Functional maps of direct electrical stimulation-induced speech arrest and anomia: a multicentre retrospective study

Junfeng Lu,1,2,3† Zehao Zhao,1,3† Jie Zhang,1,3 Bin Wu,3 Yanming Zhu,3 Edward F. Chang,4 Jinsong Wu,1,2,3,5 Hugues Duffau6 and Mitchel S. Berger4

†These authors contributed equally to this work.

Abstract

Direct electrical stimulation, the transient "lesional" method probing brain function, has been utilized in identifying the language cortex and preserving language function during epilepsy and neuro-oncological surgeries for about a century. However, comparison of functional maps of the language cortex across languages/continents based on cortical stimulation remains unclear. We conducted a retrospective multi-center study including four cohorts of direct electrical stimulation mapping from four centers across three continents, where three indigenous languages (English, French, and Mandarin) are spoken. All subjects performed the two most common language tasks: Number counting and picture naming during stimulation. All language sites were recorded and normalized to the same brain template. Next, Spearman's correlation analysis was performed to explore the consistency of the distributions of the language cortex across centers, a kernel density estimation to localize the peak coordinates, and a hierarchical cluster analysis was performed to detect the crucial epicenters. A total of 598 subjects with 917 speech arrest sites (complete interruption of ongoing counting) and 423 anomia sites (inability to name or misnaming) were included. Different centers presented highly consistent distribution patterns for speech arrest (Spearman's coefficient r ranged from 0.60 – 0.85, all pair-wise correlations p<0.05), and similar patterns for anomia (Spearman's coefficient r ranged from 0.37 – 0.80). The combinational speech arrest map was divided into four clusters: cluster...
1 mainly located in ventral precentral gyrus and pars opercularis, which contained the peak of speech arrest in ventral precentral gyrus; cluster 2 in ventral and dorsal precentral gyrus; cluster 3 in supplementary motor area; cluster 4 in the posterior superior temporal gyrus and supramarginal gyrus. The anomia map revealed two clusters: One was in the posterior part of the superior and middle temporal gyri, which peaked at the posterior superior temporal gyrus; the other within the inferior frontal gyrus, peaked at the pars triangularis. This study constitutes the largest series to date of language maps generated from direct electrical stimulation mapping. The consistency of data provides evidence for common language networks across languages, in the context of both speech and naming circuit. Our results not only clinically offer an atlas for language mapping and protection, but also scientifically provide better insight into the functional organization of language networks.

Author affiliations:

1 Glioma Surgery Division, Neurologic Surgery Department, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai, China

2 Shanghai Key Laboratory of Brain Function Restoration and Neural Regeneration, Shanghai, China

3 Brain Function Laboratory, Neurosurgical Institute of Fudan University, Shanghai, China

4 Department of Neurological Surgery, University of California, San Francisco, CA, USA

5 Institute of Brain-Intelligence Technology, Zhangjiang Lab, Shanghai, China

6 Department of Neurosurgery, Gui de Chauliac Hospital, Montpellier University Medical Center, Montpellier, France

Correspondence to: Prof. Jinsong Wu, MD, PhD

Glioma Surgery Division, Neurologic Surgery Department, Huashan Hospital, Shanghai Medical College, Fudan University, No. 12 Middle Wulumuqi Rd., Shanghai, China
Running title: Functional maps of the human language cortex

Keywords: direct cortical electrostimulation; language; functional mapping; speech arrest; anomia

Abbreviations: DES = direct electrical stimulation; KDE = kernel density estimation; MNI = Montreal Neurological Institute; UCSF = University of California, San Francisco; GCH = Gui de Chauliac Hospital; HSH = Huashan Hospital; SFG = superior frontal gyrus; MFG = middle frontal gyrus; pOrb = pars orbitalis; pTri = pars triangularis; pOp = pars opercularis; vPrCG = ventral precentral gyrus; dPrCG = dorsal precentral gyrus; PoCG = postcentral gyrus; SPL = superior parietal lobule; SMG = supramarginal gyrus; AG = angular gyrus; STG = superior temporal gyrus; MTG = middle temporal gyrus; ITG = inferior temporal gyrus

Introduction

Direct cortical electrical stimulation (DES) under awake craniotomy has been utilized to map and preserve the language cortex during epilepsy and neuro-oncological surgeries, for about 100 years since the establishment of the “Montreal procedure” (Penfield and Roberts, 1959). It was optimized by Penfield (Penfield and Roberts, 1959), then subsequently introduced into the modern era by Ojemann (Ojemann et al., 1989), and popularized in recent years (Sanai et al., 2008; Tate et al., 2014; Wu et al., 2015b; Chang et al., 2017; Sarubbo et al., 2020). The reliability and effectiveness of this technique in reducing persistent aphasia, has been exhibited during brain surgeries (Sanai et al., 2008; De Witt Hamer et al., 2012; Pallud et al., 2017; Müller et al., 2019). Although the mechanism of
stimulation effects is poorly understood (Borchers et al., 2011), intraoperative language mapping has been widely considered as the “gold standard” in localizing language areas and achieving maximal safe resection, based on its clinical relevance. It probes the causal relationship between brain regions and language function by directly, transiently and repeatedly interrupting language processing (Pallud et al., 2017). Therefore, the brain regions identified by DES represent the critical and essential areas, about a given function.

Several medical centers have separately established their native language maps using different templates (Penfield and Roberts, 1959; Tate et al., 2014; Wu et al., 2015b; Chang et al., 2017). However, whether the language maps across different languages or geographical continents shared the same distribution pattern remains unclear. Furthermore, the precise distribution may be affected by the potential differences in stimulus parameters, mass effects (tumor or epileptic lesion), brain plasticity (low-grade, high-grade glioma, and epilepsy) (Yuan et al., 2020), normalization methods (from intraoperative individual cortex to standard space). Hence, there is an urgent need to establish a language map by integrating a large sample of DES data around the world.

In the present study, we included four DES cohorts during the past century from four centers across three continents, where three indigenous languages (English, French, and/or Mandarin) are spoken. All centers performed the two most commonly used paradigms (counting and picture naming). The corresponding stimulation-induced disruptions were identified as speech arrest and anomia. Speech arrest is defined as the complete interruption of the ongoing number counting or continuous speaking, without apparent oral, facial, jaw, or tongue movements. Anomia is defined as the DES-induced inability to name an object or misnaming using the wrong word, while still being able to speak the initial words, (i.e., “This (picture) is a(n)...”). Finally, a total of 598 subjects were included into further analysis, and the spatial distribution maps of speech arrest and anomia were generated respectively using the same template. The goal of the present study was to draw the functional map of the language cortex. Relevant questions included: (1) How the distributions of language cortex correlate across centers? (2) Which and how many regions are essential to stimulation-induced speech arrest? (3) Which and how many regions are
critical to stimulation-induced anomia? (4) What is the spatial relationship between speech arrest and anomia?

