



OPEN ACCESS

EDITED BY

Roozbeh Behroozmand,
The University of Texas at Dallas, United States

REVIEWED BY

Laura Amaya Pascasio,
Complejo Hospitalario Torrecárdenas, Spain
Ziwen Yuan,
The First Affiliated Hospital of Xi'an Jiaotong
University, China

*CORRESPONDENCE

Nan Yan

✉ nan.yan@siat.ac.cn

Lan Wang

✉ lan.wang@siat.ac.cn

Shaofeng Zhao

✉ zhaosf1@163.com

RECEIVED 26 June 2024

ACCEPTED 16 December 2024

PUBLISHED 07 January 2025

CITATION

Liu J, Ruzi R, Jian C, Wang Q, Zhao S, Ng ML,
Zhao S, Wang L and Yan N (2025) Mapping
subcortical brain lesions, behavioral and
acoustic analysis for early assessment of
subacute stroke patients with dysarthria.
Front. Neurosci. 18:1455085.
doi: 10.3389/fnins.2024.1455085

COPYRIGHT

© 2025 Liu, Ruzi, Jian, Wang, Zhao, Ng, Zhao,
Wang and Yan. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Mapping subcortical brain lesions, behavioral and acoustic analysis for early assessment of subacute stroke patients with dysarthria

Juan Liu^{1,2}, Rukiye Ruzi¹, Chuyao Jian³, Qiuyu Wang⁴,
Shuzhi Zhao¹, Manwa L. Ng⁵, Shaofeng Zhao^{3*}, Lan Wang^{1*} and
Nan Yan^{1*}

¹Guangdong-Hong Kong-Macao Joint Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, ²University of Chinese Academy of Sciences, Beijing, China, ³Department of Rehabilitation Medicine, The Eighth Affiliated Hospital of Sun Yat-sen University, Shenzhen, China, ⁴Department of Radiology, The Eighth Affiliated Hospital of Sun Yat-sen University, Shenzhen, China, ⁵Speech Science Laboratory, Faculty of Education, University of Hong Kong, Hong Kong, Hong Kong SAR, China

Introduction: Dysarthria is a motor speech disorder frequently associated with subcortical damage. However, the precise roles of the subcortical nuclei, particularly the basal ganglia and thalamus, in the speech production process remain poorly understood.

Methods: The present study aimed to better understand their roles by mapping neuroimaging, behavioral, and speech data obtained from subacute stroke patients with subcortical lesions. Multivariate lesion-symptom mapping and voxel-based morphometry methods were employed to correlate lesions in the basal ganglia and thalamus with speech production, with emphases on linguistic processing and articulation.

Results: The present findings revealed that the left thalamus and putamen are significantly correlated with concept preparation ($r = 0.64$, $p < 0.01$) and word retrieval ($r = 0.56$, $p < 0.01$). As the difficulty of the behavioral tasks increased, the influence of cognitive factors on early linguistic processing gradually intensified. The globus pallidus and caudate nucleus were found to significantly impact the movements of the larynx ($r = 0.63$, $p < 0.01$) and tongue ($r = 0.59$, $p = 0.01$). These insights underscore the complex and interconnected roles of the basal ganglia and thalamus in the intricate processes of speech production. The lateralization and hierarchical organization of each nucleus are crucial to their contributions to these speech functions.

Discussion: The present study provides a nuanced understanding of how lesions in the basal ganglia and thalamus impact various stages of speech production, thereby enhancing our understanding of the subcortical neuromechanisms underlying dysarthria. The findings could also contribute to the identification of multimodal assessment indicators, which could aid in the precise evaluation and personalized treatment of speech impairments.

KEYWORDS

subacute stroke, dysarthria, speech production, linguistic processing, articulation, basal ganglia, thalamus

1 Introduction

Dysarthria refers to a motor speech disorder affecting respiration, phonation, resonance, articulation, and prosody (Enderby, 2013; Darley et al., 1969). It is the most common expressive communication deficit in post-stroke patients, with a prevalence of approximately 30–75% (Mackenzie, 2011; Liu et al., 2023; Ali et al., 2015). Recent studies suggested that dysarthria is frequently associated with pathologies in subcortical areas (Enderby, 2013; Rampello et al., 2016; Sperber and Karnath, 2016). For example, Summaka (Summaka et al., 2022) described how damage to the basal ganglia, thalamus, and cerebellum after a stroke led to dysarthria. The subacute phase exhibits distinct pathological characteristics compared to other stages of stroke or other conditions (e.g., Parkinson's disease) (Liu et al., 2023; Koyuncu et al., 2016). This phase is crucial for early and accurate diagnosis and assessment of dysarthria (Mackenzie, 2011; Spencer and Brown, 2018). Although the neural mechanisms underlying dysarthria have garnered significant attention, few studies focused on subcortical damage. Lesions in the basal ganglia and thalamus present different pathological manifestations of speech compared to other regions, such as the motor cortex and the Broca's area (Kearney and Guenther, 2019; Lim et al., 2014; Rampello et al., 2016). Therefore, investigating the impacts of these subcortical structures on speech production is essential for advancing our understanding of the pathological mechanisms of dysarthria following stroke.

Speech production is a complex process that integrates linguistic and motor functions, involving the precise control of muscles responsible for vocalization and articulation, ultimately resulting in the generation of speech signals (Enderby, 2013; Kearney and Guenther, 2019). As outlined by the lexical access model (Wilson et al., 2009), speech is produced through the following stages: conceptual preparation, lexical selection, phonological code retrieval and encoding, and articulation, all of which are influenced by various cerebral and subcortical structures (Levelt, 2000). The early stages of speech production, such as conceptual preparation and word selection, involve processes related to speech or language processing (Lieberman and Whalen, 1975; Pennington and Bishop, 2009). Conceptual preparation activates a lexical concept, which links to a lemma - an abstract lexical entry with syntactic but no phonological details. Subsequently, a phonological word form is retrieved and encoded through levels: phonological, then phonetic, leading to a motor program for articulation (Wilson et al., 2009; Levelt et al., 1999; Levelt, 2000). In recent years, many studies have revealed that the basal ganglia are associated with the early stages of speech processing outlined above (Guenther, 2007; Brown and Marsden, 1998; Graybiel, 2000), specifically, the basal ganglia play a critical role in initiating and organizing linguistic concepts and in selecting and retrieving words from the mental lexicon (Duffy et al., 2012; Silveri, 2021). Lesions in this region can result in anomia, characterized by difficulty in retrieving specific words or names (Camerino et al., 2024; Magee et al., 2019).

The stages of linguistic processing stages are succeeded by the phonological encoding and articulation phases, controlled by a coordinated set of motor activities involving the respiratory,

laryngeal, and articulatory systems, which are critical for producing different speech sounds (Wilson et al., 2009). Currently, the DIVA model is the most widely accepted neural network model for elucidating the neural and motor processes involved in speech production (Kearney and Guenther, 2019). Its input corresponds to the output of the phonological encoding stage, and it provides detailed neural and computational explanations of both the phonological encoding and articulation stages (Wilson et al., 2009), as well as the neural processes carried out by specific neuroanatomical regions (e.g., motor cortex, basal ganglia, thalamus) (Guenther, 2007). For instance, the DIVA model suggests that damage to the basal ganglia can disrupt both feedforward and feedback systems, leading to dyskinesia of the vocal organs (Tourville and Guenther, 2011; Guenther, 2007). Studies have shown that the basal ganglia are crucial for the planning and execution of fluent speech motor sequences, ensuring the precise timing and coordination of articulatory muscle movements (Duffy et al., 2012; Silveri, 2021). They also aid in the selection and initiation of speech-motor programming (Kempler and Van Lancker, 2002). Lesions to the basal ganglia can result in dysarthria, characterized by slurred speech, imprecise articulation, and irregularities in pitch, volume, and rhythm (Guenther, 2007), ultimately impairing speech muscle control and reducing vocal clarity (Silveri, 2021).

The thalamus serves as a critical conduit between subcortical nuclei and the cerebral cortex, playing a key role in speech and motor functions, with pathologies frequently leading to dysarthria (Graybiel, 2000; Bostan and Strick, 2018). Studies have demonstrated that the thalamus is essential for conceptual preparation, coordinating the integration of sensory and cognitive information (Pennington and Bishop, 2009). Its extensive connections with other brain regions enable the integration of diverse inputs, thereby influencing decision-making during lexical selection (Pickett et al., 1998). Thalamus also supports the fluency and consistency of phonological encoding processes (Wallesch et al., 1983). However, speech impairments resulting from thalamic pathologies, such as strokes and Parkinson's disease, are often overlooked or misdiagnosed as cognitive or motor deficits (Camerino et al., 2024), leaving the specific contributions of the thalamus to speech production insufficiently understood (Guenther, 2007; Johnson and Ojemann, 2000). Stroke patients with basal ganglia or thalamic damage frequently exhibit mild cognitive decline, complicating efforts to fully isolate the influence of cognitive factors - an enduring limitation in prior research (Lieberman and Whalen, 1975).

To further explore the neuromechanisms of speech processing, researchers have designed a series of classic speech behavioral experimental tasks, including speech fluency tasks and picture-naming tasks (Levelt and Meyer, 2000; Kempler and Van Lancker, 2002). Recent research has integrated neurophysiological and behavioral approaches, offering a comprehensive view of speech-related brain functions and their impairments (Chow et al., 2023). For articulation, the Frenchay Dysarthria Assessment (FDA) is a widely used clinical tool for diagnosing and assessing dysarthria, as well as the motor function of each vocal organ (Enderby, 1980). Additionally, acoustic analysis has been utilized to quantify the variability of pathological speech (Paja and Falk, 2012). However,

studies describing speech impairments often conflate measures of speech processing and articulation (Gagnon et al., 2018), leading to an incomplete understanding of the distinct contributions of each nucleus within the basal ganglia and thalamus to speech functionality, vocal movement, and speech quality (Nambu, 2008).

To address these issues, we first focused on subacute stroke patients with basal ganglia or thalamic damage and attempted to clarify the relationship between subcortical mechanisms and speech impairments by integrating neuroimaging, behavioral, and acoustic analyses. Voxel-Based Morphometry (VBM) and Voxel-Based Lesion-Symptom Mapping (VLSM) techniques were employed to precisely quantify the location and extent of these lesions (Bates et al., 2003). A series of speech tasks assessing fluency, picture association, naming, and color identification were administered to evaluate the linguistic functionality of speech production. Cognitive ability was also measured as an independent variable to determine the direct and independent effects of basal ganglia or thalamic lesion. The findings enhance our understanding of the neural basis of dysarthria, help identify multimodal potential biomarkers, and may lead to more precise and personalized treatments for speech impairments.

2 Materials and methods

2.1 Participants

All data were obtained from 42 participants, including 20 subacute stroke patients and 22 healthy controls. Patient recruitment followed specific inclusion and exclusion criteria. Inclusion criteria were as follows: (1) subacute stroke patients (first stroke within 6 months); (2) diagnosis of dysarthria; (3) subcortical damage; (4) age over 18; (5) at least a primary school education; (6) no prior speech or language therapy. Exclusion criteria included: (1) vision or hearing impairment; (2) dementia or psychiatric disorders; (3) other neurological conditions unrelated to the stroke. Based on demographic information of patients, 22 healthy individuals were selected as the control group. Detailed participant information is presented in Table 1. This study was approved by the Shenzhen Institute of Advanced Technology and the Eighth Affiliated Hospital of Sun Yat-sen University (Ethics Code: SIAT-IRB-220415-H0598). Written informed consent was obtained from all participants or their legal representatives prior to data collection.

