The Neuroanatomy of Speech Processing: A Large-Scale Lesion Study

Corianne Rogalsky,[1*] Alexandra Basilakos,[2] Chris Rorden,[2] Sara Pillay,[3] Arianna N. LaCroix,[1]

Lynsey Keator,[2] Soren Mickelsen,[1] Steven W. Anderson,[4] Tracy Love,[5] Julius Fridriksson,[2]

Jeffrey Binder,[3] Gregory Hickok[6]

Running Title: Neuroanatomy of Speech Processing

[1]Arizona State University
[2]University of South Carolina
[3]Medical College of Wisconsin
[4]University of Iowa
[5]San Diego State University
[6]University of California, Irvine

*corresponding author:
Corianne Rogalsky, PhD
Arizona State University
P.O. Box 870102
Tempe, AZ 85287-0102
(480)965-0576
Corianne.rogalsky@asu.edu

Abstract

The neural basis of language has been studied for centuries, yet the networks critically involved in simply identifying or understanding a spoken word remain elusive. Several functional-anatomical models of the critical neural substrates of receptive speech processes have been proposed, including auditory-related regions in the (1) left *mid-posterior* superior temporal lobe, (2) left *anterior* superior temporal lobe, or (3) *bilateral* in the mid-posterior superior temporal areas, and (4) *motor*-related regions in the left frontal lobe (in normal and/or noisy conditions). One difficulty in comparing these models is that they often focus on different aspects of the sound-to-meaning pathway and different types of stimuli and tasks. Two auditory tasks that are typically used in separate studies—nonword discrimination and word comprehension—often yield different conclusions. We assessed word/nonword discrimination and clear/noisy word comprehension in 158 individuals with focal brain damage: left (n=113) or right (n=19) hemisphere stroke, left (n=18) or right (n=8) anterior temporal lobectomy, and 26 neurologically-intact controls. Discrimination and comprehension tasks are doubly dissociable both behaviorally and neurologically. In support of a bilateral model, clear speech comprehension was near ceiling in 95% of left stroke cases, and right temporal damage was, for the first time, found to impair phoneme discrimination. Lesion-symptom mapping analyses for the word discrimination, non-word discrimination, and noisy word comprehension tasks each implicated most of the left superior temporal gyrus (STG) except the anterior most regions. Comprehension tasks also implicated the left pMTG, while discrimination tasks also implicated more dorsal sensorimotor regions in posterior perisylvian cortex.

The neural basis speech perception—defined here as the processes up to and including the successful activation of a phonological word form—has been a contentious topic for decades, if not centuries, and remains so today. Hickok & Poeppel, 2000, 2004, 2007 have reviewed the literature and argued that task effects are a significant source of the lack of agreement: depending on what task is used, researchers can get different answers. In particular, they pointed out that syllable discrimination, judging whether two syllables are the same or different, can yield misleading results because they do not necessarily tap into the same processes involved in processing speech sounds under more normal, ecologically valid listening conditions, such as during comprehension. Aiming to test this claim and to map the neural correlates of speech perception, we compared stimulus matched word comprehension tasks with discrimination tasks in a large sample of unilateral stroke survivors. In what follows, we provide a brief summary of the major hypotheses regarding the neural basis of speech perception and then describe the present experiment.

Carl Wernicke (Wernicke, 1874) proposed that the left superior temporal lobe serves as the substrate for what we today generally call speech perception. The proposal is intuitively appealing in that the region is in auditory-related cortex and in the language dominant hemisphere. But the claim has been challenged from several angles, leading to at least three current competitors to Wernicke's hypothesis. The first, based mostly on studies of phoneme perception using tasks such as syllable discrimination, challenges the view that sensory cortex alone is sufficient for perceptual processes, as these tasks invoke motor cortex involvement (Pullvermuller *et al.* 2005, 2006; Rizzolatti & Arbib, 1998; Watkins *et al.* 2003; Wilson *et al.*

2004). This view was first proposed in the form of the motor theory of speech perception (Liberman, 1957, 1967) and indirectly linked the process to motor cortex. It was subsequently challenged by several examples of normal speech perception ability in the face of impaired or non-existence motor speech ability (Eimas *et al.* 1971; Kuhl & Miller, 1975; Bishop *et al.* 1990), but the motor theory claim was revived with the discovery of mirror neurons in macaque monkeys (Gallese *et al.* 1996). Consistent with this motor-centric view of speech perception, modern neuroimaging studies have shown that motor speech regions indeed activate during listening to phonemes and syllables (Pulvermuller *et al.* 2006; Wilson *et al.* 2004; Buchsbaum *et al.* 2001; Hickok *et al.* 2003), but strong versions of motor centrism remain improbable for reasons given above (Hickok, 2014). A more moderate proposal in this vein is that the motor system augments speech perception only under noisy listening conditions (Wilson, 2009).

A second challenge to Wernicke's model came from neuroimaging studies at the beginning of this century showing a left dominant and specifically *anterior* superior temporal region that appeared to respond selectively to intelligible speech (Scott *et al.* 2000). Although this anterior superior temporal region has been proposed to be critical for "speech-specific processing" more broadly, the stimuli in these studies have typically been sentences (Scott *et al.* 2000; Evans *et al*. 2014; Narain *et al.* 2003), raising the possibility that the activation in these regions is at a post-phonological stage of processing (Rogalsky et al. 2011a; Brennan & Pylkkanen, 2012).

Finally, a third challenge came from the study of people with language disorders (aphasia) following brain injury: it was found that left hemisphere damage generally produced only mild deficits in speech perception (as measured by word to picture matching tasks with phonological

foils), whereas bilateral superior temporal lesions could produce profound deficits even when pure tone hearing thresholds are normal, i.e., "word deafness" (Buchman et al. 1986; Miceli et al. 1980; Poeppel, 2001; Henschen, 1926). In addition, studies of the isolated right hemisphere in callosotomy patients found that the right hemisphere exhibited good comprehension of most auditory words and simple phrases (Zaidel, 1985; Gazzaniga & Sperry, 1967). These findings led to the proposal that speech perception can be mediated in the superior mid to posterior temporal lobes *bilaterally* (Gazzaniga & Sperry, 1967; Buchman et al. 1986; Zaidel, 1985; Bachman & Albert, 1988; Goodglass, 1993; Hickok & Poeppel 2000, 2004, 2007; Hickok et al. 2008) which would explain the substantial effects of bilateral damage, while individuals with unilateral damage (to either hemisphere) are typically unimpaired on auditory word comprehension tasks.

Much of the modern research on this topic has been carried out using neuroimaging methods, which cannot assay the causal involvement of activated networks. There have been attempts to assess the competing proposals more directly using neuropsychological (Bak & Hodges, 2004) or neural interference methods (e.g., transcranial magnetic stimulation, TMS) (D'Ausilio *et al.* 2009; Meister *et al.* 2007; Mottonen & Watkins, 2009), but methodological concerns have rendered them equivocal (Hickok, 2014). Thus, despite more than a century of study, the neural basis of speech perception remains controversial. Here we present the first large scale lesion study of the ability to recognize and discriminate single words and syllables using a set of tightly controlled stimuli and populations that can adjudicate between the dominant hypotheses.

We employed two auditory word comprehension tasks and two speech discrimination tasks. Both auditory word comprehension tasks involved matching a spoken word to a picture in an array of

four where one picture was the target and the other three were phonological, semantic, or unrelated distractors. In one version of the task, speech was presented in a quiet background (clear speech) and in the other version, speech was presented in white noise. The speech-in-noise task was included for two reasons, first to mitigate the possibility of ceiling effects and, second, to allow an assessment of the claim that motor speech systems play a role in speech perception under noisy listening conditions.

The speech discrimination tasks involved the presentation of pairs of monosyllabic speech stimuli with minimally contrasting onsets and a 1 second inter-stimulus interval. The listener was asked to decide whether the two stimuli were the same or different. One version of the task involved real words and the other involved nonwords. Despite the intuitive assumption that both discrimination and auditory word comprehension tasks measure an ability to extract sound-based phonological information from a speech signal, these two tasks are reported to be doubly-dissociable (Miceli et al. 1980; Hickok & Poeppel, 2004). Particularly paradoxical are cases in which the ability to discriminate pairs of syllables (e.g., *da* from *ba*) is impaired while the ability to differentiate words with these same sounds (e.g., *doll* versus *ball*) is spared in comprehension tasks. This has confounded models of the neural basis of speech perception and recognition for decades because one can reach different conclusions depending on which task is used (Hickok & Poeppel, 2000; Hickok, 2014). We sought to resolve this issue in the present study by employing both tasks in a large sample. We studied people with either left or right hemisphere chronic stroke, left or right anterior temporal lobectomy, and neurologically intact controls. For our largest group (the left hemisphere stroke group) we use lesion-behavior mapping to identify brain regions responsible for receptive speech deficits.

Predictions. Based on their evaluation of previously published studies, Hickok & Poeppel (2000, 2004, 2007) predict that word comprehension and syllable discrimination tasks are doubly dissociable both behaviorally and neurally, even when the syllables in the discrimination task are words. This would argue strongly that discrimination tasks add some process to speech perception that is not necessarily used during normal word comprehension. Assuming that researchers are aiming to understand how speech is processed "in the wild" (this is indeed our own goal) and assuming that the word comprehension task is a closer approximation to ecological validity, such a result would seriously question the validity of discrimination tasks in studies of speech perception. Turning to the neuroanatomical models, the *Wernicke's classical model* predicts that damage to the posterior superior temporal lobe should yield substantial deficits on both tasks. *The motor-centric models* predict that damage to frontal, motor-related regions should cause speech perception deficits on both word comprehension and discrimination tasks. More *nuanced motor-based models* predict that motor involvement should be evident at least on the speech-in-noise task. The *left anterior temporal model predicts* that unilateral mid-to-anterior superior temporal lobe damage should substantially interfere with word comprehension via disruption to systems critical for recovering intelligibility. It is less clear what predictions this model has for non-word discrimination. The *bilateral speech perception models* generally predict minimal impairment on speech perception tasks, perhaps of any type, following unilateral damage. The *Hickok & Poeppel task-dependent bilateral model* predicts minimal impairment on word comprehension tasks following unilateral damage but notable impairment on discrimination tasks, even when they involve words.

Methods

Participants

Consent from all participants (described below) was obtained according to the Declaration of Helsinki and the experiments were approved by the Institutional Review Boards of all the institutions at which the work was performed.