Materials and methods

Literature screening strategy

We aimed to identify all studies that reported the two most common intraoperative DES-induced disturbances (speech arrest and anomia), with a sample size of more than 50 cases and had available coordinates of positive language sites on the left hemisphere, irrespective of the language and study design. We searched the PubMed, Web of Science and EMBASE for studies between January 1, 1946 and February 15, 2020, using a query combining keywords "language OR arrest OR anarthria OR anomia OR dysnomia OR semantic OR speech" OR "naming error" (response-related) with "intraoperative stimulation mapping" OR "direct cortical stimulation" OR "direct electrical stimulation" NOT "transcranial" (treatment-related). Relevant reviews, editorials, books and reference lists were also assessed for potential interest, where one study was identified. The identified studies were reviewed by two independent observers (Z.Z and J.L), according to the screening flowchart (Fig. 1). In the case of multiple studies from the same cohort, the most recent study was selected.

Intraoperative language mapping strategy

Four cohorts from the following four medical centers were included: (1) Montreal Neurological Institute (MNI), Canada, between 1946 to 1956 (Penfield and Roberts, 1959); (2) Department of Neurological Surgery, University of California, San Francisco (UCSF), USA, between 1999 to 2014 (Chang et al., 2017); (3) Gui de Chauliac Hospital (GCH), France, between 2004 to 2019 (Sarubbo et al., 2020); (4) Huashan Hospital (HSH), China, between 2011 to 2018 (see Supplementary material for subject information for Huashan Hospital). All the included subjects were performed with intraoperative DES language mapping under awake craniotomy (Penfield and Roberts, 1959; Wu et al., 2015b; Chang et al., 2017).
et al., 2017; Sarubbo et al., 2020). Three stimulators were used for mapping in the MNI (Penfield and Roberts, 1959), including the thyratron stimulator, square wave generator, and Rahm stimulator. The three other centers used the paradigm of Ojemann stimulators (5-mm interval bipolar electrode, current-constant bipolar square wave, 1-ms wave width, and 60 Hz stimulation frequency) (Tate et al., 2014; Wu et al., 2015b; Chang et al., 2017). The detailed strategy for determining the current intensity for each center is presented in Table 1. The exposed cortex was stimulated at an interval of 1 cm, during which patients were instructed to perform number counting and picture naming tasks. The definitions of speech arrest and anomia are described above and shown schematically in Fig. 2A and 2B. Each site was discontinuously stimulated for at least three times and was determined as a positive site when at least two of them induced language disturbances, without afterdischarge or epileptic seizure. These positive sites were marked with sterile labels. The recording methods for these positive sites on the individual brain surface were, as follows: Intraoperative photos with labels were taken by all four centers (Fig. 2C); neuro-navigation snapshots were recorded by USCF and HSH (Fig. 2D); intraoperative hand-painted brain surfaces were recorded by MNI (Penfield and Roberts, 1959).

Normalization and generation of speech maps

The MNI 152 (2009, asymmetric) was chosen as the common, standard template. In order to project all positive sites from different centers onto the same brain template, different techniques were utilized to obtain the coordinates of the standard space, according to the obtained data. Since the coordinates of the positive sites from the MNI center were recorded on the hand-drawn brain template (Fig. 2F), these sites were plotted on the MNI 152 template using the MRICron software (https://www.nitrc.org/projects/mricron), according to the anatomical relationship between the sulci and gyri. An affine transformation was used to normalize the coordinates of the Colin 27 template from UCSF (Fig. 2G) to the MNI 152 template. This affine transformation matrix was built by the "Old Normalise" Tool of SPM software (version 12), which described the spatial correspondence from the Colin 27 image to the MNI 152 template image. The positive sites
of GCH and HSH were directly projected from individual brain surfaces onto the MNI 152 according to anatomical landmarks (Fig. 2H and 2I). All positive sites from the four centers were aligned to the pial surface of the MNI 152 template (Hamilton et al., 2017). The anatomic classification of the positive sites was based on a standard atlas, Desikan-Killiany Atlas (Desikan et al., 2006). Since the ventral part of the precentral gyrus is a hot language mapping region in the previous literature (Tate et al., 2014; Wu et al., 2015b; Chang et al., 2017), we used the z = 43 plane in the MNI standard space (at the level of the intersection of the inferior frontal sulcus and the precentral sulcus) as the boundary to subdivide the precentral gyrus into ventral and dorsal portions (Fig. 2J). In order to compare the distribution patterns of language sites across the four medical centers, the percentages of language sites in different regions were separately calculated and tested between medical centers, using Spearman's correlation analysis. In order to assess the distribution of positive sites, a kernel density estimation (KDE) was then performed (Eickhoff et al., 2009). Taking into account the uncertainty caused by the spatial resolution (1cm*1cm) (Szelenyi et al., 2010), all positive sites were smoothed using three-dimensional Gaussian smoothing, with 10-mm full-width at half maximum (FWHM), which were merged to build the density map. Additionally, in order to reduce the potential bias caused by the number of times of stimulation across the different regions, the probabilistic maps of speech arrest and anomia were also generated, based on available data obtained from MNI, UCSF, and HSH (Supplementary material).

Cluster analysis and permutation test

Cluster analysis was conducted using R Project to evaluate the group-level distribution pattern of the functional sites. Initially, the Duda-Hart test (Murtagh and Farid, 2001) was performed to determine whether each set of positive sites should be divided into two or more clusters. Merely those with Duda Hart statistics > 1.645 ($P<0.05$) were further analyzed using cluster analysis, while those with Duda Hart statistics < 1.645 were considered as a single cluster. Sites with less than eight adjacent points (an average of two points per center) in the 10-mm range were considered outliers and excluded. The remaining points were then subjected to hierarchical cluster analysis (Kisilevich et al.,
The cluster number was iterated from 2 to 10, where the one with the highest silhouette score (Rousseeuw, 1987) was selected as the optimal number. Finally, the speech arrest sites were grouped into four clusters, while the anomia sites were grouped into two clusters. The characteristics of these clusters were statistically described as centroid ± standard deviation (SD).

Since there was spatial overlapping between the speech arrest and anomia maps, in both the lateral frontal cortex and superior temporal gyrus (STG), a permutation test (Good, 2006) was performed to determine whether the overlapped clusters of two different disruptions had the same spatial distribution (Supplementary material). The Euclidean distance between the centroids (ECD) of the two clusters was calculated. A P-value of <0.05 indicated significant differences between clusters.

**Comparison with the functional MRI datasets**

In order to clarify whether our DES results are comparable to the functional MRI (fMRI) datasets, we constructed the resting-state functional connectivity maps from Wu-Min Human Connectome Project (HCP) Data dataset (see Supplementary material and Supplementary Figs. 1-2 for detailed methods and results).

**Data availability**

Data involved in this study is available upon reasonable request.

