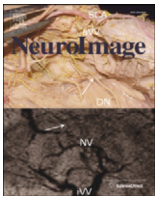
NeuroImage 62 (2012) 1825–1832

Contents lists available at SciVerse ScienceDirect

NeuroImage

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



Multisensory speech perception without the left superior temporal sulcus

Sarah H. Baum [a], Randi C. Martin [b], A. Cris Hamilton [b], Michael S. Beauchamp [a,*]

[a] *Department of Neurobiology and Anatomy, University of Texas Medical School at Houston, USA*
[b] *Department of Psychology, Rice University, USA*

ARTICLE INFO

Article history:
Accepted 14 May 2012
Available online 22 May 2012

Keywords:
Audiovisual
Speech
McGurk effect
STS
Multisensory integration

ABSTRACT

Converging evidence suggests that the left superior temporal sulcus (STS) is a critical site for multisensory integration of auditory and visual information during speech perception. We report a patient, SJ, who suffered a stroke that damaged the left tempo-parietal area, resulting in mild anomic aphasia. Structural MRI showed complete destruction of the left middle and posterior STS, as well as damage to adjacent areas in the temporal and parietal lobes. Surprisingly, SJ demonstrated preserved multisensory integration measured with two independent tests. First, she perceived the McGurk effect, an illusion that requires integration of auditory and visual speech. Second, her perception of morphed audiovisual speech with ambiguous auditory or visual information was significantly influenced by the opposing modality. To understand the neural basis for this preserved multisensory integration, blood-oxygen level dependent functional magnetic resonance imaging (BOLD fMRI) was used to examine brain responses to audiovisual speech in SJ and 23 healthy age-matched controls. In controls, bilateral STS activity was observed. In SJ, no activity was observed in the damaged left STS but in the right STS, more cortex was active in SJ than in any of the normal controls. Further, the amplitude of the BOLD response in right STS response to McGurk stimuli was significantly greater in SJ than in controls. The simplest explanation of these results is a reorganization of SJ's cortical language networks such that the right STS now subserves multisensory integration of speech.

Introduction

Speech can be understood through the auditory modality alone, but combining audition with vision improves speech perception (Grant and Seitz, 2000; Stein and Meredith, 1993; Sumby and Pollack, 1954). One striking behavioral example of audiovisual multisensory integration in speech perception is the McGurk effect (McGurk and MacDonald, 1976) in which an auditory syllable paired with a video clip of a different visual syllable results in the percept of a distinct new syllable (*e.g.* auditory "ba" + visual "ga" results in the percept "da"). Because the fused percept is different than either the auditory or visual stimulus, it can only be explained by multisensory integration.

A number of studies suggest that the left superior temporal sulcus (STS) is an important site of audiovisual multisensory integration. The left STS exhibits a larger BOLD response to multisensory stimuli as compared to unisensory stimuli (Beauchamp et al., 2004; Calvert et al., 2000; Stevenson and James, 2009). Tracer studies in rhesus macaque monkeys reveal that the STS is anatomically connected both to auditory cortex and extrastriate visual cortex (Lewis and Van Essen, 2000; Seltzer et al., 1996). There is a correlation between the amplitude of activity in the left STS and the amount of McGurk perception in both individual adults (Nath and Beauchamp, 2012) and children (Nath et al., 2011). Inter-individual differences in left STS activity have also been linked to language comprehension abilities (McGettigan et al., 2012). When the left STS is temporarily inactivated with transcranial magnetic stimulation (TMS) in normal subjects, the McGurk effect is reduced (Beauchamp et al., 2010). Unlike the transient disruptions created by TMS, lesions caused by brain injury can give insight into the results of brain plasticity that occur after a stroke. In particular, damage to areas in the language network can result in brain reorganization, with increased activity in the areas homologous to the damaged tissue (Blasi et al., 2002; Buckner et al., 1996; Cao et al., 1999; Thomas et al., 1997; Winhuisen et al., 2005).

We describe a patient, SJ, with a lesion that completely ablated her left posterior STS. Following her stroke, SJ underwent intensive behavioral therapy. In the years following her stroke, her speech perception abilities improved. Five years after her stroke SJ demonstrated multisensory speech perception similar to 23 age-matched controls when tested with two independent behavioral measures. To understand the neural substrates of this ability, we examined patient SJ and age-matched controls with structural and functional MRI.

---

* Corresponding author at: 6431 Fannin St. Suite G.550, Houston, TX 77030, USA. Fax: +1 713 500 0623.
E-mail address: Michael.S.Beauchamp@uth.tmc.edu (M.S. Beauchamp).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.05.034

Materials and methods

Patient SJ

All subjects provided informed consent under an experimental protocol approved by the Committee for the Protection of Human Subjects of the University of Texas Health Science Center at Houston. All participants received compensation for their time. Patient SJ is a 63 year-old female who presented with a language impairment following a stroke, which destroyed a large portion of her left temporal lobe, including the left STS (Fig. 1 and Table 1). Patient SJ was 58 years old when she suffered a stroke in the left tempo-parietal area in September 2006. Prior to her stroke SJ worked in public relations and had completed one year of college. SJ's performance on the Western Aphasia Battery indicated a classification of anomic aphasia. Her auditory comprehension was impaired 3 years after the stroke (48% on auditory lexical decision and 86% for CV minimal pairs, compared with expected 95–100% for controls). 5 years after the stroke, her auditory recognition had improved to near normal range (87% on auditory lexical decision and 95% for CV minimal pairs). SJ was scanned two times, once for structural MRI in February 2010, and again for structural and functional MRI in March 2011.

