Department of Computer Science and Engineering, Texas A&M University, Technical Report tamu-cs-tr-2012-8-1

Architecture of an Automated Therapy Tool for Childhood Apraxia of Speech

Avinash Parnandi*, Youngpyo Son*, Mostafa Shahin†, Beena Ahmed†, Ricardo Gutierrez-Osuna*

*Department of Computer Science and Engineering, Texas A&M University
†Department of Electrical and Computer Engineering Texas A&M University at Qatar

*{parnandi, yson, rgutier}@cse.tamu.edu †{mostafa.shahin, beena.ahmed}@qatar.tamu.edu

***Abstract*— We present a multi-tier architecture for automating the administration of speech therapy to children suffering from apraxia of speech. This architecture follows a client-server model and facilitates task-oriented remote therapeutic training in home settings. The therapy regimen is remotely assigned to the child by a speech therapist based on a standardized protocol. We utilize tablet PCs to provide stimuli to the children and record their speech response. The speech data is then streamed to a back-end server running a specialized speech-processing module to identify errors and quantify the progress of the child. These automated results allow the therapist to closely monitor the performance of each child, provide relevant feedback, and adapt the training program as needed. Our proposed architecture can accommodate a variety of interaction modalities that can serve as a complement to traditional face-to-face speech practice. In this paper we describe the client-server architecture, the middleware tools upon which the system has been built, and the speech-processing tools for automatically scoring the patients' speech.**

I. INTRODUCTION

With the advent of new technologies there is an increasing demand for innovative healthcare interventions. Children suffering from childhood apraxia of speech (CAS) are one group that can benefit immensely from these interventions. CAS is a neurological pediatric speech sound disorder (SSD) that debilitates oro-motor planning, and execution. CAS can delay acquisition of skills including the control of tone, breathing, intensity, and vocalization [3]. It also impairs the child's ability to correctly pronounce sounds, syllables, and words. CAS can thus render the child unable to start articulating the first sounds and words and can lead to a serious communicative disability.

CAS can be difficult to diagnose and monitor due to high co-morbidity and a lack of specific tools [4]. It is known that, by working intensely with a trained speech therapist, those suffering from CAS can overcome their motor planning and motor programming difficulties (articulation capabilities) [3]. However, the ratio of people suffering from CAS and requiring personalized care to the clinicians who can provide the assistance is growing at a high rate. According to the literature, current estimates of children suffering from CAS fall between 5- 6% [5]. Due to the increasing number of children needing intervention and the shortage of trained therapist, there is an ever-increasing gap between the quality and duration of needed therapeutic interventions and what is available (because of time constraints and expenses) [3]. This has created an opportunity for the development of novel technological interventions that may offer practical and cost-effective tools to complement and augment face-to-face therapy sessions. Further, as speech therapy for children with CAS usually comprises of auditory and visual interaction between a therapist and child using game-like activities [6], it may be a good candidate for technology-based alternative solutions, as these can provide not only remote and automatic monitoring but also interactive training. In this paper we present the architecture of a multi-tier system for automating the administration of CAS therapy and describe the components that comprise the framework.

Therapy for childhood apraxia comprises of two components: assessment and intervention. The assessment tests provide a baseline score of the child's current speech skills. These tests also indicate the extent to which an intervention may be effective in helping a child regain a dimension of oro-motor functionality. The therapy sessions consist of extensive practice of speech sound skills under the supervision of a speech therapist. In the system presented here, we use the Nuffield Dyspraxia Program (NDP3) for speech therapy and describe how its administration can be automated for use in in-home settings. NDP3 program is an assessment and intervention package for children with severe speech sound disorders including CAS [1, 6]. It comprises of a therapy manual, assessment procedure, and a set of picture-based therapy materials. The

Figure 1: System overview

NDP3 protocol follows a bottom-up approach and provides task-oriented speech therapy, which makes it ideal for our purposes. NDP3 can also be adapted to the child's individual needs and progress.

Our proposed system architecture consists of three main modules (Fig 1): a server (for task management, storage, and speech processing to analyze patients utterances), tablet-based mobile clients (for remote therapy), and a clinician's interface (for creating exercises and reporting). The tablets run a specially designed client application that administers the therapy session. This application prompts the child to complete predetermined exercises and records the child's utterances. Upon completion of the activity, the recorded utterance is uploaded to the server. The speech-processing unit analyzes the utterance and uploads the results onto the child's profile, which is visible to both the child's parents and the therapist. In particular, the therapist is able to analyze the performance of each child and adjust the therapy regimen as needed. To this aim, the interface allows the therapist to create new exercises from the NDP3 protocol and assign them remotely.