**Results**

In the present study, a total of 598 subjects were included to generate functional maps. Three different languages are spoken, with 90 subjects from MNI in Canada, mostly speaking English or French; 98 subjects from UCSF in the United States, mostly speaking English; 155 subjects from GCH in France, mostly speaking French; and 255 subjects from
HSH in China, mostly speaking Mandarin. Detailed clinical characteristics of these subjects and the stimulation strategy of each center are presented in Table 1.

Spatial distribution of speech arrest sites

A total of 917 speech arrest sites (219 sites from MNI; 75 sites from UCSF; 198 sites from GCH; and 425 sites from HSH) were included (Fig. 3A-3E and Supplementary Table 4), with overlap across different centers. A Spearman's correlation analysis showed that speech arrest sites had similar patterns across the four centers and the pair-wise correlation coefficients ranged from 0.6 to 0.85 (P<0.05) (Fig. 3H and 3I). After combining all of the sites from different centers, the peak of the speech arrest was found to be localized to the ventral part of the precentral gyrus (vPrCG) (-66, 4, 17) (Fig. 3E, 3F and 3J), not the "classic" Broca's area (pOp, pars opercularis; pTri, pars triangularis). Furthermore, the hierarchical clustering analysis revealed that all of the speech arrest sites could be divided into four clusters (Fig. 3G and Supplementary Fig. 4A). The anterior cluster was the largest (cluster 1, Fig. 3G), containing 597 sites (accounting for 65.1% of the total speech arrest sites). Its centroid was located in the vPrCG, which mainly covered the ventral portion of the central lobe and pOp. The middle cluster (cluster 2) contained 203 sites (22.1%), which were superiorly centered in the vPrCG, while also covering the dorsal precentral gyrus (dPrCG) and posterior middle frontal gyrus. The superior cluster (cluster 3, Fig. 3G) contained 30 sites (3.2%) in the supplementary motor area. The posterior cluster (cluster 4, Fig. 3G) contained 41 sites (4.4%), which was centered in the posterior part of the superior temporal gyrus (pSTG), just inferior to the Sylvian fissure. This cluster mainly covered the cerebral regions around the posterior Sylvian fissure, which included the pSTG and supramarginal gyrus. The coordinates of the centroids and peak points for each cluster were illustrated in Supplementary Table 1. Also, the probabilistic map showed a similar distribution pattern to the KDE analysis (Supplementary Fig. 3A and 3D).

Spatial distribution of anomia sites
A total of 423 anomia sites from four centers (68 sites from MNI; 99 sites from UCSF; 126 sites from GCH; and 130 sites from HSH) were included for further analysis (Fig. 4A-4E, Supplementary Table 5). They were mainly distributed in the inferior frontal gyrus and posterior temporal lobe. The Spearman's correlations showed that there was significant consistency ($r=0.81$, $P=0.0039$) between the two centers with the most sites (HSH and GCH), as well as between GCH and UCSF ($r=0.60$, $P=0.022$) (Fig. 4H and 4I). As the sites reported by any single center were limited and sparse, extracting the peak coordinate and centroid would be unreliable. The combined results showed a more straightforward pattern at the inferior frontal gyrus and STG (Fig. 4E). The KDE analysis revealed two peaks: One peak at (-70, -34, 11) in the posterior part of the STG, with a maximum density of 0.052, and another peak at (-62, 22, 11) in the pars triangularis (pTri), with a maximum density of 0.056 (Fig. 4F and 4J). Besides, the probabilistic map based on the stimulation times also presented a similar distribution pattern as the KDE analysis (Supplementary Fig. 3B and 3E). These anomia sites can be divided into two clusters through hierarchical clustering analysis (Fig. 4G and Supplementary Fig. 4B). The posterior cluster (cluster 1, Fig. 4G) consisted of 194 sites (accounting for 45.9% of the total anomia sites), and its centroid was located in the pSTG. This cluster covered the pSTG and middle temporal gyrus (MTG), as well as the inferior portions of the supramarginal (SMG). The second cluster (cluster 2) consisted of 158 sites (37.4%), which centered at the pTri. This cluster mainly contained sites from the pOp, pTri, and posterior middle frontal gyrus (MFG). The coordinates of the centroids and peak points for each cluster are illustrated in Supplementary Table 2.

The spatial relationship between speech arrest and anomia sites

Since speech arrest and anomia sites appeared to partially overlap in the inferior frontal lobe and pSTG (Fig. 5A and Supplementary Fig. 3C), a permutation test was performed to test the spatial discrepancy. The permutation test revealed that the anomia cluster was significantly more anterior than the speech arrest cluster in the lateral frontal cortex (ECD=16.6 mm, $P<0.001$). For the pSTG, the anomia cluster was significantly more anterior and inferior than the speech arrest sites (ECD=11.8 mm, $P<0.001$) (Fig. 5B).
In order to evaluate the distribution of speech arrest and anomia in each cerebral region, the percentage (% of all language sites) of these two responses were calculated and tested using either the chi-square or Fisher's exact probability test. It was found that more speech arrest responses were induced in the central lobe \((P<0.001)\), pOp \((P=0.001)\), and superior frontal gyrus \((P<0.001)\), compared to anomia. More anomia responses were induced in the pTri \((P=0.007)\), STG \((P<0.001)\), and MTG \((P<0.001)\) (Fig. 5C).

**Subgroup analysis**

Compared with the other three series, the data from MNI was heterogeneous in terms of electrical stimulators, stimulation strategies, brain templates for recording the stimulation sites, and disease types of subjects (Table 1). Therefore, we performed a subgroup analysis by excluding the MNI data and re-analyzing the distribution of the positive sites from HSH, GCH and UCSF (Fig. 6A and 7A).

In terms of speech arrest, the data of HSH, GCH and UCSF consistently showed the absolute dominant distribution of speech arrest sites in vPrCG (Fig. 6D and 6E). Cluster analysis only revealed 2 clusters. The ventral cluster (cluster II, Fig. 6C) contained 496 (71.1%) sites, the centroid of which was located in the vPrCG. The dorsal cluster (cluster I, Fig. 6C) contained 180 (25.8%) sites, which also centroided in the vPrCG. The centroids of these two clusters were similar to those of the two largest clusters (cluster 1 and 2, Fig. 3G) based on all four data sets. However, the original SMA cluster and the posterior peri-Sylvian cluster (cluster 3 and 4, Fig. 3G) failed to form discrete clusters due to insufficient density. The density map showed that the peak of speech arrest was located at (-66, 4, 17) in the ventral premotor cortex within cluster II (Fig. 6B and 6F).

