



AMERICAN JOURNAL OF AGRICULTURAL SCIENCE, ENGINEERING, AND TECHNOLOGY (AJASET)

ISSN: 2158-8104 (ONLINE), 2164-0920 (PRINT)

VOLUME 6 ISSUE 3 (2022)



PUBLISHED BY: E-PALLI, DELAWARE, USA

Interactive Learning Tool Using a Neural Network Algorithm for Speech-Delayed Child

Ram Eujohn J. Diamante^{1*}

Article Information

Received: November 16, 2022

Accepted: November 21, 2022

Published: November 29, 2022

Keywords

*Speech Delay, Speech Therapy,
Child, Voice Training,
Phonology*

ABSTRACT

Interactive Learning Tool using a Neural Network Algorithm for Speech-Delayed Child is a system that was developed mainly for speech-delayed children that utilized a multi-sensory approach and interactive learning technique in assessing the speech disorder. Phonology was focused as the content of the learning tool the lessons are based on the consonant-acquisition chart which contains the rules and system for speech sound production. Speech recognition was implemented to analyze, recognized, and matched the child's speech to the exact word in the lesson. The system's functionality is internet based and Google's voice recognition is implemented to carry out the matching process. Neural Network Algorithm methodology is used to calculate the child's test result. The application software is developed using HTML, Java, and PHP as the main programming language and Android Studio with Android Developer Tools as the working environment. It uses MySQL in managing its database. The software is developed for the android operating system, with the version starting which stored the child's information and results. The set of lessons is categorized by age namely: AGES 3-4, AGES 4-5, AGES 5-6, and AGES 6-7. Each category has a unique set of lessons categorized by: syllables with 67 lessons, words with 57 lessons, and phrases which contains 26 lessons. When the user finishes all the lessons prepared the system will generate the result, and the submit button transfers the data result into the teachers' record in a separate system installed on a desktop/laptop computer.

INTRODUCTION

Primary speech delay disorder is a common developmental difficulty that, if unresolved, could last into adolescence and beyond. It can cause difficulties in both learning and socialization. The likelihood of persistent difficulties for young children with speech and language problems appears to be directly related to the range of language functions that are impaired, with the best prognosis for children who have a developmental speech delay.

Concerns for delay arise if there are no verbalizations by the age of 1 if speech is not clear and if speech is different from that of other children of the same age. Most formal instruments were designed for diagnostic purposes and had not been widely evaluated for screening. A specific diagnosis is made most often by a speech and language specialist using a battery of instruments. There are no universal guidelines on what type of intervention to offer children with primary speech delay disorder nor is there consistent evidence upon which to base a decision, meaning that the decision is often left to individual therapists and services.

In this connection, this study aims to develop an android interactive learning tool that will help the child with speech delay cope with his speech disability by implementing the use of multimedia in assessing the child. The set of lessons is based on a consonant-acquisition chart which describes the approximate order of the development of the child's speech sound. With the implementation of speech recognition, the system will be able to recognize, analyze and match the speech to the

exact word in the lesson. With the integration of a neural network algorithm, the system has the ability to calculate the child's test result and produce the assessment in a tabular and graphical representation.

LITERATURE REVIEW

Nemours Children's pediatric speech and language specialists give children the power to effectively interact with their world. We're here for kids who have trouble speaking, understanding, writing, feeding, or swallowing. Our team is highly experienced and child-focused. We work closely with renowned Nemours programs and specialists in ear, nose, and throat, hearing, physical rehabilitation, and more.

