and 50% correct performance was
 10 achieved on the speech-in-noise pre-test in Experiment 2, and of the SNRs from
 11 Experiment 1 for comparison. SNRs from Experiment 1 ranged widely due to being
 12 matched to performance on the motor-distorted speech pre-test. In Experiment 2,
 13 the two SNRs associated with 50% and 75% correct performance for each subject
 14 were subsequently used in the two speech-in-noise TMS conditions. *** = $p < 0.001$;
 15 * = $p < 0.05$.
 16
 17

18 3.2.2 Motor evoked potentials

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

22 The mean difference in lip MEP size as a function of speech type and
 23 articulator is reflected in Figure 5. On average, there were no observable differences
 24 between MEPs elicited during perception of speech-in-noise at varying levels of

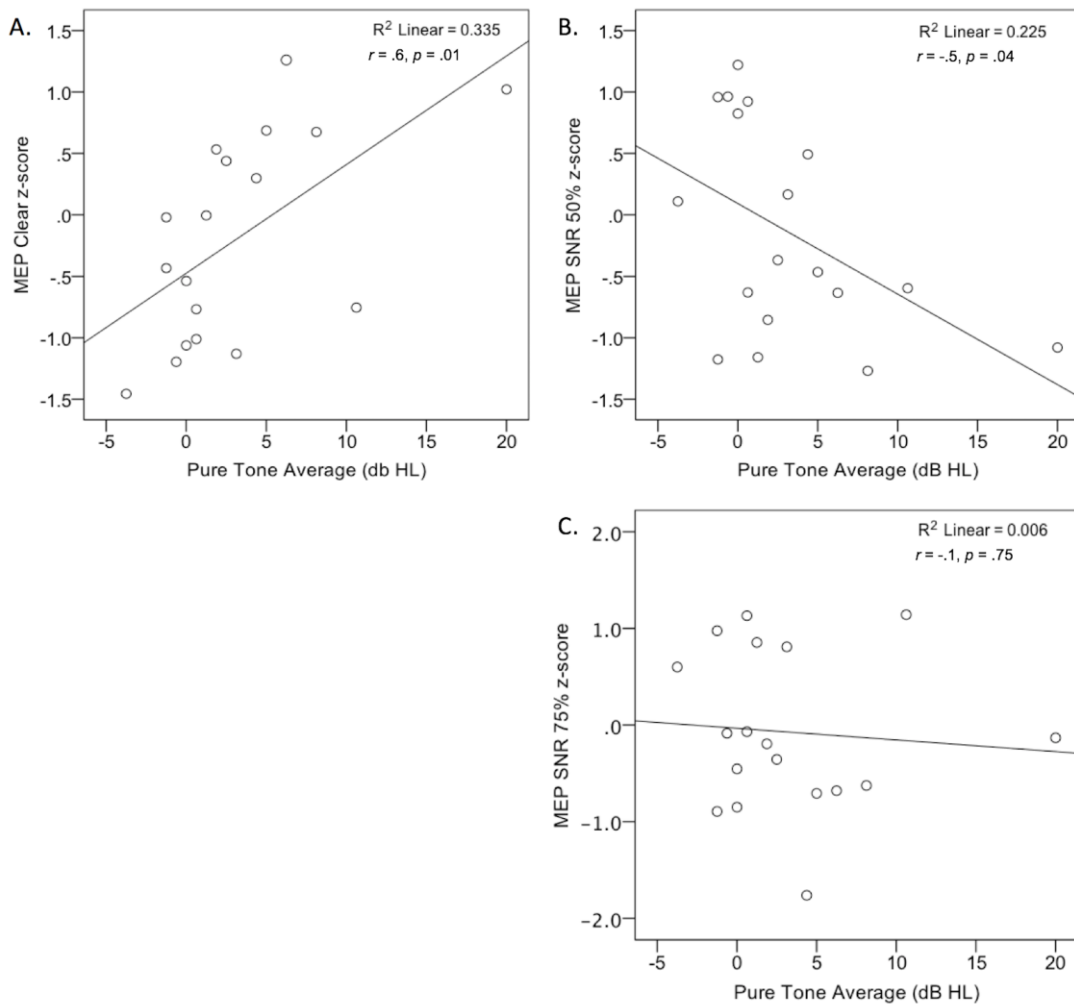
1 intelligibility, or differences depending on place of articulation of the speech
 2 stimulus. After controlling for the effect of hearing ability, a two-way repeated
 3 measures ANCOVA found that the main effect of speech type was non-significant
 4 ($F(2,32) = 0.5; p = 0.6$). However, the covariate, hearing ability, was significantly
 5 related to the effect of speech type ($F(2,32) = 3.6, p = 0.03$, partial $\eta^2 = .2$), suggesting
 6 that MEPs were differentially modulated by speech type in a manner dependent on
 7 the listener's peripheral hearing acuity. All other main effects and interactions were
 8 non-significant (articulator: $p = 0.8$; speech type \times articulator: $p = 0.9$; articulator \times
 9 PTA: $p = 0.8$; speech type \times articulator \times PTA: $p = 0.9$).