All patients underwent a thorough clinical evaluation by certified Speech Language Therapists (SLTs) and Neurologists. Cognitive and speech ability of participants were assessed using the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) and the Frenchay Dysarthria Assessment (FDA) (Enderby, 1980; Ghio et al., 2019), respectively. The FDA evaluates 28 distinct dimensions of speech production, categorized into eight aspects: reflex, respiration, lips, jaw, velum, laryngeal, tongue, and intelligibility. The intelligibility category comprises word, sentence, conversation, and speed. Each dimension was rated on a scale of 0 (no function) to 4 (normal function). For more detailed FDA scores of all patients, refer to Supplementary Table 2.

TABLE 1 Demographic information and clinical evaluation results of all subjects.

	Patient	Normal	Statistic (t/p)
Number(N)	20	22	42
Gender(M:F)	16:4	16:6	0.58/0.30
Age(year)	59.65(14.02)	60.68(8.79)	-0.28/0.77
Education(year)	10.95(3.75)	11.82(3.76)	-0.74/0.45
MoCA(30)	20.90(5.22)	27.48(2.09)	-5.45**/0.00
FDA(0-4)	3.24(0.47)	4.00(0.00)	-7.42**/0.00

The statistics for Gender are χ^2 and p values (Male, M; Female, F). * indicates p-value, and the following is detailed representation: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.2 Data collection

2.2.1 Speech data

Speech materials were designed with consideration of the phonetic features of Chinese and the diverse impairment profiles observed in dysarthria, including syllables and characters. (1) Syllables: a total of 200 distinct syllables (monosyllabic, disyllabic, and multisyllabic) were included, covering high-frequency syllables, low-frequency syllables, and easily confused syllables. (2) Characters: a total of 90 high-frequency characters were selected, including numbers, direction, and basic verbs. Audio signals were recorded in a soundproof room using a professional-grade microphone (Takstar, MS400), with a sampling rate of 16 kHz, 16-bit encoding, and single-channel recording. The microphone was positioned approximately 8-10 cm from the speaker's mouth. Further details of each task and the experimental paradigm can be found in previous studies (Liu et al., 2023).

2.2.2 Behavioral data

The linguistic processing of speech production was evaluated by using four classic tasks, which involving the stages of conceptual preparation and lexical selection (Levelt, 1992; Indefrey and Levelt, 2004). The tasks included a speech fluency task (Wang et al., 2019), a picture association task (Lin et al., 2020), a picture naming task (Liu et al., 2011), and a color naming task (Monsell et al., 2001). Each behavioral task involved common objects and was conducted in separate sessions using the E-prime program and Chronos system (Babjack et al., 2015) for precise measurement of reaction times of sound onset on each trial. The order of item presentation was randomized for each task and consistent across subjects. There was a 1-minute response limit for the speech fluency task and a 6-second limit for the other tasks. Testing was conducted individually in a quiet room, with each session lasting for no more than half an hour, including rest breaks as needed. The schematic diagram of the behavior experiment procedure is shown in Figure 1. In addition to recording the reaction time of each behavioral task, synchronized audio data were also obtained using Praat software (Boersma, 2001). The behavioral audio data were recorded using the same equipment as the speech data mentioned above.

Speech fluency task (SF). The SF task was designed to examine linguistic functioning and speech fluency (Wang et al., 2019). In the SF task, participants were instructed to produce as many words

as they could within one minute which were under each of the ten semantic categories such as fruits, animals, vehicles, etc. A practice block was provided with a different category (weather) to familiarize participants with the task, but was not included in the analysis (see Figure 1A). To avoid potential memory effects, no further instructions were given during the experimental trials.

Picture association task (PA). Following Lin et al. (2020), this task involved naming randomly presented black-and-white outlined drawings on a screen. The stimuli consisted of 40 black-and-white outline drawings displayed on a white background. Each trial was displayed for 6 seconds, following a 50 ms beep (see Figure 1B). Meanwhile, participants were expected to verbally name the picture. The 40 stimuli were randomly arranged in a sequence, with 10 trials comprising a block. Participants were trained to name six additional pictures aloud before the experiment to ensure readiness for the formal task. The task was repeated to enhance the quality of the behavioral data.

Picture naming task (PN). This task followed the same format as the PA task, and the stimulus material was based on Liu et al. (2011). A total of 100 items were selected from all 435 object line drawings, with the concept familiarity, subjective word frequency, image agreement, image variability, and visual complexity of these pictures balanced. Participants were instructed to name each picture quickly and accurately.

Color naming task (CN). This task followed the same format as the PA and PN tasks (see Figure 1B). A total of 36 most common object words were selected as experimental stimuli, all of which possessed obvious color characteristics, such as banana and milk. Participants were instructed to quickly and accurately name the color of each item (Monsell et al., 2001).

2.2.3 Imaging data

The imaging data of the participants were acquired from the Eighth Affiliated Hospital of Sun Yat-sen University using a 3T Siemens scanner. We collected three types of images: (i) high-resolution 3D T1-weighted images; (ii) FLAIR T2-weighted images; and (iii) diffusion-weighted images. The 3D images were T1-weighted 3D MPRAGE images on the sagittal plane with parameters: matrix size = 512×512 , voxel size = $0.5 \times 0.5 \times 0.5 \text{ mm}^3$, repetition time = 12 ms, echo time = 4.2 ms, inversion time = 400 ms, field of view = $256 \times 256 \text{ mm}^2$, flip angle = 15° , slice number = 320 slices. The FLAIR T2 images were FLAIR T2-weighted images on the axial plane with parameters: matrix size = 512×512 , voxel size = $0.5 \times 0.5 \times 5 \text{ mm}^3$, repetition time = 8000 ms, echo time = 122 ms, inversion = 2 s, field of view = $250 \times 250 \text{ mm}^2$, flip angle = 90° , slice number = 28 slices. Diffusion-weighted imaging had two separate sequences with different diffusion weighting direction sets so 32 directions were covered in total. The first acquisition had the following parameters: 12 diffusion weighting directions, matrix size = 128×128 , voxel size = $2.0 \times 2.0 \times 2.6 \text{ mm}^3$, repetition time = 13 000 ms, echo time = 69 ms, inversion time = 0 s, field of view = $250 \times 250 \text{ mm}^2$, flip angle = 90° , slice number = 53 slices. The other acquisition had the same parameters except that it included 20 different directions. The first two volumes were b_0 volumes and the b-value of other volumes was $1,000 \text{ s/mm}^2$ in each sequence.

2.3 Data analysis

2.3.1 Speech data annotation and acoustic feature extraction

All the speech data were segmented and transcribed verbatim by two experienced researchers using Praat (Boersma, 2001). Relevant speech materials were manually transcribed in Chinese characters on the first tier. Subsequently, the speech content of each participant was transcribed in Pinyin (a phonetic representation of Chinese) on the second tier. The character and syllable tasks were manually segmented into vowel and consonant segments, which were marked on the third tier based on their spectrograms and auditory judgments. The method for annotating vowels and consonants is presented in Supplementary Figure 1.

Since vowel distortion and consonant errors are characteristic manifestations of dysarthria, acoustic features were extracted from the vowel segments. A comprehensive 12-dimensional set of acoustic features was derived from the medial 80% of the steady-state vowel segment, encompassing the following parameters: vowel duration, F1 and F2 variability, jitter, shimmer (Hernandez et al., 2020), harmonics-to-noise ratio (HNR) (Karan et al., 2020), tongue distance, jaw distance (Sapir et al., 2010), Vowel Space Area (VSA), Formant Centralization Ratio (FCR) (Banks et al., 2019), degree of movement (F2i/F2u ratio) (GE et al., 2021), and Vowel Articulation Index (VAI).

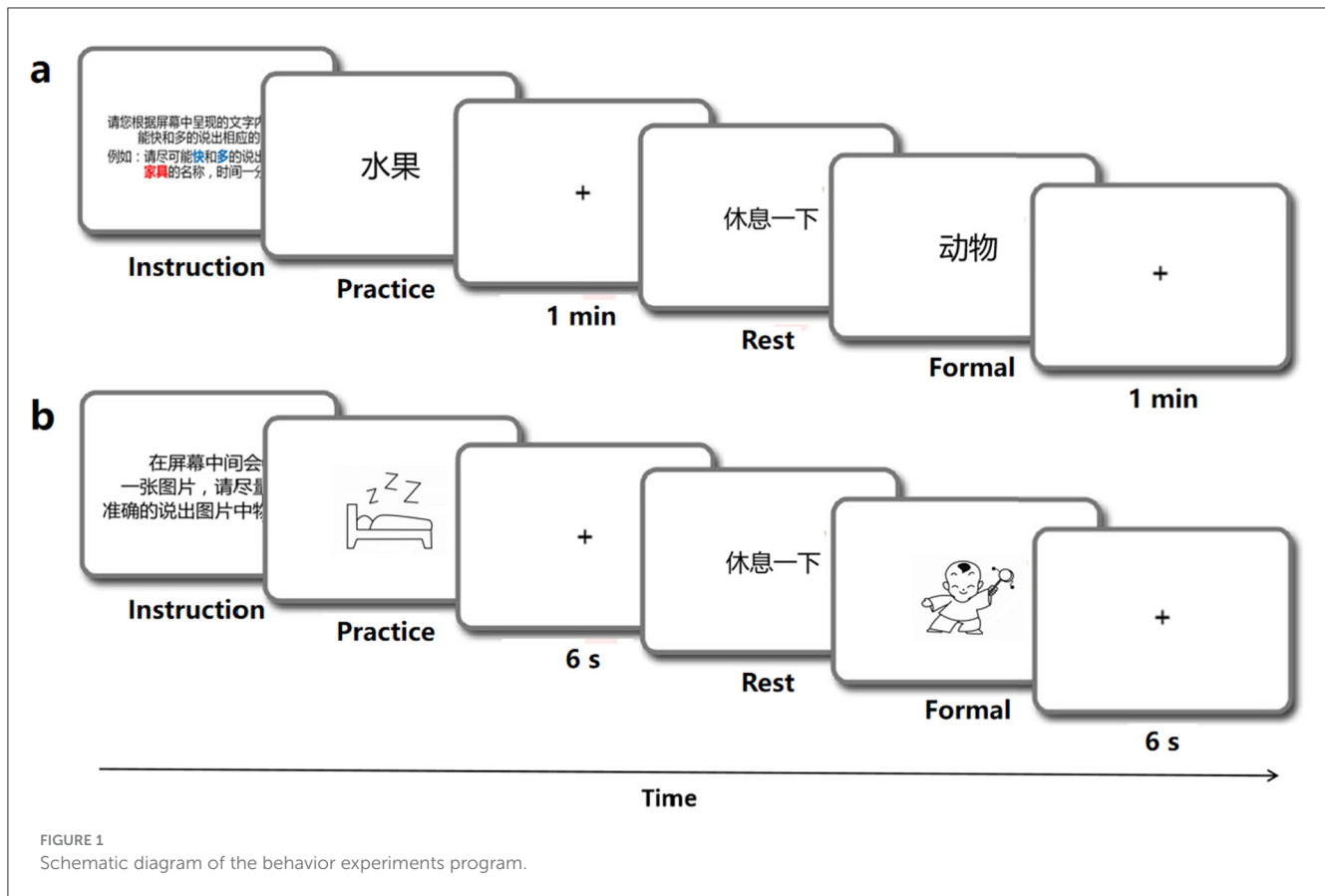
2.3.2 Behavioral data annotation and feature extraction

To ensure the precise calculation of both the number of valid responses and the reaction times for each behavioral task, all audio data underwent manual annotation. For the SF task, correct answers were annotated within one minute and counted within each 15-s time window, including four periods: 0–15 s (T1), 15–30 s (T2), 30–45 s (T3), and 45–60 s (T4). For the PA, PN, and CN tasks, each “beep” voice and effective answer were also labeled on the first tier. The reaction time was determined by the interval between the “beep” and the response, as depicted in Supplementary Figure 2.