For anomia, the data from UCSF, GCH and HSH showed even more consistency after removing MNI data \((r=0.81, P<0.001\) between GCH and HSH; \(r=0.60, P<0.05\) between GCH and UCSF; \(r=0.44, P=0.11\) between HSH and UCSF) (Fig. 7D and 7E). Furthermore, cluster analysis also showed 2 clusters similar to those reported from four data sets (cluster 1 and 2, Fig. 4G). The anterior cluster (cluster II, Fig. 7C) consisted of 127 (35.8%) sites with a centroid in pTri, peaked at (-60, 24, 11) (Fig. 7B and 7F). The posterior cluster
Discussion

Comparison of the language cortex distributions across languages

Our results showed a similar distribution of speech arrest and anomia among English, French, and Mandarin (Fig. 3I and 4I), which indicates that different languages might share a similar neural basis of speech motor control/output and lexical access/semantics. Despite the cross-continental differences presented in this study, humans generally use similar brain regions to process speech output/speech motor control or lexical retrieval/semantics while learning different languages (alphabetic or ideographic languages) and phonemic structures. Although not surprising, it provides essential causal evidence across English, French, and Mandarin, for “the minimal common language network” (Ius et al., 2011; Brouwer and Hoeks, 2013). In addition, there were still some subtle differences. For example, more speech arrest sites from HSH (Mandarin-speaking) and MNI (English or French) were elicited than UCSF and GCH in the pOp and posterior middle frontal gyrus, which is consistent with the previous study (Wu et al., 2015a; Wu et al., 2015b). This discrepancy might be explained by a higher stimulation intensity in the MNI center, which may lead to a greater probability of the local spread of current and the remote effect via white matter pathway (Borchers et al., 2011). Another interpretation is that Mandarin involves more orthography-to-phonology conversion in the posterior middle frontal gyrus, compared to alphabetic languages (Tan et al., 2005). As for anomia, significant correlations between GCH and HSH, and between GCH and UCSF were found, suggesting that they had similar distributions. At the same time, other pair-wise comparisons were not significantly correlated. A possible explanation of why MNI was poorly correlated with others could be that there were more extensive cortex exposures and a higher stimulation intensity for epilepsy surgeries in MNI, which induced more sites at the parietal lobe. After excluding the most heterogeneous data from MNI center, the subgroup analysis showed more consistent bimodal pattern with the pTri and the pSTG as two discrete peaks of
anomia. Based on these findings, it can be inferred that different languages share the same distribution pattern of anomia.

**Spatial distribution and underlying connectivity of speech arrest**

We found the largest cluster of speech arrest was situated in a broader area of the lateral frontal cortex, which included the vPrCG, pOp, and pTri. However, the centroid and coordinate of maximum density were in the vPrCG, instead of the "classic" Broca's area (pOp and pTri). This finding is consistent with previous studies, which demonstrated that the ventral premotor cortex is the epicenter of stimulation-induced speech arrest (Tate *et al.*, 2014; Wu *et al.*, 2015b). This area was posited as a "speech sound map" or "syllabary" in speech production models. It might contain motor programmes that are probably involved in the motor programming stage of speech production (Indefrey, 2011; Guenther, 2016). However, a recent study found stimulating some speech arrest sites in vPrCG also elicited the arrest of ongoing manual movements (negative motor responses) (Breshears *et al.*, 2019). It might suggest that speech arrest sites in the lateral frontal cortex could be categorized into two different groups (speech-specific sites and general negative motor response sites) (Rech *et al.*, 2019; Chang *et al.*, 2020).

Another cluster was identified in the dorsal precentral gyrus and posterior middle frontal gyrus at the vertical level of the middle frontal gyrus. In addition to the potential interpretation as negative motor areas, the cluster is also spatially similar to the re-emerged cortical area 55b (Glasser *et al.*, 2016; Rech *et al.*, 2019; Chang *et al.*, 2020). Structurally, area 55b is isolated from the surrounding cortex, according to the gradients of the myelin content. Functionally, this is universally activated in both covert speech production and listening tasks (Glasser and Van Essen, 2011; Glasser *et al.*, 2016; Chang *et al.*, 2020). Furthermore, the resection of this area resulted in pure apraxia of speech (a disorder of the articulatory coordination and planning in speech sound production, leading to deficits in articulation, prosody, and fluency), which also explains the DES-induced speech arrest response in this region (Duffy, 2013; Chang *et al.*, 2020).
Furthermore, an isolated cluster was identified in the SMA. This area was posited as an "initial map" in speech production models and is involved in the initiation and coordination of speech production (Guenther, 2016). Diffusion tensor imaging studies show that the SMA is connected to the ventral premotor area and pOp via the frontal aslant tract (FAT), and DES of this bundle induces speech arrest (Fujii et al., 2015; Kinoshita et al., 2015; Budisavljevic et al., 2017). Most importantly, besides the frontal regions, speech arrest was also elicited in the Wernicke area of the pSTG, which is approximately centered in the Sylvian-parietal-temporal area. However, previous studies largely overlooked this area as a speech production center, due to a low positive rate and relatively fewer chances of exposure during surgeries. The Sylvian-parietal-temporal area was assumed to subserve the transformation between auditory or visual information, and articulatory representations of speech, according to the dual-stream model proposed by Hickok and Poeppel (Hickok and Poeppel, 2007).

Moreover, the evidence obtained from fiber tractography was also in line with the present findings. The inferior and middle clusters of the frontal lobe and the pSTG cluster overlapped with the terminations of the superior longitudinal fasciculus-III (SLF-III)/arcuate fasciculus (AF) (Maldonado et al., 2011b; Martino et al., 2013; Moritz-Gasser and Duffau, 2013). The DES of the SLF-III was found to induce anarthria or speech arrest, while the DES of the AF induced phonemic paraphasia (Maldonado et al., 2011a), providing subcortical connective evidence (Maldonado et al., 2011b; Duffau, 2015). It should be noted that the components of SMA and pSTG clusters mainly came from MNI, so that in the subgroup analysis, the speech arrest sites from the other three centers were too sparse to form these two clusters. One possible explanation of the sites in these two clusters might be due to the remote effects via subcortical fibers caused by the excessive current intensity (Mandonnet et al., 2010). However, even if SMA and pSTG clusters were caused by remote effects (Mandonnet et al., 2010), the positive sites induced in these regions might still be essential for speech production because it indicated they were the cortical regions of the dorsal phonological streams (pSTG by SLF-III/AF, and SMA by FAT).
Spatial distribution and underlying connectivity of anomia

It was found that anomia sites can be split into two clusters with equivalent density. The finding that the pTri is the frontal epicenter of lexical-semantic processing is consistent with neuroimaging (Price, 2010) and DES studies (Herbet et al., 2018). In addition, our results reported a gradual transitional pattern from speech output function to a higher-level lexical/semantic processing aspect in the direction from ventral premotor cortex to the pTri of Broca’s area (Fig. 5A). This pattern fits well to the modern understanding of the multi-receptor and co-activation based parcellation of the anterior language region, where the caudal part of the Broca’s area (pOp) and the ventral premotor cortex have predominantly motor functions, while the rostral part of the Broca’s area (predominantly pTri) have more prefrontal functions (Amunts et al., 2010; Clos et al., 2013). Concerning the role of the pOp, this has been contentious for several decades. Although the present data suggest that the pOp is implicated in both lexical retrieval (eliciting anomia) and speech output (eliciting speech arrest), converging evidence from DES and neuroimaging studies have also revealed that syntactic processing is subserved specifically by the pOp, at the level of sentence production and comprehension (Wilson et al., 2011; Newhart et al., 2012; Zaccarella and Friederici, 2015; Chang et al., 2018; Friederici, 2018). Thus, more intraoperative tasks including counting, naming and sentence production/comprehension might be required during mapping the function of pOp.