Lesion Participants. The following groups of participants were examined: 113 individuals who had previously experienced a left hemisphere stroke (Figure 1), 19 individuals who had previously experienced a right hemisphere stroke (Figure 2 A,B), 18 individuals who had previously underwent a unilateral left temporal lobectomy, and 8 individuals who had previously underwent a unilateral right temporal lobectomy (Figure 2 C,D). See Table 1 for demographic information of each group. The left hemisphere stroke participants were tested at six institutions: University of Iowa (n=40), University of South Carolina (n=32), Medical College of Wisconsin (n=24), San Diego State University (n=9), Arizona State University (n=7) and University of California, Irvine (n=1). The right hemisphere stroke participants were tested at University of Iowa (n=17), University of South Carolina (n=1) and Arizona State University (n=1). All lobectomy participants were tested at the University of Iowa. Participants were included in the present study based on the following criteria: (i) a chronic focal (6 months or more post-onset) lesion due to a single stroke in either the left or right cerebral hemisphere or a unilateral temporal lobectomy (ii) no significant anatomical abnormalities other than the signature lesion of their vascular event (or evidence of surgery for the lobectomy subjects) nor signs of multiple strokes,

(iii) an absence of a history of psychological or neurological disease other than stroke (or seizure disorder for the lobectomy subjects), (iv) native English speaker, (v) right handed pre-stroke, and (v) ability to follow task instructions. The temporal lobectomy patients all had a seizure disorder that required lobectomy surgery for the treatment of their seizures[1]. All stroke participants had ischemic strokes with the exception of two participants in the left hemisphere group with hemorrhagic strokes.

	n	*Age (years)*	*Gender (M/F)*	*Education (years)*	*time post (years)*
Left Hemi. Stroke	113	61 (31-86)	67 / 46	15 (8-20)	6.1 (0.5-32)
Right Hemi. Stroke	19	52 (25-75)	13 / 6	15 (10-20)	11 (0.5-35)
Left Lobectomy	18	48 (29-62)	8 / 10	13.5 (10-16)	4 (0.7-14)
Right Lobectomy	8	41 (23-60)	3 / 5	14 (12-18)	9 (4-18)
Control	26	59 (12-16)	5 / 21	15 (12-16)	n/a

Table 1. Demographic information for each participant group. Note that "time post" refers to the number of years since a stroke for the left hemisphere and right hemisphere stroke groups, and refers to the number of years post-resection for the left lobectomy and right lobectomy groups.

Aphasia type was assessed in the stroke and lobectomy groups via site-specific protocols that included the Western Aphasia Battery, Boston Diagnostic Aphasia Examination, and clinical ratings. The left hemisphere stroke group consisted of the following aphasia diagnoses: Broca's

[1] The participants with temporal lobectomies have lesions localized in the anterior temporal lobe (Fig 2), a region that is often left intact by all but the most severe middle cerebral artery strokes (Holland & Lambdon Ralph, 2010) and therefore of particular value to the present study. However, there is evidence that epileptic brains have variable functional organization, including laterality of speech processing (Swanson et al. 2002, 2007; Hertz-Pannier et al. 2002). Thus, findings in this population should be interpreted carefully.

(n=30), anomic (n=15), conduction (n=9), residual aphasia (n=8), Wernicke's (n=7), transcortical motor (n=5), mixed non-fluent (n=5), global (n=3), mild Broca's / anomic (n=2), and no aphasia (n = 29). None of the participants in the lobectomy groups or in the right hemisphere stroke group were diagnosed with aphasia.

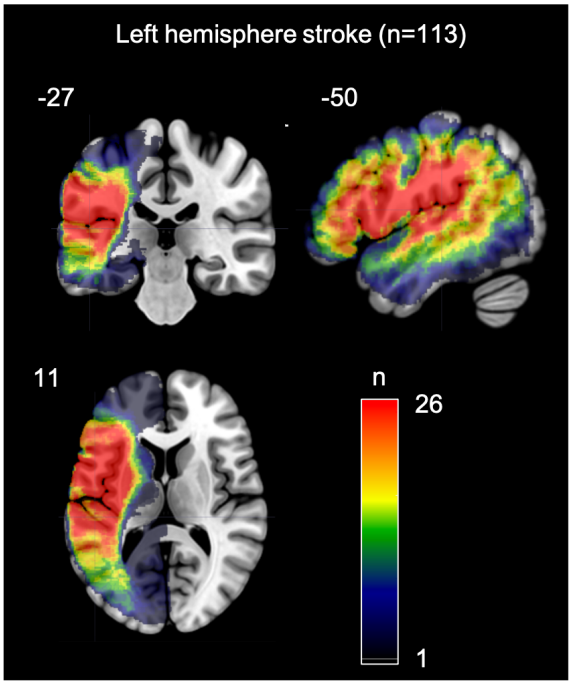

Figure 1. Overlap map of the areas of damage in the 113 participants with a left hemisphere stroke.

Given the proposed role of the right mid-posterior STG in speech perception, the right hemisphere stroke patients were divided into two groups: those with right mid-posterior STG damage (n=11) and those with no mid-posterior STG damage (n=8; max overlap is 3/8 subjects in portions of the inferior frontal gyrus, precentral gyrus and insula) (Figure 2A,B). Mid-posterior STG was defined as any gray or white damage in the STG adjacent or posterior to

Heschl's gyrus as determined by visual inspection by one of the authors (Rogalsky) with extensive training in cortical neuroanatomy. The area of maximum lesion overlap for the RH stroke participants with mid-posterior STG damage is located in white matter underlying the posterior temporal lobe more generally.

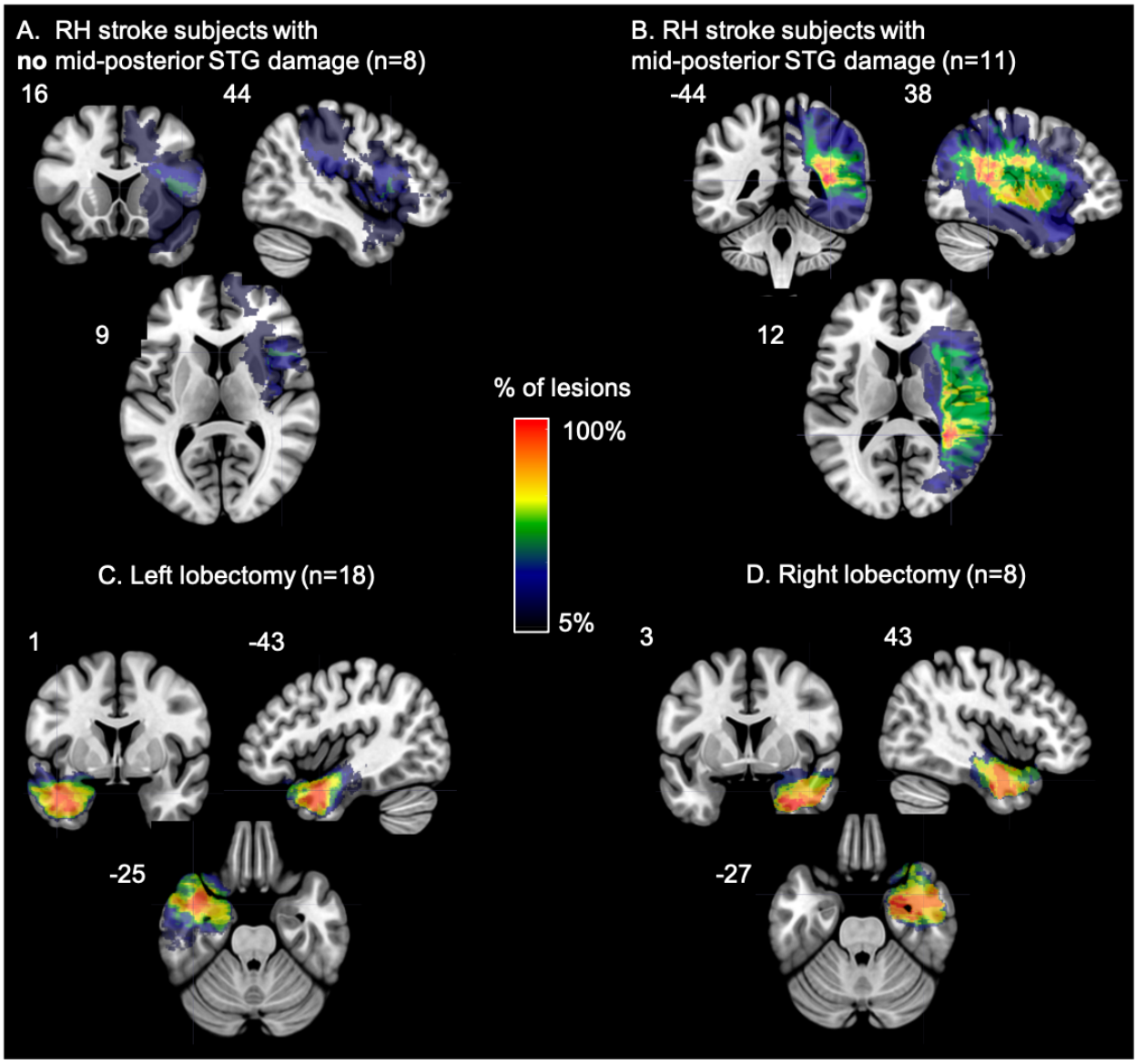

Figure 2. (A,B): Overlap map of the areas of damage in the participants with a right hemisphere stroke (A) without and (B) with mid-posterior STG damage. (C,D): Overlap map of the areas resected in the participants who underwent a (C) left or a (D) right temporal resection.

Control Subjects. Twenty-six adults were recruited and tested at San Diego State University. All control subjects were native English speakers, right-handed, and self-reported no history of psychological or neurological disease. See Table 1 for demographic information.

Informed consent was obtained from each participant prior to participation in the study, and all procedures were approved by the Institutional Review Boards of UC Irvine, University of South Carolina, San Diego State University, Arizona State University, Medical College of Wisconsin and University of Iowa.

Standard hearing and vision screenings were performed during the recruitment of all participants to ensure adequate hearing and vision abilities to complete the experimental tasks and aphasia assessments. The exact hearing and vision screenings varied across sites, but also all participants were provided with several practice trials for each task during which presentation volume could be adjusted to each participant's preferences.

Materials

The present study focuses on two types of tasks: auditory word comprehension and speech discrimination. The tasks were administered as part of an extensive psycholinguistic test battery to assess receptive and productive speech abilities. Within the battery, individual tests

themselves were presented in a non-fixed pseudorandom order; The order and proximity within the test battery of the tasks described below was random across subjects. Items within each test described below were presented in a fixed random order at all testing sites, except for Medical College Wisconsin (n=24), where items were presented in a random order for each subject.

