

Article

Development of Computer-Aided Semi-Automatic Diagnosis System for Chronic Post-Stroke Aphasia Classification with Temporal and Parietal Lesions: A Pilot Study

Bhagya Nathali Silva ¹, Murad Khan ¹, Ruchire Eranga Wijesinghe ^{2,3}, Samantha Thelijjagoda ⁴ and Kijun Han ^{1,*}

¹ School of Computer Science and Engineering, Kyungpook National University, Daegu 41566, Korea; nathalis@netopia.knu.ac.kr (B.N.S.); mkhan@knu.ac.kr (M.K.)

² Department of Biomedical Engineering, College of Engineering, Kyungil University, Gyeongsangbuk-do 38428, Korea; eranga@kiu.kr

³ Department of Autonomous Robot Engineering, College of smart Engineering, Kyungil University, Gyeongsangbuk-do 38428, Korea

⁴ Faculty of Computing, Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka; samantha.t@slit.lk

* Correspondence: kjhan@knu.ac.kr; Tel.: +82-940-8697

Received: 8 April 2020; Accepted: 22 April 2020; Published: 24 April 2020



Abstract: Survivors of either a hemorrhagic or ischemic stroke tend to acquire aphasia and experience spontaneous recovery during the first six months. Nevertheless, a considerable number of patients sustain aphasia and require speech and language therapy to overcome the difficulties. As a preliminary study, this article aims to distinguish aphasia caused from a temporoparietal lesion. Typically, temporal and parietal lesions cause Wernicke's aphasia and Anomic aphasia. Differential diagnosis between Anomic and Wernicke's has become controversial and subjective due to the close resemblance of Wernicke's to Anomic aphasia when recovering. Hence, this article proposes a clinical diagnosis system that incorporates normal coupling between the acoustic frequencies of speech signals and the language ability of temporoparietal aphasias to delineate classification boundary lines. The proposed inspection system is a hybrid scheme consisting of automated components, such as confrontation naming, repetition, and a manual component, such as comprehension. The study was conducted involving 30 participants clinically diagnosed with temporoparietal aphasias after a stroke and 30 participants who had experienced a stroke without aphasia. The plausibility of accurate classification of Wernicke's and Anomic aphasia was confirmed using the distinctive acoustic frequency profiles of selected controls. Accuracy of the proposed system and algorithm was confirmed by comparing the obtained diagnosis with the conventional manual diagnosis. Though this preliminary work distinguishes between Anomic and Wernicke's aphasia, we can claim that the developed algorithm-based inspection model could be a worthwhile solution towards objective classification of other aphasia types.

Keywords: objective diagnosis; aphasia; hybrid aphasia diagnosis; acoustic frequency analysis

1. Introduction

Aphasia is a language disorder mostly acquired after a stroke. Broca's, Wernicke's, and Anomic aphasia are the most common subtypes of aphasia [1]. Broca's aphasia is a non-fluent aphasia type caused from a lesion site in the frontal lobe. As a preliminary work this study considered Anomic

and Wernicke's aphasia types, which are fluent and caused from a lesion site in the temporoparietal lobe. Precise classification of Wernicke's and Anomic aphasia among the patients beyond the acute period could be fuzzy, since Wernicke's aphasia eventually resembles Anomic aphasia with recovery progression [2]. Anomia is a fluent type of aphasia, which results after damage in the left hemisphere. The lesion site of Anomia is the least defined, since it is caused by damage throughout the peri-sylvian region [1]. Generally, Wernicke's aphasia is encountered after a stroke in the left temporoparietal region, which is responsible for language comprehension [3]. Word finding difficulties may obstruct satisfactory communication of Anomic patients. Even though Wernicke's aphasia belongs to the fluent type, the conversations tend to be unintelligible to listeners due to word finding difficulties. Aphasia is associated with difficulties in both comprehension and expression means of verbal and written language [4–8]. Even though a stroke may cause extreme communication difficulties due to necrotic damage in certain brain areas, the recovery can be exacerbated by overlooking residual communication abilities. Follow-up assessments conducted manually in clinical settings may misinterpret the level of residual language competencies as a matter of subjectivity of a clinician's bias [9]. For better rehabilitation, aphasia is further classified into subcategories considering lesion site, naming ability, fluency, auditory comprehension, and ability to repeat [10].

In general, a majority of the standardized aphasia assessments follow a manual procedure conducted by a speech and language pathologist (SLP). Conventional aphasia assessments take a considerable amount of time and require expert knowledge on linguistic and cultural background [11]. In conventional assessments, SLPs do not evaluate skill levels quantitatively, and thus, they become statistically unreliable. In this regard, acoustic frequency analysis of pathological speech signals has potential as an accurate, non-invasive, inexpensive, and objectively administered method, which assesses and differentiates aphasia subtypes.

Advancements in automatic speech recognition (ASR) and digital signal processing (DSP) lead speech and language assessments towards a new realm. During the past few decades, numerous groups have suggested substituting manual language assessment procedures with effective automated assessment processes. Pakhomov et al. developed a semi-automated picture description task to determine objective measurements of speech and language characteristics of patients with frontotemporal lobar degeneration including aphasia [9]. The practicability of the test was confirmed through measured magnitude of speech errors among patients with left hemisphere disorders (LHD) [9]. Fraser et al. used ASR to generate transcripts of spoken narratives of healthy controls and patients with semantic dementia and non-fluent progressive aphasia [11]. Although this work assures the feasibility of an ASR-assisted assessment process in diagnosing aphasia, interactions between features, i.e., text and acoustic features, were considered less. Haley et al. examined basic psychometric properties to develop new measures for monosyllabic intelligibility of individuals with aphasia or apraxia of speech (AOS) [12]. The production validity and construct validity were assessed for each person by means of phonetic transcription and overall articulation ratings. The system for automated language and speech analysis (SALSA) is one of the commercially available automated speech assessments [13]. It automates the quantitative analysis of spontaneous speech and language to assess a wide range of neuropsychological impairments including aphasia, dementia, schizophrenia, autism, and epilepsy [13]. However, most of the proposed software solutions are bound to one language, which hinders generalizability in multi-lingual communities [11,14,15]. Thus, there is a need for a generalizable automated software solution competent of making efficient and accurate diagnosis decisions. Hence, developing an automated solution that classifies temporoparietal aphasia using acoustic frequencies would be a promising approach towards the betterment of clinical diagnostics. In fact, the analysis of acoustic frequencies of speech has been considered as a promising technique to implement non-invasive and objective speech and language assessments [16]. In addition, ASR free feature analysis improves accuracy, while reducing language dependency bound with ASR techniques [16]. Nevertheless, the clinical application of automated aphasia assessment is minimal due to the lack of accessibility, appropriateness, availability, and monetary restrictions.

Hence, this article demonstrates a hybrid software solution that incorporates acoustic frequencies of speech signals to determine and distinguish aphasia. In fact, the utilization of ASR free techniques augments the generalizability, since acoustic frequencies are not language dependent. The developed system consists of three diagnosis components, i.e., confrontation naming, single word repetition, and comprehension analysis. The hybrid system automates naming and repetition tasks. Meanwhile, quantitative measurements of the patients' comprehension levels were taken into account for the comprehension analysis. Introduced mathematical relationships among identified parameters derive a score for each assessment component. Accordingly, the proposed algorithm determines presence of aphasia or/and a differential diagnosis for each participant. Finally, a diagnosis report is generated for the reference of SLPs. As a pilot study, this work aims to distinguish aphasia occurrence with temporoparietal lesions, which are basically Anomic aphasia and Wernicke's aphasia. The conceptual breakthrough of the proposed work is in distinguishing two types of aphasia using acoustic frequencies of pathological speech production and comprehension analysis. To the best of our knowledge, previous works have not developed a language independent hybridized software solution to classify subtypes of aphasia using acoustic frequencies of speech signals.

