

Effect of Folic Acid Therapy on the Acoustic Behaviour Analysis of Dysarthria in Subacute Stroke

YANBO WANG*, JUAN LIU AND JIANGUO WEI¹

College of Intelligence and Computing, Tianjin University, Yaguan, Tianjin 300354, ¹Chinese Academy of Sciences (CAS) Key Laboratory of Human Machine Intelligence Synergy Systems, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

Wang *et al.*: Acoustic-Behaviour Analysis of Subacute Stroke with Dysarthria

Speech production involves highly complex coordinated articulatory movements and is modulated by the cerebral cortex and neurobehavioral processes. Impairments to different structures in the nervous system lead to differing pathological speech and behavioural manifestations of dysarthria. However, the relationship between language cognitive ability and speech quality remains unclear. The current study explored significant speech and behavioural symptoms of dysarthria and revealed the relationships between acoustic features and speech-cognitive behaviour parameters. Audio and speech behavioural data were collected from 20 subacute patients with subacute dysarthria and 22 healthy controls. A series of speech behavioural tasks were designed to quantify functional speech production ability effectively. The experimental group of patients received folic acid treatment, with an initial dose of 4 mg twice daily, adjusted up to a maximum of 8 mg twice daily based on response and tolerance. Multidimensional acoustic-behaviour features were then extracted and analysed. Finally, significant abnormalities and correlations were found between acoustic (e.g., vowel space area, mel frequency cepstral coefficients) and behavioural features (e.g., reaction time). These results verify that dysarthria is affected not only by the motor control processes but also modulated by speech behaviour levels. In conclusion, the present study can contribute to understanding the speech-behavioural mechanisms of dysarthria and suggest the influence of subacute stroke on speech production. It has important implications for future refined diagnosis and targeted treatment of post-stroke speech impairments.

Key words: Pharmacological treatment, speech impairment, acoustic-behaviour analysis, dysarthria, speech cognition, stroke, rehabilitation

Dysarthria is common post-stroke sequelae, resulting in impaired strength, speed, range, stability, tone, or accuracy of movements required for respiration, phonation, resonance, articulation, or prosody during speech production. Dysarthria speech displays slow, effortful utterance, hypernasality, and reduced intelligibility^[1]. Communication disorders severely impact patients' daily lives, work, and social interactions, often leading to social difficulties, isolation, and depression. They also affect patients' rehabilitation process^[2]. Therefore, in recent years, the study of the pathogenesis, of dysarthria has received increasing international attention.

However, current dysarthria research predominantly focuses on acoustic characteristics, the neurological speech disorder of dysarthria manifests in degraded speech acoustics. such as a speaker's overall intelligibility in dysarthric speech is represented well

by the overlap degree among vowels, result from speakers with dysarthria (Parkinson's disease and stroke) have shown reduced second-Formant (F2) slopes is when compared to healthy controls^[3], and often, impaired intelligibility. Speech intelligibility is related to the magnitude of multiple acoustic feature, such as articulatory precision, vowel duration, F2 Vowel Inherent Spectral Change (VISC), etc.,^[4]. With few studies on the impact of language processing in the brain on the final acoustic expression, the function of speech production is a complex cognitive process. Scientists have proposed neural computational models to study the neural mechanism of the speech production processing. The word production and Directions into Velocities of Articulators (DIVA) models are most accepted^[5]. The DIVA model establishes a mathematical and neurological computational framework for speech

*Address for correspondence
E-mail: 13132250817@163.com

production. The lexical model establishes a linguistics and neuropsychology-based framework for the speech production process, involving conceptual preparation, lexical selection, phonological encoding, phonetic encoding, articulation, and auditory feedback stages^[6]. Here, the conceptual preparation stage transforms information into lexical concepts, maximally activating the target concept and lexical items. The activated lexical then spread activation to the phonological encoding stage, where activated phonological encoding and connections constitute the output. The lexical model is the current mainstream model of speech production behaviour, with most related research based on this theoretical model. For example, Steurer *et al.*^[7] demonstrate the relationship between Parkinson's disease-induced dysarthria, language fluency, speech execution, and overall cognitive function, along with associated brain structural changes, elucidating the cerebrum's important role in speech processing and lexical representation. Based on DIVA model assumptions, MIT groups investigated cerebellar and cerebral activity biases during mono and disyllabic utterances, respectively^[8]. They found left-lateralized premotor cortex activity area contains cell populations representing syllable motor programs, and compared with vowels, the superior preverbal cerebellum is more active for consonant-vowel syllables.

The above studies focus more on links between local brain changes and specific utterance. However, there has been limited research on the correlation between the speech motor control process and the cognitive processing mechanisms, with their interaction yet to be clearly elucidated. Hennessey *et al.*^[9] investigated the relationship between speech motor features and behaviours in children's Picture Naming (PN) tasks, discovering that language behaviour ability development coincides with laryngeal speech production system maturation that they somehow inform communication loss in adults with dysarthria post-stroke, which motivates the choice of acoustic variables and the cognitive-linguistic tasks of the current study explored the role of lexical information in initiating naming responses in word naming tasks, along with suprasegmentally factors in naming. They documented that the naming task durations are more affected by pre-speech production variables, indicating some continuity in the transfer processes of speech motor control. Schmitt *et al.*^[10] implied that there was an initial lexical selection stage in speech production followed by phonetic encoding, where

only selected items undergo phonological encoding.

Few studies investigated that the correlation between the speech motor control process and the cognitive processing mechanisms on dysarthria of subacute stroke, with their interaction yet to be clearly elucidated. Using picture-naming tasks, Walker *et al.*^[11] investigated the functional cortical activities' relationships between language cognition and lexical variables in stroke patients. They found significant linear relationships between lexical attributes and overall difficulty across items. Word frequency uniquely contributes to all potential processing decisions, where phonological length and density significantly impact ultimate phonological selection and production. Based on these results, the authors built a more complex phonological production model.