We evaluate your child and create a care plan based on their needs and abilities. In addition to individual and group therapy sessions, we offer at-home activity training for families. We use advanced technologies and the latest assistive communication devices. Our goal is to build your child's skills and confidence to express themselves. We work with you to help your child succeed and reach their full potential. (The Nemours Foundation Report, 2022)

According to the reports, the now-toddlers, who began attending preschool in 2020 needed additional support in the areas of speech, language, and communication. The increased screen time coupled with restricted real-life interactions significantly affected their progress toward age-appropriate speech-language milestones. 55 out of the 57 schools surveyed by the Education Endowment

¹ College of Technology Iloilo State College of Fisheries, Dumangas Campus Iloilo, Philippines

* Corresponding author's e-mail: diamanteram@gmail.com

Foundation in 2021 said they were “extremely concerned” about the children’s language & communication skills development. They also stated that they were very concerned about the child’s social, emotional, and personal development. These are all skills that depend upon the speech, language, and communication abilities of a child. When you are saying a new word, describe its sound. It may sound weird, but it can help a child distinguish between the “/p/,” “/b/,” and “/d/” sounds. For example, you say, “the dog is going for a walk. D-for dog.” Or, “we are going to have nuggets for dinner. Nugget-T-T-T” It is common for children up to the age of 4 to confuse different consonant sounds. This simple exercise will help them learn the correct pronunciation as well as their spelling. (Samurai Report, 2021)

A mobile learning’s key features are personalized learning and independence in time and place. Those are the collaboration with students and teachers in both formal and informal settings. Wijayanto (2018) stated that some challenges to preparing the English course materials are to investigate, categorize, select, construct, modify, and develop the materials that apply to the learners’ needs. The interactivity of mobile devices makes mobile learning more efficient. MALL could provide easy access for any learner without the constraints of both place and time. In mobile learning, devices like smartphones, iPods, tablets, laptops, iPads are implemented to support language learning. Even though mobile learning seems to like to be effective overall, it is necessary to design, plan and implement it with carefulness, according to students’ needs, and to deliver multiple language skills in authentic learning environments. (Kacetl & Klímová, 2019)

Many studies confirmed that research has been conducted into targeting the adaptations of adult speech recognition systems toward children’s speech recognition. In the above section, we have discussed the various acoustic differences that exist between children’s and adults’ speech, which will be an important area to investigate when moving forward. A number of researchers have worked on compensating for the various acoustic variations influenced by short vocal tract length in children using vocal tract length normalization (VTLN). Due to these variations, an ASR system trained using adult speech delivered a poor recognition rate when tested using children’s speech and mismatched acoustic conditions. Thus, in this section, we discuss the approaches used for better children’s SR under mismatched conditions. (Bhardwaj V, Ben et.al., 2022)

METHODOLOGY

Documentary research data. Data gathering is the most important factor in developing the system. It focuses on gathering information needed for the development of the system. These are sources of information that are essential for software development. This includes web research, particularly, tutorials and techniques. Some information is derived from special education related research papers and some cases of observation in a speech learning center. Interview is conducted with the

Speech Pathologist, Ms. Revrose Selyne D. Pison, CSP-PASP. Set of words for each lesson in the application was taken from home-speech-home.com referred by the speech pathologist. Software evaluation form. A standard evaluation form was used to determine the quality of the proposed system. The evaluation form has nine criteria: reliability, efficiency, usability and understandability, appropriateness of feedback to user, navigation and organization, correctness and integrity. Each criterion is evaluated as “Excellent,” “Very Good,” “Good,” “Fair,” “Poor,” “Very Poor”. For statistical purposes, numerical weights were assigned as follows:

Weight	Description
6	Excellent
5	Very Good
4	Good
3	Fair
2	Poor
1	Very Poor

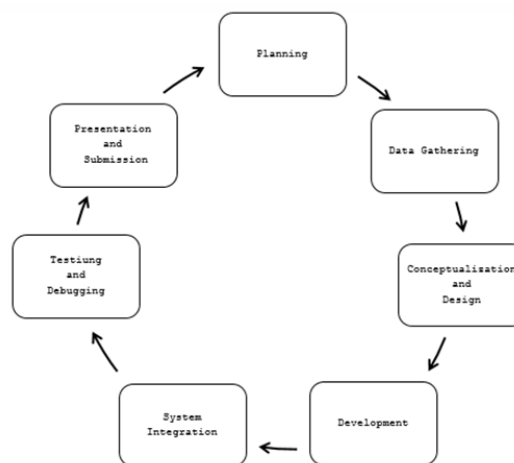


Figure 1: Conceptual Model of the system.