Auditory Comprehension Measures

Auditory word comprehension in a clear background. A four alternative forced choice word-to-picture matching task previously used by Rogalsky et al. (2011b) and similar to Baker, Blumstein, and Goodglass' (1981; Blumstein *et al.* 1977) paradigm was administered. The task contained 20 trials. In each trial an auditory word was presented one second after the appearance of a picture array depicting four objects on a computer screen. The words were recorded by a male native speaker of American English. The participant was instructed to point to the picture that matched the auditorily-presented word. The picture array contained the target and three distractors. Ten of the 20 trials included a one syllable target word and its matching picture, as well as a picture representing the phonemic distractor that was a minimal pair to the target (4 place of articulation contrasts, 4 voicing contrasts, and 2 manner contrasts), a semantic distractor, and an unrelated distractor, which was semantically related to the phonological distractor. For example, the target *goat* was presented with distractor images of a coat, a pig, and a shoe. The auditory stimuli were counterbalanced such that each picture in each array was presented once as the target. All of the words presented have an estimated average age of acquisition under six years old (range: 2.6- 5.7 years old; Kuperman et al. 2012) and ranged in frequency from 1.33 to

86.86 per million words (Brysbaert & New, 2009). The frequency of each error type was calculated, as well as proportion correct for each subject's overall performance in each task.

Auditory word comprehension in noise. The auditory comprehension in noise task contains the same word and picture array pairings as the auditory comprehension word-picture matching task described above, except that the auditory stimuli were presented in 14 dB Gaussian white noise. The noise began one second before the onset of the word and ended one second after the end of the word.

Speech Discrimination Measures

We employed two speech discrimination measures: a real word discrimination task and a non-word discrimination task. These two speech discrimination tasks were modeled on previous speech discrimination tasks used with individuals with aphasia (Baker et al. 1981; Blumstein et al. 1977; Caplan & Waters, 1995). Each of these speech discrimination tasks contained 40 trials. The real word speech discrimination task contained pairs of one syllable real words and the non-word discrimination task contained one-syllable non-words, with neighborhood density and phonotactic probability between the real words and non-words matched. The Irvine Phonotactic Online Dictionary (Vaden et al. 2009) was used to select the real words and non-words, and to calculate neighborhood density and phonotactic probability for each word and non-word. Each word and its corresponding non-word have the same onset consonant phoneme. Paired-samples t-tests comparing each word with its matched non-word indicate that there is no significant difference between the two stimulus types for density $t(39) = .14$, $p = .88$ or phonotactic

probability $t(39) = .18$, $p = .86$. The words used in the real word discrimination task do not overlap with the words in the auditory comprehension tasks described above, but the minimal pairs in the speech discrimination tasks are the same minimal pairs present in the auditory word comprehension tasks.

The real word and non-word discrimination tasks both were structured in the following manner: There were 40 trials in each task, and in each trial, a pair of one-syllable words (or non-words) were presented via headphones while a fixation cross was presented on the computer screen. Each trial presented the two words (or non-words) in one of four arrangements: A-B, B-A, A-A, and B-B such that for half of the trials the correct answer was "same" and for the other half the correct answer was "different". In the "same" trials, two different tokens of the same syllable were presented. The same male native English speaker of the words in the comprehension task described above recorded all of the stimuli. There was a one second interval between each word (or non-word) in a pair, and five seconds between each pair (more time between pairs was given, if necessary, for the patient's comfort). Patients were instructed to determine if the two words (or non-words) presented were the same or different. “Different” trials contained two syllables that differed by one feature of the onset consonant (e.g. “puff” vs “cuff” in the real word discrimination task, or “pag” vs “kag” in the non-word discrimination task). Participants made their response either by speaking their answer or by pointing to the correct answer written in large print on a paper adjacent to the computer screen. Signal detection methods were used to determine how well subjects could discriminate between the same and different pairs by calculating the measure d’ for the real word and non-word speech discrimination tasks, respectively (Swets, 1964).

Neuroimaging

All stroke and lobectomy participants (n=158) underwent MRI scanning using a 3T or 1.5T MRI system at the respective testing site. T1-MRIs and T2-MRIs with 1 mm^3 resolution were collected and used to demarcate the lesion; the lesion demarcation was conducted manually by well-trained and experienced individuals in lesion-mapping studies. The lesion maps were smoothed with a 3mm full-width half maximum Gaussian kernel to remove jagged edges associated with manual drawing. Enantiomorphic normalization (Nachev *et al.* 2008) was conducted using SPM12 and 'in house' Matlab scripts in the following way: a mirrored image of the T1 image (reflected around the midline) was coregistered to the native T1 image. Then, we created a chimeric image based on the native T1 image with the lesioned tissue replaced by tissue from the mirrored image (using the smoothed lesion map to modulate this blending, feathering the lesion edge). SPM12's unified segmentation-normalization (Ashburner & Friston, 2005) was used to warp this chimeric image to standard space, with the resulting spatial transform applied to the actual T1 image as well as the lesion map. The normalized lesion map was then binarized, using a 50% probability threshold.

Lesion-Symptom Mapping (LSM)

All of the LSM routines used here are integrated into the NiiStat toolbox for Matlab (http://www.nitrc.org/projects/niistat). The following routines were performed: Univariate

analyses fitting a general linear model (GLM) were completed to identify anatomical regions in which the percentage of lesioned voxels was associated with performance on the auditory comprehension in noise, word speech discrimination, or non-word speech discrimination tasks. A GLM was fit for performance on each task within each of the 94 left hemisphere regions of interest (ROIs) defined by the John's Hopkins University (JHU) atlas (Mori *et al.* 2008; Faria *et al.* 2012; den Ouden *et al.* 2019). An ROI approach, compared to voxel-based statistics, increases statistical power by reducing the severity of multiple comparison correction (Rorden et al. 2009). Univariate analyses were conducted for the following variables: auditory word comprehension in noise (proportion correct), word discrimination (d') and non-word discrimination (d'). Clear speech auditory word comprehension was not used for LSM due to a ceiling effect with insufficient variability; we discuss this further below. The behavioral data were significantly skewed (i.e. Z-Skew > 1.96) thus de-skewing was performed in NiiStat prior input into the GLM by computing the standard error of the skew and applying a square root transform to the data.

Permutation thresholding included 4,000 permutations to correct for multiple comparisons ($p < 0.05$ controlled for familywise error). As ROIs that are infrequently damaged will have low statistical power while increasing the number of comparisons, only ROIs where at least 11 participants (i.e. approximately 10% of the largest left hemisphere patient sample size in a task) had damage were included in the analyses. Variability due to overall lesion size was regressed out of each LSM. Data collection site also was used as a covariate to remove variance accounted for by collection site. All reported coordinates are in MNI space. Figure 3 displays the extent of

coverage of the ROIs included in the lesion-symptom mapping analyses; the coverage does not vary between the analyses.

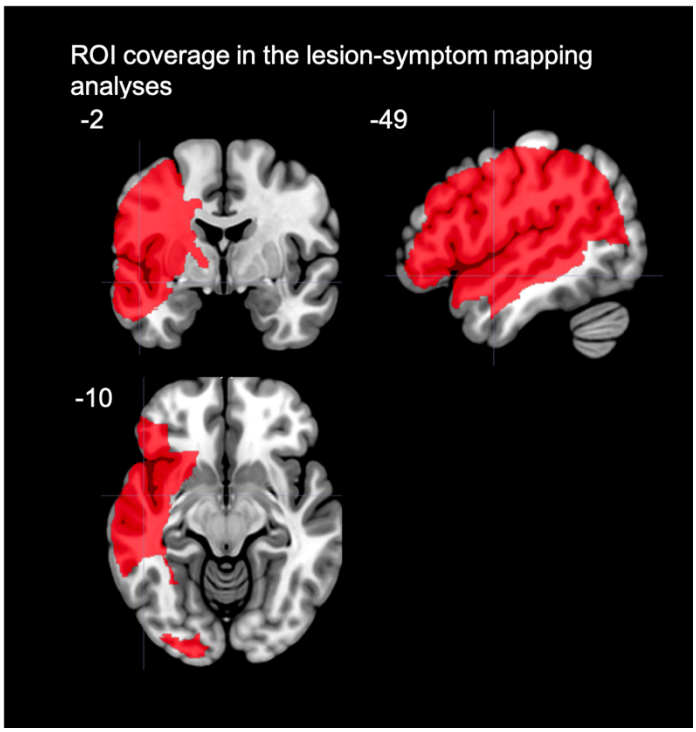

Figure 3. Map of the extent of coverage by the ROIs from the JHU atlas included in the lesion-symptom mapping analyses of the left hemisphere group.

Post-hoc comparisons of LSMs

Based on the close proximity and partial overlap of the significant ROIs identified by the LSMs described above, we decided to use conjunction and independence tests to further investigate the relationships between (1) the brain regions supporting real word speech discrimination versus those regions supporting non-word speech discrimination, and (2) the brain regions supporting

non-word discrimination versus auditory comprehension in noise. The conjunction tests were computed in each ROI using an in-house Matlab script following the valid conjunction test procedure recommended by Nichols et al. (2005). Code for the conjunction analysis can be found here: https://raw.githubusercontent.com/rordenlab/spmScripts/master/nii_thresh_conjunction.m. This conjunction analysis procedure generates, in each ROI, the p-value of the z-values under the conjunction null, i.e. the union of the null hypotheses for the two tasks' z-scores. ROIs with $p < .05$ indicate that damage in that ROI is associated with lower performance in both tasks. The independence tests were computed in each ROI using a custom contrast in the nii_stat toolbox (e.g. 1 -1) to determine if damage in each ROI is significantly more related to one task than the other; in other words, if an ROI's z-score for one task is significantly different than the z-score for another task ($p < .05$, Freedman-Lane permutation as described by Winkler et al. 2014).

Due to a combination of time limitations and site-specific testing protocols, not all participants completed all speech perception tasks. Table 1 indicates the sample sizes for each task within each participant group. Within each lesion group, there are no qualitative differences in the lesion distributions or lesion sizes of the participants who did and did not complete each task.

Results

Overall Behavioral Results

A summary of the descriptive statistics for each group on each of the tasks is presented in Table 2.