The current study aimed to investigate the pathogenesis of cognitive impairment and pathological speech in post-stroke dysarthria using joint analysis of the pathological acoustic features combining four classic behavioural experiments involving different levels of speech processing. The present study can fill the gap of the traditional dysarthria acoustic research that does not consider higher-level cognitive processing impacts on speech output to explore correlations between pathological speech and speech behaviours. The study intends to verify speech cognitive processing effects mechanism on final speech production underlying post-stroke dysarthria, which is significant for enabling refined diagnosis and targeted treatment of dysarthria patients.

MATERIALS AND METHODS

General information:

The experimental subjects included 20 subacute stroke patients with dysarthria (16 males and 4 females) and 22 age-matched healthy controls (16 males and 6 females). All were recruited based on the following inclusion and exclusion criteria.

Inclusion criteria: Every participant must be above 18 y of age; they must have received at least primary school education; dysarthria patients must be diagnosed with stroke within 6 mo (no history of prior stroke) and they had no prior professional treatment in speech or language therapy.

Exclusion criteria: Vision or hearing impairment; dementia and psychiatric disorder; other neurologic diseases not related to the subacute stroke. The study was approved by the Ethics Committees of the Shenzhen

Institutes of Advanced Technology and the Eighth Affiliated Hospital of Sun Yat-sen University. Written informed consent was obtained from all participants.

Treatment protocol:

Grouping: Participants were divided into two groups; control group and an experimental group. The control group did not receive any pharmacological treatment. The experimental group received folic acid treatment with the following dosing regimen; initial dose was 4 mg of folic acid twice daily. Maintenance dose was adjusted based on patient response and tolerance, up to a maximum of 8 mg twice daily.

All patients received comprehensive clinical evaluations including Frenchay Dysarthria Assessment (FDA) provided by professional Speech Language Therapists (SLTs) and Montreal Cognitive Assessment (MoCA) by neurologists. Cognitive abilities were assessed based on MoCA scores; dysarthria was assessed using the current most widely used scale, the FDA. The FDA

measures 28 speech-related audio-visual dimensions across eight categories; reflex, respiration, lips, mandible, larynx, tongue, soft palate, and speech. Each dimension receives a score from 0 (no function) to 4 (normal function). Dysarthria severity is determined by the percentage of dimensions scored as standard out of 28 by SLTs. Clinical assessment results and other demographic information was shown in Table 1.

Audio data collection: Speech data were recorded in sound-proof rooms at the Eighth Affiliated Hospital of Sun Yat-sen University and Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. The equipment included a Dell laptop for displaying prompts, text, and picture materials and a Taskstar MS400 microphone recording audio at a 16 kHz sampling rate (16-bit, single channel). Participants sat with lips (8-15) cm from the microphone, directly facing the computer screen. One trained experimenter accompanied participants throughout the collection process (fig. 1).

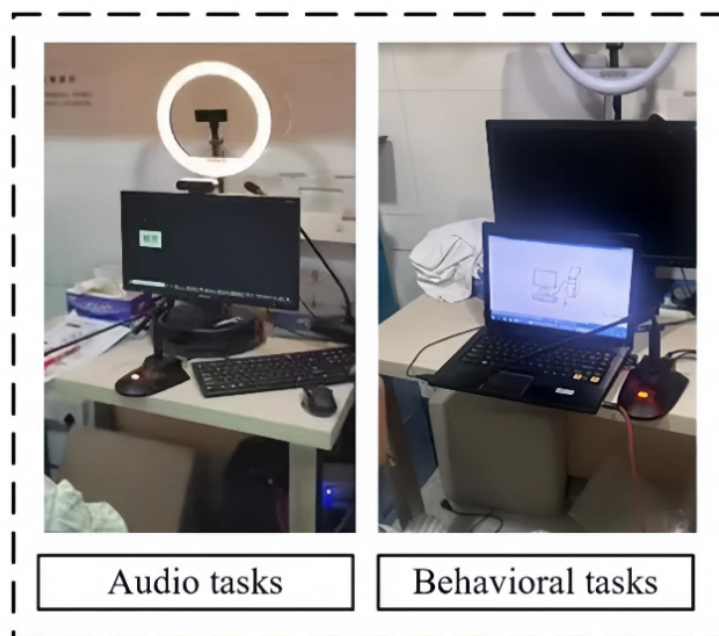


Fig. 1: Experimental collection environment

TABLE 1: DEMOGRAPHIC INFORMATION AND CLINICAL EVALUATION RESULTS OF ALL SUBJECTS

	Patient	Normal	Statistic	p
Gender	16:4	16:6	$\chi^2=0.03$	0.85
Age	59.65 (14.02)	60.68 (8.79)	t=0.28	0.78
Education (years)	12.95 (3.75)	12.82 (3.76)	t=0.74	0.46
MoCA (30)	20.90 (5.22)	27.48 (2.09)	t=5.26**	<0.001
Frenchay index (0-4)	2.87 (0.91)	4	t=5.38**	<0.001

Note: **p<0.01, indicates extremely significant difference

Speech materials were designed according to Chinese phonetic characteristics and different impairment profiles in dysarthria. Considering typical dysarthric pathological characteristics of vowel errors and consonant distortions, tasks with varying difficulties were designed, including syllables, characters, words, and short sentences. The corpus comprised five tasks: Syllables, characters, words, sentences, and self-introduction. The details of each task and the data collection can be found in the previous studies.

Cognitive behavioural data experiments: Cognitive ability in speech production was assessed through four classic behavioural experiments involving conceptual preparation, lexical selection, phonetic encoding/retrieval, and articulation at varying difficulties. The four experiments combined Speech Fluency (SF) tasks, Picture Association (PA) tasks^[12], PN tasks^[13], and Color Naming (CN) tasks to determine the different stages of the cognitive processing across language production^[14].