The iteration process was followed for the basic phases of achieving the goals and objectives of the proposed system. The process was composed of the following stages:

The first stage was the planning stage, which includes identifying the subject and the problem research and resources, scheduling of activities, and identifying the individuals to be consulted and hardware and software components to be used by the system.

The second stage was the data gathering stage in which the researchers gathered information about the subject matter to check its feasibility. The result of this stage is the raw information about the subject. A thorough examination was done of all the gathered data to ensure the relevance and significance of the information.

The third stage was the conceptualization and design which includes examination, preparation, validation, and presentation of the proposal, the conceptualization of software design, and building initial hardware design.

The fourth stage was developing the system which involves the initiation of the interface, database, system,

architecture, coding and programming language. The prototype system was designed from the data acquired in the analysis stage. The designs are also made in a child-friendly way.

The fifth stage was the system integration where the process of assembling and configuring takes place. The sixth stage was testing and debugging which was applied so that the application will be working properly in accordance to its specification and design. It was also tested repeatedly to guarantee that the process will be executed properly.

The last stage was the presentation and submission of the final output working in full functionality. Reality is one of the easiest applications to make but the Unity 3D application made it easier with the use of Cordova Android which is made especially for Augmented Reality. Prototypes are much easier compared to other tasks. In this kind of application, the researchers make the design simple yet it still aims to impress users as to what the application could do. Minimal buttons were placed to avoid making the screen messy or cluttered. Troubleshooting was done every time after errors were fixed and maintenance was done. Assembling of the document followed and every progress was recorded for observation and recording. Deployment of the application included a medium amount of coding since it focused more on its design. The application makes some of the codes function automatically. It runs mainly HTML and Javascript program files or hybrid applications.

The System

The technique converts the 3D model of Landmark tourist Spots in the municipality of Dumangas and popular tourist attractions into virtual augmented reality. Its features include using a 3D model projected in augmented reality to create a user-friendly educational application. Additionally, the system has markers that help it locate the image target. The 3D model is automatically displayed as soon as the input location is selected. As soon as the user chooses to enter the mastery game or the virtual map, the system begins. When one of the available alternatives is selected, the camera application instantly launches and scans the marker.

Technical Specifications

Software Specifications

The software is developed for the android operating system version starting from Android Nougat and up.

Hardware Specifications

The application requires an android device that has a 1 GB or higher Random Access Memory (RAM) 1.0 GHz or higher processor. The higher the RAM and Processor, the faster and better the result is obtained. A minimum of 300MHz stock or custom GPU (Graphic Processing Unit).

User Specifications

The application is accessible to android smartphone users, especially to people who want to travel to and know more about Dumangas or who those want to install the application.

Systems Implementation

Implementing the system involved testing some prototypes of the application, starting with the Dumangas' Town Plaza. The researcher assessed and evaluated applications for any Landmarks or Tourist Spots for the improvement of the system in terms of rendering display of the Landmarks or Tourist Spots and developing accuracy of marker identification to avoid lags and post errors and to perfectly provide the 3D model of the Landmarks and Tourist Spots in Dumangas using Augmented Reality.

Systems Inputs and Outputs

The system application is conceivably an optical input application. It uses the camera to input identified markers and objects relating to the system input. Input markers are logos. The output is the 3D model of the landmark, rendered on the display, which is a visual output that serves information to the user.

System Design

The system design demonstrates the interaction of the software, hardware, and the technology used in developing the system. The researchers presented the architectural design, procedural design, context diagram and database design to fully understand the proposed system.

Architectural Design

Figure 2 shows the architectural design of the system. The process starts when the user utters the word for the image shown in the device, and the system analyzes and matches the speech. If the child pronounces the word correctly, the system allows the child to proceed to the next lesson. If not, the child will repeat the process until the speech matches with the image. The application features a help option; in case the child is having a hard time determining the image. By pressing the help button, it will produce the correct pronunciation of the word. The result is based on the child's ability which includes reaction time and the number of tries.