In addition to the frontal epicenter, it is not surprising that the posterior cluster was located in the "classic" Wernicke's area, which also expanded to the mid-posterior STG and MTG, and inferior SMG, with the centroid and peak point located in the pSTG. Activation of the mid-posterior STG in semantic processing has been supported by multiple fMRI studies, which utilized auditory and visual naming tasks among healthy volunteers (Vigneau et al., 2006; González et al., 2016; Trimmel et al., 2018). From the perspective of fiber connectivity, these two clusters overlapped with the frontal and temporal terminations of the inferior fronto-occipital fasciculus (IFOF) (Martino et al., 2010a; Martino et al., 2010b; Sarubbo et al., 2013), respectively. This location is an important part of the ventral semantic stream. Stimulation to this bundle can elicit semantic paraphasia (Duffau et al., 2005; Moritz-Gasser et al., 2013), further supporting our findings.
Limitations

Our study had some limitations. Firstly, only data on the left hemisphere was included, due to the lack of large sample studies on the right hemisphere. Therefore, we were unable to compare the relationship of language distribution between the two hemispheres. Secondly, the numbers of stimulations in each brain region were highly variable; that is, the exposure of the cortex was unbalanced. Hence, the distribution of the speech disturbance sites was likely biased. However, the available probability maps from the three medical centers demonstrated similar distribution patterns (Supplementary Fig. 3). There is thus reason to suggest that these maps represent ground truth that vPrCG is the most critical cortical region of the minimal common network of speech output, while the pTri and the pSTG are those of the semantic/lexical processing. Thirdly, the subject’s handedness information is not fully available (see “Handedness information” in Supplementary Material). However, we don’t think handedness would significantly affect our results. In most people, the left hemisphere of the brain is dominant for language, irrespective of the handedness (Penfield and Roberts, 1959; Knecht et al., 2000; Floel et al., 2005). Besides, the previous direct electrical cortical stimulation study showed that data in the left hemispheres of left-handed/ambidextrous patients suggested an analogous pattern of speech output and naming to right-handed patients (Tate et al., 2014). Finally, we only focused on the consistency of the language cortex of cross-lingual spoken language, and the possible discrepancy between spoken and non-spoken language was not covered. Interestingly, an recent electrocorticography study showed that the cerebral regions for non-spoken sign language might differ from those of monolingual spoken language in terms of language production (Shum et al., 2020).

Conclusion

This is by far the first multi-center DES language mapping study with the largest sample size. We highlighted the critical role of ventral precentral gyrus in the speech circuit, which
challenged the dogmatic localizationist view (i.e., pars opercularis/triangularis are the speech output center). We also highlighted two distinct functional peaks in posterior superior temporal gyrus and pars triangularis, respectively, during semantic/lexical processing. The patterns of language maps were consistent across the three languages and provided evidence for common networks across languages, in the context of both speech and naming circuit. This data can not only offer clinical neurologists and neurologic surgeons high-level, evidence-based guidelines for the definition of functional boundaries of language, but also provide better insight into fundamental language organization.

Acknowledgements

We would like to thank Yuanning Li, Han Zhang, Binke Yuan, Ye Yao, Shelley Tong and Lingyun Zhao for critically reading the manuscript and for helpful suggestions.

Funding

Dr J.W receives funding from Shanghai Municipal Science and Technology Major Project (2018SHZDZX01), Shanghai Zhangjiang National Innovation Demonstration Zone Special Funds for Major Projects (ZJ2018-ZD-012) and Shanghai Shenkang Hospital Development Center (SHDC12018114). Dr J.L receives funding from Shanghai Rising-Star Program (19QA1401700) and Shanghai Municipal Health and Family Planning Commission (2017YQ014). Dr J.Z receives funding from the National Natural Science Foundation of China (81701289).

Competing interests

The authors report no competing interests.
Supplementary material

Supplementary material is available at Brain online.

References

Amunts K, Lenzen M, Friederici AD, Schleicher A, Morosan P, Palomero-Gallagher N, et al. Broca's region: novel organizational principles and multiple receptor mapping. PLoS Biol 2010; 8(9): e1000489.

Borchers S, Himmelbach M, Logothetis N, Karnath HO. Direct electrical stimulation of human cortex - the gold standard for mapping brain functions? Nat Rev Neurosci 2011; 13(1): 63-70.

Breshears JD, Southwell DG, Chang EF. Inhibition of Manual Movements at Speech Arrest Sites in the Posterior Inferior Frontal Lobe. Neurosurgery 2019; 85(3): E496-e501.

Brouwer H, Hoeks JC. A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. Front Hum Neurosci 2013; 7: 758.

Budisavljevic S, Dell'Acqua F, Djordjilovic V, Miotto D, Motta R, Castiello U. The role of the frontal aslant tract and premotor connections in visually guided hand movements. Neuroimage 2017; 146: 419-28.

Chang EF, Breshears JD, Raygor KP, Lau D, Molinaro AM, Berger MS. Stereotactic probability and variability of speech arrest and anomia sites during stimulation mapping of the language dominant hemisphere. Journal Of Neurosurgery 2017; 126(1): 114-21.

Chang EF, Kurteff G, Andrews JP, Briggs RG, Conner AK, Battiste JD, et al. Pure apraxia of speech after resection based in the posterior middle frontal gyrus. Neurosurgery 2020.
Chang EF, Kurteff G, Wilson SM. Selective interference with syntactic encoding during sentence production by direct electrocortical stimulation of the inferior frontal gyrus. Journal of cognitive neuroscience 2018; 30(3): 411-20.

Clos M, Amunts K, Laird AR, Fox PT, Eickhoff SB. Tackling the multifunctional nature of Broca's region meta-analytically: co-activation-based parcellation of area 44. NeuroImage 2013; 83: 174-88.

De Witt Hamer PC, Robles SG, Zwinderman AH, Duffau H, Berger MS. Impact of intraoperative stimulation brain mapping on glioma surgery outcome: a meta-analysis. J Clin Oncol 2012; 30(20): 2559-65.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 2006; 31(3): 968-80.

Duffau H. Stimulation mapping of white matter tracts to study brain functional connectivity. Nature reviews Neurology 2015; 11(5): 255-65.