13
 14
 15
 16 Figure 5. Mean standardized area under the curve (AUC) of MEPs elicited during
 17 perception of speech-in-noise when speech was 50% intelligible (white bars), 75%
 18 intelligible (grey bars), and when speech was presented without noise (black bars),
 19 for syllables articulated by either the lips or the tongue (\pm SEM).

20
 21
 22
 23 A correlation analysis was conducted in order to characterize the significant
 24 interaction. To test the relationship between hearing ability and motor activity, we

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



14
 15 Figure 6. Correlations between pure tone average (PTA) hearing sensitivity, and MEP
 16 area during perception of clear speech (A), speech-in-noise when intelligibility was
 17 50% (B), and speech-in-noise when intelligibility was 75% (C).

18 Taken together, these data suggest that listeners with less sensitive hearing
 19 exhibit more motor activity during perception of clear speech relative to speech-in-

1 noise compared to listeners with better hearing, who tend to show more motor
2 cortex activation when listening to speech-in-noise relative to clear speech. As such,
3 this difference appears to obfuscate the effect of speech type on motor activity,
4 which averages out when considered at the level of the group mean.
5
6
7
8

9 **4. Discussion**

10
11 The present study first aimed to establish whether motor facilitation
12 resulting from degraded speech perception is modulated by the type and extent of
13 degradation (form dependent), or is independent of the form of degradation (form
14 independent). Second, we aimed to establish whether hearing ability is associated
15 with articulatory motor activation during speech perception. Experiment 1 replicated
16 earlier findings that showed that activity in the motor system is enhanced when
17 listening to distorted speech (Murakami et al., 2011; Nuttall et al., 2016). Experiment
18 1 also demonstrated that both speech external distortions (noise), and speech
19 internal distortions (speech produced with a motor perturbation), led to equivalent
20 increases in MEP area. This finding suggests that the motor system does not
21 differentiate between speech external and speech internal types of distortion.
22 Experiment 2 used significantly less adverse SNRs compared to Experiment 1, and
23 did not find evidence of motor system facilitation when speech could be identified
24 with either 50% or 75% accuracy in background noise. However, there was a
25 significant interaction between noise condition and hearing ability, demonstrating
26 that motor facilitation during perception of speech-in-noise is modulated by hearing
27 ability, but only when the SNR is moderate. Here, young adults with less sensitive
28 hearing showed significantly more motor activation when listening to clear speech.
29 Conversely, listeners with better hearing showed greater motor activation
30 (facilitation) during perception of speech-in-noise.

31
32 Our observation in Experiment 1 that the speech motor system did not
33 differentiate between different forms of speech distortion supports a form
34 independent role of the primary motor system during speech perception. Similar
35 studies on orofacial neurophysiology are limited, but our findings are in line with
36 studies on observation of degraded visual hand stimuli, which highlight a sensitivity
37 of the motor cortex when presenting more difficult to produce finger-tapping
38 actions relative to simple finger-tapping actions (Roosink and Zijdwind, 2010).
39 Similarly, the motor cortex has been shown to remain sensitive to hand action
40 observation even when the hand grasping action cannot be fully observed (Valchev
41 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