Clear Auditory Word Comprehension					
	Control	**LH Stroke**	**RH Stroke**	**Left ATL**	**Right ATL**
n	21	109	19	18	8
Mean	1	0.97	0.99	1	1
Median	1	1	1	1	1
Std. Deviation	0	0.08	0.03	0	0
Range	1-1	0.4-1	0.9-1	1-1	1-1

Auditory Word Comprehension in Noise					
	Control	**LH Stroke**	**RH Stroke**	**Left ATL**	**Right ATL**
n	22	109	19	17	8
Mean	0.91	0.78	0.88	0.87	0.91
Median	0.90	0.80	0.90	0.85	0.90
Std. Deviation	0.05	0.14	0.05	0.07	0.05
Range	0.85-1	0.3-1	0.8-0.95	0.75-1	0.85-1

Word Speech Discrimination					
	Control	**LH Stroke**	**RH Stroke**	**Left ATL**	**Right ATL**
n	25	78	18	17	8
Mean	5.05	4.20	4.72	5.02	5.15
Median	5.15	4.22	5.15	5.15	5.15
Std. Deviation	0.29	1.06	0.78	0.38	0
Range	4.22-5.15	1.71-5.15	2.68-5.15	3.86-5.15	5.15-5.15

Non-word Speech Discrimination					
	Control	**LH Stroke**	**RH Stroke**	**Left ATL**	**Right ATL**
n	23	77	18	17	8
Mean	4.79	3.75	4.29	4.58	4.76
Median	5.15	4.22	4.22	5.15	5.15
Std. Deviation	0.57	1.27	1.02	0.86	0.56
Range	3.61-5.15	0-5.15	2.07-5.15	2.68-5.15	3.86-5.15

Table 2. Descriptive statistics for the clear auditory word comprehension (proportion correct), auditory word comprehension in noise (proportion correct), word speech discrimination (d-prime values), and non-word speech discrimination (d-prime values) tasks.

Auditory word comprehension task: all groups were at ceiling on the clear speech condition. In the left hemisphere stroke group, the mean performance was 97% correct (Figure 4) with both the left and right anterior temporal lobectomy groups scoring at 100% correct. More variance in performance was found on the speech in noise task, enabling statistical analyses. A one-way ANOVA across the 5 groups (control, LH stroke, RH stroke, L lobectomy, R lobectomy) revealed a significant group effect in performance on the comprehension in noise test ($F(4, 170) = 10.6$, $p < .001$) with the LH Stroke group differing from all other groups (*p*s ≤ .002, Bonferroni correction of $\alpha = .05/10 = .005$) and no other group differences (Figure 5, left panel).

Word and non-word speech discrimination tasks: A 5 x 2 mixed ANOVA was computed for the performance on the two speech discrimination tasks in each of the five groups revealing significant main effects of Task ($F(1,138) = 17.80$, $p < .001$) with higher scores on the word than nonword discrimination, and Group ($F(4, 138) = 9.20$, $p < .001$). The interaction was not significant ($F(4,138) = 0.38$, $p=.83$). The LH Stroke group performed significantly worse than the control, L lobectomy, and R lobectomy groups (*p*s ≤ .002, Bonferroni correction of $\alpha = .05/10 = .005$) with no other significant group differences. The LH stroke group numerically performed worse than the RH stroke group across both speech discrimination tasks, but the comparison did not withstand multiple comparison correction ($p = .015$) (Figure 5, right panel).

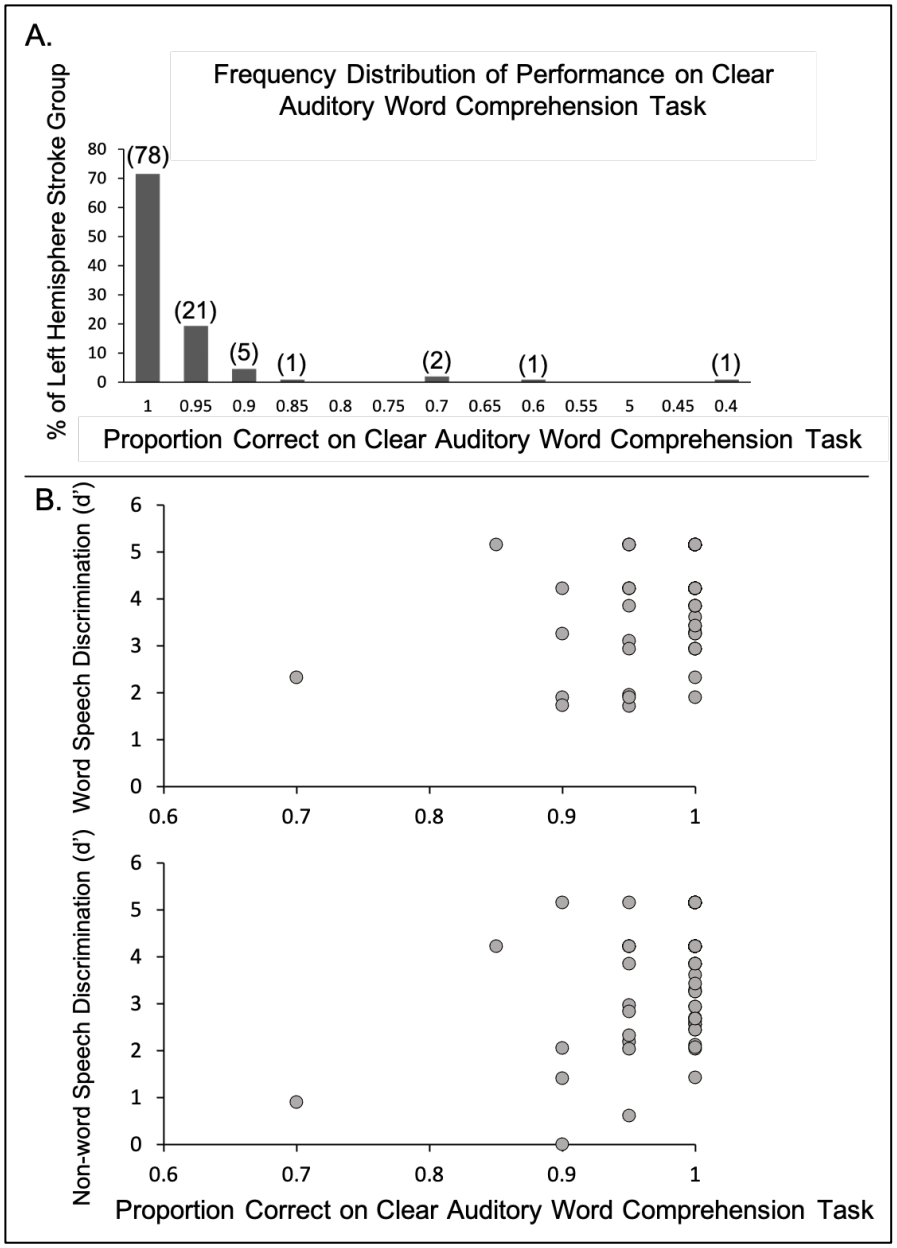

Figure 4. A. Distribution of performance of the left hemisphere stroke group on the clear auditory comprehension task. Numbers in parentheses indicate the actual number of left hemisphere stroke participants who performed at that accuracy level (out of a total 109 who completed the task). Note that chance performance would be .25 proportion correct (not shown). B. Distribution of performance of the left hemisphere stroke group on the clear auditory

comprehension task as a function of word discrimination (top) and non-word discrimination (bottom) task performance.

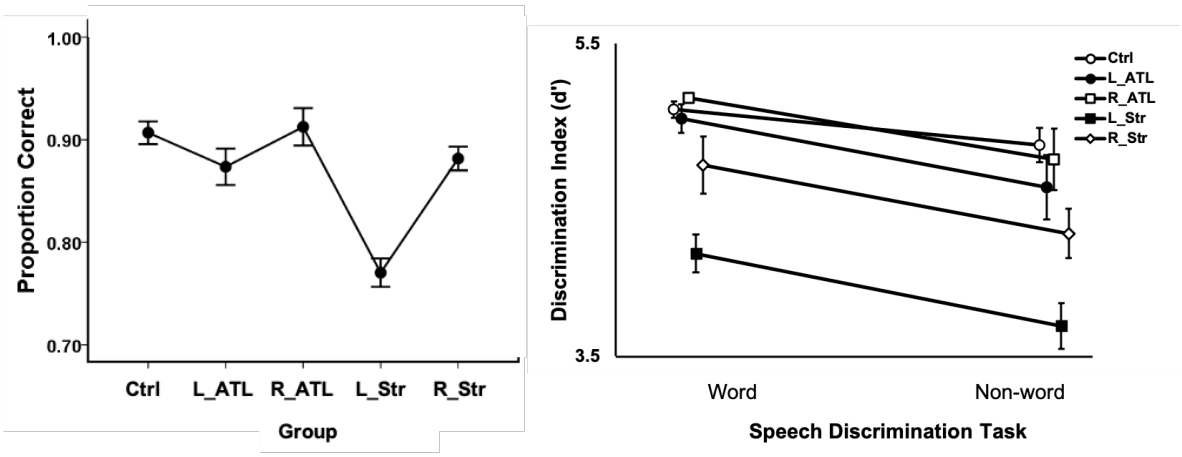

Figure 5. Performance on the auditory word comprehension in noise (left) and speech discrimination tasks (right) in each participant group. Error bars represent standard error. Overlapping points in the discrimination task graph have been offset slightly to make them visible. Ctrl = control group, ATL = lobectomy, Str = stroke.

As noted in the introduction, previous observations indicated that performance on auditory word comprehension and syllable discrimination tasks are doubly dissociable. We assessed this in our sample by categorizing performance as impaired or spared on the auditory word comprehension in noise task and the word speech discrimination task. "Impaired" was defined as performance greater than two standard deviations below the control mean; "spared" was defined as performance that is within or above .5 standard deviations of the control mean. Seventy-five LH stroke participants completed both of these tasks. Table 3 presents participant counts that fall into the 4 possible spared-impaired categories on the two tasks. It is apparent that these two tasks

are dissociable, with 13 cases (17.3%) exhibiting impaired performance on the auditory word comprehension task but spared performance on the speech discrimination task, and 10 cases (13.3%) showing the reverse pattern. There are fewer cases with spared auditory word comprehension in noise in the face of impaired word discrimination but recall that performance on *clear* word comprehension is at ceiling for the majority of LH Stroke participants. There are 59 such participants who were both at ceiling on the clear auditory word comprehension task and completed the word speech discrimination task; 30 (50.8%) were impaired on the latter, which was also presented as clear speech. This provides substantial evidence that these tasks are dissociable.

