Mu wave suppression during the perception of meaningless syllables: EEG evidence of motor recruitment

Stephen Crawcour, Andrew Bowers, Ashley Harkrider, Tim Saltuklaroglu*

Department of Audiology and Speech Language Pathology, University of Tennessee, 553 S. Stadium Hall, Knoxville, TN 37996-0740, United States

ARTICLE INFO

ABSTRACT

Motor involvement in speech perception has been recently studied using a variety of techniques. In the current study, EEG measurements from Cz, C3 and C4 electrodes were used to examine the relative power of the mu rhythm (i.e., 8–13 Hz) in response to various audio-visual speech and non-speech stimuli, as suppression of these rhythms is considered an index of 'mirror neuron' (i.e., motor) activity. Fourteen adult native English speaking females watched and listened to nine audio-video stimuli clips assembled from three different auditory stimuli (speech, noise, and pure tone) combined with three different video stimuli (speech, noise, and kaleidoscope—made from scrambling an image from the visual speech). Relative to the noise–noise (baseline condition), all visual speech conditions resulted in significant levels of suppression, a finding that is consistent with previous reports of mirror activity to visual speech and mu suppression to 'biological' stimuli. None of the non-speech conditions or conditions in which speech was presented via audition only resulted in any significant suppression of the mu rhythm in this population. Thus, visual speech perception appears to be more closely associated with motor activity than acoustic speech perception. It is postulated that in this study, the processing demands incurred by the task were insufficient for inducing significant mu suppression via acoustic speech only. The findings are discussed in theoretical contexts of speech perception and the mirror system. We suggest that this technique may offer a cost-efficient, non-invasive technique for measuring motor activity during speech perception.

1. Introduction

The processes underlying human speech perception have been widely examined and debated over the last six decades or so. By some accounts, speech is perceived as a function of its acoustic constituents and their impact on the auditory system (Klatt, 1979; Kuhl & Miller, 1975; Massaro & Cohen, 1990; Ohala, 1996; Stevens, 1981; Sussman, 1989). Additionally, these acoustic theories generally hold that speech perception and production are distinct processes. Criticisms of these perspectives were propagated by studies showing the lack of acoustic invariance in similar speech percepts (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), the influence of visual stimuli on speech percepts (e.g., McGurk effect; McGurk & MacDonald, 1976), the phenomenon of categorical perception (e.g., Mann & Liberman, 1983), and the limited temporal resolution of the auditory system for processing rapidly changing acoustic stimuli (Liberman, 1957; Liberman, Delattre, & Cooper, 1952). As such, Liberman and Mattingly (1985) proposed an alternative viewpoint. They suggested that speech is perceptually coded as a sequence of dynamic 'gestures', that are representative of the manner in which

they are produced and co-articulated in the vocal tract. It is these gestures that are thought to form the invariants for both perception and production, linking the two processes and allowing them to operate efficiently and effectively, together as one, in a specialized linguistic manner. As such, under this 'motor theory', the dynamic architecture of the mechanism employed to produce speech plays an essential role in its perception.

Conceptually, motor theory appeared to answer many questions regarding the nature of speech, yet one of its downfalls was a relative dearth of physiological evidence for the neural connectivity between speech perception and production. The discovery of mirror neurons in the ventral premotor cortex (area F5) of the macaque monkey (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996) provides compelling evidence for motor involvement in sensory processes and therefore, a central linking of perception and production. Neurons in this motor region, which is considered to be a homolog of Broca's area in humans (Rizzolatti & Craighero, 2004), were found to fire both when monkeys performed or observed goal directed actions (e.g., grasping). The location and firing patterns of these 'mirror neurons' helped support the notion of motor involvement in speech perception (Liberman & Whalen, 2000) and theoretical perspectives of human communication evolving from this observation/action matching system. This neural matching system is thought to

* Corresponding author.
E-mail address: tsaltukl@utk.edu (T. Saltuklaroglu).

create a biological link between senders and receivers of gestural goals, which may have served in the evolution of communication (Rizzolatti & Arbib, 1998). These notions were further bolstered by the discovery of a subclass of mirror neurons that fired not only in response to seeing an action, but also to hearing sounds associated with a specific action such as paper ripping (Kohler et al., 2002) or peanut breaking (Keysers et al., 2003). Hence, the authors suggested that mirror neurons code the intended goal of an action in an abstract amodal manner rather than a specific action itself, eliciting strong parallels to the nature of speech 'gestures' as described by motor theorists.

A growing body of research suggests that in humans, the mirror system may be involved in action recognition, imitation, empathy and theory of mind. Its role in speech perception also has been scrutinized using various measures. In the auditory modality, transcranial magnetic stimulation (TMS) has been used to show that listening to lingual speech sounds could evoke stronger motor evoked potentials (MEPs) in the tongue relative to non-speech sounds (Fadiga, Craighero, Buccino, & Rizzolatti, 2002) and stronger MEPs in lip muscles when listening to speech while watching white noise (Watkins & Paus, 2004; Watkins, Strafella, & Paus, 2003). In addition, using functional magnetic resonance imaging (fMRI), listening to meaningless speech has been found to bilaterally activate portions of the ventral premotor cortex (though not Broca's area), portions of the motor cortex, and the supplementary motor area (Wilson, Saygin, Sereno, & Iacoboni, 2004) and motor cortical regions in a somatotopic manner (Pulvermüller et al., 2006) relative to non-speech stimuli. In the visual modality, Nishitani and Hari (2002) used magnetoecepahlography (MEG) to reveal that observing still pictures of lips could activate Broca's area and the motor cortex. Similar bilateral motor activation patterns have also been found to silent speech lip movements using fMRI (Campbell et al., 2001) and stilled speech (Calvert & Campbell, 2003). Buccino et al. (2004) also used fMRI to discover significantly higher activations in portions of the left inferior frontal gyrus in response to viewing speech reading and lip smacking, but not to viewing a dog barking. They suggested that activation of one's own motor system via action observation occurs when the action in question is part of one's own motor repertoire, again suggesting a biological underpinning for the mirror system.