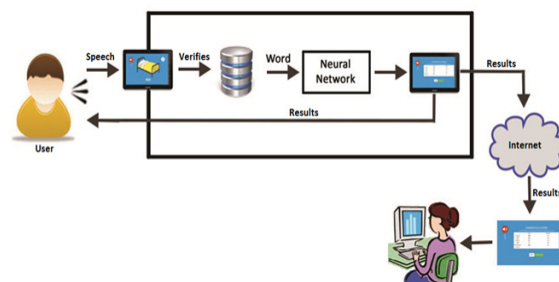


Figure 2: Architectural Design of the system.

Procedural Design

The procedural design presents the procedural description of the proposed system that shows sequence and repetition of the processes. The procedural design of the system is shown in Figures 3 and 4.

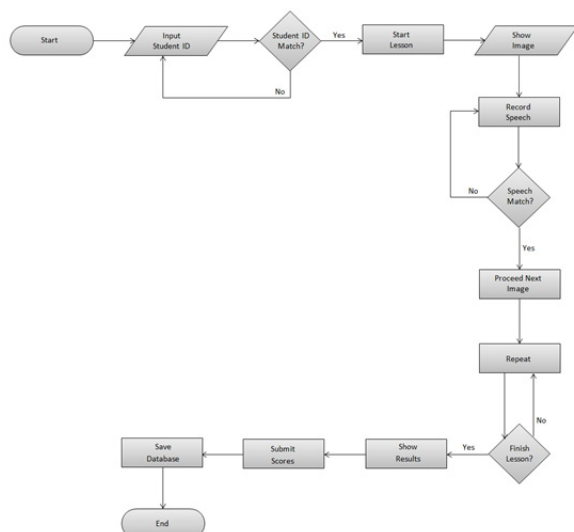


Figure 3: Procedural Design for the child.

RESULTS

The proposed system is entitled “Teacher-Aided Android Interactive Learning for Speech Delayed (SD) Child using Neural Network Algorithm and Speech Recognition”. The system's operation is divided into three sections: Speech Recognition, Speech Matching, and Speech Confirmation. Speech Recognition has an important role in the system. It is the key element in the process because it translates the speech into text. Speech matching analyzes the user's voice and checks if the speech matches the data stored in the speech database.

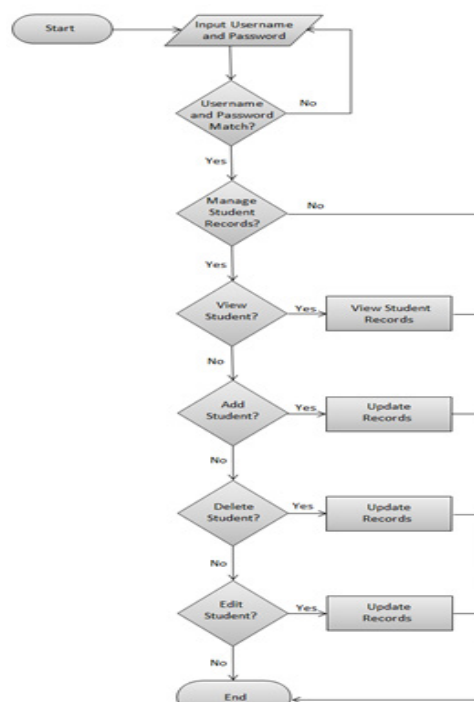


Figure 4: Procedural Design for the teacher.