1 activity increases in order to provide supplementary information and improve the
2 detection of the speech signal. Such functionality would not need to be sensitive to
3 differences in the acoustic form of the distortion; it would simply attempt to
4 compensate for the missing information. Form independent motor facilitation
5 complements contemporary ideas on the neural hierarchy involved in spoken
6 language processing, whereby temporal auditory areas at the origin of cortical
7 speech processing are sensitive to surface acoustic form, such as speaker identity
8 and intelligibility. Somatomotor regions are considered to be at the pinnacle of the
9 processing hierarchy, and potentially most involved with higher-level linguistic
10 abstraction, and not explicitly considered to be sensitive to variations in acoustic
11 form (Evans and Davis, 2015). Indeed, Evans and Davis (2015) observed blood
12 oxygen level dependent (BOLD) changes to indicate that the somatomotor region is
13 sensitive to syllable identity but not to surface acoustic form, with adjacent motor
14 areas demonstrating greater activity for degraded relative to clear speech. Whilst
15 motor areas may be sensitive to the presence of distortion in the signal, it appears
16 that they are insensitive to the nature or quality of the distortion.

17 Unlike Evans and Davis (2015), and other TMS studies (Bartoli et al., 2015;
18 Fadiga et al., 2002; Roy et al., 2008), we did not find evidence that the motor system
19 is sensitive to syllable identity (somatotopic for speech). This may be due to the fact
20 that we recorded MEPs from one articulator only, orbicularis oris, which in
21 combination with the sensitivity of our design did not emphasize somatotopic
22 speech activity during passive speech perception. Contrasting activity in M1 lip with
23 M1 tongue, for example, during perception of the same lip- and tongue-articulated
24 stimuli, presents a more optimal test of somatotopy. Although Evans and Davis
25 (2015) did not specifically demarcate different articulators in their fMRI analyses, it is
26 possible that they observed identity-specific changes during perception of syllables
27 due to the detection of BOLD activity associated with more than one articulator.
28 Furthermore, engaging in an active task, as opposed to passive listening as in the
29 present study, requires an explicit need to differentiate places of articulation during
30 speech perception, and this may also increase the likelihood of detecting subtle,
31 somatotopic responses. Lastly, new evidence has come to light that suggests that the
32 sensorimotor cortex activates along an acoustic, and not somatotopic, dimension
33 during speech perception. Cheung and colleagues (2016) found that sounds that
34 have similar acoustic properties but different motor configurations, such as 'b' and
35 'd', activate the motor cortex in similar ways when perceiving speech. This may also
36 relate to why we did not find somatotopic MEP differences.

37 In Experiment 2, we did not find a significant main effect of noise on MEPs
38 elicited during perception of speech-in-noise, in contrast to Experiment 1. This may
39 be in part related to the significantly less adverse SNRs used in Experiment 2 relative
40 to Experiment 1. It should be noted that the results of Experiment 2 do not
41 constitute a total replication of Experiment 1. Experiments 1 and 2 use significantly

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

25 Nonetheless, in listeners with better hearing, we did observe evidence of
26 motor facilitation during perception of speech in the most adverse noise condition
27 compared to clear speech. This does not preclude the possibility that other levels of
28 intelligibility, SNR, or types of noise masking may differentially modulate auditory-
29 motor activity. Interestingly, we have previously observed that individuals with
30 better hearing show less facilitation when listening to speech which is internally-
31 distorted relative to listeners with less sensitive hearing (Nuttall et al., 2016). These
32 diverging relationships are most likely linked to the influence of different auditory
33 processes during speech perception, as the effect of background noise has a very
34 different impact on the auditory system compared to the effect of speech that is
35 difficult to recognize due to the manner of speaking. For example, background noise
36 causes neural adaption at the inner hair cell-auditory nerve junction, and as a result
37 is thought to engage auditory efferent fibers to provide relief from background noise
38 masking (Guinan, 2006). Therefore, different speech distortions do not affect the
39 auditory system equally, and thus correlations between hearing and cortical motor
40 facilitation during different types of speech distortion may be mediated by different
41 auditory mechanisms. We also cannot establish if exposure to the numerous stimuli