Healthy age-matched control subjects

23 healthy older adults ranging in age from 53 to 75 years (14 female, mean age 62.9 years) served to provide a healthy age-matched comparison to patient SJ. Participants were recruited through word-of-mouth and flyers distributed around the greater Houston area. 21 subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All subjects were fluent English speakers.

Table 1
Anatomical regions impacted by stroke lesion. Column 1 shows the FreeSurfer automatic parcellation anatomical label. Column 2 shows the t-value of the volume difference between SJ and controls. All differences are statistically significant at a level of $p<0.01$ corrected for multiple comparisons. Column 3 shows the difference between the gray matter volume in SJ and the average gray matter volume in 23 age-matched controls (column 4–column 5).

Label	t-value	Delta (mm^3)	Volume in SJ (mm^3)	Mean ± SD volume in controls (mm^3)
Supramarginal gyrus	6.8	−4984	438	5422 ± 714
Superior temporal sulcus	5.5	−4872	3038	7910 ± 867
Postcentral sulcus	4.1	−2023	1376	3399 ± 482
Inferior segment of the circular sulcus of the insula	4.7	−1853	547	2400 ± 385
Temporal plane of the superior temporal gyrus	4.4	−1732	4	1736 ± 389
Posterior segment of the lateral fissure	6.6	−1376	14	1390 ± 203
Anterior transverse temporal gyrus	4.8	−856	30	886 ± 174
Long insular gyrus and central sulcus of the insula	4.5	−747	287	1034 ± 163
Transverse temporal sulcus	4	−429	4	433 ± 287

Stimuli used for testing

Stimuli consisted of a digital video recording of a female native English speaker speaking “ba”, “ga”, “da”, “pa”, “ka” and “ta”. Digital video editing software (iMovie, Apple Computer) was used to crop the total length of each video stimulus such that each clip both started and ended in a neutral, mouth-closed position. Each video clip ranged from 1.7 to 1.8 s.

Auditory-only stimuli were created by extracting the auditory track of each video and pairing it with white visual fixation crosshairs

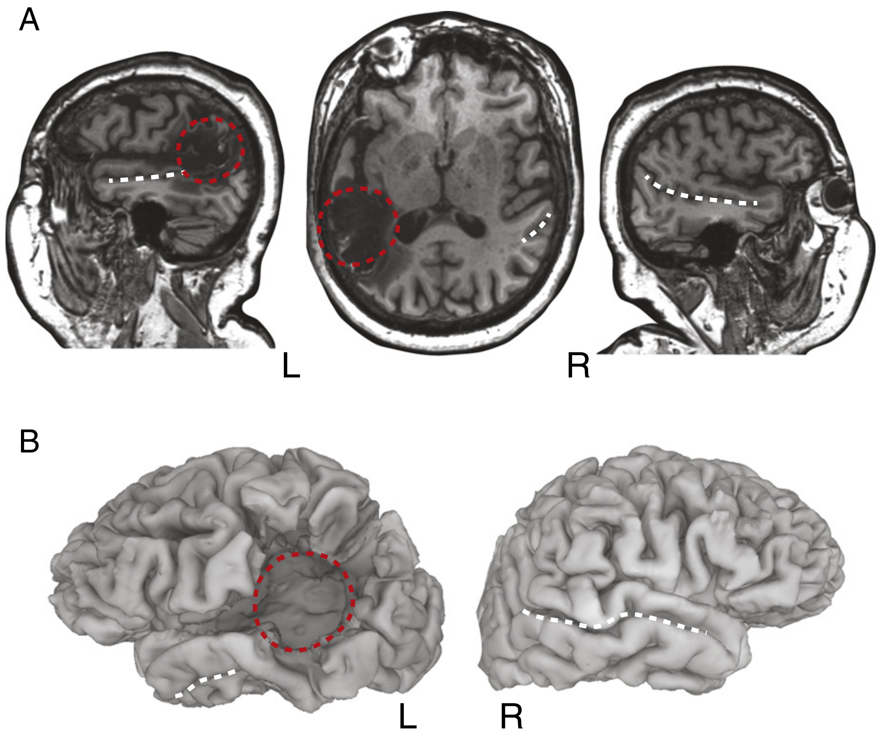


Fig. 1. Anatomical MRI of SJ. A. Sagittal and axial slices of SJ's structural MRI. White dashed lines indicate the location of the STS. Red dashed circle indicates stroke-damaged cortex in left hemisphere (left is left on all images). B. Cortical surface reconstruction of SJ's brain from the structural MRI.