Speech confirmation checks and confirms if the speech matches or not in the speech database. In the end of the learning module, the output was the result of the child's learning progress that was saved on a separate system managed by the teacher. The proposed system uses Neural Network Algorithm which falls under unsupervised learning which focuses on posterior probability distribution of unknown quantity treated as unknown variable. Posterior, in this context, is taking into account the relevant evidence related to the particular case being examined. Table 1 also shows the computation

Table 1: shows the number of days the child had taken the test

Day	Ss	sw	sp	spd	spr	rs	Rw	rp	Rpd	Rpr
1	7650	7490	0	81.6	15140	8.5	8.3	0	0	8.4
2	14529	12389	0	57.2	26910	9.1	8.3	0	0	8.7

for the scores and remarks of syllables, words, phrases, speed, and performance. Symbols were used where the “ss” variable is the total score for the syllables, the “sw” variable is the total score for the words, the “sp” variable is the total score for the phrases, the “spd” variable is the total score for speed, and the “spr” is the total score for performance. The “rs” variable is the remarks for the syllables, the “rw” variable is the remarks for the words, the “rp” variable is the remarks for the phrases, the “rpd” variable is the remarks for speed, and the “rpr” is the remarks for performance.

Each item in the different lessons is classified in three different data types, the syllables, words, and phrases. To get the score for each item, the timer is set to 1000 which is equivalent to 50 seconds. If the child pronounces the correct word, the timer will stop and the remaining time will be the score of the child.

To calculate the total score for the syllables, words, and phrases; the first item is added to the second and to the third item and so on. Below shows how the total score is calculated where “i” variable is the item for each lessons total score = $i_1 + i_2 + i_3 + \dots + i_n$

The performance is the sum of syllables, words, and phrases. To calculate the performance, “ss”, “sw”, and “sp” were added as shown in the formula below.

$$spr = ss + sw + sp$$

Speed determines the responsiveness of the child. To calculate the speed, the total number of items in a lesson is subtracted by the performance's score divided by 50 which is the allotted time for the timer as shown in the formula below.

$$spd = \frac{spd - s}{50}$$

The remarks are the ratings which determine if the child is above the passing rate or not. To calculate the remarks

for syllables, words, phrases, and performance; the total score of each data type is multiplied to 100 which is the perfect percentage of score. Then the result is divided by the total number of items of each data type in a particular lesson, which is then divided by 10 which is the perfect rating for the remarks. The result, if in decimal, is then rounded off to the nearest hundredths. The formula below shows how to calculate the said formula, where “t” variable is the total number of items in a lesson, “t1” variable is the total items for syllables, “t2” variable is the total items for words, and “t3” variable is the total items for phrases.

$$rs = | ((ss * 100 / t1) / 10) |$$

$$rw = | ((ss * 100 / t1) / 10) |$$

$$rp = | ((ss * 100 / t1) / 10) |$$

$$rpr = | ((ss * 100 / t1) / 10) |$$

To determine the remarks for speed, the “spd” is multiplied to 100 which is the value assigned for the perfect percentage of score, it is then divided by 50 which is the allotted time for the timer, and divided by 10 which is the perfect score for the ratings and subtracted by 10 which is the perfect score for the ratings. Finally, it is rounded off by the nearest hundredths place. The formula is shown below.

$$spd_remarks = | 10 - ((spd * 100 / 50) / 10) |$$

Table 2: The Child’s Progress Analysis

Variable	
Crs	9.1
Crw	8.3
Crp	0
Crspd	0
Crpr	8.7
Prs	8.5
Prw	8.3
Prp	0
Prspd	0
Prpr	8.4
Lrs	3%
Lrw	0%
Lrp	0%
Lrspd	0%
Lrpr	1.50%
Llr	0.90%

Table 2 shows the progress analysis of the child’s performance. The “crs” variable is the current remarks for the syllables, the “crw” variable is the current remarks for the words, the “crp” variable is the current remarks for the phrases, the “crspd” variable is the current remarks for the speed, and the “crpr” variable is the current remarks for the performance. Moreover, the “prs” variable is the previous remarks for the syllables, the “prw” variable is the previous remarks for the word, the “prp” variable is the previous remarks for the phrases, the “prspd” variable

is the previous remarks for the speed, and the “prpr” variable is the previous remarks for the performance.