Duffau H, Gatignol P, Mandonnet E, Peruzzi P, Tzourio-Mazoyer N, Capelle L. New insights into the anatomo-functional connectivity of the semantic system: a study using cortico-subcortical electrostimulations. Brain 2005; 128(4): 797-810.

Duffy J. Motor Speech Disorders: Substrates, Differential Diagnosis, and Management, 3rd edn. St. Louis, MO: Elsevier, Mosby 2013.

Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum Brain Mapp 2009; 30(9): 2907-26.
Floel A, Buyx A, Breitenstein C, Lohmann H, Knecht S. Hemispheric lateralization of spatial attention in right- and left-hemispheric language dominance. Behav Brain Res 2005; 158(2): 269-75.

Friederici AD. The neural basis for human syntax: Broca's area and beyond. Current Opinion in Behavioral Sciences 2018; 21: 88-92.

Fujii M, Maesawa S, Motomura K, Futamura M, Hayashi Y, Koba I, et al. Intraoperative subcortical mapping of a language-associated deep frontal tract connecting the superior frontal gyrus to Broca's area in the dominant hemisphere of patients with glioma. J Neurosurg 2015; 122(6): 1390-6.

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, et al. A multimodal parcellation of human cerebral cortex. Nature 2016; 536(7615): 171-8.

Glasser MF, Van Essen DC. Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI. Journal of Neuroscience 2011; 31(32): 11597-616.

Gonzálvez GG, Trimmel K, Haag A, van Graan LA, Koepp MJ, Thompson PJ, et al. Activations in temporal areas using visual and auditory naming stimuli: a language fMRI study in temporal lobe epilepsy. Epilepsy research 2016; 128: 102-12.

Good PI. Permutation, parametric, and bootstrap tests of hypotheses: Springer Science & Business Media; 2006.

Guenther FH. Neural control of speech: Mit Press; 2016.

Hamilton LS, Chang DL, Lee MB, Chang EF. Semi-automated Anatomical Labeling and Inter-subject Warping of High-Density Intracranial Recording Electrodes in Electrocorticography. Front Neuroinform 2017; 11: 62.
Herbet G, Moritz-Gasser S, Duffau H. Electrical stimulation of the dorsolateral prefrontal cortex impairs semantic cognition. Neurology 2018; 90(12): e1077-e84.

Hickok G, Poeppel D. Opinion - The cortical organization of speech processing. Nature Reviews Neuroscience 2007; 8(5): 393-402.

Indefrey P. The spatial and temporal signatures of word production components: a critical update. Front Psychol 2011; 2: 255.

Ius T, Angelini E, de Schotten MT, Mandonnet E, Duffau H. Evidence for potentials and limitations of brain plasticity using an atlas of functional resectability of WHO grade II gliomas: towards a “minimal common brain”. NeuroImage 2011; 56(3): 992-1000.

Kinoshita M, de Champfleur NM, Deverdun J, Moritz-Gasser S, Herbet G, Duffau H. Role of fronto-striatal tract and frontal aslant tract in movement and speech: an axonal mapping study. Brain Struct Funct 2015; 220(6): 3399-412.

Kisilevich S, Mansmann F, Nanni M, Rinzivillo S, Clustering S-T. Data mining and knowledge discovery handbook. Springer US, Boston, MA; 2010.

Knecht S, Drager B, Deppe M, Bobe L, Lohmann H, Floel A, et al. Handedness and hemispheric language dominance in healthy humans. Brain : a journal of neurology 2000; 123 Pt 12: 2512-8.

Maldonado IL, Moritz-Gasser S, Duffau H. Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. Brain Struct Funct 2011a; 216(3): 263-74.

Maldonado IL, Moritz-Gasser S, Duffau H. Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. Brain Structure and Function 2011b; 216(3): 263-74.
Mandonnet E, Winkler PA, Duffau H. Direct electrical stimulation as an input gate into brain functional networks: principles, advantages and limitations. Acta neurochirurgica 2010; 152(2): 185-93.

Martino J, Brogna C, Robles SG, Vergani F, Duffau H. Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. cortex 2010a; 46(5): 691-9.

Martino J, Hamer PCDW, Berger MS, Lawton MT, Arnold CM, de Lucas EM, et al. Analysis of the subcomponents and cortical terminations of the perisylvian superior longitudinal fasciculus: a fiber dissection and DTI tractography study. Brain Structure and Function 2013; 218(1): 105-21.

Martino J, Vergani F, Robles SG, Duffau H. New insights into the anatomic dissection of the temporal stem with special emphasis on the inferior fronto-occipital fasciculus: implications in surgical approach to left mesiotemporal and temporoinsular structures. Operative Neurosurgery 2010b; 66(suppl_1): ons-4-ons-12.

Moritz-Gasser S, Duffau H. The anatomo-functional connectivity of word repetition: insights provided by awake brain tumor surgery. Frontiers in human neuroscience 2013; 7: 405.

Moritz-Gasser S, Herbet G, Duffau H. Mapping the connectivity underlying multimodal (verbal and non-verbal) semantic processing: a brain electrostimulation study. Neuropsychologia 2013; 51(10): 1814-22.

Müller DMJ, Robe PAJT, Eijgelaar RS, Witte MG, Visser M, de Munck JC, et al. Comparing Glioblastoma Surgery Decisions Between Teams Using Brain Maps of Tumor Locations, Biopsies, and Resections. JCO Clin Cancer Inform 2019; 3.
Murtagh F, Farid MM. Pattern Classification, by Richard O. Duda, Peter E. Hart, and David G. Stork. 2001; 18(2): 273-5.

Newhart M, Trupe LA, Gomez Y, Cloutman L, Molitoris JJ, Davis C, et al. Asyntactic comprehension, working memory, and acute ischemia in Broca’s area versus angular gyrus. Cortex 2012; 48(10): 1288-97.

Ojemann G, Ojemann J, Lettich E, Berger M. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. J Neurosurg 1989; 71(3): 316-26.

Pallud J, Rigaux-Viode O, Corns R, Muto J, Lopez Lopez C, Mellerio C, et al. Direct electrical bipolar electrostimulation for functional cortical and subcortical cerebral mapping in awake craniotomy. Practical considerations. Neurochirurgie 2017; 63(3): 164-74.

Penfield W, Roberts L. Speech and brain mechanisms: Princeton University Press; 1959.

Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann N Y Acad Sci 2010; 1191: 62-88.

Rech F, Herbet G, Gaudeau Y, Mezieres S, Moureau JM, Moritz-Gasser S, et al. A probabilistic map of negative motor areas of the upper limb and face: a brain stimulation study. Brain 2019.

Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics 1987; 20: 53-65.

Sanai N, Mirzadeh Z, Berger MS. Functional outcome after language mapping for glioma resection. N Engl J Med 2008; 358(1): 18-27.
Sarubbo S, De Benedictis A, Maldonado IL, Basso G, Duffau H. Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. Brain Structure and Function 2013; 218(1): 21-37.