2. Methodology

2.1. Operational Procedure of the Developed Aphasia Classification System

The implementation of the proposed aphasia classification system was initiated by compiling a manual assessment. The manual speech assessment was developed to assess confrontational naming, single word repetition, single word comprehension, and simple command comprehension tasks. In manual assessment, a live model carried out the aforementioned tasks. Herein, manual assessment was tested with 5 neurologically healthy adults and 5 adults who had experienced a stroke but had been assessed as non-aphasic to validate the assessment materials, tasks, and instructions. Accordingly, the materials, tasks, and instructions used in the manual assessment were replicated in the software solution. Thus, this guarantees that the hybrid aphasia assessment tool is free from ambiguities. The assessment procedures were conducted by a speech and language pathologist certified at the medical council of Sri Lanka. The study followed a non-invasive, non-destructive, and non-harmful procedure, which only recorded human speech samples via a microphone. The ethical committee of the National Hospital Sri Lanka (Colombo) approved the study, and speech sample recording was performed in accordance with the guidelines of the of ethical committee of National Hospital Sri Lanka (Colombo) and the Institutional Care and Ethical Use Committees of Sri Lanka Institute of Information Technology, Sri Lanka and Kyungil University, Korea. All participants were included with approved consent of participant and/or guardian. Informed consent covered the participant's approval to participate in the language assessment procedure, to share demographic information, to share clinical history, and to audio record assessment speech samples.

Speech recognition is highly influenced by the frequency and the amplitude of a signal, which are the main components of an acoustic signal. However, these two components alone are insufficient to analyze and suggest a differential diagnosis of a pathological speech signal. Thus, the proposed system requires additional parameters that facilitate differential diagnosis of aphasia. The analysis focused on the number of pathological speech errors, cues given, and time taken to respond. The cues given are categorized into phonemic cues and semantic cues, where these cues as well as test words were included in our laboratory customized user manual. In order to determine a comprehensive relationship between language skills and errors, each error type and cue were handled separately. The clinical diagnosis was taken as the baseline to evaluate the accuracy of the developed system. All participants were clinically diagnosed and clinical diagnosis was blinded to the assessor who evaluated the participants using the proposed algorithm and vice versa. The proposed workflow began with the manual assessment compilation. Manual assessment materials were validated with a pilot test involving 5 neurologically healthy adults and 5 non-aphasic stroke patients. In the software

design process, influential parameters were defined and assigned with scores to derive the math model. Simultaneously, a testing sample and control sample were identified and a clinical diagnosis was performed to use as the baseline for system validation. The materials and instructions of the manual assessment and math model were implemented as a software solution to evaluate participants. Finally, system generated diagnosis was tested with the clinical diagnosis to validate the accuracy of the system. Figure 1 outlines the afore described workflow of the proposed hybrid aphasia inspection tool. Moreover, the numerical algorithm as well as the automated software interface was developed using MATLAB software (Mathworks, MA, USA).

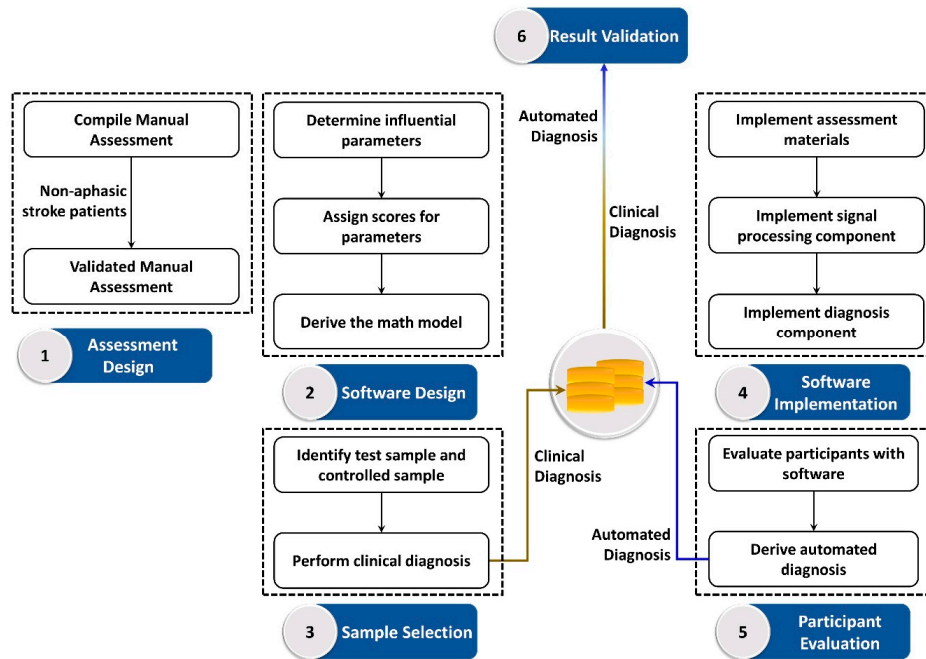


Figure 1. Workflow of the proposed quantitative aphasia assessment tool.

2.2. Graphical Description of the System Overview

Elicited speech samples were directly recorded on the laptop using a microphone mounted to a headphone. Each response from confrontation naming and single word repetition tasks were recorded separately. Pre-recorded words were presented to the subject in the word repetition task. The auditory models were presented via external speakers connected to a laptop. Initially, a trial test for confrontation naming and single word repetition tasks was conducted, in order to introduce the instructions. The pictures and words included in the trial test were not included in the actual assessment. For a better diagnosis, all attempts of a participant were recorded to determine the nature of word elicitation. The assessment was conducted twice using manual assessment and software. However, speech samples of both instances were recorded and analyzed. Data obtained from all participants were evaluated using the algorithm. According to scores obtained in each task, the proposed hybridized aphasia diagnosis process objectively derives the potential diagnosis. Figure 2 depicts the overview of the software architecture. The system consists of confrontation naming, repetition, and comprehension components. The confrontational naming task consists of 30 pictures. In confrontation naming, a single picture stimulus is presented to the participant at a time. The participants were asked to identify the picture and elicit the corresponding word. The repetition task consists of 15 single words. The participants were instructed to repeat the presented word. Elicited speech signals for confrontation naming and repetition tasks were recorded and stored in .wav format. The comprehension task consists of two subcomponents, i.e., single word comprehension and simple command comprehension. The single word comprehension task contains 10 frequent words. The participants should identify the target object

among pictures presented on the laptop screen. The simple command comprehension task evaluates the ability to follow simple instructions. Each instruction consists of one or more objects, sequences, and actions to be performed. In order to avoid the subjective perception of SLPs, the objects, actions, and sequences are included in the assessment. The implemented software autonomously process input speech signals to evaluate acoustic frequencies in confrontation naming and repetition tasks. Initially, the analog speech signal was digitized by sampling repeatedly and encoding the results into a set of bits. Sequentially, the acoustic parameters of the speech signals were determined by applying mel-frequency cepstral coefficients (MFCCs). The similarity level with the template was calculated using dynamic time warping (DTW). Although the comprehension task was not automated at this release and uses inputs from a SLP, all the parameters and desired responses are readily available in the assessment manual to avoid subjectivity and ambiguity. Upon completion, respective scores, i.e., naming score, repetition score, and comprehension score, were calculated using acoustic frequencies and SLP input.