1 in the pre-test in Experiment 2, which always preceded the TMS experiment, may
2 have affected listening strategy, motor activation, or adaptation during MEP
3 recording, as subjects were not exposed to any speech stimuli prior to TMS in the
4 previous study (Nuttall et al. 2016). Notably, hearing ability did not interact with
5 motor activity in Experiment 1, which used significantly more adverse speech
6 distortions than in Experiment 2. There was also no correlation between MEPs
7 recorded during perception of clear speech in Experiment 1 and hearing ability ($r = -$
8 0.2 , $p = 0.5$). This is most likely due to the presence of significantly smaller and less
9 variable MEPs in Experiment 1 than in Experiment 2 ($t(34) = -2.7$, $p = 0.01$; equal
10 variances not assumed). Listeners' hearing abilities were not significantly different
11 across the two experiments ($t(34) = -0.4$, $p = 0.66$), and therefore it is unlikely that
12 differences in hearing ability led to the difference in correlations between
13 Experiments 1 and 2.

14 In light of the current data, two outstanding questions remain: 1) how much
15 distortion needs to be present in the speech signal before motor activity is engaged,
16 and 2) how does the extent of motor activity interact with encoding of the auditory
17 signal at the ear, and beyond. The findings of Experiment 1 and 2 give some
18 indication of these relationships. In Experiment 1, the mean SNR used in the noise
19 condition was -16 dB SPL, whereas in Experiment 2, the mean SNR for the mild noise
20 condition (SNR 75%) was -6 dB SPL, and for the moderate noise condition (SNR 50%),
21 -11.5 dB SPL. This difference in SNRs between experiments arises from the
22 intelligibility matching in Experiment 1. Participants achieved a wide range of scores
23 on the motor-distorted pre-test, ranging from 15% to 87.5%. If participants
24 performed around or below chance (25%) on the motor-distorted pre-test, this
25 would result in a considerably adverse SNR being used in the TMS speech-in-noise
26 condition, to equate performance to around chance in the speech-in-noise pre-test.
27 Conversely, in Experiment 2, the SNRs were chosen based solely on the 50% and 75%
28 performance points in the speech-in-noise pre-test. Therefore, a greater range of
29 more adverse SNRs were used in Experiment 1 compared to Experiment 2.

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

37 Related work from Du and colleagues (2014) using fMRI has demonstrated
38 that speech motor areas exhibit significant phonetic discrimination activation for
39 SNRs above -6 dB SPL, whereas bilateral auditory cortices encode phoneme-specific
40 information only when the noise is absent or extremely weak (SNR > 8 dB). Notably,
41 the most adverse SNR tested by Du and colleagues was -8 dB SPL, and how to equate

1 the BOLD signal and MEP is not straightforward. However, Du et al's findings, as well
2 as our findings, support the possibility that the motor system assists the auditory
3 system during speech perception under challenging listening conditions. This is in
4 line with recent work from Murakami and colleagues (2015), who found that
5 disrupting motor function caused subtle but significant phonological errors, the
6 extent of which was dependent upon the level of background noise. Notably,
7 however, the authors observed substantial neurophysiological changes during
8 passive speech perception, which were linked only to limited effects on behavioural
9 speech perception. Therefore, we concur with Murakami et al. (2015) that
10 sensorimotor mapping via the dorsal stream may not be essential for speech
11 perception, but that it contributes to subtle phonological analyses when sensory
12 information is insufficient.

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

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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 <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