Meanwhile, behavioral features were extracted corresponding to different tasks. For the SF task, SF_T1, SF_T2, SF_T3, and SF_T4 represented the number of words within 0–15 s, 15–30 s, 30–45 s, and 45–60 s windows. Reaction times (RT) for each trial in the PA, PN, and CN tasks were also directly extracted for each subject.

2.3.3 Imaging data preprocessing

All DICOM files were converted into NifTi format using SPM12 software, and the data quality, including sharpness, whole-head coverage, and orientation, was carefully checked. We first coregistered the two T1-weighted structural imaging data in the same native space using the trilinear interpolation method implemented in SPM12, resulting in an averaged structural image. Subsequently, we coregistered and resliced the FLAIR T2 images to the averaged structural image using the same trilinear interpolation method in SPM12. The lesion contours of each patient were drawn on the T1 structural image by one trained person, slice-by-slice, visually referring to the FLAIR T2 images in FSL (<http://fsl.fmrib.ox.ac.uk>). These lesion boundaries were further validated by a



neuroradiologist (Q.W.) from the Eighth Affiliated Hospital of Sun Yat-sen University, utilizing T2 and DWI images. Patients with diffuse damage or cortical lesions that could not be accurately bounded were excluded from the study. Then, the structural images were resliced into $1 \times 1 \times 1\text{mm}^3$ voxel sizes.

For normalization, a manual registration technique was applied to minimize the impact of lesions on brain image distortion. This approach has been documented by Li et al. (2017) and Zhao et al. (2017). Structural images from each patient were manually registered into Talairach space using the “3D Volume Tools” in BrainVoyager QX v2.0 (www.brainvoyager.com). ANTs software was utilized to estimate the affine transformation matrix between the native and Talairach spaces. Using this matrix, we then transformed the lesion maps into Talairach space using the “WarpImageMultiTransform” program. Finally, the lesion maps were further transformed into MNI space for the next analysis.

2.3.4 Lesion-symptom mapping analysis

The present study utilized structural MRI to evaluate brain damage severity through two primary indices: the lesion status, expressed as a lesion percent value, and the volume of gray matter (VGM value). The lesion percent is calculated as the number of voxels with lesions divided by the total number of voxels in each area. Consistent with the methodology of Han et al. (2013), identical parameters for scanning and preprocessing were used to generate the lesion map. The lesion index reflects physical damage

to voxels, encompassing both white and gray matter. It serves as a dichotomous variable, classifying each voxel as either intact or lesioned, and is widely regarded as a classical indicator of lesion severity (Meyer et al., 2016). In the present study, each voxel in every patient was assigned a lesion value (categorical variable) derived from the lesion map, along with a structure volume (continuous variable). It is noted from prior research that brain regions with lesions often show reduced VGM values compared to undamaged areas (Fox et al., 2016), suggesting a correlation between these variables.

2.3.5 Voxel-based morphological analysis

Structural image analysis was performed using SPM12 software. The technical details of primary cortical reconstruction and volumetric segmentation procedures have been previously described (Chen et al., 2020; Weisstein, 2004). In brief, the processing included removal of nonbrain tissue using a hybrid watershed/surface deformation procedure, automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures, intensity normalization, tessellation of the gray matter/white matter boundary, automated topology correction, and surface deformation following intensity gradients to optimally place the gray/white and gray/cerebrospinal fluid (CSF) borders at the location where the greatest shift in intensity defined the transition to the other tissue class. Cortical thickness was calculated as the closest distance from

the gray/white matter boundary to the gray/CSF boundary at each vertex. Adjustments were then made for differences in head size and volumes for each region (white matter and subcortical structure volume) were adjusted to intracranial volume (ICV).

2.3.6 Statistical analysis

First, statistical analyses were conducted on behavioral and acoustic features to assess speech impairments. A linear mixed-effects regression model (Liu et al., 2023) was employed to investigate the differences in linguistic processing and speech quality between the dysarthria and control groups. Each model incorporated fixed effects for the group (dysarthria vs. normal), cognitive level as measured by Montreal Cognitive Assessment (MoCA) scores, and participants and age as random intercepts. The slope estimates of the model, which indicate the average impact of dysarthria or cognitive status on speech, are expressed as beta coefficients (B) with standard errors (SE). A positive B value suggests an increase in the measured parameter, while a negative value suggests a decrease.

To ascertain the direct impact of subcortical nuclei on articulatory movement and linguistic functions, the Pearson correlation analysis was conducted separately for each acoustic feature, behavioral feature, FDA score, and VGM. This included comparisons between specific acoustic features (e.g., tongue distance, Vowel Space Area [VSA]), FDA scores, and behavioral features (e.g., Speech Fluency task at 0–15 s [SF_T1], Reaction Time of PA task [RT_PA]), as well as the volume of gray matter (VGM) and the percentage of brain lesions in the basal ganglia and thalamus. Correlation coefficients were calculated between each objective acoustic-behavioral feature, subjective FDA scores, and the VGM, as well as the lesion percentages. A Bonferroni correction was applied to adjust for multiple comparisons (Weisstein, 2004), with the threshold for statistical significance set at an alpha level of 0.05. Statistical models were developed using the R programming language and the lme4 package (Bates, 2010).

3 Results

3.1 Distribution and proportion of subcortical structure damage

Under the guidance of professional radiologists, the precise extent and boundaries of brain lesions for each patient were carefully outlined. The distribution and proportion of lesion locations within subcortical nuclei were then assessed. The distribution of subcortical lesions in the study cohort is illustrated by the lesion prevalence map shown in Figure 2. The subsequent analysis focused on subcortical nuclei where the damage ratio surpassed 20% in all patients (at least 4 patients) to ensure the statistical validity and reliability of our findings. This included an evaluation of the caudate nucleus, putamen, globus pallidus, and thalamus. For more detailed lesion information of the 20 subacute stroke patients, refer to Supplementary Tables 1, 7.

3.2 Behavioral and acoustic results

The results of the behavioral tasks reflect the significant differences on linguistic processing between the dysarthric and normal group. As shown in Figure 3A, there are significant differences in the T1, T2, and ALL windows of SF task between dysarthria and normal groups ($t = 4.68, p < 0.001$; $t = 2.11, p = 0.03, t = 3.71, p < 0.001$). Figure 3B suggests that the RTs of CN, PN, and PA tasks of patients were longer than normal controls ($t = -7.13, p < 0.001$; $t = -3.19, p = 0.001$; $t = -3.60, p < 0.001$), and the significance of differences between two groups gradually diminishes with the difficulty of tasks. In addition, comparing to the influence of cognition, we found significant differences in the T1, T2, T3, T4, and ALL windows ($t = 2.59, p = 0.009$; $t = 2.29, p = 0.02$; $t = 2.92, p = 0.003$; $t = 2.53, p = 0.011$; $t = 2.39, p = 0.016$). The detailed statistical results of behavioral features are presented in Supplementary Tables 2, 3.

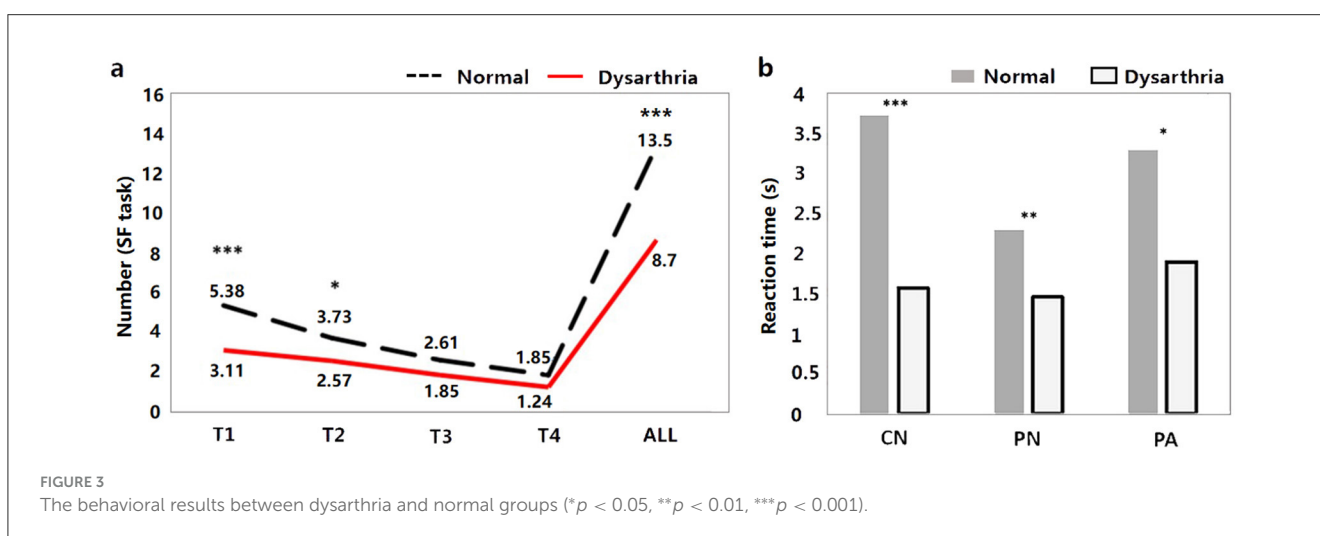
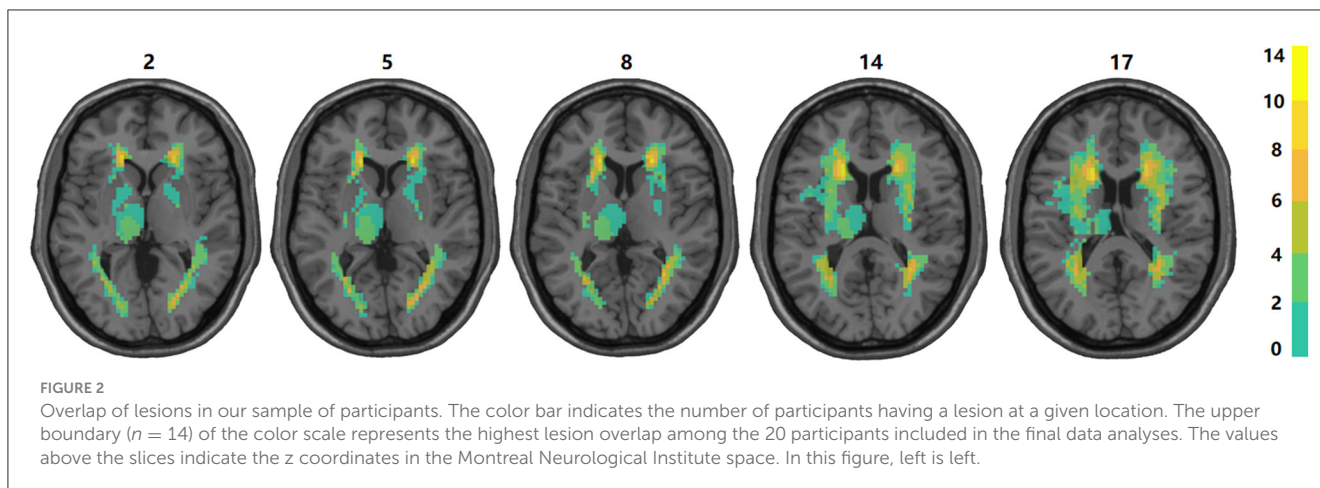
The FDA and acoustic results represent the significant difference on articulation stage between the dysarthric and normal group. As shown in Figure 4, dysarthria patients had significantly longer vowel durations ($t = 6.81, p < 0.001$), higher F1/2 variability vs. controls ($t = -2.74, p = 0.005$; $t = -8.43, p < 0.001$). In voice quality, patients had higher jitter, shimmer, and HNR ($t = 0.77, p = 0.43$; $t = 1.50, p < 0.13$; $t = 2.02, p = 0.04$). In articulation, patients showed significantly lower jaw distance, tongue distance, and movement degree ($t = 1.74, p = 0.08$; $t = 4.22, p < 0.001$; $t = 3.43, p < 0.001$). In vowel space measures, patients had significantly lower vowel space area (VSA) and vowel articulation index (VAI), and higher vowel centralization (FCR) than controls ($t = 4.60, p < 0.001$; $t = 3.14, p = 0.001$; $t = -2.91, p = 0.003$). In addition, when comparing the impact of cognitive factors, significant differences were observed only in vowel duration and F2 variability between dysarthria patients and normal controls ($t = -5.02, p < 0.001$; $t = 3.03, p = 0.002$). The complete statistical values of acoustic features are referred in Supplementary Table 1.