SF task: The task examines verbal functioning and SF. Participants were asked to produce as many words as possible within 1 min for each of the 10 semantic categories, such as festivals, cities, and fruits.

PA task: The task followed Lin *et al.*^[12] format. Dozens of black and white outline drawings were presented randomly. After a beep, a picture was shown while participants performed overt naming.

PN task: The task followed Liu *et al.*^[13] format. 100 items were selected from all the 435 objects, balanced for concept familiarity, subjective frequency, image agreement, variability, and complexity. Participants were instructed to quickly and accurately name each picture. According to the word length, divide the tasks into PN1, PN2, and PN3 (the number represents the word length of stimulus material).

CN task: The task followed the PA and PN formats. Dozens of familiar objects with noticeable colour characteristic (e.g., cloud and tomato) were selected. Participants were asked to quickly and accurately name the colour of each item.

Each behaviour language task included a number of common objects and was executed in separate sessions using the E-prime program and Chrono's system. Item presentation order was randomized within each task but identical across subjects. The response deadline was 1 min for the SF task and 6 s for the other three tasks. Each test took no more than 1 h, with breaks as needed. The behavioural experiment procedure was shown in fig. 2.

Data pre-processing:

Speech data annotation: Recorded audio files were segmented and transcribed verbatim by two experienced researchers using TextGrid in Praat. Unrelated utterances were deleted, such as questions about the task or conversations with the examiner. All relevant speech materials were manually transcribed on the first tier using Chinese characters, and the contents of each subject's speech were transcribed on the second tier in the form of Chinese pinyin, a pronunciation system for Chinese. Character and syllable tasks were manually divided into vowel and consonant segments and marked on the third tier according to their spectrograms and auditory judgments by two linguists, in a way similar to Liu *et al.*^[13].

Behavioural data annotation: To accurately calculate the number of practical answers and reaction time of each behavioural task, all behavioural audio data has also been manually annotated. For the SF task, all correct answers were annotated in 1 min, and the number was counted in each 15 s time window. For the PA, PN, and CN tasks, each beep voice and effective answer were also labelled on the first tier, and the time interval between them was used as the reaction time, as shown in fig. 3.

Feature extraction:

Acoustic feature extraction: Based on vowel steady states, articulators' movements, vowel space, and vowel stability features were extracted^[15]. Poor voice quality arises from abnormal vocal fold vibration. Thus, phonation, stability and voice features were extracted, reflecting glottal/vocal fold control capabilities. Due to dysarthria relate to articulators (tongue, lips) position changes, formant frequencies F1/F2/F3, jaw distance, tongue distance, and articulatory movement change were examined for articulation defects. F1/F2 variability was represented by the standard deviations of values. Vowel Space Area (VSA), Formant Centralization Ratio (FCR), and Vowel Articulation Index (VAI) quantified space and movement range/position changes were also calculated^[16]. Mel Frequency Cepstral Coefficients (MFCC) effectively simulates the human auditory system, reflecting vocal organ issues in dysarthria.

Behavioural feature extraction: Based on the lexical production model containing conceptual preparation, lexical selection, phonological encoding, and articulation stages, behavioural features were extracted. For the SF task, SF_N1, SF_N2, SF_N3, and SF_N4 represented words uttered within (0-15)

s, (15-30) s, (30-45) s, and (45-60) s windows. These reflect reaction time and conceptual understanding. For CN and PA tasks, reaction times (RT-CN, RT-PA) were directly extracted. For PN, reaction times were grouped by noun character counts referring to pictured objects, denoted RT-PN1, RT-PN2, and RT-PN3. Reaction time, accuracy, and other parameters were extracted per subject per stimulus. Analysing these parameters revealed speech production behaviour differences in dysarthria.

Statistical analysis:

Statistical analyses were performed on all acoustic and

behavioural features. Linear mixed-effects regression models^[6] examined each feature to explore behaviour and speech performance factors and determine acoustic and behavioural difference between stroke patients with dysarthria and healthy controls. Models comprised group (dysarthria vs. standard), cognitive level (MoCA score) as fixed effects, and participants as random factors. For all dysarthric patients, Pearson's correlation between acoustic features, behavioural, and FDA scores has also been applied for all dysarthric patients. For all analyses, statistical significance alpha (α) was set to 0.05 with Bonferroni correction for multiple comparisons, implemented in R using the lme4 package.