Sarubbo S, Tate M, De Benedictis A, Merler S, Moritz-Gasser S, Herbet G, et al. Mapping critical cortical hubs and white matter pathways by direct electrical stimulation: an original functional atlas of the human brain. NeuroImage 2020; 205.

Shum J, Fanda L, Dugan P, Doyle WK, Devinsky O, Flinker A. Neural correlates of sign language production revealed by electrocorticography. Neurology 2020; 95(21): e2880-e9.

Szelenyi A, Bello L, Duffau H, Fava E, Feigl GC, Galanda M, et al. Intraoperative electrical stimulation in awake craniotomy: methodological aspects of current practice. Neurosurg Focus 2010; 28(2): E7.

Tan LH, Laird AR, Li K, Fox PT. Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. Human brain mapping 2005; 25(1): 83-91.

Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. Probabilistic map of critical functional regions of the human cerebral cortex: Broca's area revisited. Brain 2014; 137: 2773-82.

Trimmel K, van Graan AL, Caciagli L, Haag A, Koepp MJ, Thompson PJ, et al. Left temporal lobe language network connectivity in temporal lobe epilepsy. Brain 2018; 141(8): 2406-18.

Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, et al. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. Neuroimage 2006; 30(4): 1414-32.
Wilson SM, Galantucci S, Tartaglia MC, Rising K, Patterson DK, Henry ML, et al. Syntactic processing depends on dorsal language tracts. Neuron 2011; 72(2): 397-403.

Wu J, Lu J, Zhang H, Zhang J, Mao Y, Zhou L. Probabilistic map of language regions: challenge and implication. Brain 2015a; 138(Pt 3): e337.

Wu J, Lu J, Zhang H, Zhang J, Yao C, Zhuang D, et al. Direct evidence from intraoperative electrocortical stimulation indicates shared and distinct speech production center between Chinese and English languages. Hum Brain Mapp 2015b; 36(12): 4972-85.

Yuan B, Zhang N, Yan J, Cheng J, Lu J, Wu J. Tumor grade-related language and control network reorganization in patients with left cerebral glioma. Cortex 2020.

Zaccarella E, Friederici AD. Merge in the human brain: a sub-region based functional investigation in the left pars opercularis. Frontiers in psychology 2015; 6: 1818.

Figure Legends

Figure 1 Flowchart for the literature screening

Figure 2 The technical route of acquisition and normalization of the language sites. (A-B) Sketch maps for the number counting task (red timeline) and picture naming task (blue timeline). (A) Speech arrest (marked in grey) is defined as the DES-induced complete interruption while counting, without obvious oral, facial, mandibular, and laryngeal muscle movement. (B) Anomia (marked in grey) is defined as the DES-induced inability to name the object in a picture, or misnaming by using the wrong word (e.g., using "CAT" instead of "DOG"), while still being able to speak the leading word (e.g., "This is a(n)"). (C-E)
Intraoperative photographs with sterile labels to mark the positive sites (C) and/or neuro-navigation images (D) were used to record the locations of these sites on the individual surfaces (E). (F-I) The original functional maps from the four centers were projected onto different templates. These include the line graph template of MNI (F), Colin 27 template of UCSF (G), and MNI 152 template of GCH (H) and HSH (I). (J) All positive sites were normalized to the common MNI 152 template.

Figure 3 Spatial distribution of the speech arrest sites. (A-E) The speech arrest sites (red) from MNI (A), UCSF (B), GCH (C), and HSH (D) were separately plotted on the same brain template (MNI 152 template). A total of 917 speech arrest sites from the four centers were gathered on the same template (E). (F) The density map of the speech arrest sites shows that they peaked in the ventral precentral gyrus (vPrCG). (G) The speech arrest sites were split into four clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled numbers of the cluster labels. (H) The percentage of speech arrest sites in the different regions, across the four medical centers, are shown. (I) The heatmap shows the correlations of spatial distributions between different medical centers. The numbers indicate the Spearman's correlation coefficients, while the labels indicate the significance of the correlations (*: P<0.05; **: P<0.01; ***: P<0.001). (J) The maximum density of the major regions in each cluster are shown (colored by cluster labels).

Figure 4 Spatial distribution of anomia sites. (A-E) Anomia sites (blue) from MNI (A), UCSF (B), GCH (C), and HSH (D) were separately plotted on the same brain template (MNI 152 template). A total of 423 anomia sites from the four centers were gathered on the same template (E). (F) The density map of the anomia sites is shown. (G) The cluster analysis revealed that the anomia sites could be divided into two clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled numbers of the cluster labels. (H) The percentage of anomia sites in the different regions across the four medical centers is shown. (I) The heatmap shows the correlations between different medical centers. The numbers indicate the Spearman's correlation coefficients, while the
labels indicate the significance of the correlations (*: P<0.05; **: P<0.01; ***: P<0.001; NS: nonsignificant). (J) The maximum density of the major regions in each cluster (colored by cluster label).

**Figure 5 Spatial relationship between speech arrest and the anomia sites.** (A) The merged image for the speech arrest and anomia density maps is shown. (B) The merged cluster map for the speech arrest (red) and anomia (blue) sites is shown (plotted as centroid ± SD in the y-z axis of the MNI space). (C) The percentages (% of total language sites) for speech arrest and anomia in each cerebral region. **: P<0.01; ***: P<0.001; ns, not significantly different.

**Figure 6 Subgroup analysis of the speech arrest sites.** (A) A total of 698 speech arrest sites (red) from HSH, GCH and UCSF were pooled together on MNI 152 template. (B) The density map of the speech arrest sites showed that they peaked in the ventral precentral gyrus (vPrCG). (C) The speech arrest sites were split into two clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled Roman numerals of the cluster labels. (D) The percentage of speech arrest sites in the different regions, across the three medical centers, are shown. (E) The heat map shows the correlations of spatial distributions between different medical centers. The numbers indicate the Spearman's correlation coefficients, while the labels indicate the significance of the correlations (**: P<0.01; ***: P<0.001). (F) The maximum density of the major regions in each cluster are shown (colored by cluster labels).