3.3 The lesion-behavioral-acoustic results of basal ganglia

3.3.1 Caudate nucleus

Figure 5 illustrates the correlation results between the caudate nucleus and various aspects of the speech production process, including each speech behavioral task (SF, PA, PN, CN tasks), the motor ability of each vocal organ (the FDA score of each category), and acoustic features (e.g., vowel duration, jaw distance, VSA). The detailed results associated with the caudate nucleus are presented in Supplementary Table 4.

Left caudate nucleus results. (1) Behavioral tasks: there were significant correlations observed between damage to the left caudate nucleus and the SF_T2 window, PN task ($r = 0.40, p = 0.07$; $r = -0.50, p = 0.02$), as well as a moderate correlation with the SF_T3 window and PA task ($r = 0.30, p = 0.19$; $r = -0.39, p = 0.08$). (2) FDA score of each category: there were significant correlations observed between damage to the left caudate nucleus and the Tongue ($r = 0.56, p = 0.02$), as well as a mild correlation with the Laryngeal and the speed of intelligibility ($r = 0.24,$



$p = 0.35$; $r = 0.22$, $p = 0.40$). (3) Acoustic features: there were mild correlations between the left caudate nucleus and the vowel duration, F1/F2 variability ($r = -0.22$, $p = 0.35$; $r = -0.23$, $p = 0.34$; $r = -0.20$, $p = 0.39$).

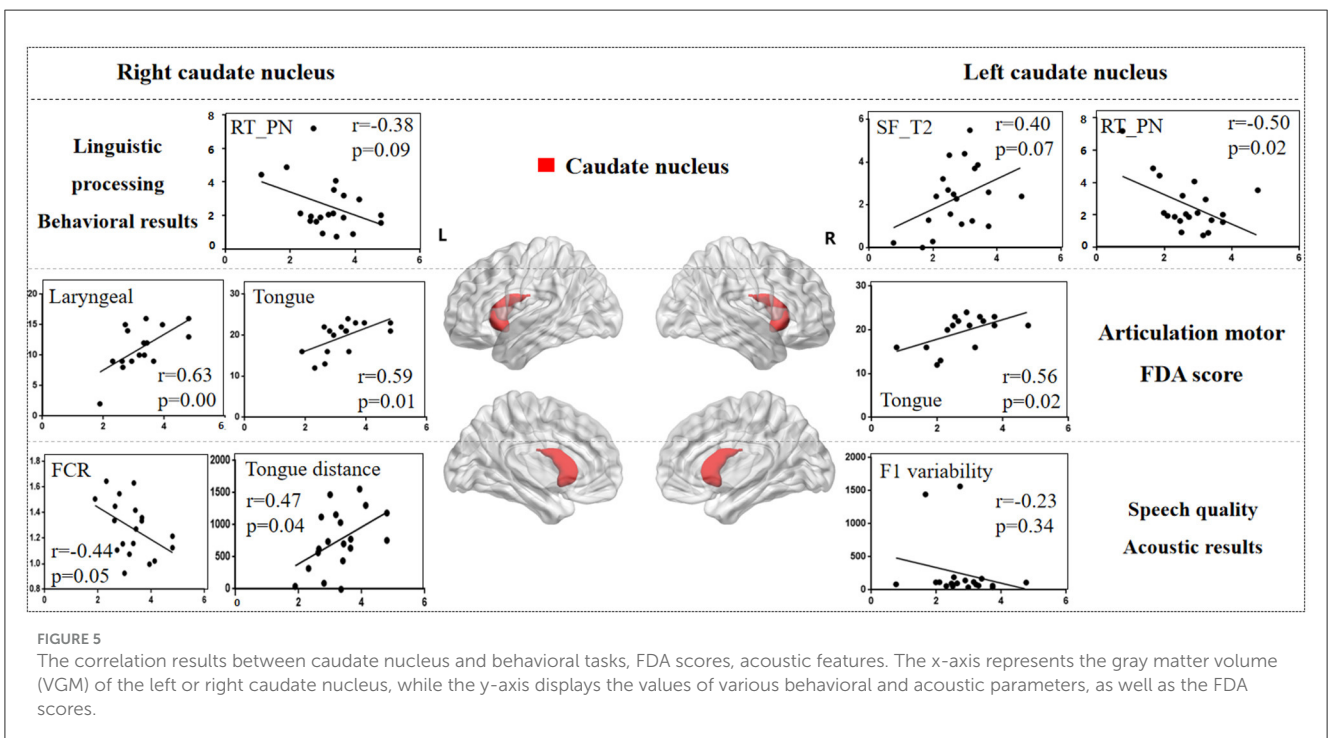
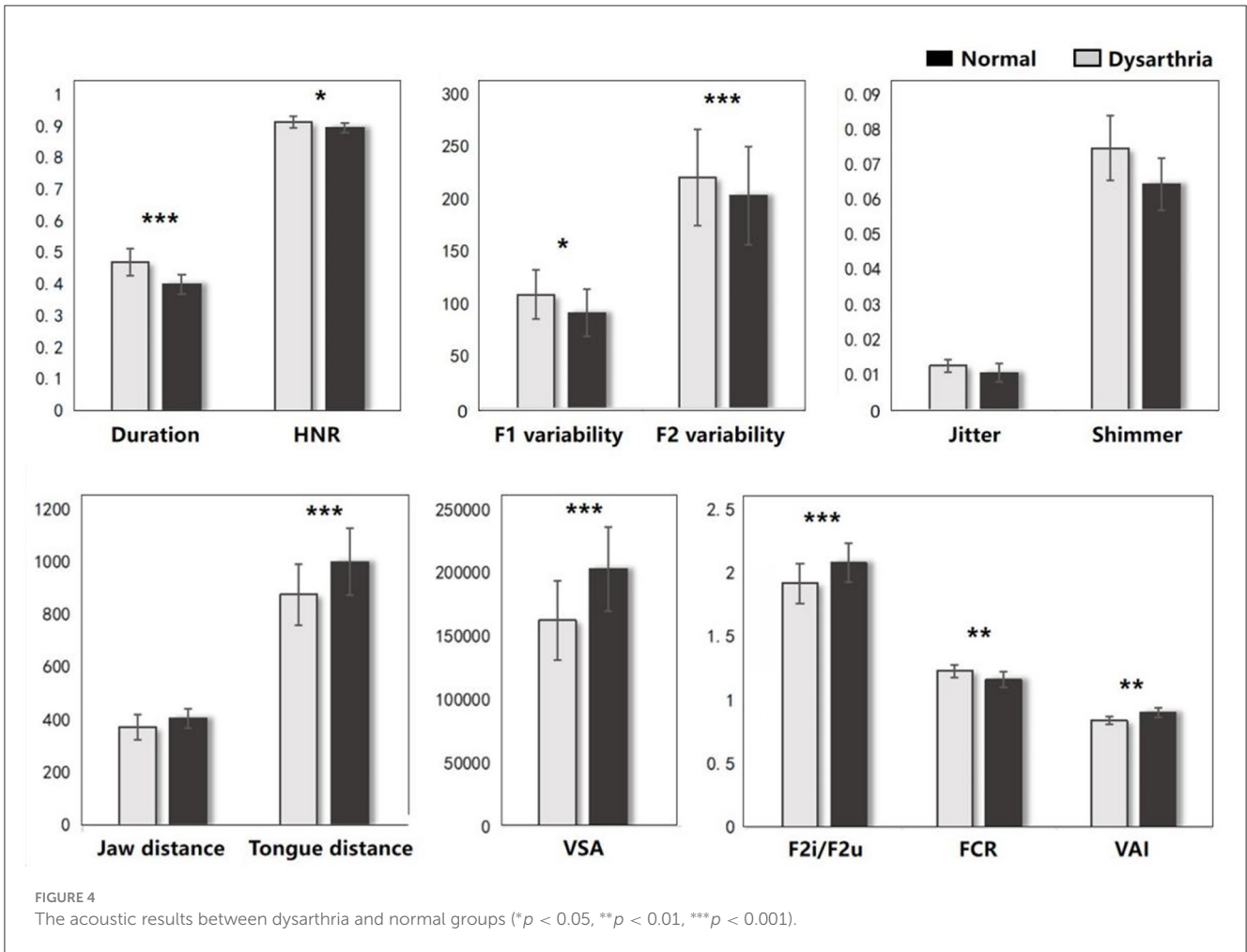
Right caudate nucleus results. (1) FDA score of each category: significant correlations have been observed between the right caudate nucleus and the Laryngeal, Tongue, word and speed of intelligibility ($r = 0.63$, $p = 0.008$; $r = 0.59$, $p = 0.01$; $r = 0.61$, $p = 0.01$; $r = 0.56$, $p = 0.02$), and moderate correlations with the Lips, sentence and conversation intelligibility ($r = 0.46$, $p = 0.06$; $r = 0.42$, $p = 0.10$; $r = 0.42$, $p = 0.10$). (2) Behavioral tasks: there were mild correlations observed between the right caudate nucleus and the PN, PA tasks, and SF_T2 window ($r = -0.38$, $p = 0.09$; $r = -0.19$, $p = 0.41$; $r = 0.19$, $p = 0.40$). (3) Acoustic features: meanwhile, there also were mild correlations between the right caudate nucleus and the Duration, F1/2 variability ($r = -0.22$, $p = 0.35$; $r = -0.23$, $p = 0.34$; $r = -0.20$, $p = 0.39$).