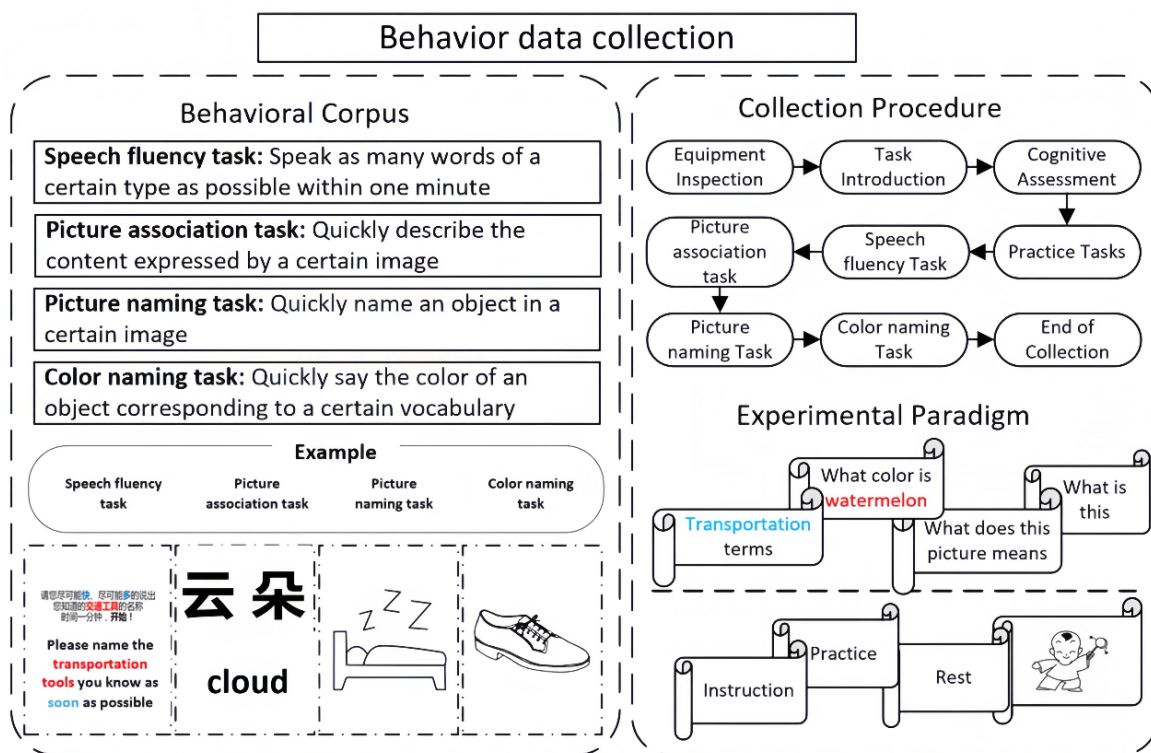


Fig. 2: Speech behaviour data collection programme

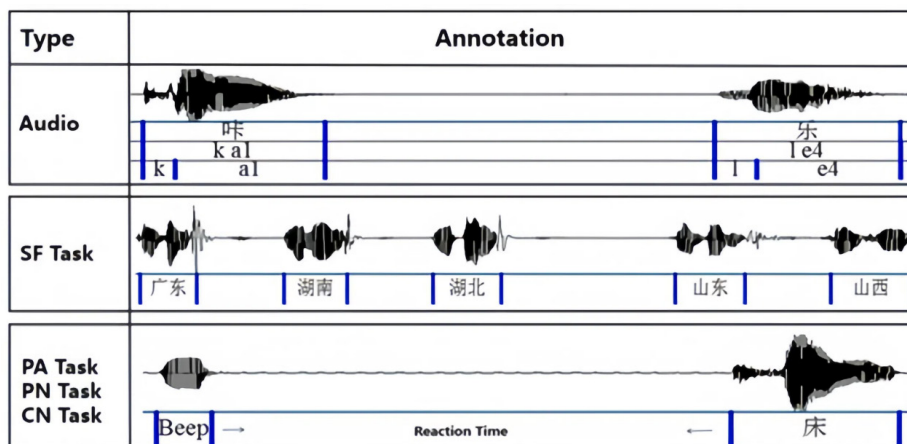


Fig. 3: Schematic diagram of audio and behaviour data annotation

RESULTS AND DISCUSSION

Statistical analysis using regression model revealed significant differences between dysarthric patients and healthy controls (group as the fixed effect we observe in results). In prosody, patients had significantly longer vowel durations ($t=6.81$, $p<0.001$), slightly higher intensity ($t=0.81$, $p<0.001$), and higher F1/2/3 variance vs. controls ($t=-2.74$, $p<0.001$; $t=-8.4$, $p<0.001$; and $t=-6.02$, $p<0.001$). In articulation, patients showed significantly lower jaw distance, tongue distance, and movement degree ($t=1.74$, $p<0.001$; $t=-4.22$, $p<0.001$; and $t=-6.02$, $p<0.001$). In vowel space measures, patients had significantly lower VSA and VAI, and higher vowel centralization (FCR) than controls ($t=4.61$, $p<0.001$; $t=3.14$, $p<0.001$; and $t=-2.92$, $p<0.001$). In phonation, patients had higher mean jitter, shimmer, and harmonicity ($t=0.12$, $p<0.05$; $t=1.39$, $p<0.05$; and $t=2.76$, $p=0.07$). MFCC (0/1/2.12) values were significantly lower in patients ($t=4.61$, $p<0.001$; $t=5.73$, $p<0.001$; $t=11.74$, $p<0.001$; $t=22.44$, $p<0.001$; $t=23.09$, $p<0.001$; $t=7.16$, $p<0.001$; $t=-9.16$, $p<0.001$; $t=-4.59$, $p<0.001$; $t=6.59$, $p<0.001$; $t=10.78$, $p<0.001$; and $t=27.90$, $p<0.001$), potentially due to substitutions and

disruptions altering speech signal spectral distributions, changing MFCC features. The additional movements or efforts patients require to introduce unnecessary noise components into speech reduce MFCC eigenvalues.

Statistical results revealed that dysarthric patients had longer reaction times (RT_CN, RT_PA, RT_PN1, RT_PN2, RT_PN3) vs. healthy controls ($t=11.4$, $p<0.001$; $t=8.89$, $p<0.001$; $t=6.29$, $p<0.001$; $t=5.09$, $p<0.001$; and $t=4.45$, $p<0.001$), and more words (SF_N1, SF_N2, SF_N3, SF_N4) of SF tasks ($t=-9.78$, $p<0.001$; $t=-5.35$, $p<0.001$; $t=-3.78$, $p<0.001$; $t=5.09$, $p<0.001$; and $t=-3.48$, $p<0.001$). Patients took longer on the colour naming vs. PA and PN tasks. Among PN tasks, reaction times from longest to shortest were RT-PN3, RT-PN2, and RT-PN1 for both groups. Patients uttered fewer words per 15 s interval in the semantic vocabulary task than controls, with more significant differences initially that decreased over time.

Pearson analysis examined links between extracted acoustic features and Frenchay subtask scores for reflex, respiration, lip, mandible, soft palate, larynx, tongue, reading, sentence reading, and conversation speed. The results are shown below in Table 2.