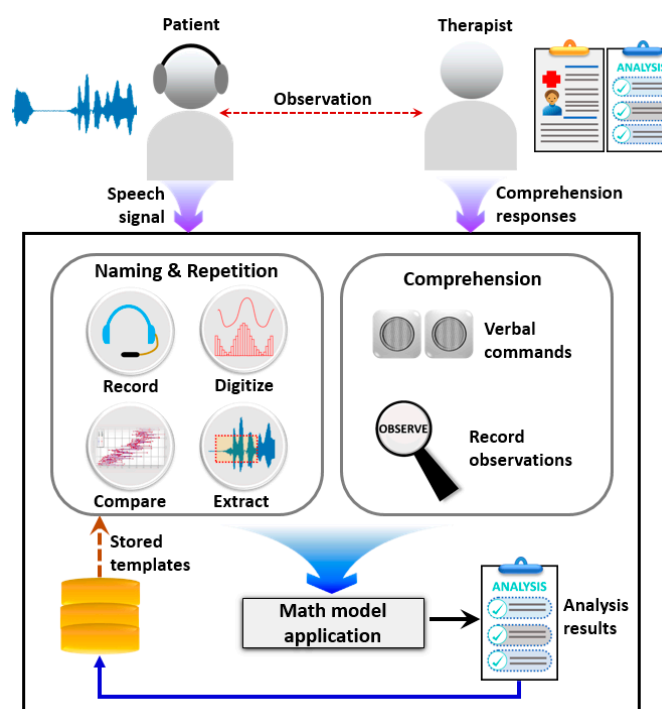


Figure 2. Pictorial overview of the proposed semi-automated aphasia diagnosis system.

2.3. Algorithm of the Developed Inspection Procedure

Confrontation naming is a key component in neurological assessments, since word retrieval stages are commonly affected by inappropriate neurological changes [17]. All responses for the confrontational naming task were audio recorded separately in .wav format with the subject’s consent. The auditory model provided instructions and cues for the confrontation naming task. Accordingly, the confrontation naming score was calculated considering time taken to word production, formant frequencies, and pathological speech characteristics. Formants are the frequency peaks, which have a high degree of energy in the spectrum. In general, formants are prominent in vowels.

$$Naming\ Accuracy\ (Sp_{NA}) = \frac{\sum_{i=0}^{i=4} F_i}{4 + T_{avg} + Sp_{con}} \quad (1)$$

Sp_{NA} is the achieved score for the confrontation naming task. F_i denotes the average of the corresponding formant frequency achieved for all target words. In this study, we considered the

first four formant frequencies. Moreover, the evaluation considers the average time taken for the confrontation naming task, which is denoted by T_{avg} . The quantitative measurements of the speech characteristics were taken into account by calculating $Sp_e, Sp_{pa}, Sp_c, Sp_r$ as mentioned below.

$$Sp_{con} = Sp_e + Sp_{pa} + Sp_c + Sp_r \tag{2}$$

Each speech characteristic was evaluated based on the potential impact on language skills of a person. Hence, identified pathological speech characteristics were considered to calculate speech errors, partial attempts, cues, and responses, which are denoted by $Sp_e, Sp_{pa}, Sp_c,$ and Sp_r , respectively. The Sp_e evaluated the number of occurrences of circumlocution, elicitation of an unrelated word, and absence of any response. If a participant is using many words to express the target word [18], it encounters circumlocution. Whenever the participant produces a word that neither has phonetic nor semantic relation to the target word that is counted as an unrelated word. A no response scenario is encountered when participants do not attempt at all to elicit the target word. Sp_{pa} calculates the partial attempts made by a participant. In phonemic attempts, phonemes are produced instead of the target word. Semantic attempts rely on the verbal meaning of the target word [5], e.g., “eat” instead of “plate”. The cueing comes in two types, (1) auditory cues (phonemic) and (2) verbal meaning of the target word (semantic). Whenever a participant produces a related word or produces the target word after the expected time, they are taken into account to calculate Sp_r .

$$Sp_e = \frac{\text{Circumlocution}}{2} + \text{Unrelated words} + \text{No responses} \tag{3}$$

$$Sp_{pa} = \frac{\text{Phonemic attempts} + \text{Semantic attempts}}{4} \tag{4}$$

$$Sp_c = \frac{\text{Phonemic cues} + \text{Semantic cues}}{4} \tag{5}$$

$$Sp_r = \frac{\text{Related words} + \text{Delayed responses}}{6} \tag{6}$$

The calculation process is designed to reduce the achieved score when the numbers of speech errors are increasing. In addition, the impact of the error is considered to get a more realistic confrontation naming score. For example, the severity of a phonemic attempt has less impact compared to circumlocution, whereas no responses and unrelated words are highly influential for a diagnosis of aphasia.

As a prominent language feature, repetition plays a vital role in differential diagnosis of aphasia subtypes. In this study, we propose a quantitative differential diagnosis tool for Anomic aphasia and Wernicke’s aphasia. Although Wernicke’s present deficits in ability to repeat spoken language, Anomic aphasia demonstrates significantly preserved repetition skills. DTW was used to determine the similarity between the repeated word and the presented auditory model to automatically determine the nature of the repetition. Herein, we defined similarity threshold values to differentiate the nature of the repetition. Accordingly, a repetition score was calculated as given below.

$$\text{Repetition Accuracy } (Sp_{RA}) = \frac{\sum_{i=0}^{i=4} F_i}{4 + T_{avg} + Sp_{rep}} \tag{7}$$

$$Sp_{rep} = \text{Not repeated} + \frac{\text{Phonemic attempts}}{2} + \frac{\text{Partial attempts}}{4} \tag{8}$$

Sp_{RA} determines the achieved score for a single word repetition task. F_i is the average of the corresponding formant frequency for 15 words used in the repetition task. Similar to confrontation naming, T_{avg} indicates the average time taken to repeat a single word. Sp_{rep} calculates the score

corresponding to the nature of the repetition, which is determined by the similarity index of the produced speech signal.

Competency in auditory comprehension is widely used as a key language feature to classify aphasia. The comprehension assessment is not fully automated in this software solution. Nevertheless, it utilizes the manually entered quantitative SLP perceptive decisions on the comprehension competency data inputs to evaluate the auditory comprehension level as it is essential to classify aphasia into subtypes. The entire comprehension component identifies 10 objects in the single word comprehension, 20 objects in simple command comprehension, 15 action sequences, and 15 actions. The score is calculated using the following relationship.

$$\text{Comprehension Accuracy } (Sp_{CA}) = I_{obj} + I_{seq} + I_{act} \tag{9}$$

I_{obj} indicates the number of accurately identified objects followed by I_{seq} and I_{act} denoting the number of accurate sequences and actions. Finally, the conceptual algorithm evaluates obtained scores for naming, repetition, and comprehension tasks to determine a potential differential diagnosis of aphasia. The algorithm for the differential diagnosis process is presented in Figure 3. The scheme identifies Anomic aphasia and Wernicke’s aphasia by considering the strengths and weaknesses of the performed language skills. For evaluation we considered right handed (RH) and non-right handed (NRH) participants with a left hemisphere lesion (LHL) and right hemisphere lesion (RHL).