	Word Speech Discrimination	
Auditory Word Comprehension in Noise	**Impaired**	**Spared**
Impaired	26	13
Spared	10	12

Table 3. Frequency counts of LH Stroke cases that fall into the four categories of impaired and spared performance. Impaired is defined as performance that is > 2 standard deviations below the control mean; spared is defined as performance that is within or above .5 standard deviations of the control mean.

Within the LH stroke group, correlation analyses performed between the auditory word comprehension in noise and word speech discrimination tasks revealed a significant relationship ($r(73) = .39$, $r^2 = .15$, $p < .001$). However, it is evident from the scatterplot in Figure 6, that variation in performance on one task is a rather poor predictor of variation on the other, accounting for only 15% of the variation. A partial correlation between auditory word

comprehension in noise and word speech discrimination tasks controlling for lesion size was also significant ($r(72) = .0.32$, $p = .005$), suggesting that overall lesion size is not accounting for the relationship between the two tasks.

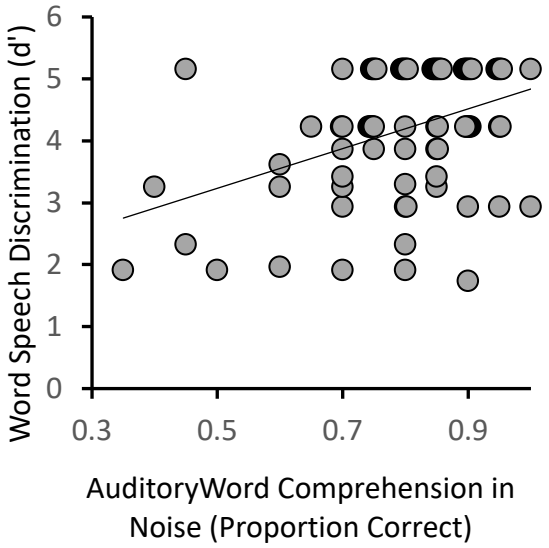

Figure 6: Relation between Auditory Word Comprehension in Noise and Word Speech Discrimination tasks within the LH stroke group. Overlapping points have been offset slightly to make them visible.

Lesion-behavior mapping results

Lesion-symptom mappings (LSM) were computed within the left hemisphere participants for auditory word comprehension in noise, word speech discrimination and non-word speech discrimination. The auditory word comprehension in noise task identified five adjacent ROIs covering much of the superior temporal lobe (Figure 7, Table 4). These ROIs included the superior and middle temporal gyri and underlying white matter, as well as extending into the

inferior portion of the supramarginal gyrus; Heschl's gyrus also was implicated as it is part of the anterior superior temporal ROI identified (anterior portion of ROI cluster: peak *z* of the ROI cluster = 3.08, center of mass = MNI coordinates -51 -12 2; posterior portion of ROI cluster: peak *z* of the ROI cluster = 3.67, center of mass = MNI coordinates -56 -47 5) (Figure 8A). The LSM of word speech discrimination identified two significant ROIs ($p < .05$, corrected; Table 4) adjacent to one another (Figure 7, 8B), implicating most of the STG, as well as Heschl's gyrus and the planum temporale (peak $z = 3.67$, center of mass = -49 -24 3 MNI), as well as an additional significant ROI, consisting of a portion of the superior longitudinal fasciculus (SLF) underlying the supramarginal gyrus (peak $z = 3.46$; center of mass = -35 -24 29 MNI). The LSM of non-word speech discrimination identified the ROIs implicated in real word speech discrimination (peak $z = 3.50$ in the superior temporal ROIs, and 3.30 in the SLF), as well as an additional ROI in the supramarginal gyrus extending into the post-central gyrus (peak $z = 2.87$, center of mass = -51 -29 33 MNI), which is dorsal to the supramarginal gyrus region implicated in the auditory word comprehension in noise LSM (Figure 7, 8B). All three LSMs overlap in the ROI that includes posterior portions of the superior temporal gyrus as well as more mid-temporal regions including Heschl's gyrus (orange, Figure 8C; Table 4).

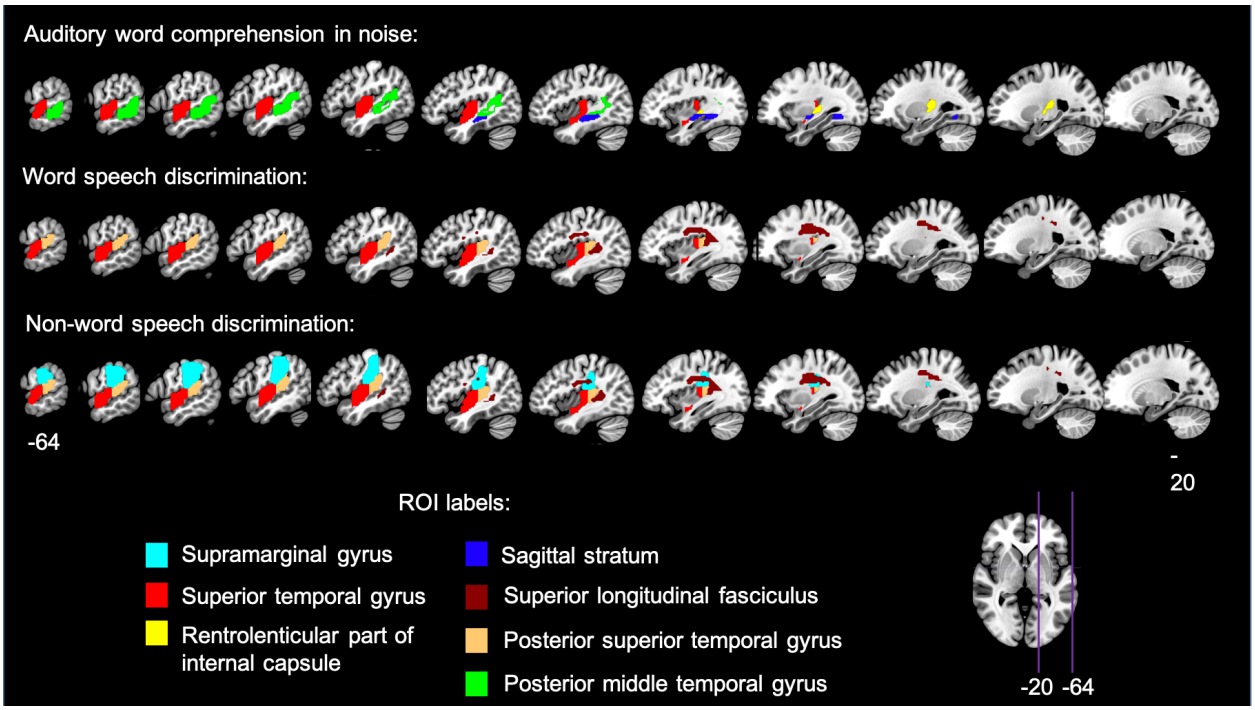

Figure 7. Lesion-symptom mapping results in the left hemisphere stroke group. Significant ROIs are displayed for each of the three tasks (p < .05, permutation-based family-wise error corrected) on sagittal slices through the left hemisphere.

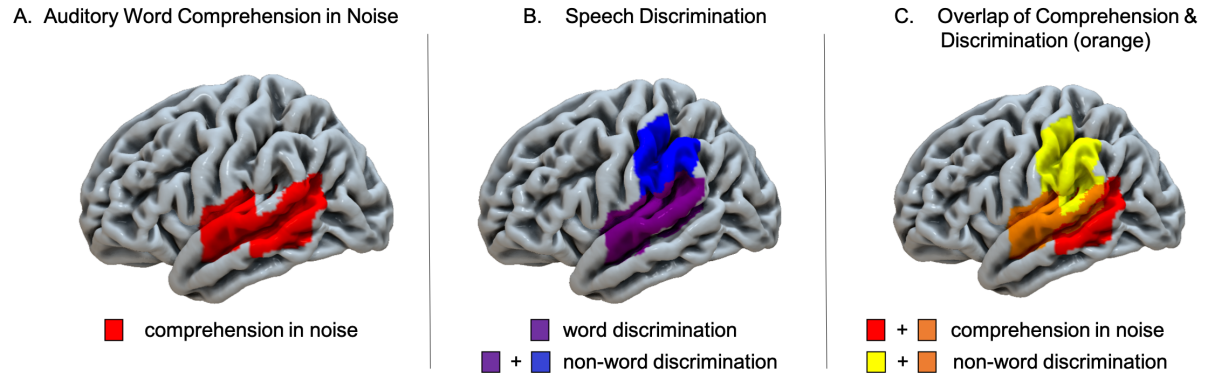

Figure 8. Lateral view summary of brain regions implicated in each task in the left hemisphere stroke group, p < .05, permutation-based family-wise error corrected. (A) Left lateral view of the regions significantly implicated in auditory word comprehension in noise (B) Left lateral view of

the regions significantly implicated in word and non-word speech discrimination. Note that the word speech discrimination regions are a subset of the regions implicated in non-word speech discrimination. (C) Overlap of auditory word comprehension in noise and non-word speech discrimination indicated in orange.

JHU atlas ROI description (ROI number and label)	Auditory Word Comprehension in Noise	Word Speech Discrimination	Non-word Speech Discrimination
Left supramarginal gyrus (29 SMG_L)	--	--	2.87
Left superior temporal gyrus (35 STG_L)	3.04	3.27	3.49
Left rentrolenticular part of internal capsule (135 RLIC_L)	3.08	--	--
Left Sagittal stratum* (151 SS_L)	3.67	--	--
Left superior longitudinal fasciculus** (155 SLF_L)	--	3.46	3.30
Left posterior superior temporal gyrus (184 PSTG_L)	--	3.67	3.60
Left posterior middle temporal gyrus (186 PSMG_L)	3.29	--	--

Table 4. Peak z-scores in each significant ROI in the auditory word comprehension in noise and speech discrimination tasks.