The paper is organized as follows. Section II provides background material on childhood apraxia of speech, reviews previous work on computer-based CAS therapy tools, and describes the NDP3 protocol. Section III describes the system components: server, mobile clients, clinician interface, and speech processing unit. Finally, Section IV discusses system validation through a series of clinic-based and in-home experimental studies.

II. BACKGROUND

A. *Childhood apraxia of speech*

The American Speech-Language-Hearing Association (ASHA) defines Childhood apraxia of speech (CAS) as a "*neurological childhood (pediatric) speech sound disorder in which the precision and consistency of movements for underlying speech are impaired in the absence of neuromuscular deficits (e.g., abnormal reflexes, abnormal tone)*" [3]. Although children with CAS usually have no damage to muscles or nerves, the area of the brain sending signals to the muscles is damaged or not fully developed, thus they have marked difficulty in motor programming, motor planning, and correctly producing sounds, syllables, and words [7]. In addition, such children often have an oro-motor dyspraxia, which is a difficulty in coordinating precise and consistent movements of the articulators (tongue, lips, jaw, and palate) required to produce speech (and to achieve an acceptable pronunciation of a given word) [8]. The speech of children with CAS is usually unintelligible to unfamiliar listeners due to phonemic speech errors and articulatory abnormalities. In the absence of treatment, this neuromuscular developmental disability can delay the acquisition of speech skills and phonological abilities, thus causing severe communicative disability [9]. Hence accurate and timely intervention is critical for children with CAS.

B. *Previous work on computerized CAS therapy*

Automated assessment and therapy is a subcategory of technological approaches to health care known as "tele-medicine" or "tele-practice". Traditional CAS therapy requires a child to undergo extended therapy sessions with a trained speech therapist in a clinic. This can be both logistically and financially prohibitive, thus paving the way for remote and automated therapy tools. Waite et al. investigated the feasibility of remote assessment of childhood SSDs and compared it with face-to-face interaction. They found high level of agreement between the two methods (single-word articulation (92%), speech intelligibility (100%) and oro-motor tasks (91%)) [10]. Other studies have shown that children with SSD/CAS have higher levels of engagement and reduced error response with computer-based intervention compared to traditional therapy [11]. Further, researchers have also shown the effectiveness of internet-based tele-rehabilitation sessions and found it comparable to clinic-based sessions [12]. These studies emphasize the impact an interactive, automated, and remote speech therapy system can have.

A number of tools have been developed for general speech assessment and therapy. General diagnostic and therapy tools include Dragon Dictate (for diagnosis of speech disorder) [13], IBM's SpeechViewer (provides parameters such as pitch, loudness, voice production, delay) [14], Speech Training, Assessment, and Remediation system: STAR (to assist therapists in treating children with articulation problems) [15], Ortho-Logo-Paedia (OLP) (intended to be used in in-home settings) [16].

In the specific context of CAS and SSD, a few software programs and tele-rehabilitation tools are available for diagnosis and treatment of neurological speech and language disorders [17]. Phoneme Factory Sound Sorter (PFSS) [18], Sound Contrasts in Phonology (SCIP) [19], and Speech Assessment and Interactive Learning Systems (SAILS) [20] are some examples. These tools train the child to develop phonological patterns and phonemic contrasts. The main drawback of these systems is the absence of automatic feedback, which makes it hard to adapt the therapy regimen on-the-fly based on the specific needs of the child.

Further, the recent advances in mobile technologies have given us great opportunities to gain richer data and provide better experience to the patients undergoing clinical intervention. Devices such as touch screen tablet-PCs are intuitive and engaging as compared to the tablet-top alternatives and present a very viable solution for in-home therapy sessions. This has led to the development of generic speech therapy applications for mobile platforms. Examples of such apps include PocketSLP [21], ArtikPix [22], and Speech with Milo [23]. These apps usually focus on articulation problems and include features such as voice recording, self-scoring, and animated stimuli. Although mobile tools can lead to greater levels of engagement, the current apps are simplified tools with no assessment or feedback capabilities. To the best of our knowledge, there are no mobile tools available which specifically targets the population suffering from CAS.