TABLE 2: STABILITY PROFILE OF 1 mg/ml BACLOFEN IN X-Temp® ORAL SUSPENSION

R value	Frenchay	Reflex	Respiration	Lips	Jaw	Larynx	Tongue	Word	Sentence	Talk	Speed
Duration	-0.24		-0.4			-0.48*		-0.28	-0.14	-0.15	-0.24
Pitch_mean	-0.29	-0.24	-0.02	-0.37	-0.54*	-0.18	-0.13	-0.08	-0.15	-0.49*	-0.42
Pitch_sd	-0.77**	-0.07	-0.66**	-0.73**	-0.38	-0.60*	-0.73**	-0.54*	-0.48*	-0.51*	-0.66**
Loudness_mean	0.14	0.73*			0.45	0.18				0.56*	0.54*
f1_mean		-0.58*							-0.19	-0.25	-0.15
f1_sd	-0.04	-0.01	-0.05	-0.06		-0.4		-0.33		-0.22	
f2_mean		-0.51*			-0.13	-0.27	0.47*	-0.31	-0.18	-0.44	-0.28
f2_sd	-0.18	-0.11	-0.14		-0.07	-0.35	-0.11	-0.31		-0.2	-0.1
f3_mean		-0.2			-0.32					-0.2	-0.1
f3_sd	-0.45		-0.36	-0.37*	-0.07	-0.47*	-0.43	-0.31		-0.19	-0.04
Jaw_dist	0.12		0.08		0.12		0.45			0.04	
Tongue_dist	0.48*	0.63**	0.33	0.14		0.55*	0.39	0.50*	0.42	0.57*	0.53*
Mov_deg	0.4	0.62**	0.25	0.05		0.48*	0.36	0.47*	0.40*	0.53*	0.46
VSA	0.16	0.39	0.06			0.19	0.32			0.32	0.2
FCR	-0.48*	-0.51*	-0.29*	-0.15	-0.06	-0.54*	-0.44	-0.44	-0.25	-0.56	-0.44
VAI	0.40*	0.54*	0.23	0.04	0	0.49*	0.4	0.4	0.24	0.55*	0.42
Jitter		-0.25					-0.11				
Shimmer		-0.25									-0.04

Harmonicity	-0.44		-0.40*	-0.27	-0.18	-0.41	-0.21	-0.33	-0.25	-0.2	-0.21
MFCC0	0.25	0.16		0.22	0.47		0.32			0.46	0.41
MFCC1		0.25			0.45		-0.29			0.31	0.28
MFCC3		0.45								0.29	0.24
MFCC4		0.41				0.23			0.49*	0.44	0.38
MFCC6			0.55*	0.28		0.47		0.51*	0.47*		0.42
MFCC7	0.60**		0.54**	0.56**		0.61**					0.43
MFCC8			0.59**	0.46			0.35	0.23	0.39		
MFCC9		0.27	0.41	0.41		0.25					

Note: **p<0.01, indicates extremely significant correlation and *p<0.01, indicates significant correlation

In prosody, vowel duration negatively correlated with total Frenchay score and respiration, soft palate, larynx, and speech subtasks. The mean and standard deviation of pitch negatively correlated with all subtasks. Mean intensity positively correlated with reflex, mandible, conversation, and speed subtasks. In phonation, harmonicity negatively correlated with all subtasks, likely due to poorer articulatory coordination in patients disrupting the harmonic structure and weaker, hoarser voices with reduced high-frequency components. Jitter and shimmer negatively correlated with reflex and some articulator subtasks. In articulation, the mean and standard deviation of first/second/third formants negatively correlated with most subtasks. Jaw distance, tongue distance, movement degree, VSA, and VAI negatively correlated with most subtasks, while FCR positively correlated. MFCC positively correlated with breathing, lips, throat and speech subtasks. Patient behavioural task features mildly correlated with Frenchay scores.

Correlations between acoustic and semantic task features were shown in Table 3. Semantic task features SF_N1, SF_N2, SF_N3, and SF_N4 positively correlated with intensity and pitch. In contrast, the number of words uttered positively correlated with F3, tongue distance, articulatory movement degree, and negatively with FCR but positively correlated with VAI, indicating that articulation clarity is positively correlated with conceptual preparation and lexical selection abilities during speech production, while negatively correlating with tongue position.

Correlations between acoustic features and semantic words uttered are more remarkable for the first 15 s (SF_N1) than the last 45 s (SF_N2, SF_N3, and SF_N4). Semantic features are correlated with tongue distance, articulatory movement degree, and VAI. As

shown in Table 3, colour selection, PA and PN reaction times are strongly negatively correlated with pitch, and with F3. Jitter positively correlated with colour naming and PA reaction times. Harmonicity negatively correlated with colour and PA times. MFCC features positively correlate with semantic features SF_N1-4 while negatively correlate with RT_CN, RT_PA, and RT_PN1-3 reaction time features.

This study investigated the impact of stroke-impaired brain language processing on pathological speech characteristics in dysarthric patients. The results show that abnormal movements of organs like the larynx, tongue, and jaw lead to reduced stability and accuracy of articulation, resulting in distorted vowels and consonant errors that severely affect speech intelligibility. The correlation between acoustic features and behavioural features reflects that damage to the brain's language network can affect acoustic features, reflected in behavioural features at different stages of speech production.

Significant acoustic differences between dysarthric patients and controls reflected articulation, phonation, and prosody abilities. Patients showed longer vowel durations, unstable pitch and intensity, and higher F1/F2/F3 variance in prosody than controls. In articulation, patients exhibited significantly lower jaw distance, tongue distance, and movement degree. In vowel space measures, patients had significantly lower VSA and VAI but higher vowel centralization (FCR) vs. controls. Patients also demonstrated higher F1 and F2 variability in vowel stability. In phonation, patients presented with higher mean jitter, shimmer, and harmonicity.