Speech can be perceived unimodally via either audition or vision in isolation, or bimodally (audio-visually). Though the studies above provide evidence that speech perception through either audition or vision can activate the human mirror system, they did not examine the relative strength of each modality for inducing mirror activation or their relative strengths compared to audio-visual speech perception. Because the mirror system is thought to have close connections to the somatosensory (SI) system, Möttönen, Järveläinen, Sams, and Hari (2004) used MEG to examine how viewing and hearing speech modulated activity in the left SI mouth cortex. Whereas viewing speech induced significant SI modulation, hearing speech did not. Similarly Sundara, Namasivayam, and Chen (2001) found that visual and audio-visual presentations of the syllable/ba/both yielded significant increases in MEP amplitudes, whereas the MEP increase produced from auditory perception alone did not reach significance. Activation levels of motor areas during speech perception have also been examined using fMRI. Skipper, Nusbaum, and Small (2005) found that audio-visual speech activated the inferior frontal gyrus and premotor cortex to a greater extent than audio or visual speech alone. They also found that the activation level of the premotor cortex was modulated by a number of phonemes that participants could visually identify.

It has been suggested that the visually perceived gestures may play the stronger role in activating the motor system during speech perception (Skipper, von Wassenhove, Nusbaum, & Small, 2007), especially in conditions in which auditory speech is absent or degraded (Callan et al., 2003). In addition, it is not clear how brain regions involved in speech production are differentially activated during various perception tasks. For example, it has been suggested that motor recruitment might be influenced by the degree of linguistic processing necessary to a task (Callan, Jones, Callan, & Akahane-Yamada, 2004; Ojanen et al., 2005; Wilson & Iacoboni, 2006). Thus, the numerous measures of motor recruitment in speech perception and the variety of stimuli employed, combined with the diversity of findings, make it difficult to reconcile the discrepancies within the current body of research and explain the extent to which motor recruitment may be necessary in speech perception. As such, further investigation in this area is warranted.

Electroencephalography (EEG) has been suggested as a promising, cost-efficient and non-invasive means of indirectly examining the mirror neuron activity in humans. In particular, measurements of oscillation amplitudes in the mu frequencies (8–13 Hz) measured across the sensorimotor cortices acquired via surface level electrodes are thought to provide a valid index of mirror activity (Altschuler, Vankov, Wang, Ramachandran, & Pineda, 1997). Mu 'rhythms' are influenced by both motor activity and attention (see Pineda, 2005 for full review). When a person is at rest, amplitudes of waves in this band are highest because sensorimotor neurons responsible for generating these waves fire synchronously. Conversely, when a person performs an action, the pattern of firing is asynchronized, resulting in suppression of the mu wave and smaller amplitudes. However, a number of studies also have found that mu waves are suppressed when normal adults observe human hand movements (Muthukumaraswamy & Johnson, 2004; Muthukumaraswamy, Johnson, & McNair, 2004; Oberman et al., 2005; Virji-Babul et al., 2008) and implied point-light human biological animation (Saygin, Wilson, Hagler, Bates, & Sereno, 2004; Ulloa & Pineda, 2007), and even when participants imagine biological motion (Pineda, Allison, & Vankov, 2000).

Because mu suppression can occur in these passive observation/imagination conditions in the absence of motor activity, the level of suppression is thought to provide an index of mirror neuron activity. When employing these paradigms, recordings from Cz, C3 and C4 are thought to be indirect measures of cortical activity in the supplementary motor areas and left and right sensorimotor cortices (S1-M1; Babiloni et al., 1999), respectively. Hence, EEG recordings from these electrodes are considered to be measuring "downstream modulation of sensorimotor areas by mirror neurons" (Oberman et al., 2005, p. 191). As the recordings are made from the scalp, it is difficult to map the sources of suppression to cortical landmarks. However, Nishitani and Hari (2000), in a study using MEG, found that the sources of mirror activity may be further 'upstream' in the primary motor cortex and in the inferior frontal cortex (e.g., BA 44). As these regions are often activated in speech perception tasks, it seems plausible that EEG recordings of mu rhythms at electrode sites Cz, C3 and C4 might also be suppressed when speech is perceived.

Though Muthukumaraswamy, Johnson, Gaetz, and Cheyne (2004) examined mu suppression to oro-facial movements (i.e., teeth-baring, blowing), to our knowledge, EEG has not yet been used to examine differential levels of mirror neuron activity to the perception of speech and non-speech stimuli. Previous studies that have identified selective mirror neuron functioning in response to observed dynamic biological stimuli have employed visual noise and 'non-biological' conditions as bases for comparison. Oberman et al. (2005) used a bouncing ball, whereas Ulloa and Pineda (2007) used scrambled versions of their point-light biological animations. As speech can be conveyed through both auditory and visual channels, in order to differentiate the effects of speech from noise and non-biological stimuli, it seems logical in the current study to employ all three types of stimulus (noise, speech, and non-biological) conditions for both input modalities. Based