NeuroImage 128 (2016) 218–226

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

The effect of speech distortion on the excitability of articulatory motor cortex

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

[a] Department of Speech, Hearing and Phonetic Sciences, University College London, Chandler House, 2 Wakefield Street, London, UK, WC1N 1PF
[b] Department of Experimental Psychology, University College London, 26 Bedford Way, London, UK, WC1H 0AP

ARTICLE INFO

Article history:
Received 21 August 2015
Accepted 21 December 2015
Available online 28 December 2015

Keywords:
Speech perception
Motor cortex
Transcranial magnetic stimulation
Motor evoked potentials

ABSTRACT

It has become increasingly evident that human motor circuits are active during speech perception. However, the conditions under which the motor system modulates speech perception are not clear. Two prominent accounts make distinct predictions for how listening to speech engages speech motor representations. The first account suggests that the motor system is most strongly activated when observing familiar actions (Pickering and Garrod, 2013). Conversely, Wilson and Knoblich's account asserts that motor excitability is greatest when observing less familiar, ambiguous actions (Wilson and Knoblich, 2005). We investigated these predictions using transcranial magnetic stimulation (TMS). Stimulation of the lip and hand representations in the left primary motor cortex elicited motor evoked potentials (MEPs) indexing the excitability of the underlying motor representation. MEPs for lip, but not for hand, were larger during perception of distorted speech produced using a tongue depressor, relative to naturally produced speech. Additional somatotopic facilitation yielded significantly larger MEPs during perception of lip-articulated distorted speech sounds relative to distorted tongue-articulated sounds. Critically, there was a positive correlation between MEP size and the perception of distorted speech sounds. These findings were consistent with predictions made by Wilson & Knoblich (Wilson and Knoblich, 2005), and provide direct evidence of increased motor excitability when speech perception is difficult.

Introduction

Human listeners are adept at perceiving speech in a variety of listening conditions, which can differ widely in the difficulty they pose to the listener. Indeed, most of us will have engaged in a conversation affected by a poor telephone connection, an unfamiliar speech style, or a distracting discussion taking place nearby, but despite such limitations we remain remarkably good at extracting the meaning from our interlocutor's speech. It is surprising, therefore, that the neural architecture underlying this success remains little understood. Current models outlining the neural organization of speech processing (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009) propose that the locus of intelligible speech understanding is the temporal lobe within the ventral stream of speech processing. However, the neural pathway of the ventral stream differs in these two models; Rauschecker & Scott suggest that speech processing has its center of gravity in left anterior STS (Superior Temporal Sulcus), while Hickok & Poeppel propose that recognising intelligible speech is bilaterally organized and located both anteriorly and posteriorly to Heschl's Gyrus. Both models also feature a dorsal stream, which is thought to translate acoustic speech signals into articulatory representations for speech production, though the models also posit differences in the integration of dorsal stream function (for details see Hickok and Poeppel, 2007; Rauschecker and Scott, 2009, and Turkeltaub and Coslett, 2010; Adank et al., 2012 for more in-depth reviews).

Evidence suggests that temporal areas in the ventral stream form part of a functional hierarchy for speech processing, where primary auditory cortex is sensitive to the acoustic features of speech, but higher-order temporal, and frontal, sites (middle and superior temporal gyri, left inferior frontal gyrus) are sensitive to the intelligibility of speech, but insensitive to the acoustic form of the stimuli (Davis and Johnsrude, 2003). Accordingly, it has been demonstrated that when listening to speech in challenging conditions, activity increases in peri-auditory and frontal regions relative to when listening to intelligible speech, and it is thought that such activity may support processing in primary auditory areas and help compensate for the acoustic distortion (Davis and Johnsrude, 2003, 2007; Shahin et al., 2009; Wild et al., 2012a). Concurrently, Wild and colleagues (Wild et al., 2012b) have recently shown that motor regions and left inferior frontal gyrus exhibit elevated responses when attending to degraded speech in the presence of auditory distractors. As such, it is becoming increasingly apparent that speech perception may also involve areas beyond those classic temporal sites already identified.

In particular, recent years have seen renewed interest in the idea that cortical motor systems involved in producing speech may also contribute to perceiving it. Originally proposed by Lieberman (Liberman

* Corresponding author at: Department of Speech, Hearing and Phonetic Sciences, University College London, Chandler House, 2 Wakefield Street, London, WC1N 1PF, UK.
E-mail address: h.nuttall@ucl.ac.uk (H.E. Nuttall).

http://dx.doi.org/10.1016/j.neuroimage.2015.12.038
1053-8119/

and Mattingly, 1985; Liberman et al., 1967), the Motor Theory of Speech Perception was roundly criticized for its claim that motor cortex, rather than auditory cortex, was the key site for speech comprehension (Diehl et al., 2004; Jusczyk et al., 1981; Scott et al., 2009). Nonetheless, accumulating evidence from Transcranial Magnetic Stimulation studies (TMS) suggests that regions of primary motor cortex (M1), important for accurate control of articulatory gestures, activate during speech comprehension, and are also involved in the precise categorization of complex acoustic signals (D'Ausilio et al., 2009; Möttönen and Watkins, 2009; Sato et al., 2010). These TMS findings resonate with functional Magnetic Resonance Imaging (fMRI) observations of motor cortex activation during speech perception (Osnes et al., 2011; Hervais-Adelman et al., 2012; Szenkovits et al., 2012). Furthermore, this activation is modulated in a somatotopic way, whereby speech articulators in left motor cortex are more responsive when listening to speech produced using the same articulator, compared to when listening to speech produced using a different articulator. Indeed, Fadiga and colleagues (Fadiga et al., 2002) demonstrated that passively listening to words that involve tongue articulation results in the automatic facilitation of the tongue region in primary motor cortex. Such facilitation does not result in overt movement generation, but can be observed as changes in the potentiation of the tongue muscle resulting from increased corticobulbar excitation.