- 1 the interface between audition and speech perception. *Hear. Res.* 229, 132–47.
2 doi:10.1016/j.heares.2007.01.014
- 3 Davis, M.H., Johnsrude, I.S., 2003. Hierarchical processing in spoken language
4 comprehension. *J. Neurosci.* 23, 3423–31.
- 5 Devlin, J.T., Watkins, K.E., 2007. Stimulating language: insights from TMS. *Brain* 130,
6 610–22. doi:10.1093/brain/awl331
- 7 Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts
8 phoneme representations in the auditory and speech motor systems. *Proc.*
9 *Natl. Acad. Sci. U. S. A.* 111, 7126–31. doi:10.1073/pnas.1318738111
- 10 Evans, S., Davis, M.H., 2015. Hierarchical Organization of Auditory and Motor
11 Representations in Speech Perception: Evidence from Searchlight Similarity
12 Analysis. *Cereb. Cortex.* doi:10.1093/cercor/bhv136
- 13 Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G., 2002. Speech listening specifically
14 modulates the excitability of tongue muscles: A TMS study. *Eur. J. Neurosci.* 15,
15 399–402.
- 16 Gambi, C., Pickering, M.J., 2013. Prediction and imitation in speech. *Front. Psychol.* 4,
17 340. doi:10.3389/fpsyg.2013.00340
- 18 Grush, R., 2004. The emulation theory of representation: motor control, imagery,
19 and perception. *Behav. Brain Sci.* 27, 377–96–442.
- 20 Guinan, J.J., 2006. Olivocochlear efferents: anatomy, physiology, function, and the
21 measurement of efferent effects in humans. *Ear Hear.* 27, 589–607.
22 doi:10.1097/01.aud.0000240507.83072.e7
- 23 Harris, K.C., Dubno, J.R., Keren, N.I., Ahlstrom, J.B., Eckert, M. a, 2009. Speech
24 recognition in younger and older adults: a dependency on low-level auditory
25 cortex. *J. Neurosci.* 29, 6078–87. doi:10.1523/JNEUROSCI.0412-09.2009
- 26 Hervais-Adelman, A.G., Carlyon, R.P., Johnsrude, I.S., Davis, M.H., 2012. Brain regions
27 recruited for the effortful comprehension of noise-vocoded words. *Lang. Cogn.*
28 *Process.* 27, 1145–1166. doi:10.1080/01690965.2012.662280
- 29 Hickok, G., Houde, J., Rong, F., 2011. Sensorimotor integration in speech processing:
30 computational basis and neural organization. *Neuron* 69, 407–22.
31 doi:10.1016/j.neuron.2011.01.019
- 32 Keil, J., Timm, J., Sanmiguel, I., Schulz, H., Obleser, J., Schönwiesner, M., 2014.
33 Cortical brain states and corticospinal synchronization influence TMS-evoked
34 motor potentials. *J. Neurophysiol.* 111, 513–9. doi:10.1152/jn.00387.2013
- 35 Kiers, L., Cros, D., Chiappa, K.H., Fang, J., 1993. Variability of motor potentials evoked
36 by transcranial magnetic stimulation. *Electroencephalogr. Clin. Neurophysiol.*
37 *Potentials Sect.* 89, 415–423. doi:10.1016/0168-5597(93)90115-6
- 38 Locatelli, M., Gatti, R., Tettamanti, M., 2012. Training of manual actions improves
39 language understanding of semantically related action sentences. *Front.*
40 *Psychol.* 3, 547. doi:10.3389/fpsyg.2012.00547
- 41 Londei, A., D’Ausilio, A., Basso, D., Sestieri, C., Gratta, C. Del, Romani, G.-L.,
42 Belardinelli, M.O., 2010. Sensory-motor brain network connectivity for speech
43 comprehension. *Hum. Brain Mapp.* 31, 567–80. doi:10.1002/hbm.20888
- 44 Lotto, A.J., Hickok, G.S., Holt, L.L., 2009. Reflections on mirror neurons and speech
45 perception. *Trends Cogn. Sci.* 13, 110–4. doi:10.1016/j.tics.2008.11.008
- 46 Meister, Wilson, Deblieck, Wu, 2007. Report The Essential Role of Premotor Cortex
47 in Speech Perception. *Curr. Biol.* 17, 1692–1696. doi:10.1016/j.cub.2007.08.064