**Figure 7 Subgroup analysis of the anomia sites.** (A) A total of 355 anomia sites (blue) from HSH, GCH and UCSF were pooled together on MNI 152 template. (B) The density map of the anomia sites showed that they peaked in the pars triangularis and posterior superior temporal gyrus. (C) The anomia sites were split into two clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled Roman numerals of the cluster labels. (D) The percentage of anomia sites in the different
regions, across the three medical centers, are shown. (E) The heat map shows the correlations of spatial distributions between different medical centers. The numbers indicate the Spearman's correlation coefficients, while the labels indicate the significance of the correlations (*: P<0.05; ***: P<0.001; NS: nonsignificant). (F) The maximum density of the major regions in each cluster are shown (colored by cluster labels).
Figure 1 Flowchart for the literature screening
Figure 2 The technical route of acquisition and normalization of the language sites. (A-B) Sketch maps for the number counting task (red timeline) and picture naming task (blue timeline). (A) Speech arrest (marked in grey) is defined as the DES-induced complete interruption while counting, without obvious oral, facial, mandibular, and laryngeal muscle movement. (B) Anomia (marked in grey) is defined as the DES-induced inability to name the object in a picture, or misnaming by using the wrong word (e.g., using "CAT" instead of "DOG"), while still being able to speak the leading word (e.g., "This is a(n)"). (C-E) Intraoperative photographs with sterile labels to mark the positive sites (C) and/or neuro-navigation images (D) were used to record the locations of these sites on the individual surfaces (E). (F-I) The original functional maps from the four centers were projected onto different templates. These include the line graph template of MNI (F), Colin 27 template of UCSF (G), and MNI 152 template of GCH (H) and HSH (I). (J) All positive sites were normalized to the common MNI 152 template.
Figure 3 Spatial distribution of the speech arrest sites. (A-E) The speech arrest sites (red) from MNI (A), UCSF (B), GCH (C), and HSH (D) were separately plotted on the same brain template (MNI 152 template). A total of 917 speech arrest sites from the four centers were gathered on the same template (E). (F) The density map of the speech arrest sites shows that they peaked in the ventral precentral gyrus (vPrCG). (G) The speech arrest sites were split into four clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled numbers of the cluster labels. (H) The percentage of speech arrest sites in the different regions, across the four medical centers, are shown. (I) The heatmap shows the correlations of spatial distributions between different medical centers. The numbers indicate the Spearman's correlation coefficients, while the labels indicate the significance of the correlations (*: P<0.05; **: P<0.01; ***: P<0.001). (J) The maximum density of the major regions in each cluster are shown (colored by cluster labels).
Figure 4 Spatial distribution of anomia sites. (A-E) Anomia sites (blue) from MNI (A), UCSF (B), GCH (C), and HSH (D) were separately plotted on the same brain template (MNI 152 template). A total of 423 anomia sites from the four centers were gathered on the same template (E). (F) The density map of the anomia sites is shown. (G) The cluster analysis revealed that the anomia sites could be divided into two clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled numbers of the cluster labels. (H) The percentage of anomia sites in the different regions across the four medical centers is shown. (I) The heatmap shows the correlations between different medical centers. The numbers indicate the Spearman’s correlation coefficients, while the labels indicate the significance of the correlations (*: P<0.05; ***: P<0.001; NS: nonsignificant). (J) The maximum density of the major regions in each cluster (colored by cluster label).
Figure 5 Spatial relationship between speech arrest and the anomia sites. (A) The merged image for the speech arrest and anomia density maps is shown. (B) The merged cluster map for the speech arrest (red) and anomia (blue) sites is shown (plotted as centroid ± SD in the y-z axis of the MNI space). (C) The percentages (% of total language sites) for speech arrest and anomia in each cerebral region. **: P<0.01; ***: P<0.001; ns, not significantly different.
Figure 6 Subgroup analysis of the speech arrest sites. (A) A total of 698 speech arrest sites (red) from HSH, GCH and UCSF were pooled together on MNI 152 template. (B) The density map of the speech arrest sites showed that they peaked in the ventral precentral gyrus (vPrCG). (C) The speech arrest sites were split into two clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled Roman numerals of the cluster labels. (D) The percentage of speech arrest sites in the different regions, across the three medical centers, are shown. (E) The heat map shows the correlations of spatial distributions between different medical centers. The numbers indicate the Spearman’s correlation coefficients, while the labels indicate the significance of the correlations (**: P<0.01; ***: P<0.001). (F) The maximum density of the major regions in each cluster are shown (colored by cluster labels).
Figure 7 Subgroup analysis of the anomia sites. (A) A total of 355 anomia sites (blue) from HSH, GCH and UCSF were pooled together on MNI 152 template. (B) The density map of the anomia sites showed that they peaked in the pars triangularis and posterior superior temporal gyrus. (C) The anomia sites were split into two clusters (colored by cluster labels) after eliminating outlier (grey), with the centroids marked with circled Roman numerals of the cluster labels. (D) The percentage of anomia sites in the different regions, across the three medical centers, are shown. (E) The heat map shows the correlations of spatial distributions between different medical centers. The numbers indicate the Spearman’s correlation coefficients, while the labels indicate the significance of the correlations (*: P<0.05; ***: P<0.001; NS: nonsignificant). (F) The maximum density of the major regions in each cluster are shown (colored by cluster labels).
Table 1 Clinical characteristics of subjects and mapping strategies

|                         | MNI         | UCSF        | GCH         | HSH         |
|-------------------------|-------------|-------------|-------------|-------------|
| Patients, n             | 110         | 102         | 256         | 256         |
| Hemisphere, n(%)        |             |             |             |             |
| Left                    | 90 (81.8)   | 98 (96.1)   | 155 (60.5)  | 255 (99.6)  |
| Right                   | 20 (18.2)   | 4 (3.9)     | 101 (39.5)  | 1 (0.4)     |
| Language                | Indo-European (mostly English and French) | Indo-European (mostly English) | Indo-European (mostly French) | Sino-Tibetan (Mandarin) |
| Disease                 | Mostly focal epilepsy | Tumour | Tumour | Tumour |
| WHO tumour grade, n(%)  |             |             |             |             |
| I or II                 | Not applicable | 41 (40.2)  | 256 (100.0) | 165 (64.5)  |
| III                     | Not applicable | 37 (36.3)  | 0 (0.0)     | 50 (19.5)   |
| IV                      | Not applicable | 24 (23.5)  | 0 (0.0)     | 41 (16.0)   |
| Positive sites on the left hemisphere, n | | | | |
| Speech arrest           | 219         | 75          | 198         | 425         |
| Anomia                  | 68          | 99          | 126         | 130         |
| Stimulator              |             |             |             |             |
| (i) Thyratron stimulator (before 1945) | | | Ojemann stimulator (5-mm interval bipolar electrodes, current-constant bipolar square wave, 1 ms wave width, 60 Hz frequency) |
| (ii) Rahm stimulator (1945-1951) | | | |
| (iii) Square wave generator (after 1951); (bipolar, mostly 60 Hz frequency) | | | |
| Language mapping strategy | Stimulate the postcentral gyrus until sensory responses were induced (mostly 1–3 V). Language mapping voltage was set twice the voltage. | The after-discharge threshold was induced before mapping. The current intensity was set to the maximum intensity without after-discharges. | Stimulate the central lobe until reliable motor/sensory responses were induced. Language mapping current was set to the same intensity. | Stimulate the precentral gyrus until motor responses were induced. Language mapping current was set to the same intensity. |
| Stimulation intensity range | ≤8 mA | 1–3.5 mA | 2–4 mA | 1–3 mA |