The precise contribution of this observed motor activation, however, is under active debate, and the field is still divided in opinion as to whether the articulatory motor system is essential for speech comprehension. What has been acknowledged is that the motor system may have a modulatory influence on perceptual systems (Hickok et al., 2011). However, the conditions under which the motor system has cause to modulate audition are not clear. Evidence from fMRI argues for a preferential engagement of the motor system when listening to speech that is difficult to understand. In a recent study, Du and colleagues (Du et al., 2014) tested the hypothesis that motor activation contributes to categorical speech perception under adverse, but not quiet, listening conditions. The authors observed a negative correlation between neural activity and perceptual accuracy in left premotor cortex, which contributed to phoneme categorization specifically at moderate-to-adverse signal-to-noise ratios. Using TMS and a highly similar phoneme categorization tasks to assess categorical perception, Möttönen and Watkins (Möttönen and Watkins, 2009) demonstrated that primary motor cortex makes a significant contribution to phoneme judgments in quiet listening conditions, though specifically at the ambiguous phonetic category boundary. As such, evidence corroborating the activation of motor processes during speech perception is compelling, but our understanding of the listening conditions that preferentially engage the speech motor system is uncertain.

One prominent interpretation of motor activation during perception is provided by motor simulation accounts (Pickering and Garrod, 2013a; Wilson and Knoblich, 2005). These accounts posit that perception of another person's actions results in activation of the corresponding motor plan in the perceiver, leading to covert motor simulation. Although motor simulation accounts are not based specifically on speech, speech production is a form of motor activity and thus these accounts are equally relevant to the processing of speech actions. Under this account, the articulatory plans stored in the speech motor system for production are automatically activated during speech perception, although this activation does not result in overt articulation due to presumed suppression of activity in the subcortical motor system (Baldissera et al., 2001). These motor plans are then used to inform forward models of upcoming articulatory gestures in the incoming speech stream.

However, two different forms of the motor simulation account make dissociable predictions about how the perceiver's motor activity is modulated during action perception. The first proposes greater motor involvement when the observer is familiar with the perceived action (Pickering and Garrod, 2013a), for instance when observing actions that the observer can also perform or easily understand. Indeed, Calvo-Merino and colleagues (Calvo-Merino et al., 2005) found greater bilateral activation in motor areas when expert dancers viewed movements that they had been trained to perform compared to movements they had not, indicating that the action-observation system integrates observed actions with the motor repertoire of the observer. In line with this possibility, Swaminathan and colleagues (Swaminathan et al., 2013) found that when subjects observed visual speech movements from a known language, motor excitability in the lip area of M1 was higher than when subjects observed speech movements from an unknown language. The authors interpreted these results to suggest that activity in articulatory motor cortex is enhanced when perceiving speech movements that the perceiver is already experienced in producing and perceiving themselves. Similarly, Bartoli and colleagues observed that the effect of TMS to speech motor areas was related to the listener-speaker perceived acoustic distance, such that response times were facilitated for smaller acoustic distances (Bartoli et al., 2015). By this account, activity in M1 speech areas should be greatest when listening to familiar, natural speech, relative to less familiar, motor-perturbed distorted speech, which is difficult to understand, suggesting somatotopic differences would also be most distinct during perception of natural, unperturbed, speech. Under this account, comprehension of degraded/perturbed speech is assumed to be subserved by increased utilization of auditory, but not motor, resources (Pickering and Garrod, 2013a). Notably, Pickering and Garrod also claim that motor simulation and prediction-by-association, driven by co-occurrence in the auditory input, can also be combined.

In contrast to this view, the second account claims that the motor system is most strongly activated when perception is challenging, predicting greater involvement of M1 speech areas under difficult listening conditions (Wilson and Knoblich, 2005). Under this account, although challenging perceptual conditions would catalyse greater motor activation, the success of the resultant predictive signaling would depend on the degree of similarity between what the observer can perform motorically, and what is being perceived. In turn, this would suggest that articulator-specific effects would be maximally dissociable in terms of M1-activation when listening is difficult. Indeed, TMS in combination with motor evoked potentials (MEPs) has been found to suggest increased motor processing when perceiving spoken sentences in noise (Murakami et al., 2011), although importantly the effect of speech-internal distortion, and somatotopic responsiveness, are unknown. In addition, it has also been shown that TMS to motor areas can significantly affect accuracy (Meister et al., 2007) and response times for speech stimuli in noise (D'Ausilio et al., 2009), but not for speech presented without noise (D'Ausilio et al., 2012). These data support the latter version of the simulation account, and suggest that speech motor activity may be necessary when comprehending speech that is degraded.

In the present study, we aimed to disambiguate between these two accounts by using TMS to elicit a direct measure of motor excitability during speech perception. To this end, MEPs were elicited during perception of natural speech, and speech distorted via a lip and tongue-perturbation during production. Stimulation was thus used to probe the excitability of M1 lip muscle representation to determine whether activation was greater when listening to normal versus distorted speech sounds. In addition, by using speech sounds with two different places of articulation, we tested whether somatotopic facilitation enhanced MEPs in line with predictions made by motor simulation accounts. Lastly, we explored the relationship between individual ability in perceiving distorted speech and motor system activity, to assess the extent to which motor activity is associated with listening performance.

Methods

Subjects

Eighteen subjects took part in this study (six males; average age: 23 years 9 months (± S.D. 3.5 months); age range: 19–30 years).