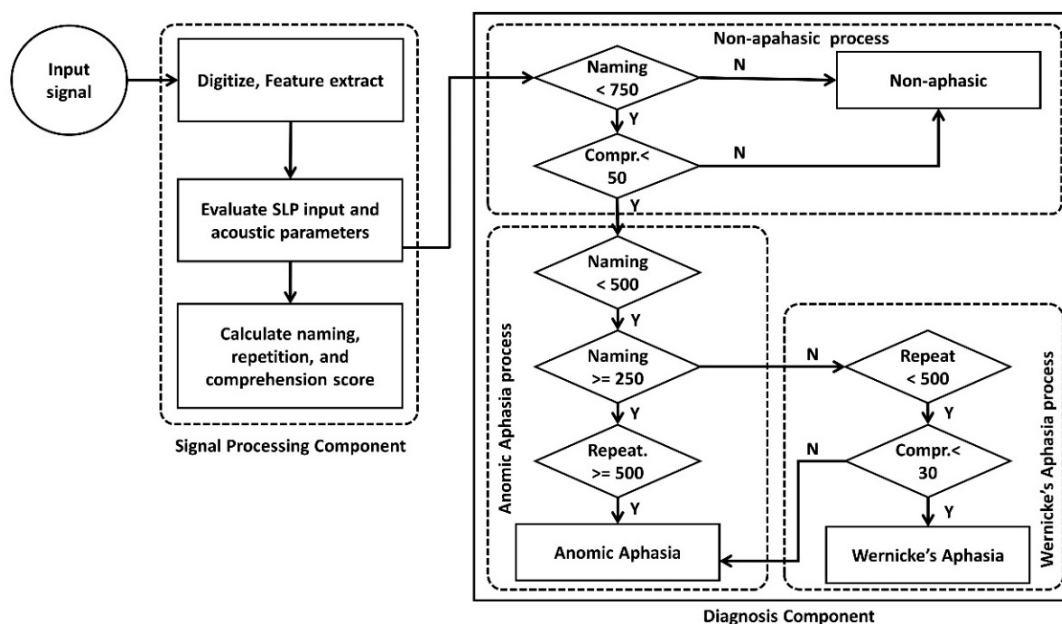


Figure 3. Algorithm for the differential diagnosis of Anomic and Wernicke’s aphasia.

2.4. Patient Information

After validating the manual assessment, 30 participants [female (n = 13), Anomic (n = 18), Wernicke’s (n = 12), RH (n = 24), NRH (n = 6)] who had experienced a single hemisphere stroke [LHL (n = 28) and RHL (n = 2)] at least 12 months before were recruited for the evaluation. The inclusion criteria was absence of any other psychiatric or neurological disorder except the stroke condition. Selected participants should be able to follow simple commands. Simultaneously, 30 participants who had also experienced a single hemisphere stroke (similar with test group) but had been assessed as non-aphasic were evaluated as the control cohort. In general, adults with aphasia are prone to motor speech disorders, i.e., dysarthria [19] and AOS [20]. It commonly occurs with aphasia and neurodegenerative disorders. To select the participants clinically diagnosed with either Anomic aphasia or Wernicke’s aphasia, Frenchay Dysarthria Assessment (FDA) [21] and Apraxia Battery for

Adults (ABA) [22] tests were performed to exclude the patients with dysarthria or/and apraxia. In fact, these two disorders can influence the objective diagnosis of aphasia based on ASR free properties. Therefore, participants with dysarthria and apraxia will be evaluated in future endeavors to optimize the accuracy of the proposed algorithm. Table 1 presents demographic data and initial clinical diagnosis of participants.

Table 1. Participants' demographic data and initial clinical diagnosis.

ID	Age	Gender	Aphasia Type
AH001	62	Female	Wernicke's
AH002	41	Male	Anomic
AH003	52	Male	Anomic
AH004	53	Female	Anomic
AH005	28	Male	Anomic
AH006	59	Female	Wernicke's
AH007	67	Female	Anomic
AH008	49	Male	Anomic
AH009	71	Female	Wernicke's
AH010	38	Male	Wernicke's
AH011	63	Male	Anomic
AH012	55	Female	Wernicke's
AH013	57	Female	Anomic
AH014	66	Male	Anomic
AH015	47	Male	Wernicke's
AH016	43	Female	Anomic
AH017	72	Male	Anomic
AH018	66	Male	Wernicke's
AH019	51	Male	Anomic
AH020	44	Male	Wernicke's
AH021	46	Female	Anomic
AH022	68	Female	Wernicke's
AH023	52	Male	Anomic
AH024	40	Female	Anomic
AH025	69	Male	Wernicke's
AH026	50	Male	Anomic
AH027	63	Female	Wernicke's
AH028	48	Male	Anomic
AH029	39	Female	Anomic
AH030	62	Male	Wernicke's

3. Results

3.1. Confrontational Naming Analysis

Data obtained through the proposed system for naming, repetition, and comprehension tasks from non-aphasic, Anomic, and Wernicke's participants were analyzed to determine distinguishable characteristics of each sample. Figure 4 depicts the confrontation naming performance of three samples. The average time taken to elicit each word by each sample is illustrated in Figure 4a. Considering the total word list, on average non-aphasic participants took 2.51 seconds to elicit a target word. Contrastingly, the aphasic sample (Anomic and Wernicke's) took 10.73 seconds on average to elicit a target word. The system obtained the first four formant frequencies (F_1 – F_4) of the entire confrontation naming task ($n = 30$ words) for each participant. Upon completing the total word list, averages for each formant frequency were obtained for each participant.

Consequently, derived averages were separately analyzed considering the clinical diagnosis, in order to obtain average formant frequency values for non-aphasic, Anomic, and Wernicke's samples. Accordingly, formant frequency variation for the aforementioned samples are presented in Figure 4b, which depicts a higher formant frequency average for Anomic aphasia and lower formant frequency average for Wernicke's aphasia than the non-aphasic sample. Further, paired t -test confirmed that formant frequencies of Anomic [$t(17) = 1.9977$, (P value < 0.05)] and Wernicke's [$t(11) = 2.5642$,

(P value < 0.05)] are significantly different from the non-aphasic sample. Figure 4c presents the average number of pathological speech characteristics encountered in the Anomic and Wernicke’s samples, since the control sample was free from these characteristics. Comparatively, Anomic participants had 15% fewer pathological characteristics than Wernicke’s participants. Upon the completion of the confrontation naming task, the mathematical module calculates Sp_{NA} for each participant. The algorithm uses the naming score to perform a differential diagnosis of Anomic and Wernicke’s aphasia. The Anomic sample’s mean Sp_{NA} value was always higher than mean Sp_{NA} of the Wernicke’s sample. Thus, the proposed algorithm can accurately distinguish the naming performance of patients with Anomic and Wernicke’s aphasia. A summarized analysis of confrontation naming performance is given below in Table 2.