The “lrs” variable is the learning rate for syllables, the “lrw” variable is the learning rate for words, the “lrp” variable is the learning rate for phrases, the “lrspd” variable is the learning rate for speed, the “lrpr” variable is the learning rate for performance, and the “lir” variable is the total score of learning rate. The previous score is the total remarks for the past five days that the child has taken the lesson. The current score is the total remarks for the current day that the child takes the lesson. The

Table 3: Low Learning Probability Analysis

llps	llpw	llpp	llspd	llpr
46.66%	40%	100%	100%	46.66%

“et” variable is the number of tries that the child takes the lesson. To calculate the learning probability, the current score is subtracted by the previous score, multiplied to 10, and divided by the number of tries that the child takes the lesson. The formula below shows this.

$$lrs = | (((crs - prs) * 10) / et) |$$

$$lrw = | (((crw - prw) * 10) / et) |$$

$$lrp = | (((crp - prp) * 10) / et) |$$

$$lrspd = | (((crspd - prspd) * 10) / et) |$$

$$lrpr = | (((crpr - prpr) * 10) / et) |$$

Table 3 shows the Low Learning Probability Analysis which determines the low learning probability of the child. In the table, “llps” variable is the low learning probability of the syllables, “llpw” variable is the low learning probability of words, “llpp” variable is the low learning probability of phrases, “llspd” variable is the low learning probability of speed, and “llpr” variable is the low learning probability of performance. To calculate the low learning probability, a is multiplied to, divided by a, multiplied to b, added by 1, multiplied to c, and rounded off to the nearest hundredths. This formula is shown below.

$$LLP = | ((a * b) / ((a * b) + 1) * c) |$$

In the formula, “a” variable is the score below the passing remarks, calculated as the number of scores below 7.5 which is the average multiplied to 100. Here, 100 is the perfect percentage for score divided by the number of tries that the user has taken the lesson. Moreover, “bl” variable includes the scores below 7.5 while “et” is the number of tries that the user tried the lesson. Below shows this formula.

$$a = (bl * 100) / et$$

In the formula, “b” variable is the passing score, calculated by taking the number of scores above or equal to 7.5 which is the average multiply by 100 which is the perfect percentage for score divided by the number of tries that the user has taken the lesson. On the other hand, “ab” variable includes the scores above or equal to 7.5, and “et” is the number of tries that the user tried the lesson. The formula below shows this.

$$a = (ab * 100) / et$$

The “c” variable is the score that did not pass, calculated by getting the number of mistakes, multiplied by 100 which

is the perfect percentage for score, divided by the sum of total score of the data from previous and current scores. The “np” variable includes the number of mistakes, and “ts” is the sum of total score of the data from previous and current scores. The formula below shows this.

$$a = (np * 100) / ts.$$

RESULT AND DISCUSSION

The proposed system was presented to a panel of three jurors to determine its quality. The criteria include:

1. General Technical Criteria which were based on reliability, efficiency, usability, understandability, appropriateness of feedback to user, navigation and organization.

2. General Presentation Criteria which included preparation and synthesis.

3. Specific Technical Criteria for Multimedia Technologies which were based on the content and design and the use of enhancement.

4. Specific Technical Criteria for IS and Prototype Software Systems which included correctness and integrity. The proposed system “Teacher-Aided Android Interactive Learning Tool for Speech Delayed Child using Neural Network Algorithm and Speech Recognition” was evaluated as “Passed with Revisions”. There were minor revisions recommended by the panel of jurors such as the improvement of the graphical user interface of the system, and an indicator that determines if the child has speech delay. The panel of jurors also suggested a review of the algorithm to enrich it for further enhancement of

ratings are as follows: for reliability, 4.00; efficiency, 4.00; Usability, 4.00; Understandability, 4.00; Appropriateness of Feedback to User, 3.67; Navigation and Organization, 4.00; Correctness, 4.00; and Integrity, 4.00.