**the JHU ROI atlas notes that the sagittal stratum includes inferior longitudinal fasciculus and inferior fronto-occipital fasciculus.*

***the JHU ROI atlas denotes this ROI as being part of the superior longitudinal fasciculus (SLF), but it is noteworthy that it is difficult to distinguish between SLF and arcuate fasciculus fibers in this region.*

The results for the post-hoc conjunction and independence tests of the LSMs are as follows: for the real word speech discrimination versus the non-word speech discrimination contrast (LSMs shown in Figures 7 and 8B), the conjunction test identified that damage in the left superior temporal gyrus ($z = 3.12$), the left posterior superior temporal gyrus ($z = 3.04$), and the left superior longitudinal fasciculus ($z = 2.99$) ROIs (i.e. all three ROIs that were separately significant for each task (Table 4)) were each related to significantly lower performance in both tasks ($p < .05$); the independence test did not identify any ROIs as being significantly more implicated in the non-word than the real word speech discrimination task, or vice versa (i.e. $z < 2.96$, $p > .05$ in all ROIs). For the non-word speech discrimination versus auditory comprehension in noise LSMs (LSMs shown in Figure 8C), the conjunction test identified damage in the left superior temporal gyrus ROI ($z = 3.21$) as being significantly related to lower performance in both tasks ($p < .05$); the independence test identified damage in the left supramarginal gyrus ($z = 3.21$), left posterior superior temporal gyrus ($z = 3.20$), and left superior longitudinal fasciculus ROIs ($z = 3.57$) as being significantly more related to lower performance for non-word speech discrimination than for auditory comprehension in noise, and damage in the posterior middle temporal gyrus ($z = 3.84$) and sagittal striatum ($z = 3.91$) ROIs as being significantly more related to lower performance for auditory comprehension in noise than for non-word speech discrimination ($p < .05$). In short, the findings from the post hoc conjunction and independence analysis statistically confirm the map shown in Figure 8C.

Right Hemisphere Results

To test for effects of right hemisphere damage on speech perception, we partitioned the RH Stroke group into those cases with ("Temporal +") or without ("Temporal – ") damage (Figure 2 A,B) to the posterior superior temporal lobe and compared them to controls on the auditory word comprehension in noise task and the speech discrimination tasks. The one-way ANOVA across the three groups (RH Temporal +, RH Temporal -, and control) on the auditory word comprehension in noise task yielded a borderline significant effect of group ($F(2, 40) = 3.03$, $p = .061$), with the superior temporal lobe damaged group performing numerically worse than those without damage there (Figure 8, left panel). A mixed ANOVA on the two speech discrimination tasks across the same three groups revealed a significant effect of Task ($F(1, 38) = 7.4$, $p = .01$) with nonword discrimination performance lower than word discrimination, and a significant effect of Group ($F(2, 38) = 7.5$, $p = .002$), with the RH Temporal + group performing significantly worse than the controls and the RH Temporal- group (*p*s<.002, Bonferroni correction of $\alpha = .05/3 = .017$); there was no significant difference between the control group and the RH Temporal- group (Figure 9, bottom). The Task x Group interaction was not significant ($F(2,38) = 1.5$, $p=.24$).

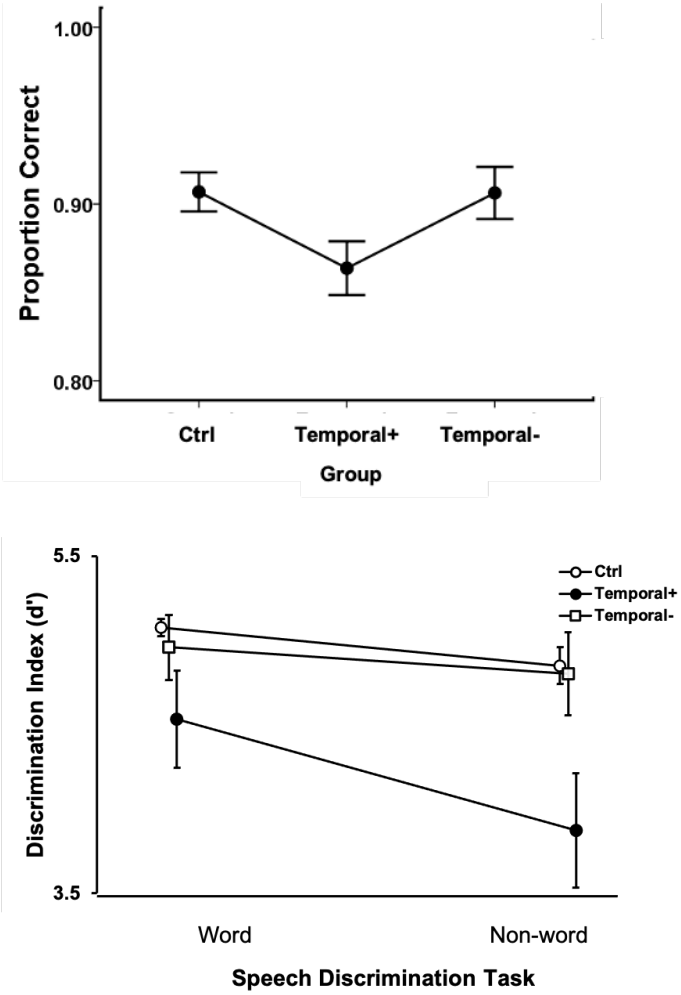

Figure 9. Right hemisphere auditory word comprehension in noise (top) and speech discrimination performance (bottom). Error bars represent standard error. Overlapping points in the speech discrimination graph have been offset slightly to make them visible.

Discussion

The present study examined auditory word comprehension and speech discrimination tasks in stroke patients, anterior temporal lobectomy patients, and unimpaired controls to examine four

prominent functional-anatomical models of speech processing regarding the critical neural substrates of single word comprehension and discrimination. Consistent with prior work, we found that performance on auditory comprehension and speech discrimination tasks are dissociable behaviorally[2]. We also found distinguishable lesion patterns on the two types of tasks, clearly showing that non-identical networks are involved in successful task performance (Hickok & Poeppel, 2000, 2004, 2007). Overall, our results provide strong support for the view that receptive speech processing is bilaterally supported and involves the mid-to-posterior superior temporal gyrus as well as the posterior middle temporal gyrus; performance on speech discrimination tasks additionally recruits dorsal stream areas in the planum temporal and inferior parietal lobe. A simple interpretation of bilateral models of speech perception (Gazzaniga & Sperry, 1967; Zaidel, 1985; Buchman et al. 1986; Bachman & Albert, 1988; Goodglass, 1993; Hickok & Poeppel 2000, 2004, 2007) predicts some degree of speech perception deficit following left *or* right hemisphere damage. We report significantly lower performance for patients with right posterior temporal lobe damage than controls or right hemisphere stroke participants with the posterior temporal lobe spared on the speech discrimination tasks, and a similar trend approaching significance for the auditory word comprehension in noise task. To our knowledge, this is the first direct evidence for a speech perception impairment following unilateral right hemisphere damage. Consistent with this finding, a recent TMS study reported similar effects when left or right superior temporal regions were interrupted in healthy participants (Kennedy-Higgins et al. 2020).

[2] One possible explanation of these tasks being dissociable might be general task difficulty differences, but there are two strong pieces of evidence against a general task difficulty effect and indicating a speech perception-related effect: (1) the ROIs identified are in auditory-related areas and not frontal regions typically implicated in task demands, and (2) the vast majority of the errors induced by adding noise were phonemic errors (i.e. selecting the phonemic foil picture).

To examine the neural correlates of speech processing within the left hemisphere, ROI-based lesion-behavior maps were generated for the auditory word comprehension in noise task and the two speech discrimination tasks in the LH Stroke sample (Figures 7,8). The major findings include the following. Impaired performance on the auditory word comprehension in noise task mapped onto a set of regions in the superior temporal and middle temporal gyri, consistent with a previous study of auditory word comprehension using lesion-connectome mapping (Bonilha *et al.* 2017). It is noteworthy that the auditory word comprehension in noise LSMs did not implicate structures in the dorsal stream including frontal motor-related areas thought to be implicated for perception of speech in noise (Wilson, 2009) nor the posterior planum temporal region likely to include auditory-motor area Spt (Hickok, 2012; Hickok *et al.* 2011; Rogalsky *et al.* 2015) nor dorsal white matter pathways. This finding provides strong support for the view that word comprehension, even under noisy listening conditions, is supported by ventral stream temporal lobe systems. Both of the speech discrimination tasks also implicated the mid-to-posterior superior temporal gyrus, providing strong support for the view that this is a critical region for phonemic perception, but did not implicate the MTG, which has been linked to higher-level word (lemma) processing (Binder et al. 2000; Hickok & Poeppel, 2007; Lau et al. 2008). The lesion maps for the speech discrimination tasks also extended dorsally into the parietal operculum as well as into the planum temporal region including auditory-motor area Spt (Hickok et al. 2003, 2011; Hickok, 2012; Rogalsky et al. 2015). The speech discrimination tasks also implicated white matter in the superior longitudinal fasciculus underlying the supramarginal gyrus and the non-word task additionally implicated adjacent portions of the supramarginal

gyrus. This lesion pattern of involvement for the different tasks was confirmed by the post-hoc conjunction and independence tests.

The lesion-behavior maps provide support for the role of the left mid and posterior superior temporal and inferior parietal regions in phonemic perception and auditory word comprehension (in noise). However, this does not preclude a bilateral representation of speech processing: it is noteworthy that all groups were at ceiling on the clear auditory word comprehension task. Even in the left hemisphere stroke group, the mean performance was 97% correct. The left anterior temporal lobectomy group performed even better, scoring 100% correct. This confirms past reports that chronic left hemisphere damage, including anterior temporal regions, does not substantially impair speech recognition even in a context that requires listeners to make subtle phonemic (e.g., bear vs. pear) or semantic (e.g., bear vs. moose) distinctions on common words (Bachman & Albert, 1988), except in rare cases as discussed below. One may wonder whether plastic reorganization is responsible for this effect with the right hemisphere taking on the speech recognition function after left hemisphere insult. However, acute stroke (Rogalsky et al. 2008) and acute left hemisphere deactivation (Hickok et al., 2008) studies report similar results.