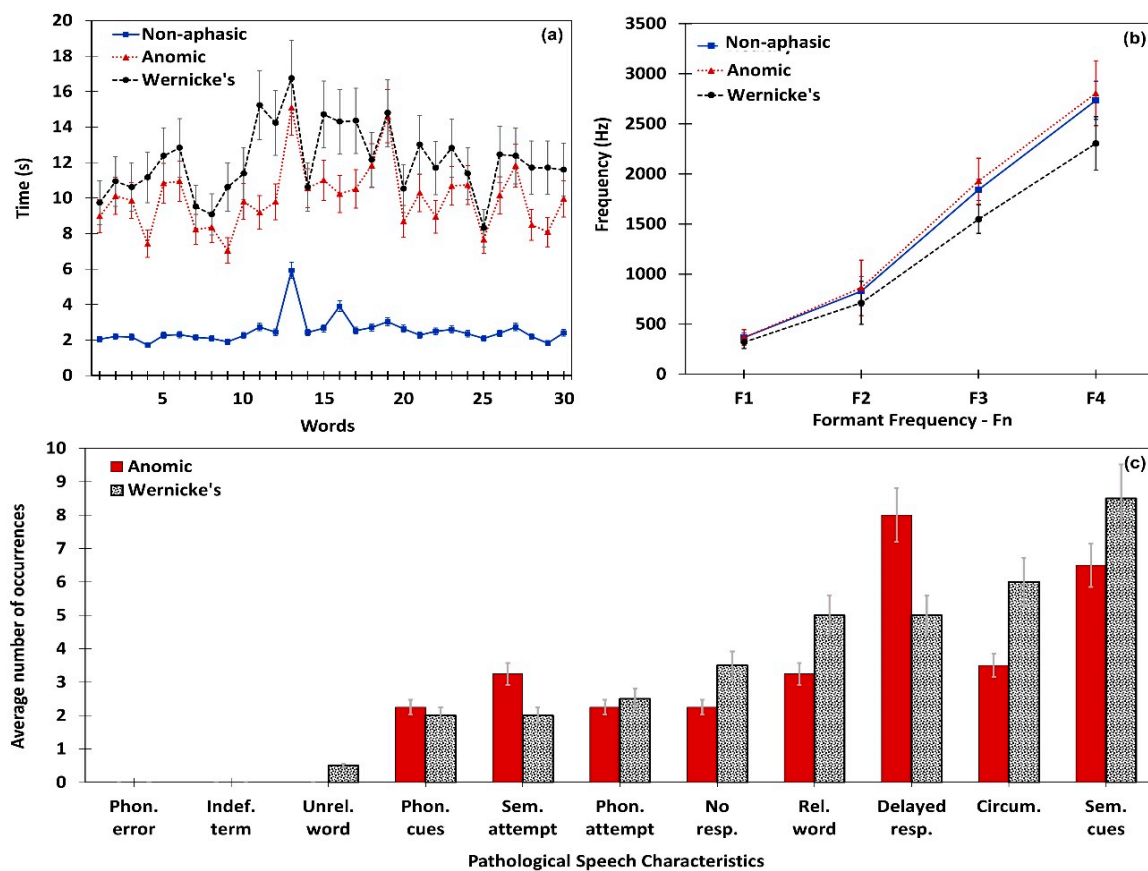


Figure 4. Confrontation naming performance of the non-aphasic, Anomic, and Wernicke’s subjects. (a) Mean time taken to elicit words in the comprehension task; (b) formant frequency variation; (c) speech characteristics identified in aphasic subjects.

Table 2. Performance analysis summary of confrontation naming task.

Sample	No. of Responses	Response Accuracy (%)	Average Time (s)	Average No. of Errors
Non-aphasic (30)	900	100	2.51	0
Anomic (18)	217	40.18	9.98	31
Wernicke’s (12)	106	29.44	11.44	36

3.2. Repetition Analysis

Similar to the confrontation naming task, aphasic participants showed significant difficulty in repeating words compared to non-aphasic participants. Time taken for word repetition, formant

frequency variation, and nature of repetition were separately analyzed for each participant. Figure 5 illustrates repetition task performances obtained through the proposed system. Non-aphasic participants repeated a target word within 2.02 seconds and aphasic participants took 9.16 seconds to repeat a word on average. Figure 5a presents each sample’s average time taken for all target words in the repetition task. It is clearly visible that anomic participants encountered lesser difficulty than Wernicke’s participants. As described in Section 3.1, a formant frequency analysis was conducted for the repetition task. Figure 5b depicts variations of average formant frequencies in non-aphasic, Anomic, and Wernicke’s samples. Paired *t*-test on formant frequency confirmed significant difference in both Anomic [$t(17) = 1.8858, (P \text{ value} < 0.05)$] and Wernicke’s [$t(11) = 2.3615, (P \text{ value} < 0.05)$] samples compared to the non-aphasic sample. As shown in Figure 5c, Anomic participants demonstrated their superior repetition ability over Wernicke’s participants. Upon completion of the repetition task, the mathematical module calculates Sp_{RA} for each participant. Since the Anomic sample’s mean Sp_{RA} is always higher than the Wernicke’s sample, the algorithm uses repetition score to quantitatively differentiate Anomic aphasia from Wernicke’s aphasia. A summarized analysis of repetition performance is given below in Table 3.