Summary of Proposed System and Research Design

The android interactive learning system was proposed to help the speech-delayed child overcome speech difficulties. This system is developed using HTML, Java, and PHP as the main programming language, and was developed using Android Developer Tools. Adobe Photoshop is used on the graphics of the learning tool. Because it could be easily understood, the users are required to have basic computer skills.

The system focuses on teaching the speech-delayed child the basic knowledge of phonetics using speech recognition. The system could also store the child’s information and could be managed by the teacher who can add, edit and delete. The system also features a visual representation of the child’s progress through a performance chart which includes the child’s data results. This report helps the better organization of data. Fixing of bugs and errors regularly is required for the maintenance of the system.

CONCLUSIONS

Speech and language development is considered by experts to be a useful indicator of a child’s overall development and cognitive ability and is related to school success. Identification of children at risk for developmental delay or related problems may lead to intervention services and family assistance at a young age when chances for improvement are best. A multisensory approach learning system that counteracts all that were mentioned must be created to overcome the speech disability. An interactive learning system developed for the speech-delayed child will surely enhance their speech ability without isolating them from learning. The game-like features of the interactive learning system combined with an entertainment approach to structured phonetics interrelated with language and vocabulary development could create an effective multisensory learning environment that is both fun and effective. The researchers, therefore, concluded that the proposed system had successfully accomplished the set of objectives that were specified in the first set of the study:

1. The system was able to develop an android interactive learning tool used for assessing the speech-delayed child.

2. The system was able to produce the module for the set of lessons of the interactive learning application tool based on the consonant-acquisition chart.

3. The system was able to monitor and analyze the progress of the child using a Neural Network Algorithm. Furthermore, the system made a good start as it met its objectives mentioned in the paper.

Recommendations

To further improve the efficiency and effectiveness of the

Table 4: Evaluation Results

Criteria	Mean	Description
Reliability	4.00	Good
Efficiency	4.00	Good
Usability	4.00	Good
Understandability	4.00	Good
Appropriateness of Feedback to User	3.67	Good
Navigation and Organization	4.00	Good
Correctness	4.00	Good
Integrity	4.00	Good
Overall Evaluation	4.00	Good

Scale	Description
5.20 – 6.00	Excellent
4.30 – 5.19	Very Good
3.50 - 4.29	Good
2.70 – 3.49	Fair
1.80 – 2.69	Poor
1.00 – 1.79	Very Poor

the system. The results of the jurors’ evaluation of the system designed in this study are presented in Table 4. Based on the criteria, the grand mean score of the system is 4.00 which is labelled as “Good”. Specifically, the

system, the researchers recommend the following.

1. The system should include in the assessment of the test result the particular letter the child has difficulty in pronouncing.
2. Modification of the system into a LAN-based connection to facilitate the time consuming to access the system.
3. Further study on Neural Network Algorithm for the implementation of reliable results for the assessment of learning of a speech-delayed child.

Acknowledgement

This thesis would not have been made possible without the guidance and help of several individuals who, in one way or another have contributed and extended their valuable assistance in the preparation and completion of this study. First and foremost I would like to thank our Almighty God for giving the blessings during this research.

REFERENCES

- Bhardwaj V, Ben Othman MT, Kukreja V, Belkhier Y, Bajaj M, Goud BS, Rehman AU, Shafiq M, Hamam H. (2022). Automatic Speech Recognition (ASR) Systems for Children: A Systematic Literature Review. *Applied Sciences*, 12(9), 4419. <https://doi.org/10.3390/app12094419>
- Kamal Singh, (22 August, 2022). Post Pandemic Speech Delay In Children: What Can Parents Do. <https://stamurai.com/blog/post-pandemic-speech-delay-in-children-advice-for-parents/>
- The Nemours Foundation. (2022). Nemours Children's Health is a registered trademark of The Nemours Foundation.
- Wijayanto, P. W., Rosely, E., & Arirang, K. D. G. (2021). Implementation of an Android-Based Application as a Thematic Interactive Learning Media for English Subjects. *KnE Social Sciences*, 524-540.