Consistent with Wernicke's 1874 proposal, the ability to perceive speech (as measured by our auditory word comprehension in noise and speech discrimination tasks) implicates the mid-to-posterior superior temporal gyrus, classical Wernicke's area. This was demonstrated directly in our left-hemisphere damaged sample using lesion-behavior mapping and was evident on all three tasks. Three additional observations from the present study indicate that the system is bilaterally organized. First, deficits following unilateral left hemisphere lesions were well above chance

level, even on the nonword speech discrimination task. Indeed, only three of the LH stroke cases performed below discrimination threshold on the nonword speech discrimination task and the lowest score on the word speech discrimination task was above threshold (d' = 1.71). (Threshold is conventionally set at d' = 1, which roughly corresponds to 75% correct.) Second, performance on the clear auditory word comprehension task, which also required differentiating between similar phonemes, was at or near ceiling (≥ 90%) for 95.4% of the LH stroke cases. This shows that unilateral damage does not substantially impair phonological analysis during clear auditory word comprehension except in relatively rare cases, approximately 5% in the present study. Interestingly, the incidence of anomalous cerebral lateralization for language in strong right handers is in that same ballpark with estimates at approximately 5% (Knecht, *et al.* 2000). This is broadly consistent with the view that rare cases of clear speech unilateral impairment may reflect some form of statistically atypical organization. Indeed, functional brain imaging work suggests the incidence of almost complete left dominance is itself atypical and with a frequency of approximately 5% (Springer, *et al.* 1999). Third, we directly documented, for the first time, impairment in phonemic discrimination following *right* posterior temporal lobe damage. Previous evidence provides further support for the bilateral model: (i) bilateral superior temporal lesions are associated with profound speech comprehension and identification deficits (word deafness) (Buchman *et al.* 1986; Poeppel, 2001); (ii) acute left hemisphere deactivation (Wada procedure) does not profoundly impair auditory word comprehension (Hickok *et al.* 2008); and (iii) the right hemisphere of split brain patients has good auditory comprehension at the word level (Zaidel, 1985; Gazzaniga & Sperry, 1967). The data point strongly to bilateral capacity of speech perception, including phonemic discrimination and auditory word comprehension. This does not, however, necessarily imply symmetry of function (Zatorre *et al.* 2002; Poeppel, 2003;

Hickok & Poeppel, 2007; Morillon *et al.* 2010). Indeed, we found greater impairment following left compared to right hemisphere damage on our auditory word comprehension in noise task as well as our discrimination tasks. This could reflect left-right computational differences within the STG/STS speech perception system (Poeppel, 2001; Zatorre *et al.* 2002) or their interaction with other more lateralized networks—lexical, sensorimotor—that augment the efficiency of the left hemisphere for processing speech for some tasks.

The lesions associated with deficits on both of the speech discrimination tasks extended dorsally into the parietal operculum, which includes auditory cortical fields (Galaburda & Sanides, 1980) as well as secondary somatosensory cortex (Ruben*, et al.* 2001). For the nonword speech discrimination task the lesion distribution extended more dorsally still into the parietal-temporal area and the supramarginal gyrus. This finding is consistent with Hickok and Poeppel's (2000) hypothesis that the superior temporal gyrus is the primary substrate for the analysis of speech sounds, but then "there is a divergence in the processing stream, with one stream involving more ventral– posterior structures, which play a role in auditory comprehension, and the other stream involving more dorsal–anterior systems, which play a greater role in sub-lexical speech tasks" (p. 135). They argued that this task-dependent divergence is rooted in the involvement of phonological working memory on "sub-lexical" (here, nonword) tasks, which cannot take advantage of semantic codes for task performance. Phonological working memory networks, in turn, were argued to be dependent on dorsal sensorimotor networks critical for speech production, which is by now fairly well-established (Hickok *et al.* 2003; Buchsbaum & D'Esposito, 2019; Buchsbaum *et al.* 2011; Schomers *et al.* 2017; McGettigan *et al.* 2011).

Thus, the task effects reported in the present study can be interpreted straightforwardly. All tasks implicate the STG because this is the primary substrate for speech perception and damage there partially disrupts acoustic-phonetic-phonological analysis of the signal, which is necessary for all receptive speech tasks. The nonword speech discrimination task additionally implicates dorsal stream structures in the posterior planum temporal region (e.g. Hickok et al. 2003, 2011), the inferior parietal lobe, and underlying dorsal stream white matter tracts because this task places additional demands on phonological working memory, which is supported by sensorimotor networks for speech production. The auditory word comprehension in noise task directly involves the ventral stream network including STG speech analysis systems and MTG lexical-semantic networks (Liebenthal et al. 2005; Hickok & Poeppel, 2007; Lau et al. 2008; Pillay et al. 2017) but not dorsal stream regions. The word speech discrimination task falls in between as it has both the possibility of using the phonological working memory route and a lexical-semantic route placing less demand on the dorsal stream.

The present study also clarifies the role of the anterior temporal lobe in speech perception and comprehension. For all tasks, the stroke-based analyses implicated mid-posterior as well as mid-anterior STG regions. The mid-anterior STG ROI extended anteriorly such that it nearly included the coordinates for fMRI activations in the ATL that have been claimed to be the foci involved in intelligible speech processing (Figure 10), as revealed in studies of intelligible versus unintelligible sentences (Scott *et al.* 2000; Narain *et al.* 2003; Evans *et al*. 2014).

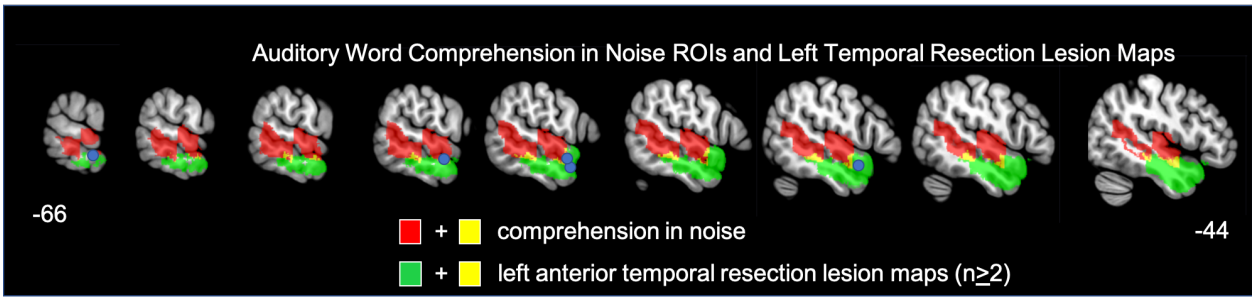

Figure 10. Red: Significant ROIs identified by the LSM analyses for the auditory word comprehension in noise task; Green: The left lobectomy group's lesion maps (areas shown indicate lesion in at least 2 subjects); Yellow: overlap between the auditory comprehension in noise ROIs and the left lobectomy lesion maps. Blue circles indicate the five anterior temporal coordinates identified by Scott et al. 2000 (MNI: -54 6 -16, -66 -12 -12) or Evans et al. 2014 (MNI: -58 2 -16, -60 -2 -12, -58 4 -20) as exhibiting main effects of speech intelligibility or a similar contrast.

These activation foci largely fall within the region impacted by ATL resection cases, which as a group scored lower than controls and similar to the group average for right posterior temporal strokes, but quite a bit better than the left hemisphere stroke group. This suggests a role for the left ATL in speech processing but, more broadly, our findings unequivocally disconfirm the claim that the *left* ATL forms the primary pathway for intelligible speech processing (Scott et al. 2000). The critical network defined by our speech discrimination and auditory word comprehension in noise tasks instead identify both anterior and posterior superior temporal regions and does so bilaterally as more recent functional imaging work on intelligible speech has shown (Evans et al. 2014; Okada et al. 2010).

In summary, the findings from the present study support a neuroanatomical model of basic auditory word identification and comprehension in which the critical region involves much of the superior temporal lobe, bilaterally: a hybrid of models (2) and (3) in the introduction. This work also confirmed the dissociability, both behaviorally and anatomically, of comprehension- and discrimination-based auditory speech tasks. Given that a large of number of studies of "speech perception" use discrimination-type tasks, particularly those that implicate dorsal, motor-related networks (D'Ausilio et al. 2009; Meister et al. 2007; Mottonen et al. 2009), the present finding calls for a critical re-evaluation of that body of work.

Funding

Funding provided by NIDCD R01 DC009659 (PI Hickok) and T32 DC0014435 (Trainee: Basilakos).

Competing Interests

None

References

Ashburner, J. & Friston, K.J. (2005). Unified segmentation. Neuroimage 26(3), 839-851.

Bachman, D.L. & Albert, M.L. (1988). Auditory comprehension in aphasia. In: Boller, F., Grafman, J. editors. Handbook of neuropsychology. Vol 1. New York: Elsevier, pp. 281-306.

Bak, T. H. & Hodges, J. R. The effects of motor neurone disease on language: further evidence. *Brain Lang* **89**, 354-361 (2004).

Baker, E., Blumstein, S.E. & Goodglass, H. (1981). Interaction between phonological and semantic factors in auditory comprehension. 19(1), 1-15.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N. & Possing, E.T. (2000). Human temporal lobe activation by speech and nonspeech sounds. Cerebral Cortex 10(5), 512-528.

Bishop, D. V., Brown, B. B. & Robson, J. The relationship between phoneme discrimination, speech production, and language comprehension in cerebral-palsied individuals. *Journal of speech and hearing research* **33**, 210-219 (1990).

Blumstein, S.E., Baker, E. & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. Neuropsychologia 15, 19-30.

Bonilha, L., Hillis, A.E., Hickok, G., den Ouden, D.B., Rorden, C. & Fridriksson, J. Temporal lobe networks supporting the comprehension of spoken words. *Brain* **140**, 2370-2380 (2017).

Brennan, J. & Pylkkanen, L. (2012). The time-course and spatial distribution of brain activity associated with sentence processing. *Neuroimage* 60(2), 1139-1148.

Brysbaert, M. & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4), 977-990.

Buchman, A. S., Garron, D. C., Trost-Cardamone, J. E., Wichter, M. D. & Schwartz, M. Word deafness: One hundred years later. *Journal of Neurology, Neurosurgury, and Psychiatry* **49**, 489-499 (1986).

Buchsbaum, B.R., Baldo, J., Okada, K., Berman, K.F., Dronkers, N., D'Esposito, M. & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory – an aggregarte analysis of lesion and fMRI data. Brain and Language 119(3), 119-128.

Buchsbaum, B.R. & D'Esposito, M. (2019). A sensorimotor view of verbal working memory. Cortex 112, 134-148.

Buchsbaum, B., Hickok, G. & Humphries, C. Role of Left Posterior Superior Temporal Gyrus in Phonological Processing for Speech Perception and Production. *Cognitive Science* **25**, 663-678 (2001).

Caplan, D. & Waters, G.S. (1995). On the nature of the phonological output planning processes involved in verbal rehearsal: evidence from aphasia. Brain and Language 48(2), 191-220.

D'Ausilio, A., Pulvermuller, F, Salmas, P., Bufalari, I., Begliomini, C. & Fadiga, L. The motor somatotopy of speech perception. *Current Biology* **19**, 381-385 (2009).

den Ouden, D-B., Malyutina, S., Basilaskos, A., Bonilha, L., Gleichgerrcht, E., Yourganov, G., et al. (2019). Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. Human Brain Mapping 40(7), 2153-2173.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. Speech perception in infants. *Science* **171**, 303-306 (1971).