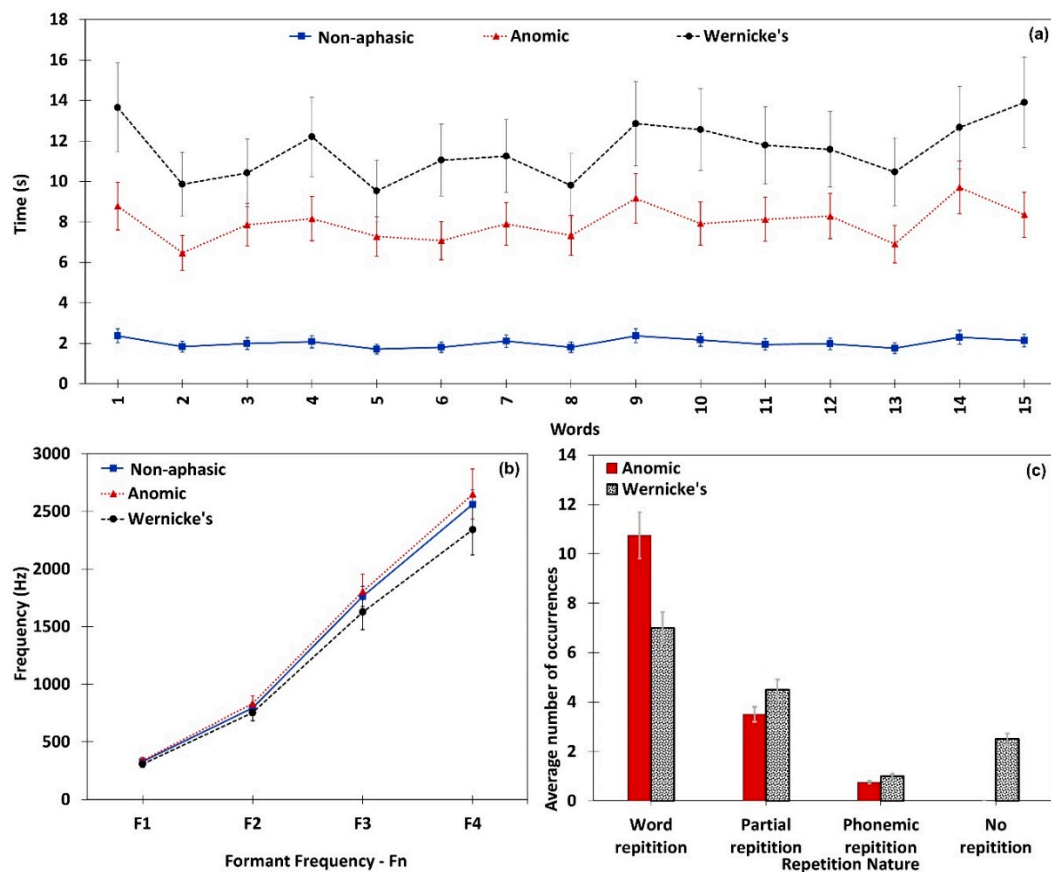


Figure 5. Repetition performance of the neurologically non-aphasic, Anomic, and Wernicke’s subjects. (a) Mean time taken to elicit words in the repetition task; (b) formant frequency variation; (c) nature of repetition in aphasic subjects.

Table 3. Performance analysis summary of repetition task.

Sample	No. of Responses	Response Accuracy (%)	Average Time (s)	Average Errors
Non-aphasic (30)	450	100	2.02	0
Anomic (18)	189	70	7.76	4.5
Wernicke’s (12)	78	43.3	10.56	8.5

3.3. Comprehension Analysis

Comprehension analysis consists of single word comprehension and simple command comprehension tasks. Figure 6 illustrates participants' comprehension levels in both tasks. However, Figure 6 does not represent non-aphasic participants, since they demonstrated 100% accuracy in both tasks. Figure 6a illustrates the number of accurate responses by Anomic and Wernicke's participants per each target word. It is clear from the results that the Anomic participants have a higher level of single word comprehension than the Wernicke's participants. The simple command comprehension task separately analyzes participants' comprehension levels of objects, sequences, and actions, and the performance results are presented in Figure 6b–d, respectively. Aligning with single word comprehension, Anomic participants secured a higher level of comprehension in the simple command task as well. A summary of the comprehension task is given in Table 4. Although the comprehension analysis was not automated, all evaluation points were enclosed with the assessment and prompt the SLP to input only quantitative observations, e.g., number of objects, sequences, and actions identified. Hence, SLP input is not subjective and maintains consistency throughout. A performance summary of the comprehension task is given below in Table 5.

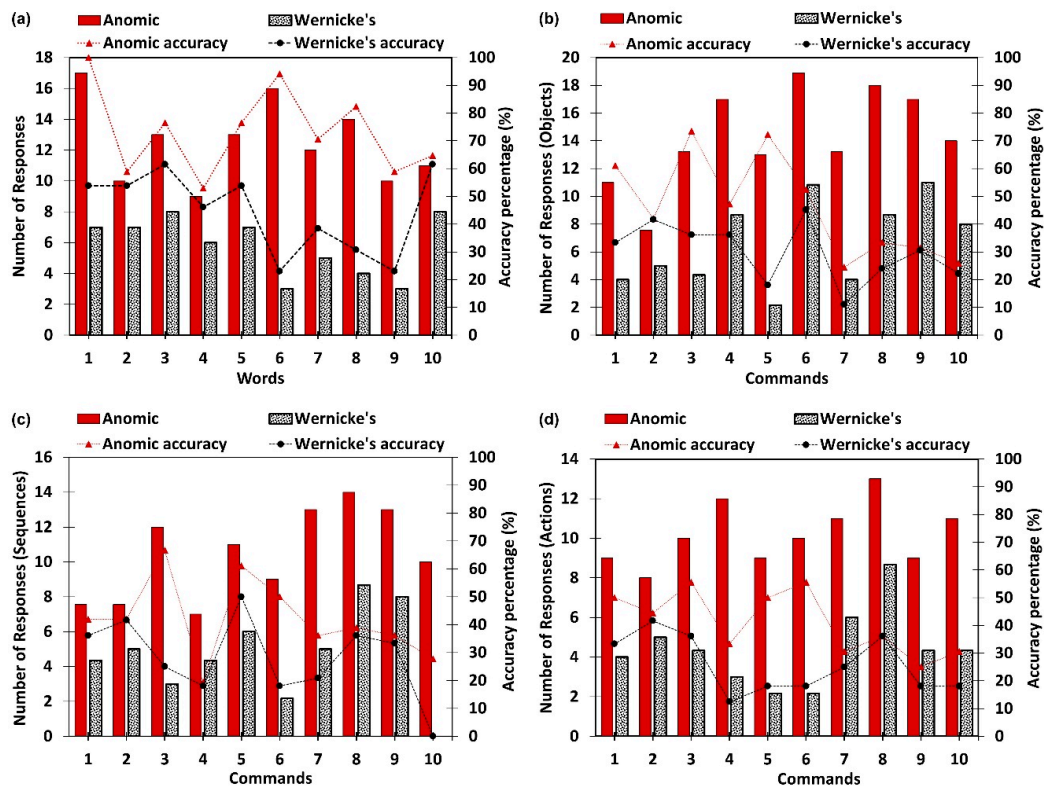


Figure 6. Comprehension performance of the Anomic and Wernicke's subjects. (a) Average accuracy of single word comprehension; (b) object identification accuracy in commands; (c) sequence accuracy in commands; (d) action identification accuracy in commands.

Table 4. Comprehension task summary (number of objects, sequences, and actions to comprehend).

Criteria	Single Word Task	Simple Commands Task									
		1	2	3	4	5	6	7	8	9	10
Objects	10	1	1	1	2	1	2	3	3	3	3
Sequences	-	1	1	1	2	1	1	2	2	2	2
Actions	-	1	1	1	2	1	1	2	2	2	2

Table 5. Performance analysis summary of comprehension task.

Sample	Single Word Accuracy (%)	Simple Command Accuracy (%)		
		Objects	Sequences	Actions
Anomic (18)	72.5	46.4	42	41.1
Wernicke’s (12)	45	29.9	27.9	25.7

4. Discussion

Experts became interested in automating speech and language assessments owing to the technological boost in ASR. In general, ASR performance is tightly coupled with the quality of the input speech signal [23]. Hence, applicability of ASR in pathological speech signal evaluation has become controversial despite its benefits. Nevertheless, ASR free features of speech signals are considered as promising alternatives when evaluating pathological speech signals [16]. Although a lot of insightful works have been performed in the autonomous language skills evaluation domain [9,11–13,23–25], there is still a lot of room left for research. Hence, herein we incorporated ASR free features of pathological speech characteristics to develop a quantitative and objective aphasia assessment tool, which differentially diagnoses Anomic and Wernicke’s aphasia. Thereby, we proposed a mathematical model to evaluate participants’ performance levels in ASR free features. Subsequently, scores obtained from the mathematical model were analyzed using the diagnosis algorithm. As a result, an objective diagnosis [9] can be made utilizing quantitative observations of pathological speech signals, instead of relying upon subjective and qualitative evaluations by SLPs. In general, a majority of the standardized aphasia assessments follow a manual procedure conducted by a SLP. Conventional aphasia assessments consume a considerable amount of time and explicitly rely on expert knowledge, cultural background, and clinical experience [26]. In this work we incorporated a live auditory model and a recorded auditory model. The average assessment duration using a live auditory model was 396.2 seconds for the naming task and 199.8 seconds for the repetition task. The recorded auditory model took only 375.6 seconds and 180.7 seconds, respectively. Paired sample *t*-test confirmed a significant difference in assessment duration for both naming [$t(29) = 10.66, (P \text{ value} < 0.05)$] and repetition [$t(29) = 11.12, (P \text{ value} < 0.05)$] tasks with the aphasic sample. Hence, we can claim that using a recorded model improves the time efficiency of the assessment process, while mitigating the potential bias of the SLP. Figure 7 depicts assessment time variation between live and auditory models for naming and repetition tasks.