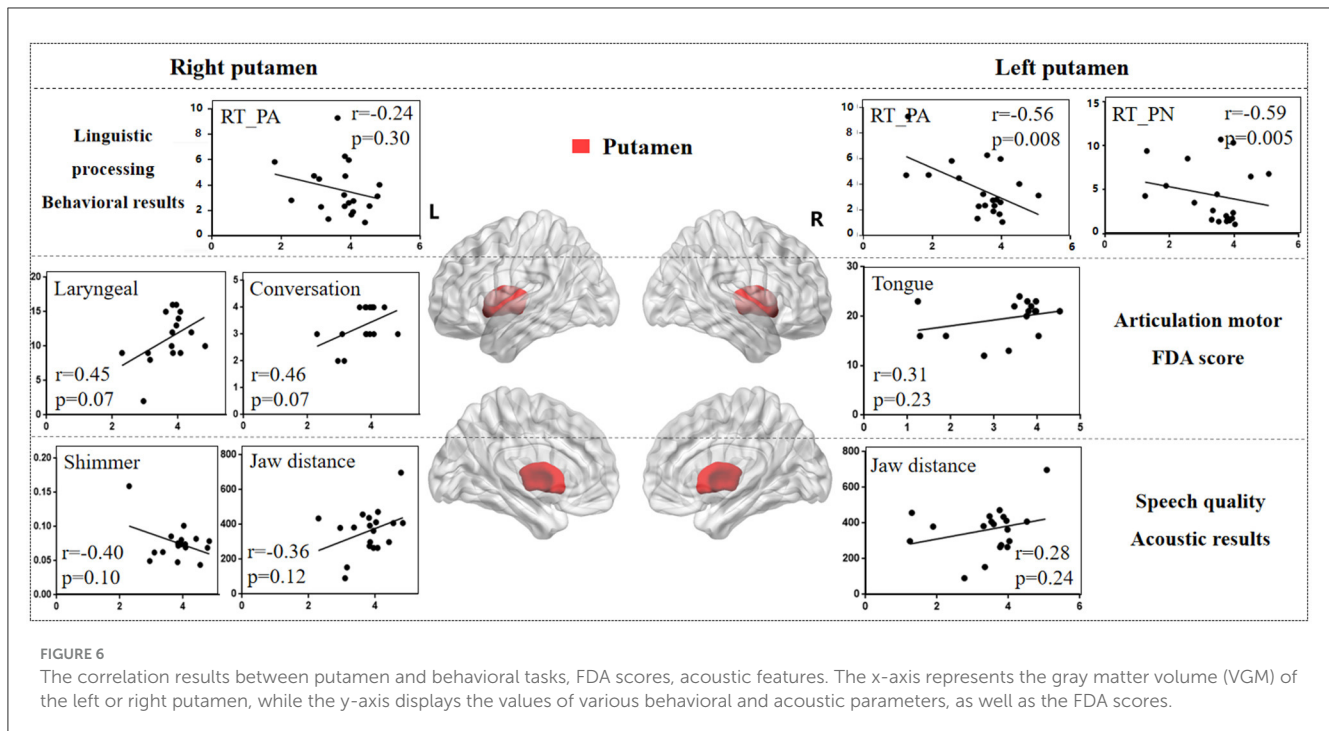
3.3.2 Putamen

Figure 6 illustrates the correlation results between the putamen and various aspects of speech production. The detailed putamen results are presented in Supplementary Table 5.

Left putamen results. (1) Behavioral tasks: there were significant correlations observed between damage to the left putamen and the PA, PN tasks, SF_T2 window ($r = -0.56$, $p = 0.008$; $r = -0.59$, $p = 0.005$; $r = 0.47$, $p = 0.03$), as well as a moderate correlation with the SF_T3 window and CN task ($r = 0.38$, $p = 0.09$; $r = -0.21$, $p = 0.37$). (2) FDA score of each category: there were only mild correlations observed between left putamen and Laryngeal, Tongue ($r = 0.31$, $p = 0.23$; $r = 0.31$, $p = 0.23$). (3) Acoustic features: there were also mild correlations between left putamen and vowel duration, F1/F2 variability ($r = -0.24$, $p = 0.32$; $r = -0.27$, $p = 0.26$; $r = -0.27$, $p = 0.25$).

Right putamen results. (1) FDA score of each category: significant correlations have been observed between the right putamen and the Laryngeal, conversation intelligibility ($r = 0.45$, $p = 0.07$; $r = 0.46$, $p = 0.07$), and mild correlations with the Respiration, Jaw, Tongue, and speed of intelligibility ($r = -0.30$, $p = 0.24$; $r = -0.23$, $p = 0.37$; $r = 0.33$, $p = 0.20$; $r = 0.38$, $p = 0.14$). (2) Acoustic features: meanwhile, there were moderate correlations between the right putamen and the Shimmer, Jaw distance ($r = -0.40$, $p = 0.10$; $r = 0.36$, $p = 0.12$), and mild correlations with the Jitter, HNR, VSA ($r = -0.29$, $p = 0.25$; $r = 0.20$, $p = 0.41$; $r = 0.27$, $p = 0.26$). (3) Behavioral tasks: there were only mild correlations





observed between right putamen and the PA, PN tasks ($r = -0.24$, $p = 0.30$; $r = -0.21$, $p = 0.36$).

3.3.3 Globus pallidus

Figure 7 illustrates the correlation results between the globus pallidus and various aspects of speech production. The detailed globus pallidus results are presented in Supplementary Table 6.

Left globus pallidus results. (1) FDA score of each category: there were significant correlations observed between damage to the left globus pallidus and the Respiration, Lips, Laryngeal, Tongue, and word, sentence, conversation intelligibility, and speed of intelligibility ($r = -0.48$, $p = 0.05$; $r = -0.68$, $p = 0.003$; $r = -0.63$, $p = 0.008$; $r = -0.60$, $p = 0.01$; $r = -0.78$, $p < 0.001$; $r = -0.52$, $p = 0.03$; $r = -0.52$, $p = 0.03$; $r = -0.59$, $p = 0.01$), as well as a moderate correlation with the Jaw ($r = -0.32$, $p = 0.22$). (2) Behavioral tasks: there were significant correlations between left globus pallidus and SF_T1 and T3 windows ($r = -0.45$, $p = 0.04$; $r = -0.42$, $p = 0.06$), and moderate correlations with SF_T2 and T4 windows, and PN task ($r = -0.33$, $p = 0.15$; $r = -0.22$, $p = 0.34$; $r = 0.23$, $p = 0.31$). (3) Acoustic features: there were only mild correlations between left globus pallidus and vowel duration, F1/F2 variability, Tongue distance, F2i/F2u, FCR, and VAI ($r = 0.30$, $p = 0.21$; $r = 0.31$, $p = 0.19$; $r = 0.30$, $p = 0.20$; $r = -0.34$, $p = 0.14$; $r = -0.31$, $p = 0.18$; $r = 0.37$, $p = 0.10$; $r = -0.31$, $p = 0.18$).

Right globus pallidus results. (1) FDA score of each category: significant correlations have been observed between the right globus pallidus and the Respiration, Lips, and word intelligibility ($r = -0.54$, $p = 0.02$; $r = -0.59$, $p = 0.01$; $r = -0.52$, $p = 0.03$), and mild correlations with the Laryngeal ($r = -0.38$, $p = 0.14$). (2) Acoustic features: meanwhile, there were significant correlations between the right globus pallidus and the Jaw distance, VSA ($r = 0.54$, $p = 0.01$; $r = 0.50$, $p = 0.02$). (3) Behavioral tasks:

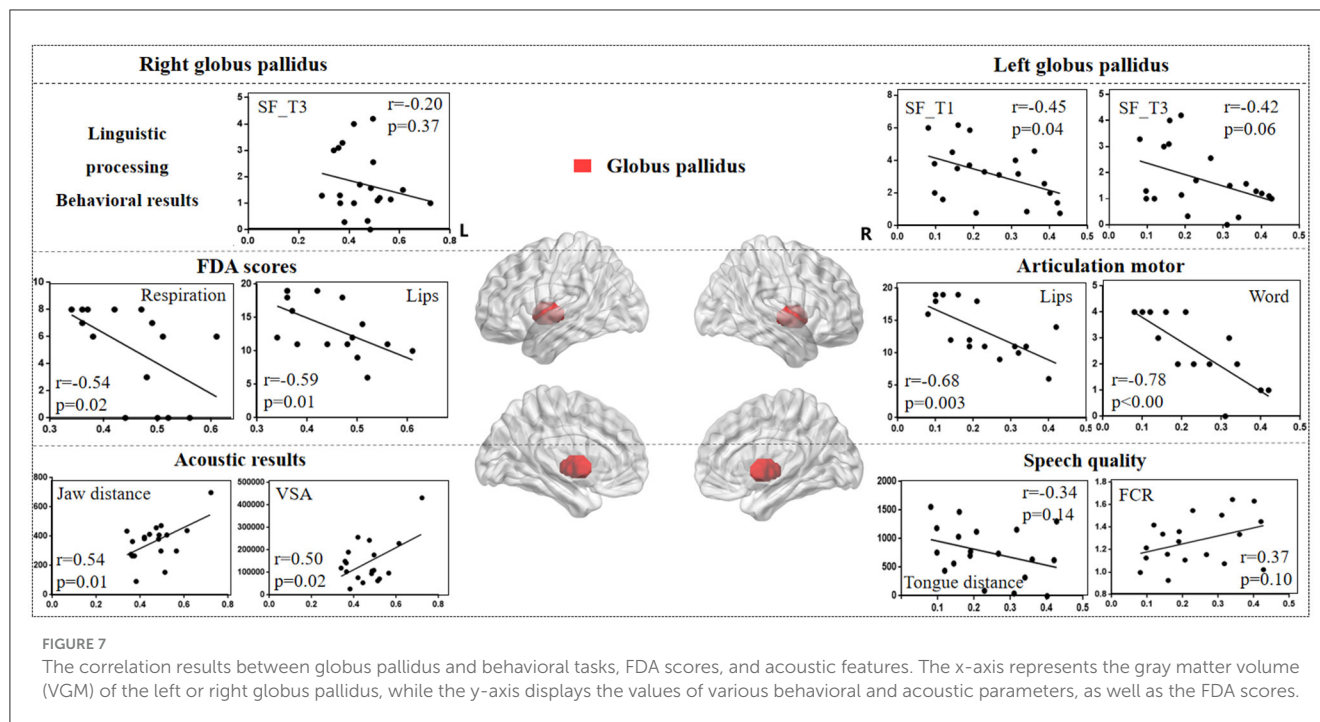
there were only mild correlations observed between the right globus pallidus and the SF_T1 and T3 windows ($r = -0.19$, $p = 0.41$; $r = -0.20$, $p = 0.37$).