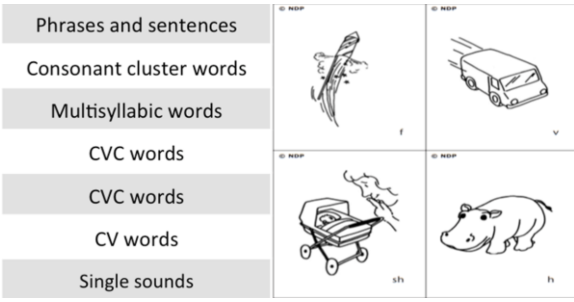

Figure 2: (a) NDP3 brick wall. (b) NDP3 exercise [1]

C. *The Nuffield Dyspraxia Programme (NDP3)*

Our proposed system is based on the Nuffield Dyspraxia Programme (NDP3), an intervention program for children with severe speech sound disorders including CAS [1, 6]. NDP3 has been designed to address the effects of CAS such as articulation of individual consonants and vowels, sequencing sounds together, and maintaining prosodic accuracy [1, 6]. In the NDP3, assessment and treatment for children with CAS is performed using a bottom-up approach; see Fig 2(a). An initial assessment provides a measure of the child's current speech skills based on which the therapy is designed, starting from isolated speech sounds and progressing to complex syllable structures, then to sentences, connected speeches, and the full range of speech [1, 6].

The NDP3 protocol requires regular therapy sessions under the supervision of a speech therapist. The NDP3 intervention approach comprises of a therapy manual, 1800 picture cards involving 450 different images and 550 line-drawn worksheets. This includes a set of picture cues to represent single consonants, vowels, diphthongs, and words at each of the phono-tactic levels; see Fig 2(b). Word creation by joining of sounds or syllables is facilitated by transition worksheets, while sequencing worksheets provide repetitive practice. Guided by the instruction manual, the picture cards are presented to the child as stimuli through tabletop games to elicit the target utterance from the child. The child is asked to produce specific sounds, syllables, or words compliant with their therapy level in these activities. On correctly completing an event the therapist usually presents the child with simple rewards. On an incorrect response, the therapist follows the instructions to assist the child so as to elicit the correct response. This approach works from the child's strengths and builds skills in incremental steps in a cumulative way [1, 6].

The NDP3 assessment procedure also follows a multi-layer approach. It relies on the production of 1) all the single consonants, vowels, and dipthongs, 2) a set of 20 single words at each phono-tactic structure (CV/VC, CVCV, CVC, CCV and multi-syllabics) through picture naming [C: consonants, V: vowels], and 3) phrases and sentences through imitation with pictures.

III. System Architecture

To enable NDP3 therapy to be administered automatically at the home, the system should be able to: (1) prompt the child with the appropriate stimuli on the tablet-pc, (2) record speech response produced by the child and stream it back to the server, (3) identify the individual consonants and vowels produced and the errors made during the therapy using the speech analysis module, (4) provide feedback to the child and the therapist, and (5) facilitate the creation or modification of therapy exercises by the therapist based on performance results. This section describes the building blocks of a system we have designed for this purpose and the software technologies that were used to build the system.

A. *Server*

The server provides logic control (task management capabilities), storage for therapy exercises and incoming speech from clients, and hosts the speech analysis module. Upon receipt of each recording, the task manager invokes the speech-processing module and stores the returned results on a centralized relational database. The database also provides storage for the profiles for each patient with their history of speech recordings, results of speech processing, clinician's assessments and annotations.

The server runs Moodle [24], an open source learning management software, which acts as the task manager in our design. Moodle's modular and object-oriented architecture facilitates information sharing, and integration with other software. It also provides course management functionalities such as user profiles, course pages, secure access by the user, and scheduling of events. Moodle is an example of a LAMP stack (LAMP: Linux, Apache, MySQL, and Perl) [2]. It comes with a web server (Apache), a database (MySQL) and a scripting interpreter (PHP), as shown in Fig 3. In our current version, the task manager and other applications run on a Linux machine with Ubuntu Server OS.