Evans, S., Kyong, J.S., Rosen, S., Golestani, N., Warren, J.E., McGettigan, C., et al. (2014). The pathways for intelligible speech: multivariate and univariate perspectives. Cerebral Cortex 24(9), 2350-2361.

Faria, A.V., Joel, S.E., Zhang, Y., Oishi, K., van Zjil, P.C., Miller, M.I., et al. (2012). Atlas-based analysis of resting-state functional connectivity: evaluation for reproducibility and multi-modal anatomy-function correlation studies.

Galaburda, A. & Sanides, F. Cytoarchitectonic organization of the human auditory cortex. *Journal of Comparative Neurology* **190**, 597-610 (1980).

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. Action recognition in the premotor cortex. *Brain* **119 (Pt 2)**, 593-609 (1996).

Gazzaniga, M.S. & Sperry, R.W. (1967). Language after section of the cerebral commissures. Brain 131-148.

Goodglass H. Understanding aphasia. San Diego: Academic Press; 1993.

Henschen, S.E. (1926). On the function of the right hemisphere of the brain in relation to the left in speech, music, and calculation. Brain 49(1), 110-123.

Hertz-Pannier, L., Chiron, C., Jambaque, I., Renaux-Kieffer, V., Van de Moortele, P.F., Delalande, .O. et al. (2002). Late plasticity for language in a child's non-dominant hemisphere: a pre- and post-surgery fMRI study. *Brain, 125*(Pt 2), 261-72.

Hickok, G. *The myth of mirror neurons: the real neuroscience of communication and cognition*. (W.W. Norton & Company, 2014).

Hickok, G. Computational neuroanatomy of speech production. *Nature Reviews Neuroscience* **13**, 135-145 (2012).

Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of cognitive neuroscience* **15**, 673-682 (2003).

Hickok, G., Houde, J. & Rong, F. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* **69**, 407-422 (2011).

Hickok, G. & Poeppel, D. Towards a functional neuroanatomy of speech perception. *Trends in cognitive sciences* **4**, 131-138 (2000).

Hickok, G. & Poeppel, D. Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition* **92**, 67-99 (2004).

Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nature Reviews Neuroscience* **8**, 393-402 (2007).

Hickok, G., Okada, K., Barr, W., Pa, J., Rogalsky, C. Donnelly, K., *et al.* Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain and language* **107**, 179-184 (2008).

Holland, R. & Lambon Ralph, M.A. (2010). The anterior temporal lobe semantic hub is a part of the language neural network: selective disruption of irregular past tense verbs by rTMS. *Cerebral Cortex, 20*(12), 2771-2775.

Kennedy-Higgins, D., Devlin, J.T., Nuttall, H.E. & Adank, P. (2020). The causal role of left and right superior temporal gyri in speech perception in noise: a transcranial magnetic stimulation study. Journal of Cognitive Neuroscience, 14, 1-11.

Knecht, S., Drager, B., Deppe, M., Bobe, L., Lohmann, H., Floel, A., *et al.* Handedness and hemispheric language dominance in healthy humans. *Brain* **123 Pt 12**, 2512-2518 (2000).

Kuhl, P. K. & Miller, J. D. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* **190**, 69-72 (1975).

Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. Behavior Resarch Methods 44, 978-990.

Lau, E.F., Phillips, C. & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. Nature Reviews Neuroscience 9, 920-933.

Liberman, A. M. Some results of research on speech perception. *Journal of the Acoustical Society of America* **29**, 117-123 (1957).

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. Perception of the speech code. *Psychological review* **74**, 431-461 (1967).

Liebenthal, E., Binder, J.R., Spitzer, S.M., Possing, E.T. & Medler, D.A. (2005). Neural substrates of phonemic perception. Cerebral Cortex 15(10), 1621-1631.

McGettigan, C., Warren, J.E., Eisner, F., Marshall, C.R., Shanmugalingam, P. & Scott, S.K. (2011). Neural correlates of sublexical processing in phonological working memory. Journal of Cognitive Neuroscience 23(4), 961-977.

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D. & Iacoboni, M. The essential role of premotor cortex in speech perception. *Current Biology* **17**, 1692-1696 (2007).

Miceli, G., Gainotti, G., Caltagirone, C. & Masullo, C. Some aspects of phonological impairment in aphasia. *Brain and language* **11**, 159-169 (1980).

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., et al. (2008). Stereotactic white matter atlas based on diffusion tensor imaging in an ICBM template. Neuroimage 40(2), 570-582.

Morillon, B., Lehongre, K., Frackowiak, R.S., Ducorps, A., Kleinschmidt, A., Poeppel, D. & Giraud, A.L. Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 18688-18693 (2010).

Mottonen, R. & Watkins, K. E. Motor representations of articulators contribute to categorical perception of speech sounds. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 9819-9825 (2009).

Nachev, P., Coulthard, E., Jager, H.R., Kennard, C. & Husain, M. (2008). Enantiomorphic normalization of focally lesioned brains. Neuroimage 39(3), 1215-1226.

Narain, C., Scott, S.K., Wise, R.J., Rosen, S., Leff, A., Iversen, S.D. & Matthews, P.M. (2003). Defining a left-lateralized response specific to intelligible speech using fMRI. Cerebral Cortex 13(12), 1362-1368.

Nichols, T., Brett, M., Andersson, J., Wager, T. & Poline, J. (2005). Valid conjunction inference with the minimal statistic. NeuroImage 25, 653-660.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hseih, I-H., Saberi, K., et al. (2010). Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. Cerebral Cortex 20(10), 2486-2495.

Pillay, S.B., Binder, J.R., Humphries, C., Gross, W.L. & Book, D.S. (2017). Lesion localization of speech comprehension deficits in chronic aphasia. Neurology 88, 970-975.

Poeppel, D. Pure word deafness and the bilateral processing of the speech code. *Cognitive Science* **25**, 679-693 (2001).

Poeppel, D. The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time.". *Speech Communication* **41**, 245-255 (2003).

Pulvermuller, F., Hauk, O., Nikulin, V. V. & Ilmoniemi, R. J. Functional links between motor and language systems. *The European journal of neuroscience* **21**, 793-797 (2005).

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O. & Shtyrov, Y. Motor cortex maps articulatory features of speech sounds. *Procedings of the National Academy of Sciences U S A* **103**, 7865-7870, doi:10.1073/pnas.0509989103 (2006).

Rizzolatti, G. & Arbib, M. Language within our grasp. *Trends in neurosciences* **21**, 188-194 (1998).

Rogalsky, C., Love, T., Driscoll, D., Anderson, S.W. & Hickok, G. (2011b). Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. Neurocase 17(2), 178-187.

Rogalsky, C., Pitz, E., Hillis, A. E. & Hickok, G. (2008). Auditory word comprehension impairment in acute stroke: relative contribution of phonemic versus semantic factors. *Brain and language* **107**, 167-169.

Rogalsky, C., Poppa, T., Chen, K.H., Anderson, S.W., Damasio, H., Love, T. & Hickok, G. (2015). Speech repetition as a window on the neurobiology of auditory-motor integration for speech: A voxel-based lesion symptom mapping study. *Neuropsychologia* **71**, 18-27, doi:10.1016/j.neuropsychologia.2015.03.012

Rogalsky, C., Rong, F., Saberi, K. & Hickok, G. (2011). Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience*, 31(10), 3843-3852.

Rorden, C., Fridriksson, J. & Karnath, H.O. (2009). An evaluation of traditional and novel tools for lesion behavior mapping. Neuroimage 44(4), 1355-1362.

Ruben, J., Schwiemann, J., Deuchert, M., Meyer, R., Krause, T., Curio, G., *et al.* Somatotopic organization of human secondary somatosensory cortex. *Cereb Cortex* **11**, 463-473 (2001).

Schomers, M.R., Garagani, M. & Pulvermuller, F. (2017). Neurocomputational consequences of evolutionary connectivity changes in perisylvian language cortex. Journal of Neuroscience, 37(11), 3045-3055.

Scott, S. K., Blank, C. C., Rosen, S. & Wise, R. J. S. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* **123**, 2400-2406 (2000).

Springer, J. A., Binder, J.R., Hammeke, T.A., Swanson, S.J., Frost, J.A., Bellgowan, P.S., *et al.* Language dominance in neurologically normal and epilepsy subjects: a functional MRI study. *Brain* **122 (Pt 11)**, 2033-2046 (1999).

Swanson, S.J., Sabsevitz, D.S., Hammeke, T.A. & Binder, J.R. (2007). Functional magnetic resonance imaging of language in epilepsy. *Neuropsychology Review, 17*, 491-504.

Swanson, S.J., Binder, J.R., Possing, E.T., Hammeke, T.A., Sabsevitz, D.S., Spanaki, M., et al. (2002). FMRI language laterality during a semantic decision task: age of onset and side of seizure focus effects. *Journal of the International Neuropsychological Society, 8*, 222.

Swets, J.A. (1964). *Signal detection and recognition by human observers, reprint.* Los Altos: Peninsula Publishing.

Vaden, K.I., Halpin, H.R. & Hickok, G. (2009). Irvine Phonotactic Online Dictionary, Version 1.4. [Data file] . Available from http://www.iphod.com

Watkins, K. E., Strafella, A. P. & Paus, T. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* **41**, 989-994 (2003).

Wernicke, C. in *Wernicke's works on aphasia: A sourcebook and review* (ed G.H. Eggert) 91-145 (Mouton, 1874/1977).

Wilson, S. M. Speech perception when the motor system is compromised. *Trends in cognitive sciences* **13**, 329-330 (2009).

Wilson, S. M., Saygin, A. P., Sereno, M. I. & Iacoboni, M. Listening to speech activates motor areas involved in speech production. *Nature neuroscience* **7**, 701-702 (2004).

Winkler, A.M., Ridgway, G.R., Webster, M.A, Smith, S.M. & Nichols, T.E. (2014). Permutation inference for the general linear model. Neuroimage 92(100), 381-397.

Zaidel, E. in *The dual brain: Hemispheric specialization in humans* (eds D.F. Benson & E. Zaidel) 205-231 (Guilford Press, 1985).

Zatorre, R. J., Belin, P. & Penhune, V. B. Structure and function of auditory cortex: music and speech. *Trends in cognitive sciences* **6**, 37-46 (2002).