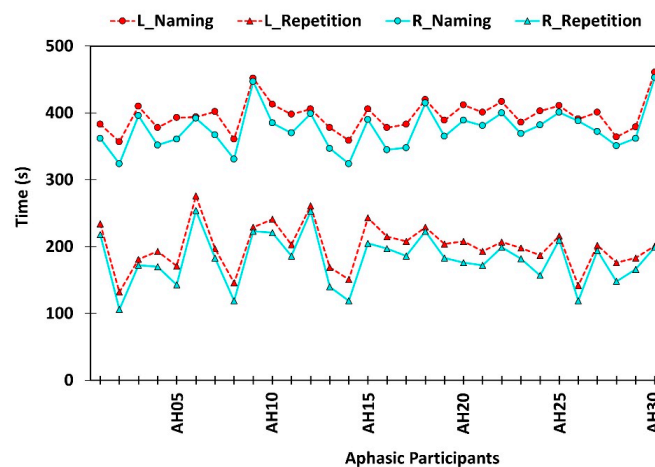


Figure 7. Time spent on confrontation naming and repetition tasks using live auditory model (L) and pre-recorded auditory model (R).

Moreover, hands on experience in clinical assessment procedures significantly influences the diagnosis. Generally, SLPs diagnose language disorders adhering to a standardized assessment kit. Nevertheless, most of the SLPs in non-English speaking countries follow non-standardized assessments due to lack of cultural appropriateness and language differences. The non-standardized assessments do not consider normative data. Instead, they determine isolated skill levels of a person despite the other contributing factors [27], subsequently reducing the benefits of the assessment procedure. Furthermore, the manual assessment process results in multiple loopholes. The major drawback of the manual assessment is the subjectivity of the diagnosis process. It is said to be subjective since the outcome highly depends on the experience, exposure, and perception of the SLP. Moreover, it has been suggested that a computerized approach could be used to minimize the subjectivity of the manual assessment procedure [9]. The participants of this study are clinically diagnosed with Anomic aphasia and Wernicke's aphasia according to Boston classification. The results shown in Figure 4 confirm the word finding difficulties of the participants diagnosed with both Anomic and Wernicke's aphasia. Moreover, the score obtained for the relationship among timing, formant frequencies, and pathological speech characteristics gave a significant boundary for differential diagnosis, as shown below in Table 6. Anomic participants were observed with minimal impairment in the repetition of words. Contrastingly, repetition ability was significantly affected in Wernicke's aphasia. Hence, near normal repetition ability of Anomic aphasia is widely used as a characteristic for differential diagnosis [1]. Evaluation scores obtained through the mathematical model identify the boundary values corresponding to Anomic and Wernicke's aphasia, as shown below. Anomic aphasia characterizes a relatively good comprehension level in general conversation. However, the difficulties appear in the comprehension of grammatically complex sentences [28]. Although visual comprehension is comparatively preserved, the lesion site of Wernicke's aphasia severely affects the verbal comprehension level [29]. By considering the correctly identified number of objects, actions, and sequences, the proposed assessment tool quantitatively evaluates the auditory comprehension level of participants.

Table 6. Average scores obtained for each assessment component.

Sample	Naming Score	Repetition Score	Comprehension Score
Non-aphasic	886.28 ± 13.32	905.99 ± 10.86	60 ± 0
Anomic	312.51 ± 24.97	585.53 ± 21.41	38 ± 4.22
Wernicke's	158.67 ± 31.53	329.75 ± 27.02	22 ± 5.71

Average scores obtained by each sample for each component are given in Table 6 with corresponding standard deviation values. A naming score less than 500 was used as the boundary to classify non-aphasic and aphasic subjects (Anomic and Wernicke's). Similarly, a repetition score less than 500 was used to distinguish Wernicke's aphasia from Anomic aphasia. Furthermore, classification accuracy was confirmed through the comprehension score. If the comprehension score is less than 30 with significantly low repetition and naming skills, it indicates the presence of Wernicke's aphasia. Hence, we can claim that the values presented in Table 6 justify the constraint selection and conditions used in the diagnosis algorithm. Since all the experimental assessments and verifications were fundamentally based on the observed 60 participants of the study, more broad assessments with multiple aspects have to be performed to further enhance the accuracy of the study.

5. Conclusions

The extensive advancements in technology have inspired the adoption of computerized software solutions to diagnose speech and language disorders in clinical settings. However, this notion is still in its infancy due to language barriers and expensive software solutions. Herein, we proposed a generalizable semi-automated system and algorithm to diagnose the presence of aphasia despite of the language restrictions, while alleviating the inefficiencies and subjectivity in the manual diagnosis process. The effective and efficient detection will result in better understanding of the efficacy of

treatments and can be helpful to improve the quality of life of post stroke patients suffering from aphasia. The obtained results are extensively analyzed to derive the relationships among the acoustic frequencies, accuracy of word retrieval, performance efficiency, and other related speech characteristics. The proposed scheme differentially diagnoses two subtypes of aphasia namely, Wernicke's aphasia and Anomic aphasia. Multi-dimensional evaluation for differential diagnosis of aphasia is an essential requirement to identify subtypes of aphasia. The fundamental scope of the study was to differentially diagnose Anomic and Wernicke's aphasia by evaluating major language evaluation components, namely, confrontation naming, repetition, and comprehension. Therefore in future endeavors, the developed algorithm will be further extended to evaluate the correlation of other language components, i.e., fluency, spontaneous speech, reading, and semantic comprehension, characteristics of dysarthria and AOS, and other speech defects, i.e., stuttering, to ensure the accuracy and efficiency of the proposed algorithm while extending its capability to diagnose the remaining subtypes of aphasia.

Author Contributions: Conceptualization, B.N.S. and R.E.W.; formal analysis, B.N.S. and M.K.; funding acquisition, B.N.S., R.E.W., and K.H.; methodology, B.N.S., M.K., and R.E.W.; project administration, K.H.; software, B.N.S.; supervision, S.T. and K.H.; writing—original draft, B.N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1F1A1042721) and BK21 Plus project (SW Human Resource Development Program for Supporting Smart Life) funded by the Ministry of Education, School of Computer Science and Engineering, Kyungpook National University, Korea (21A20131600005), National Research Foundation of Korea (NRF) grant funded by the Korea government. (MSIT) (2019R1G1A1092172), and Korea Research Fellowship program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (2019H1D3A1A01102987).

Conflicts of Interest: The authors declare no competing interests.