Significant behavioural differences existed across speech production stages. Patients showed longer reaction times throughout, especially indicating

weakened conceptual preparation and lexical selection pre-speech. Greater brain involvement and earlier processing stage disruption in tasks corresponded to more considerable reaction time differences between groups. These results suggest that strokes damaged the patient brain language areas in addition to articulatory motors^[17]. Comparing tasks, patients took no longer colour vs. PA and naming, potentially reflecting broader lexical generation coverage and more cognition in colour naming^[18]. For PN, the longest to shortest reaction times were RT-PN3, RT-PN2, and RT-PN1 for all subjects, aligning with picture lexical length effects on phonetic encoding/articulation.

In SF, patients uttered fewer words per 15 vs. controls, with more significant differences that decreased over time, further pointing to slowed initiation and conceptual preparation. Only cognitive impairment caused by stroke can have an overall impact on speech behaviour performance, and cognitive level, as the ability in the conceptual preparation stage, mainly affects the upper processing stage of speech production. The PA task involves more upper level language processing stages compared to the colour naming task. The difference between normal individuals and patients in (RT-PA) is smaller than the difference in (RT-CN), indicating that stroke induced articulation disorders mainly affect the lower level language processing stage of speech production, indirectly leading to a slight decrease in cognitive level.

Different language dysfunction may be related to different locations of brain injury^[19]. Research has found that patients with basal ganglia injuries and lesions can also experience significant speech dysfunction, with main pathological manifestations including spontaneous reduction in speech, monotony, slow speech, difficulty in initiation, lack of coherence between words, and naming disorders. Therefore, dysarthria caused by basal ganglia lesions include problems in oral expression and understanding, indicating that the basal ganglia affects both visual and auditory perceptual processing pathways. At the same time, it also affects the motor efferent pathways related to speech movement and continuous vocalization, and there is a synergistic effect between the effects of different regions and circuits^[20], this is also the reason why dysarthric differ from normal individuals in speech behaviour tasks.

The correlation analysis between objective acoustic features and subjective Frenchay scale ratings showed high significance, proving that the designed pathological speech features can diagnose and assess dysarthria types and severities^[21]. Meanwhile, behavioural features were mildly significantly correlated with Frenchay scores, especially for the reflex, speech, respiration, and lip sections, indicating that the speed of speech production behaviours is also affected by articulatory motor ability and reflex capabilities.

TABLE 3: CORRELATION ANALYSIS BETWEEN SEMANTIC TASK FEATURES AND ACOUSTIC FEATURES

	SF_N1	SF_N2	SF_N3	SF_N4	RT_CN	RT_PA	RT_PN1	RT_PN2	RT_PN3
Duration	0.28		-0.08	0.27		0.08	0.17	0.33	0.12
Pitch_mean		0.28	0.25	0.09	-0.3	-0.48*	-0.43	-0.49*	-0.37
Pitch_sd	0.15	0.08	0.12	0.32	-0.51*	-0.50*	-0.4	-0.26	-0.23
Loudness_mean	0.17			0.27	0.07	0.08			-0.08
Loudness_sd	-0.39	-0.26	-0.23		0.36	0.50*	0.53*	0.45*	0.56*
f1_mean	-0.27	-0.09	-0.13	-0.12	0.36	0.1	0.13	0.2	0.07
f3_mean	0.17	0.32	0.33		-0.19	-0.3	-0.19	-0.13	-0.11
Jaw_dist	-0.24	-0.1	-0.08		0.35	0.09	0.28	0.32	0.18
Tongue_dist	0.28	0.32	0.43	0.19	-0.15	-0.15	-0.07	-0.21	
Mov_deg	0.32	0.33	0.50*	0.28	-0.16	-0.19		-0.2	
FCR	-0.28	-0.37	-0.49*	-0.25	0.12	0.17		0.19	
VAI	0.3	0.37	0.49*	0.29	-0.14	-0.2	-0.09	-0.2	-0.03
Jitter	-0.3			-0.15	0.18	0.2		-0.07	
Shimmer		0.37	0.39	0.22	-0.14			-0.18	
Harmonicity	0.35			0.2	-0.44	-0.28	-0.12	-0.13	-0.16
MFCC3	0.35	0.36	0.45	0.26	-0.41	-0.44	-0.43	-0.53*	-0.41

MFCC4	0.34	0.21	0.3	0.3	-0.21				-0.19
MFCC5	0.69**	0.64*	0.39	0.27	-0.31	-0.16			-0.14
MFCC6	0.27	0.46*	0.36		-0.16	-0.24	-0.22	-0.3	-0.12
MFCC8	0.3	0.39	0.28		-0.38	-0.39	-0.42	-0.43	-0.29

Note: ** $p < 0.01$, indicates extremely significant correlation and * $p < 0.01$, indicates significant correlation

There were correlations between behavioural and acoustic features reflected in reaction times, especially in early speech initiation stages in the SF task. The early stages were more affected by articulatory motor capabilities, while this impact decreased over time. Cognitive ability played a decisive differentiating role throughout speech production, including conceptual preparation, lexical selection, and phonetic encoding stages. Meanwhile, it was also found that negatively correlations between colour naming, PA tasks and acoustic features. They were manifested in correlations with both articulation and phonation features. The changes in speech production behaviour and acoustic characteristics may be caused by damage to the cerebellar speech processing area in stroke patients with articulation disorders, as the cerebellum plays a crucial role in learning and issuing accurate and fluent feedforward instructions to the articulatory organs during the forward motion control process of the DIVA model. Its main characteristics are interrupted pronunciation and prosody, as well as stress errors in syllables and words, vowel distortion and excessive changes in loudness^[22]. The negative correlations may be because early lexical selection stages in speech induce vocal fold tension and cyclic perturbation, leading to abnormal jitter and harmonicity in patients. Whereas the PN task involves later speech processing by basal ganglia and cerebellum, the behavioural features primarily reflect the significant impact of dysarthria on articulatory movement features. Stroke-induced brain language processing damage has relatively less impact, as evidenced by a decreasing correlation with increasing RT-CN and increasing pictorial word length.