3.4 The lesion-behavioral-acoustic results of thalamus

Figure 8 illustrates the correlation results between the thalamus and various aspects of speech production. The detailed thalamus results are presented in the Supplementary material.

Left thalamus results. (1) Behavioral tasks: There were significant correlations between left thalamus and SF_T2 and T3 windows ($r = 0.54$, $p = 0.01$; $r = 0.45$, $p = 0.04$), PA and PN tasks ($r = -0.64$, $p = 0.002$; $r = -0.67$, $p = 0.001$), and mild correlations with SF_T4 window and CN task ($r = 0.16$, $p = 0.48$; $r = -0.36$, $p = 0.11$). (2) FDA score of each category: There were only mild correlations observed between damage to the left thalamus and the Laryngeal, conversation intelligibility, and speed of intelligibility ($r = 0.30$, $p = 0.26$; $r = -0.24$, $p = 0.38$; $r = 0.37$, $p = 0.17$). (3) Acoustic features: There were only mild correlations between left thalamus and vowel duration, F1/F2 variability, and shimmer ($r = -0.21$, $p = 0.37$; $r = -0.22$, $p = 0.34$; $r = -0.23$, $p = 0.34$; $r = 0.36$, $p = 0.15$).

Right thalamus results. (1) FDA score of each category: significant correlations have been observed between right thalamus and the Respiration, Lips ($r = -0.52$, $p = 0.04$; $r = -0.58$, $p = 0.02$), and moderate correlations with the word intelligibility ($r = -0.35$, $p = 0.19$). (2) Behavioral tasks: There were significant correlations observed between right thalamus and the PA, CN tasks ($r = -0.44$, $p = 0.04$; $r = -0.42$, $p = 0.06$), and moderate correlations with SF_T1, T2, T3 windows, and PN task ($r = 0.32$, $p = 0.15$; $r = 0.37$,



$p = 0.10$; $r = 0.38$, $p = 0.09$; $r = -0.33$, $p = 0.14$). (3) Acoustic features: Meanwhile, there were only mild correlations between the right thalamus and the Jaw distance ($r = 0.24$, $p = 0.32$).

4 Discussion

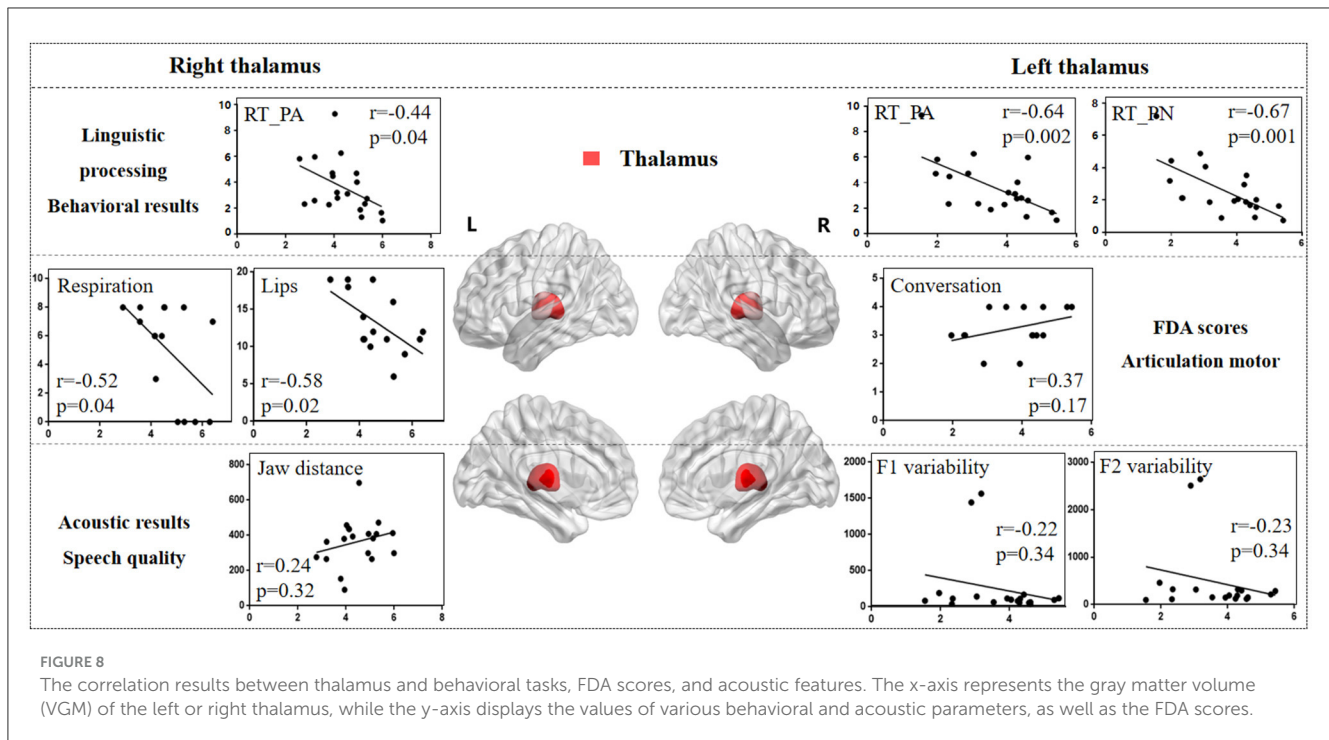
Seeing the shortcomings associated with the impact of different diseases and lesion locations in previous studies on dysarthria, the present study is the first to integrate multimodal neuroimaging, behavioral, and speech data to explore the direct effects of various subcortical nuclei on speech production (Polikoff and Bunnell, 1999; Kim et al., 2008). The present investigation uniquely focused on subacute stroke patients suffering from lesions in the basal ganglia and/or thalamus, with the goal of thoroughly uncovering the functions of these distinct nuclei throughout the linguistic processing and articulation stages. Linguistic processing encompassed a spectrum of functional stages as outlined by the lexical access model, and the articulation stage includes motor control of the vocal organs and acoustic features (Wilson et al., 2009).

Recent studies suggested that the basal ganglia are involved not only in motor execution, but also in the linguistic processing (Lim et al., 2014; Silveri, 2021). The present behavioral results confirmed this hypothesis and revealed the relationship between the basal ganglia and the early stages of speech production. The behavioral tasks used in present study collectively covered the full spectrum of stages of speech production, from conceptual preparation to articulation (Levelt, 1992). When task complexity escalated, both response time and error rate increased (as depicted in Figure 4, Supplementary Table A2), suggesting a heightened reliance on top-down and cognitive processing mechanisms (Wilson et al., 2009). For example, the SF task appeared to require

not only phonetic encoding and articulation, but also a more substantial engagement in the conceptual preparation and word retrieval phases (Levelt et al., 1999). SF results revealed significant discrepancies, particularly in the T1 and T2 windows (as shown in Figure 4A). This indicates that dysarthria may impede speech fluency, linked to the initiation and pace of speech production, and potentially associated with the feedforward control system outlined in the DIVA model (Kearney and Guenther, 2019).

Previous research has highlighted the crucial roles of the basal ganglia and thalamus in articulation movement (Lim et al., 2014; Ziegler and Ackermann, 2017). Consistent with this, the present study found that stroke patients with basal ganglia and thalamus lesions exhibited prolonged vowel durations, reduced stability, smaller vowel space area, and decreased movement distance during speech production (see Figure 3). These findings suggest that basal ganglia and thalamic damage impairs motor control, leading to less precise and more irregular speech patterns compared to healthy adults (Mou et al., 2018). Additionally, correlation analysis revealed that different basal ganglia nuclei facilitate the motor execution of articulation stages by controlling and coordinating vocal organ movements (Silveri, 2021). The left caudate nucleus primarily influences tongue movement, while the right caudate nucleus more directly affects the motor functions of the tongue, larynx, and lips, resulting in poorer speech intelligibility. Similarly, the right globus pallidus is linked to motor abilities of the vocal organs, including lips, larynx, tongue, and respiration. These findings consistent with acoustic results, where the right globus pallidus is associated with jaw distance and VSA.

Studies on Parkinson's disease (PD) patients have shown a decline in cognitive and motor functions due to basal ganglia and thalamus atrophy (Federico et al., 2015; Tang et al., 2018). However, in stroke patients, damage in the basal ganglia and thalamus more directly affects movement of vocal organs. Exploring speech



impairments caused by different diseases (e.g., PD and stroke) helps contribute to the development of targeted diagnostic and therapeutic strategies. While previous studies have linked the basal ganglia to motor functions (Lim et al., 2014; Sharma et al., 2012), our study further clarifies the correlations between different basal ganglia nuclei and motor functions of the vocal organs. Ackermann (Ackermann et al., 2014) also identified the role of the basal ganglia in executing motor programs within the feedforward system, influencing the velocity and amplitude of oral movements. Therefore, we hypothesize that, within the DIVA model, different basal ganglia nuclei contribute to varying degrees to the motor encoding and activation of each vocal organ, with the coordinated interaction between these nuclei and the thalamus playing a crucial role in the feedforward phase of speech production, including phonetic encoding, motor execution, and vocal control.

Interestingly, the present results seem to imply the presence of two distinct basal ganglia systems, or the basal ganglia is involved in two very different functions of speech production: one involved in linguistic processing and the other in articulation stage (Middleton and Strick, 2000). The right nuclei seem to primarily affect processes in motor control while the left basal ganglia are significantly associated with the linguistic aspects of speech production, such as word retrieval and selection, as well as encoding. Furthermore, the effects of different nuclei of the basal ganglia appear to be quite distinct. Prior research has identified deficits in specific aspects of speech production, including phonological decoding, lexical access, and morphological processing in individuals with extensive bilateral damage to the caudate nucleus (Pickett et al., 1998). The present findings further demonstrate that the less pronounced is the effect of damage to the left caudate nucleus during all the behavioral tasks. Specific damage to the left caudate nucleus has been linked to perseverative

speech errors (Kreisler et al., 2000). In addition, in line with Robles (2005), electrical stimulation to the caudate nucleus during a picture naming task can induce perseveration, characterized by the repetition of the previous word. Damage to the left putamen also significantly influences functional processing in speech production, albeit to a lesser degree. Groenholm et al. (2015) demonstrated that both the caudate nucleus and the adjacent putamen played a role in regulating the sequencing of articulation patterns for speech sounds. Moreover, the left globus pallidus was significantly associated with the early stages of the SF task. It is hypothesized that left nuclei may primarily affect the lower-level processing stages, such as phonological retrieval and encoding, with its impact on higher-level processes potentially being an indirect result of cognitive decline, which is different from the direct cognitive impairment observed in PD patients (Lieberman et al., 1992). These insights underscore the complex and interconnected roles of the basal ganglia in the intricate processes of speech production, with lateralization and hierarchical organization being the key factors in their contributions to these functions. To achieve early diagnosis and assessment of dysarthria, it is crucial to gain a thorough understanding of the subcortical neural mechanism and its corresponding pathological features.