References

1. Dronkers, N.; Baldo, J.V. Language: Aphasia. In *Encyclopedia of Neuroscience*; Elsevier Ltd.: Amsterdam, The Netherlands, 2010; pp. 343–348.
2. Kim, H.; Kintz, S.; Zelnosky, K.; Wright, H.H. Measuring word retrieval in narrative discourse: Core lexicon in aphasia. *Int. J. Lang. Commun. Disord.* **2019**, *54*, 62–78. [[CrossRef](#)] [[PubMed](#)]
3. Robson, H.; Zahn, R.; Keidel, J.L.; Binney, R.J.; Sage, K.; Lambon Ralph, M.A. The anterior temporal lobes support residual comprehension in Wernicke's aphasia. *Brain* **2014**, *137*, 931–943. [[CrossRef](#)] [[PubMed](#)]
4. Raymer, A.M.; Gonzalez-Rothi, L.J. *The Oxford Handbook of Aphasia and Language Disorders*; Oxford University Press: Oxford, UK, 2018.
5. Butterworth, B.; Howard, D.; McLoughlin, P. The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia* **1984**, *22*, 409–426. [[CrossRef](#)]
6. Rauschecker, J.P.; Scott, S.K. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nat. Neurosci.* **2009**, *12*, 718. [[CrossRef](#)] [[PubMed](#)]
7. Ebbels, S.H.; McCartney, E.; Slonims, V.; Dockrell, J.E.; Norbury, C.F. Evidence-based pathways to intervention for children with language disorders. *Int. J. Lang. Commun. Disord.* **2019**, *54*, 3–19. [[CrossRef](#)] [[PubMed](#)]
8. Palmer, A.D.; Carder, P.C.; White, D.L.; Saunders, G.; Woo, H.; Graville, D.J.; Newsom, J.T. The Impact of Communication Impairments on the Social Relationships of Older Adults: Pathways to Psychological Well-Being. *J. Speech Lang. Hear. Res.* **2019**, *62*, 1–21. [[CrossRef](#)] [[PubMed](#)]
9. Pakhomov, S.V.; Smith, G.E.; Chacon, D.; Feliciano, Y.; Graff-Radford, N.; Caselli, R.; Knopman, D.S. Computerized analysis of speech and language to identify psycholinguistic correlates of frontotemporal lobar degeneration. *Cogn. Behav. Neurol.* **2010**, *23*, 165. [[CrossRef](#)] [[PubMed](#)]
10. Khan, M.; Silva, B.N.; Ahmed, S.H.; Ahmad, A.; Din, S.; Song, H. You speak, we detect: Quantitative diagnosis of anomic and wernicke's aphasia using digital signal processing techniques. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; IEEE: Washington, DC, USA, 2017; pp. 1–6.

11. Lee, T.; Liu, Y.; Huang, P.-W.; Chien, J.-T.; Lam, W.K.; Yeung, Y.T.; Law, T.K.; Lee, K.Y.; Kong, A.P.-H.; Law, S.-P. Automatic speech recognition for acoustical analysis and assessment of cantonese pathological voice and speech. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 6475–6479.
12. Fraser, K.C.; Meltzer, J.A.; Graham, N.L.; Leonard, C.; Hirst, G.; Black, S.E.; Rochon, E. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex* **2014**, *55*, 43–60. [[CrossRef](#)] [[PubMed](#)]
13. Haley, K.L.; Roth, H.; Grindstaff, E.; Jacks, A. Computer-mediated assessment of intelligibility in aphasia and apraxia of speech. *Aphasiology* **2011**, *25*, 1600–1620. [[CrossRef](#)] [[PubMed](#)]
14. Pakhomov, S.V. Automatic Measurement of Speech Fluency. U.S. Patent 8494857B2, 23 July 2013.
15. Tripathi, M.; Tang, C.C.; Feigin, A.; De Lucia, I.; Nazem, A.; Dhawan, V.; Eidelberg, D. Automated differential diagnosis of early parkinsonism using metabolic brain networks: A validation study. *J. Nucl. Med.* **2016**, *57*, 60–68. [[CrossRef](#)] [[PubMed](#)]
16. Bhat, S.; Acharya, U.R.; Dadmehr, N.; Adeli, H. Clinical neurophysiological and automated EEG-based diagnosis of the Alzheimer’s disease. *Eur. Neurol.* **2015**, *74*, 202–210. [[CrossRef](#)] [[PubMed](#)]
17. Middag, C.; Bocklet, T.; Martens, J.-P.; Nöth, E. Combining phonological and acoustic ASR-free features for pathological speech intelligibility assessment. In Proceedings of the Twelfth Annual Conference of the International Speech Communication Association, Florence, Italy, 27–31 August 2011.
18. Randolph, C.; Lansing, A.E.; Ivnik, R.J.; Cullum, C.M.; Hermann, B.P. Determinants of confrontation naming performance. *Arch. Clin. Neuropsychol.* **1999**, *14*, 489–496. [[CrossRef](#)] [[PubMed](#)]
19. Stanton, I.; Jeong, S.; Mishra, N. Circumlocution in diagnostic medical queries. In Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, Queensland, Australia, 6–11 July 2014; pp. 133–142.
20. Pennington, L.; Roelant, E.; Thompson, V.; Robson, S.; Steen, N.; Miller, N. Intensive dysarthria therapy for younger children with cerebral palsy. *Dev. Med. Child Neurol.* **2013**, *55*, 464–471. [[CrossRef](#)] [[PubMed](#)]
21. Josephs, K.A.; Duffy, J.R.; Strand, E.A.; Machulda, M.M.; Senjem, M.L.; Master, A.V.; Lowe, V.J.; Jack, C.R.; Whitwell, J.L. Characterizing a neurodegenerative syndrome: Primary progressive apraxia of speech. *Brain* **2012**, *135*, 1522–1536. [[CrossRef](#)] [[PubMed](#)]
22. Enderby, P. Frenchay dysarthria assessment. *Br. J. Disord. Commun.* **1980**, *15*, 165–173. [[CrossRef](#)]
23. Dabul, B. *Apraxia Battery for Adults: Examiner’s Manual*; PRO-ED: Austin, TX, USA, 2000.
24. Young, V.; Mihailidis, A. Difficulties in automatic speech recognition of dysarthric speakers and implications for speech-based applications used by the elderly: A literature review. *Assist. Technol.* **2010**, *22*, 99–112. [[CrossRef](#)] [[PubMed](#)]
25. Vipperla, R.; Renals, S.; Frankel, J. Longitudinal Study of ASR Performance on Ageing Voices. Available online: <https://era.ed.ac.uk/handle/1842/3900> (accessed on 23 April 2020).
26. Marek-Spartz, K.; Knoll, B.; Bill, R.; Christie, T.; Pakhomov, S. System for Automated Speech and Language Analysis (SALSA). In Proceedings of the Fifteenth Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014.
27. Shipley, K.G.; McAfee, J.G. *Assessment in Speech-Language Pathology: A Resource Manual*; Nelson Education: Scarborough, ON, Canada, 2015.
28. Andreetta, S.; Cantagallo, A.; Marini, A. Narrative discourse in anomic aphasia. *Neuropsychologia* **2012**, *50*, 1787–1793. [[CrossRef](#)] [[PubMed](#)]
29. Robson, H.; Grube, M.; Ralph, M.A.L.; Griffiths, T.D.; Sage, K. Fundamental deficits of auditory perception in Wernicke’s aphasia. *Cortex* **2013**, *49*, 1808–1822. [[CrossRef](#)] [[PubMed](#)]