- 1 Möttönen, R., Rogers, J., Watkins, K.E., 2014. Stimulating the lip motor cortex with
2 transcranial magnetic stimulation. *J. Vis. Exp.* doi:10.3791/51665
- 3 Möttönen, R., Watkins, K.E., 2009. Motor representations of articulators contribute
4 to categorical perception of speech sounds. *J. Neurosci.* 29, 9819–9825.
5 doi:10.1523/JNEUROSCI.6018-08.2009
- 6 Murakami, T., Kell, C.A., Restle, J., Ugawa, Y., Ziemann, U., 2015. Left Dorsal Speech
7 Stream Components and Their Contribution to Phonological Processing. *J.*
8 *Neurosci.* 35, 1411–1422. doi:10.1523/JNEUROSCI.0246-14.2015
- 9 Murakami, T., Restle, J., Ziemann, U., 2011. Observation-execution matching and
10 action inhibition in human primary motor cortex during viewing of speech-
11 related lip movements or listening to speech. *Neuropsychologia* 49, 2045–2054.
12 doi:10.1016/j.neuropsychologia.2011.03.034
- 13 Neff, D.L., Green, D.M., 1987. Masking produced by spectral uncertainty with
14 multicomponent maskers. *Percept. Psychophys.* 41, 409–15.
- 15 Nuttall, H.E., Kennedy-Higgins, D., Hogan, J., Devlin, J.T., Adank, P., 2016. The effect
16 of speech distortion on the excitability of articulatory motor cortex.
17 *Neuroimage* 128, 218–226. doi:10.1016/j.neuroimage.2015.12.038
- 18 Nuttall, H.E., Moore, D.R., Barry, J.G., Krumbholz, K., de Boer, J., 2015. The influence
19 of cochlear spectral processing on the timing and amplitude of the speech-
20 evoked auditory brain stem response. *J. Neurophysiol.* 113, 3683–91.
21 doi:10.1152/jn.00548.2014
- 22 Osnes, B., Hugdahl, K., Specht, K., 2011. Effective connectivity analysis demonstrates
23 involvement of premotor cortex during speech perception. *Neuroimage* 54,
24 2437–45. doi:10.1016/j.neuroimage.2010.09.078
- 25 Peelle, J.E., Troiani, V., Grossman, M., Wingfield, A., 2011. Hearing loss in older
26 adults affects neural systems supporting speech comprehension. *J. Neurosci.*
27 31, 12638–43. doi:10.1523/JNEUROSCI.2559-11.2011
- 28 Pickering, M.J., Garrod, S., 2013. An integrated theory of language production and
29 comprehension. *Behav. Brain Sci.* 36, 329–47. doi:10.1017/S0140525X12001495
- 30 Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analysing
31 18 psychophysical data. Retrieved 18 February 2016 from
32 19 <http://www.palamedestoolbox.org>
- 33 Roosink, M., Zijdwind, I., 2010. Corticospinal excitability during observation and
34 imagery of simple and complex hand tasks: implications for motor
35 rehabilitation. *Behav. Brain Res.* 213, 35–41. doi:10.1016/j.bbr.2010.04.027
- 36 Roy, A.C., Craighero, L., Fabbri-Destro, M., Fadiga, L., 2008. Phonological and lexical
37 motor facilitation during speech listening: A transcranial magnetic stimulation
38 study. *J. Physiol. Paris* 102, 101–105. doi:10.1016/j.jphysparis.2008.03.006
- 39 Ruggles, D., Bharadwaj, H., Shinn-Cunningham, B.G., 2012. Why middle-aged
40 listeners have trouble hearing in everyday settings. *Curr. Biol.* 22, 1417–22.
41 doi:10.1016/j.cub.2012.05.025
- 42 Sato, M., Buccino, G., Gentilucci, M., Cattaneo, L., 2010. On the tip of the tongue:
43 Modulation of the primary motor cortex during audiovisual speech perception.
44 *Speech Commun.* 52, 533–541. doi:10.1016/j.specom.2009.12.004
- 45 Sato, M., Tremblay, P., Gracco, V.L., 2009. A mediating role of the premotor cortex in
46 phoneme segmentation. *Brain Lang.* 111, 1–7. doi:10.1016/j.bandl.2009.03.002
- 47 Scott, S.K., McGettigan, C., Eisner, F., 2009. A little more conversation, a little less

1 action--candidate roles for the motor cortex in speech perception. *Nat. Rev.*
2 *Neurosci.* 10, 295–302. doi:10.1038/nrn2603

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

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

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

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

15 Villiger, M., Chandrasekharan, S., Welsh, T.N., 2011. Activity of human motor system
16 during action observation is modulated by object presence. *Exp. brain Res.* 209,
17 85–93. doi:10.1007/s00221-010-2522-x

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

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

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

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

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

31