The thalamus is an important transmission hub for information that connects cortical regions and subcortical nuclei (Bostan and Strick, 2018; Duffy et al., 2012). The relationship between the thalamus and various behavioral tasks is notably pronounced in the left thalamus. Research indicated that thalamic damage or dysfunction often lead to speech production impairments, including articulation difficulties, fluency issues, and deficits in language comprehension (Ehlen et al., 2016). Some studies also reported that the interplay between the basal ganglia and thalamus is crucial for the initiation of movement (DeLong, 1990).

Findings from behavioral tasks suggest that, as task complexity increases, the impact of the left thalamus on these tasks gradually decreases. This implies that thalamic damage could potentially lead to a deterioration in linguistic function, affecting higher-level processing such as conceptual preparation and word retrieval. Conversely, the right thalamus has a more significant influence over the control of vocal organs, including respiration and lip movements, compared to the left thalamus. Further evidence is needed to support the role of the thalamus and basal ganglia and their connection in synchronizing the timing and sequencing of speech movements, ensuring fluent and harmonious vocal expression (Ehlen et al., 2016; Schirmer, 2004).

Previous studies failed to fully elucidate the distinct effects of basal ganglia damage on cognition and linguistic processing, or investigate speech performance after accounting for the influence of cognitive factors. The present study revealed that cognitive impairments significantly impact only vowel duration and formant variability, as detailed in [Supplementary Table 1](#). This leads to the hypothesis that dysarthria following brain injury primarily results in a reduction in the motoric capabilities of the vocal apparatus, while the associated cognitive deficits may further affect the stability and rhythm of speech. As task difficulty increased, the significance of the differences observed gradually declined across the behavioral tasks ([Figure 4B](#)). This aligns with the prior findings of [Duffy et al. \(2012\)](#). Cognitive capacity, as measured by the MoCA score, may predominantly influence the earlier stages of speech production, such as conceptual preparation and lexical selection. The divergent outcomes between dysarthria and cognitive impairment may also explain the inconsistent findings regarding characteristics of dysarthric speech previously reported. For example, dysarthria stemming from neurodegenerative conditions such as Parkinson's disease presents very different speech manifestations compared to post-stroke dysarthria ([Silveri, 2021](#)). Such discrepancies may arise because many studies did not consider cognition as a distinct variable, failing to clearly differentiate the effects of cognitive function (MoCA score) from those of dysarthria (FDA score), which could also confuse clinical treatment protocols for speech impairments following PD or stroke. This study is the first to integrate speech, behavioral, and neuroimaging data to explore the detailed pathological manifestations of speech production following basal ganglia or thalamic lesions. These findings may have significant implications for the development of multimodal physiological markers for the intelligent diagnosis and assessment of dysarthria.

However, several limitations are noted in the present study. First, the sample size of stroke patients is relatively small. The collection of comprehensive multimodal pathological data is challenging due to various factors, including the specific location of lesions, post-stroke complications, patient conditions, and cognitive abilities. Moreover, patients with isolated lesions are exceptionally rare and valuable. We believe that the multimodal pathological data and findings from this study, involving subacute stroke patients with basal ganglia and thalamic lesions, are highly significant for advancing the understanding of subcortical neural mechanisms underlying linguistic processing and articulatory movement. Second, the relationship between the basal ganglia, thalamus, and speech

production is highly intricate, involving complex structural connectivity and functional interactions with other brain regions critical for speech processing. Future investigations will focus on subcortical structural and functional networks to provide deeper insights into the interplay between the basal ganglia, thalamus, and associated regions, as well as their distinct impacts on speech production. These issues can also provide valuable guidance for future research directions. In conclusion, this study investigated the direct impact of basal ganglia and thalamic lesions on linguistic processing and articulatory movement, contributing to a deeper understanding of the subcortical neural mechanisms underlying dysarthria. Furthermore, multimodal data encompassing neuroimaging, behavioral, and acoustic features can serve as effective indicators for developing objective evaluation methods, thereby facilitating the early diagnosis and precise intervention of speech impairments.

5 Conclusion

The present study utilized multi-modal mapping analyses to provide a thorough examination of the specific impacts of damage to the basal ganglia and thalamus to the various stages of speech production. It significantly advanced our understanding of the neural mechanisms associated with dysarthria. The present innovative approach not only advances our understanding of the intricate relationship between brain damage and speech production, but also offers a more nuanced perspective on the role of the basal ganglia and thalamus in this complex process. By focusing on specific dysarthria patients and integrating multiple types of data, the present findings shed light on the multifaceted nature of dysarthria and its underlying mechanisms. The research has also offered detailed insights into the functional, behavioral, and speech-related manifestations of dysarthria, which are instrumental for precise clinical diagnosis and the development of targeted treatment strategies.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Institutional Research Ethics Committee of Shenzhen Institute of Advanced Technology and the Eighth Affiliated Hospital of Sun Yat-sen University (ethics code: SIAT-IRB-220415-H0598). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JL: Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. RR: Methodology,

Writing – review & editing. CJ: Methodology, Writing – original draft. QW: Methodology, Writing – review & editing. ShuZ: Methodology, Writing – review & editing. MN: Methodology, Writing – review & editing. ShaZ: Methodology, Writing – review & editing. LW: Writing – review & editing. NY: Investigation, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by National Natural Science Foundation of China (U23B2018, NSFC 62271477, and 82302881), Shenzhen Science and Technology Program (JCYJ20220818101411025, JCYJ20220818101217037, and JCYJ20220818102800001), Fundamental Research (Key Program) of Shenzhen (JCYJ20220818102016034), Shenzhen Peacock Team Project (KQTD20200820113106007), and Futian Healthcare Research Project (FTWS2022015).

References

- Ackermann, H., Hage, S. R., and Ziegler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* 37, 529–546. doi: 10.1017/S0140525X13003099
- Ali, M., Lyden, P., and Brady, M. (2015). Aphasia and dysarthria in acute stroke: recovery and functional outcome. *Int. J. Stroke* 10, 400–406. doi: 10.1111/ijls.12067
- Babjack, D. L., Cernicky, B., Sobotka, A. J., Basler, L., Struthers, D., Kisic, R., et al. (2015). Reducing audio stimulus presentation latencies across studies, laboratories, and hardware and operating system configurations. *Behav. Res. Methods* 47, 649–665. doi: 10.3758/s13428-015-0608-x
- Banks, R., Arjmandi, M. K., and Ghasemzadeh, H. (2019). “Formant dynamics in vowel pronunciation as acoustic correlates of concussed speech,” in *International Brain Injury Association*. Available at: https://www.researchgate.net/publication/331983220_Formant_Dynamics_in_Vowel_Pronunciation_as_Acoustic_Correlates_of_Concussed_Speech
- Bates, D. (2010). *lme4: Mixed-Effects Modeling with R*. Springer.
- Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., et al. (2003). Voxel-based lesion-symptom mapping. *Nat. Neurosci.* 6, 448–450. doi: 10.1038/nm1050
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott. Int.* 5, 341–345.
- Bostan, A. C. and Strick, P. L. (2018). The basal ganglia and the cerebellum: nodes in an integrated network. *Nat. Rev. Neurosci.* 19, 338–350. doi: 10.1038/s41583-018-0002-7
- Brown, P. and Marsden, C. (1998). What do the basal ganglia do? *Lancet* 351, 1801–1804. doi: 10.1016/S0140-6736(97)11225-9
- Camerino, I., Ferreira, J., Vonk, J. M., Kessels, R. P., de Leeuw, F.-E., Roelofs, A., et al. (2024). Systematic review and meta-analyses of word production abilities in dysfunction of the basal ganglia: stroke, small vessel disease, parkinson’s disease, and huntington’s disease. *Neuropsychol. Rev.* 34, 1–26. doi: 10.1007/s11065-022-09570-3
- Chen, Y., Zhu, G., Liu, D., Liu, Y., Yuan, T., Zhang, X., et al. (2020). Brain morphological changes in hypokinetic dysarthria of parkinson’s disease and use of machine learning to predict severity. *CNS Neurosci. Therapeut.* 26, 711–719. doi: 10.1111/cns.13304
- Chow, H. M., Garnett, E. O., Ratner, N. B., and Chang, S.-E. (2023). Brain activity during the preparation and production of spontaneous speech in children with persistent stuttering. *NeuroImage: Clin.* 38:103413. doi: 10.1016/j.nicl.2023.103413
- Darley, F. L., Aronson, A. E., and Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *J. Speech Hear. Res.* 12, 246–269. doi: 10.1044/jshr.1202.246
- DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* 13, 281–285. doi: 10.1016/0166-2236(90)90110-V
- Duffy, J. R. (2012). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*. St. Louis, MO: Elsevier Health Sciences.
- Ehlen, F., Vonberg, I., Kuehn, A. A., and Klostermann, F. (2016). Effects of thalamic deep brain stimulation on spontaneous language production. *Neuropsychologia* 89, 74–82. doi: 10.1016/j.neuropsychologia.2016.05.028
- Enderby, P. (1980). Frenchay dysarthria assessment. *Br. J. Disord. Commun.* 15, 165–173. doi: 10.3109/13682828009112541
- Enderby, P. (2013). Disorders of communication: dysarthria. *Handb. Clin. Neurol.* 110, 273–281. doi: 10.1016/B978-0-444-52901-5.00022-8
- Federico, N., Sabatini, U., Rascol, O., and Péran P. (2015). Parkinson’s disease and local atrophy in subcortical nuclei: insight from shape analysis. *Neurobiol. Aging* 36(1):424–433. doi: 10.1016/j.neurobiolaging.2014.07.010
- Fox, J., Kraemer, M., Schormann, T., Dabringhaus, A., Hirsch, J., Eisele, P., et al. (2016). Individual assessment of brain tissue changes in ms and the effect of focal lesions on short-term focal atrophy development in MS: a voxel-guided morphometry study. *Int. J. Mol. Sci.* 17:489. doi: 10.3390/ijms17040489
- Gagnon, M., Barrette, J., and Macoir, J. (2018). Language disorders in huntington disease: a systematic literature review. *Cognit. Behav. Neurol.* 31, 179–192. doi: 10.1097/WNN.0000000000000171
- Ghio, A., Giusti, L., Blanc, E., and Pinto, S. (2019). French adaptation of the “frenchay dysarthria assessment 2” speech intelligibility test. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* 137:2. doi: 10.1016/j.anorl.2019.10.007
- Graybiel, A. M. (2000). The basal ganglia. *Curr. Biol.* 10, R509–R511. doi: 10.1016/S0960-9822(00)00593-5
- Groenholm, E. O., Roll, M. C., Horne, M. A., and Lindgren, A. G. (2015). *Predominance of caudate nucleus lesions in acute ischaemic stroke patients with impairment in language and speech.* *Eur. J. Neurol.* 23, 148–153. doi: 10.1111/ene.12822
- Guenther, F. H. (2007). “Neuroimaging of normal speech production,” in *Neuroimaging in Communication Sciences and Disorders* (San Diego: Plural Publishing Inc), 1–51.
- Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., and Bi, Y. (2013). White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. *Brain* 136, 2952–2965. doi: 10.1093/brain/awt205
- Hernandez, A., Yeo, E. J., Kim, S., and Chung, M. (2020). “Dysarthria detection and severity assessment using rhythm-based metrics,” in *Interspeech*, 2897–2901. Available at: https://www.isca-archive.org/interspeech_2020/hernandez20_interspeech.pdf