This study also explored the relationship between acoustic-behavioural analysis of subacute stroke patients with dysarthria and pharmacological treatment (fig. 4 and fig. 5). Current pharmacological treatments aimed at improving post-stroke speech function recovery were examined. Existing research indicates that certain medications can enhance speech production recovery. For example, pharmacological interventions that improve neuroplasticity and neurotransmitter functions can positively impact speech and language recovery. These effects are reflected in speech cognition and behavioural parameters.

The data suggest that pharmacological treatments may significantly improve acoustic features such as VSA, pitch stability, and phonation measures (e.g., jitter and shimmer). Correlations between medication use and improvements in these acoustic features were observed, indicating that medications can facilitate better motor control and cognitive processing in speech production. This relationship underscores the potential of pharmacological treatment as an adjunctive therapy for improving speech outcomes in dysarthria patients.

Dysarthric patients may also have concurrent impairments in basal ganglia and cerebellum, reflected in longer behavioural task reaction times and fewer semantic vocabulary items. Dysarthria mainly affects early speech initiation stages and motor processes, but cognitive ability regulates speech production, including conceptual preparation, lexical selection, and phonetic encoding stages. Meanwhile, behavioural changes caused by impaired brain language processing were correlated with some pathological acoustic features. These results illustrate that brain damage (cognitive impairments) caused by strokes leads to weakened speech processing abilities, while dysarthria itself affects control of phonation and articulatory organs.

Our study established a joint analysis framework combining speech production behaviours and pathological speech to investigate different aspects of cognitive impairment, brain language areas, and articulatory motor dysfunction on speech capabilities in dysarthria. Furthermore, the study highlighted the potential role of pharmacological treatment in improving speech function recovery in post-stroke patients. The positive impact of medications on both acoustic and behavioural features suggests a promising avenue for targeted rehabilitation strategies. Further research will feature-extract language processing and cognitive deficits and apply these extracted features along with qualities of disordered speech to improve diagnosis, assessment, and targeted rehabilitation of post-stroke dysarthria patients. The integration of pharmacological treatment with acoustic-behavioural analysis holds potential clinical applications for enhancing the efficacy of dysarthria management in subacute stroke patients.

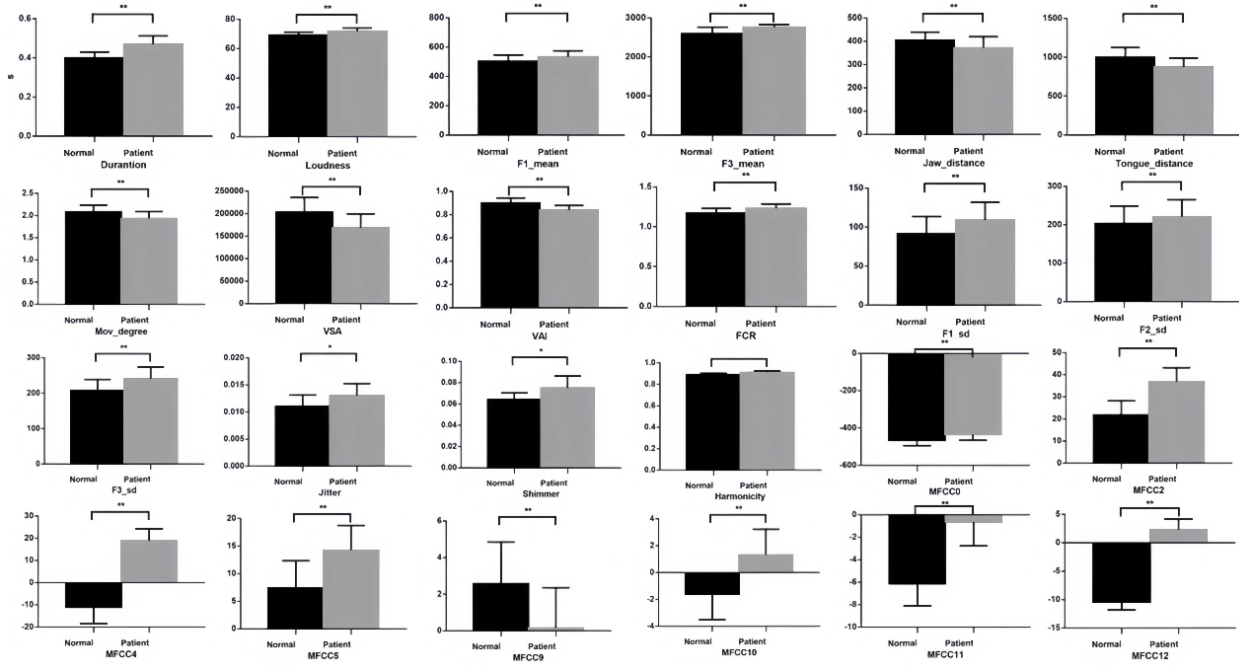


Fig. 4: Comparison of acoustic characteristics between patients and normal individuals
Note: **p<0.01, indicates extremely significant correlation

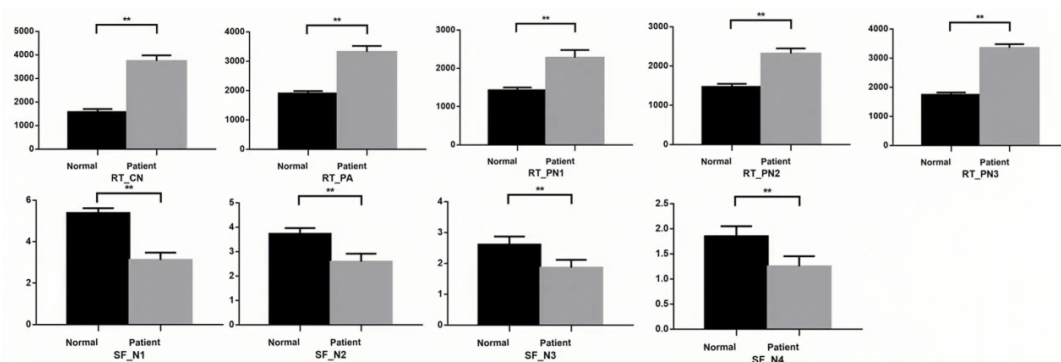


Fig. 5: Comparison of behavioral characteristics between patients and normal individuals
Note: **p<0.01, indicates extremely significant correlation

Acknowledgement:

This work is partly supported by National Natural Science Foundation of China (No: NSFC 62271477), Shenzhen Science and Technology Program (No: JCYJ20220818101411025, JCYJ20220818101217 037, JCYJ20210324115810030), and Shenzhen Peacock Team Project (No: KQTD20200820113106007).

Ethical approval:

This study was reviewed and approved by the Institutional Research Ethics Committee of 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. And there are no vulnerable patients in this study.

Authors' contributions:

Yanbo Wang and Juan Liu collected the data and performed the analysis, Nan Yan and Lan Wang contributed to the conceptualization and supervision of the project. All authors provided critical feedback and helped shape the framework of the article.

Conflict of interests:

The authors declared no conflict of interests.

REFERENCES

1. Joshy AA, Rajan R. Automated dysarthria severity classification: A study on acoustic features and deep learning techniques. *IEEE Trans Neural Syst Rehabil Eng* 2022;30:1147-57.
2. Ali M, Lyden P, Brady M. Aphasia and dysarthria in acute stroke: Recovery and functional outcome. *Int J Stroke* 2015;10(3):400-6.

3. Kim H, Hasegawa-Johnson M, Perlman A. Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatr Logop* 2011;63(4):187-94.
4. Borrie SA, Wynn CJ, Berisha V, Barrett TS. From speech acoustics to communicative participation in dysarthria: Toward a causal framework. *J Speech Lang Hear Res* 2022;65(2):405-18.
5. Levelt WJ. Spoken word production: A theory of lexical access. *Proceed Natl Academy Sci* 2001;98(23):13464-71.
6. Wilson SM, Isenberg AL, Hickok G. Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Human Brain Mapping* 2009;30(11):3596-608.
7. Steurer H, Schalling E, Franzén E, Albrecht F. Characterization of mild and moderate dysarthria in Parkinson's disease: Behavioral measures and neural correlates. *Front Aging Neurosci* 2022;14:870998.
8. Ghosh SS, Tourville JA, Guenther FH. A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *J Speech Lang Hear Res* 2008;51:1183-202.
9. Hennessey NW, Kirsner K. The role of sub-lexical orthography in naming: A performance and acoustic analysis. *Acta Psychol* 1999;103(1):125-48.
10. Schmitt BM, Meyer AS, Levelt WJ. Lexical access in the production of pronouns. *Cognition* 1999;69(3):313-35.
11. Walker GM, Hickok G, Fridriksson J. A cognitive psychometric model for assessment of picture naming abilities in aphasia. *Psychol Assess* 2018;30(6):809.
12. Lin F, Cheng SQ, Qi DQ, Jiang YE, Lyu QQ, Zhong LJ, *et al.* Brain hubs and dark functional networks: Correlation analysis between amplitude and connectivity for Broca's aphasia. *PeerJ* 2020;8:e10057.
13. Liu Y, Hao M, Li P, Shu H. Timed picture naming norms for Mandarin Chinese. *PLoS One* 2011;6(1):e16505.
14. Monsell S, Taylor TJ, Murphy K. Naming the color of a word: Is it responses or task sets that compete? *Memory Cognition* 2001;29:137-51.
15. McNeil MR, Weismer G, Adams S, Mulligan M. Oral structure nonspeech motor control in normal, dysarthric, aphasic and apraxic speakers: Isometric force and static position control. *J Speech Hear Res* 1990;33(2):255-68.
16. Caverle MW, Vogel AP. Stability, reliability, and sensitivity of acoustic measures of vowel space: A comparison of vowel space area, formant centralization ratio, and vowel articulation index. *J Acoust Soc Am* 2020;148(3):1436-44.
17. Urban PP, Wicht S, Vukurevic G, Fitzek C, Fitzek S, Stoeter P, *et al.* Dysarthria in acute ischemic stroke: Lesion topography, clinicoradiologic correlation, and etiology. *Neurology* 2001;56(8):1021-7.
18. Cope TE, Wilson B, Robson H, Drinkall R, Dean L, Grube M, *et al.* Artificial grammar learning in vascular and progressive non-fluent aphasia. *Neuropsychologia* 2017;104:201-13.
19. Kent RD, Duffy JR, Slama A, Kent JF, Clift A. Clinicoanatomic studies in dysarthria. *J Speech Lang Hear Res* 2001;44:535-51.
20. Paton JJ, Buonomano DV. The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron* 2018;98(4):687-705.
21. Kim Y, Kent RD, Weismer G. An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *J Speech Lang Hear Res* 2011;54:417-29.
22. Spencer KA, Slocumb DL. The neural basis of ataxic dysarthria. *Cerebellum* 2007;6:58-65.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms

This article was originally published in a special issue, "Clinical Advancements in Life Sciences and Pharmaceutical Research" Indian J Pharm Sci 2024;86(5) Spl Issue "123-133"