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2024.1455085/full#supplementary-material>

- Indefrey, P. and Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001
- Johnson, M. D. and Ojemann, G. A. (2000). The role of the human thalamus in language and memory: evidence from electrophysiological studies. *Brain Cogn.* 42, 218–230. doi: 10.1006/brcg.1999.1101
- Karan, B., Sahu, S. S., and Mahto, K. (2020). Parkinson disease prediction using intrinsic mode function based features from speech signal. *Biocybernet. Biomed. Eng.* 40, 249–264. doi: 10.1016/j.bbe.2019.05.005
- Kearney, E. and Guenther, F. H. (2019). Articulating: the neural mechanisms of speech production. *Lang. Cogn. Neurosci.* 34, 1214–1229. doi: 10.1080/23273798.2019.1589541
- Kempler, D. and Van Lancker, D. (2002). Effect of speech task on intelligibility in dysarthria: a case study of parkinson's disease. *Brain Lang.* 80, 449–464. doi: 10.1006/brln.2001.2602
- Kim, H., Hasegawa-Johnson, M., Perlman, A., Gunderson, J. R., Huang, T. S., Watkin, K. L., et al. (2008). Dysarthric speech database for universal access research. *Interspeech* 2008, 1741–1744. doi: 10.21437/Interspeech.2008-480
- Koyuncu, E., Çam, P., Altınok, N., Çallı, D. E., Duman, T. Y., and Özgür, N. (2016). Speech and language therapy for aphasia following subacute stroke. *Neural Regenerat. Res.* 11, 1591–1594. doi: 10.4103/1673-5374.193237
- Kreisler, A., Godefroy, O., Delmaire, C., Debachy, B., Leclercq, M., Pruvo, J. P., et al. (2000). The anatomy of aphasia revisited. *Neurology* 54, 1117–1123. doi: 10.1212/WNL.54.5.1117
- Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition* 42, 1–22. doi: 10.1016/0010-0277(92)90038-J
- Levelt, W. J. (2000). "Speech production," in *Encyclopedia of Psychology* (Oxford: Oxford University Press), 432–433.
- Levelt, W. J. and Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *Eur. J. Cognit. Psychol.* 12, 433–452. doi: 10.1080/095414400750050178
- Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. doi: 10.1017/S0140525X99451775
- Li, M., Zhang, Y., Song, L., Huang, R., Ding, J., Fang, Y., et al. (2017). Structural connectivity subserving verbal fluency revealed by lesion-behavior mapping in stroke patients. *Neuropsychologia* 101, 85–96. doi: 10.1016/j.neuropsychologia.2017.05.008
- Lieberman, A. M. and Whalen, D. H. (1975). On the relation of speech to language. *ASHA* 4, 187–196. doi: 10.1016/S1364-6613(00)01471-6
- Lieberman, P., Kako, E., Friedman, J., Tajchman, G., and Jimenez, E. B. (1992). Speech production, syntax comprehension, and cognitive deficits in parkinson's disease. *Brain Lang.* 43:169–189. doi: 10.1016/0093-934X(92)90127-Z
- Lim, S.-J., Fiez, J. A., and Holt, L. L. (2014). How may the basal ganglia contribute to auditory categorization and speech perception? *Front. Neurosci.* 8:230. doi: 10.3389/fnins.2014.00230
- Lin, F., Cheng, S.-Q., Qi, D.-Q., Jiang, Y.-E., Lyu, Q.-Q., Zhong, L.-J., et al. (2020). Brain hothubs and dark functional networks: correlation analysis between amplitude and connectivity for broca's aphasia. *PeerJ* 8:e10057. doi: 10.7717/peerj.10057
- Liu, J., Du, X., Lu, S., Zhang, Y.-M., An-ming, H., Ng, M. L., et al. (2023). Audio-video database from subacute stroke patients for dysarthric speech intelligence assessment and preliminary analysis. *Biomed. Signal Process. Control* 79:104161. doi: 10.1016/j.bspc.2022.104161
- Liu, Y., Hao, M., Li, P., and Shu, H. (2011). Timed picture naming norms for mandarin chinese. *PLoS ONE* 6:e16505. doi: 10.1371/journal.pone.0016505
- Mackenzie, C. (2011). Dysarthria in stroke: a narrative review of its description and the outcome of intervention. *Int. J. Speech Lang. Pathol.* 13, 125–136. doi: 10.3109/17549507.2011.524940
- Magee, M., Copland, D., and Vogel, A. P. (2019). Motor speech and non-motor language endophenotypes of parkinson's disease. *Expert Rev. Neurother.* 19, 1191–1200. doi: 10.1080/14737175.2019.1649142
- Meyer, S., Kessner, S. S., Cheng, B., Bönstrup, M., Schulz, R., Hummel, F. C., et al. (2016). Voxel-based lesion-symptom mapping of stroke lesions underlying somatosensory deficits. *NeuroImage: Clin.* 10, 257–266. doi: 10.1016/j.nicl.2015.12.005
- Middleton, F. A., and Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Res. Rev.* 31, 236–250. doi: 10.1016/S0165-0173(99)00040-5
- Monsell, S., Taylor, T. J., and Murphy, K. (2001). Naming the color of a word: Is it responses or task sets that compete? *Memory Cognit.* 29, 137–151. doi: 10.3758/BF03195748
- Mou, Z., Chen, Z., Yang, J., and Xu, L. (2018). Acoustic properties of vowel production in mandarin-speaking patients with post-stroke dysarthria. *Sci. Rep.* 8:14188. doi: 10.1038/s41598-018-32429-8
- Nambu, A. (2008). Seven problems on the basal ganglia. *Curr. Opin. Neurobiol.* 18, 595–604. doi: 10.1016/j.conb.2008.11.001
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x
- Paja, M. O. S. and Falk, T. H. (2012). "Automated dysarthria severity classification for improved objective intelligibility assessment of spastic dysarthric speech," in *Interspeech*, 62–65. Available at: https://www.isca-archive.org/interspeech_2012/paja12_interspeech.pdf
- Pennington, B. F. and Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annu. Rev. Psychol.* 60, 283–306. doi: 10.1146/annurev.psych.60.110707.163548
- Pickett, E. R., Kuniholm, E., Protopoulos, A., Friedman, J., and Lieberman, P. (1998). Selective speech motor, syntax and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: a case study. *Neuropsychologia* 36, 173–188. doi: 10.1016/S0028-3932(97)00065-1
- Polikoff, J. B. and Bunnell, H. T. (1999). "The nemours database of dysarthric speech: a perceptual analysis," in *XIVth International Congress of Phonetic Sciences* (Princeton: CiteSeer), 783–786.
- Rampello, L., Rampello, L., Patti, F., and Zappia, M. (2016). When the word doesn't come out: a synthetic overview of dysarthria. *J. Neurol. Sci.* 369, 354–360. doi: 10.1016/j.jns.2016.08.048
- Robles, G. S. (2005). The role of dominant striatum in language: a study using intraoperative electrical stimulations. *J. Neurol. Neurosurg. Psychiatr.* 76, 940–946. doi: 10.1136/jnnp.2004.045948
- Sapir, S., Ramig, L. O., Spielman, J. L., and Fox, C. (2010). *Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech.* *J. Speech Lang. Hear. Res.* 53, 114–125. doi: 10.1044/1092-4388(2009)08-0184
- Schirmer, A. (2004). Timing speech: a review of lesion and neuroimaging findings. *Cognitive Brain Res.* 21, 269–287. doi: 10.1016/j.cogbrainres.2004.04.003
- Sharma, K. R., Sheriff, S., Maudsley, A., and Govind, V. (2012). Diffusion tensor imaging of basal ganglia and thalamus in amyotrophic lateral sclerosis. *J. Neuroimaging* 23, 368–374. doi: 10.1111/j.1552-6569.2011.00679.x
- Sheng-nan, G., Yong-li, W., Min-min, Y., Qin, W., Hong, Q., and Zhao-ming, H. (2021). Relationship between vowel production and speech clarity in post-stroke dysarthria patients. *Chin. J. Rehabil. Theory Pract.* 27, 43–47.
- Silveri, M. C. (2021). Contribution of the cerebellum and the basal ganglia to language production: speech, word fluency, and sentence construction—evidence from pathology. *Cerebellum* 20, 282–294. doi: 10.1007/s12311-020-01207-6
- Spencer, K. A. and Brown, K. A. (2018). Dysarthria following stroke. *Semin. Speech Lang.* 39, 015–024. doi: 10.1055/s-0037-1608852
- Sperber, C. and Karnath, H.-O. (2016). Topography of acute stroke in a sample of 439 right brain damaged patients. *NeuroImage: Clin.* 10, 124–128. doi: 10.1016/j.nicl.2015.11.012
- Summaka, M., Hannoun, S., Harati, H., Daoud, R., Zein, H., Estephan, E., et al. (2022). Neuroanatomical regions associated with non-progressive dysarthria post-stroke: a systematic review. *BMC Neurol.* 22, 1–11. doi: 10.1186/s12883-022-02877-x
- Tang, X., Chen, N., Zhang, S., Jones, J. A., Zhang, B., Li, J., et al. (2018). Predicting auditory feedback control of speech production from subregional shape of subcortical structures. *Hum. Brain Mapp.* 39, 459–471. doi: 10.1002/hbm.23855
- Tourville, J. A. and Guenther, F. H. (2011). The diva model: A neural theory of speech acquisition and production. *Lang. Cogn. Process.* 26, 952–981. doi: 10.1080/01690960903498424
- Wallesch, C. W., Kornhuber, H. H., Brunner, R. J., Kunz, T., and Suger, G. (1983). Lesions of the basal ganglia, thalamus, and deep white matter: differential effects on language functions. *Brain Lang.* 20, 286–304. doi: 10.1016/0093-934X(83)90046-9
- Wang, T., Yan, Q., Pan, J., Zhu, F., Su, R., Guo, Y., et al. (2019). "Towards the speech features of early-stage dementia: Design and application of the mandarin elderly cognitive speech database," in *Interspeech*, 4529–4533. Available at: https://www.isca-archive.org/interspeech_2019/wang19q_interspeech.pdf
- Weisstein, E. W. (2004). *Bonferroni Correction*[J]. Available at: <https://mathworld.wolfram.com/>
- Wilson, S. M., Isenberg, A. L., and Hickok, G. (2009). Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Hum. Brain Mapp.* 30, 3596–3608. doi: 10.1002/hbm.20782
- Zhao, Y., Song, L., Ding, J., Lin, N., Wang, Q., Du, X., et al. (2017). Left anterior temporal lobe and bilateral anterior cingulate cortex are semantic hub regions: evidence from behavior-nodal degree mapping in brain-damaged patients. *J. Neurosci.* 37, 141–151. doi: 10.1523/JNEUROSCI.1946-16.2017
- Ziegler, W., and Ackermann, H. (2017). Subcortical contributions to motor speech: phylogenetic, developmental, clinical. *Trends Neurosci.* 40:458. doi: 10.1016/j.tins.2017.06.005