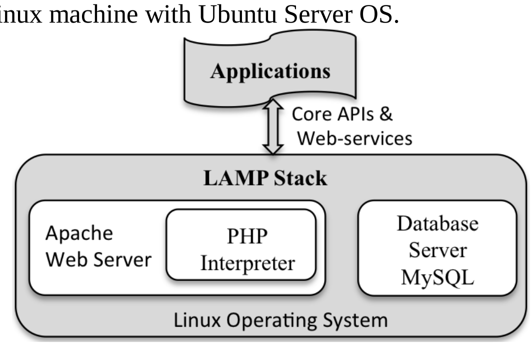

Figure 3: Moodle architecture

We extensively utilize the Moodle web-services and core APIs to integrate it in our framework; see Fig 4. Web services are functions that can be called by external applications via a number of communication protocols supported by Moodle, e.g., SOAP, XML-RPC, REST, and AMF, to mention but a few. We utilize the Representational State Transfer (REST) protocol and JavaScript Object Notation (JSON) to facilitate communication between the server and mobile clients. REST is a software architectural style for distributed hypermedia systems. REST-based web services are object (resource) oriented and implemented directly over the HTTP protocol [2]. Web services allow us to create fully parameterized generic methods and enable other systems to login to the task manager and perform operations. This is an important feature in this architecture

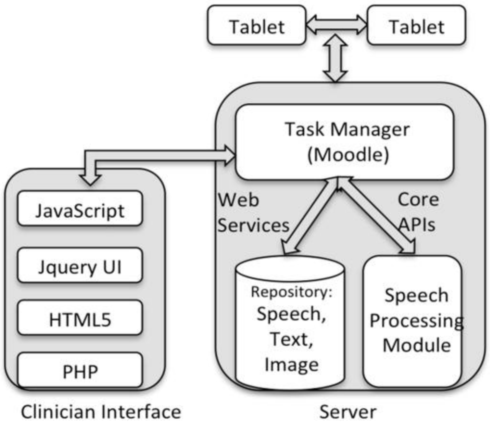

Figure 4: System architecture

because this allows the individual clients to securely transmit the speech files and invoke appropriate functions.

The Moodle core APIs provide a consistent interface to securely link it with the repositories, speech processing module, and the client systems. Our implementation uses the file, repository, external function, and data manipulation APIs to achieve the required functionalities. Using the external function and repository APIs, the repositories are exposed to external programs including the speech-processing module. We utilize the repository and data manipulation APIs for browsing, reading, and writing files. File API allows us manage files in connection to various Moodle plugins and also to list files on local file system and mount remote files to the local directory.

Our design contains a document repository, a speech repository, and an image repository. When completed, the document repository will be integrated with a version control system; it will store the XML files corresponding to the therapy exercises, results and annotation for individual students, and the version control feature will provide a logical way to store and manage XML files. Likewise, all the NDP3 therapy images will be stored in the file tables in the image repository and will be downloaded by the mobile client as required. The file table will contain a complete list of source image files ordered by a unique identifier and ground-truth IPA transcription for the word (to be used in conjunction with the speech processing algorithms).

B. *Clinician interface*

Each NDP3-based therapy exercise consists of a (fixed) number of images that serve as stimuli to the child (Fig 2(b)). These images can be combined into complete exercises by means of a simple web-based interface for the therapist; using this interface, the therapist can create new exercises or modify the existing ones and upload them to the server. The interface consists of a canvas and a table. The canvas comprises of a drag-drop box to facilitate the selection of stimuli (from a bank of stimuli) for creating therapy exercises. The second part of the interface is a dynamic HTML table that allows the clinician to specify the various therapy exercise parameters (e.g., level, duration, repetitions, sequence, frequency).

The clinician's interface is being implemented using JavaScript (Jquery), PHP, and HTML5. JavaScript is an object-oriented, client-side scripting language and it allows for the development of enhanced user interfaces and dynamic websites. Jquery is a lightweight JavaScript library that makes it easier to manipulate the document object model. The drag-drop feature has been built using HTML5 and the Jquery UI library, while the functionality for data handling and communication with the server is provided by PHP. Further, by using the HTML5 filereader API, the selected files can be stored in memory and if needed uploaded to the server using the XMLHttpRequest object.

Another component of the clinician interface provides the ability to view results and reports. This interface allows the clinician to analyze the individual performance of each patient, play recordings of individual utterances, and produce comprehensive reports with updates and summaries of the child's performance, as specified.

C. *Speech Analysis*

The speech analysis module identifies errors made in the child's utterance based upon criteria recommended by ASHA for the analysis of CAS [3, 9], and quantifies them for presentation to the therapist. The three segmental and suprasegmental features of CAS validated by ASHA are (1) inconsistent errors on consonants and vowels in repeated productions of syllables or words and differential use of a certain phoneme or sound class in different word positions (i.e. inconsistency and variability) [3], (2) lengthened and disrupted coarticulatory transitions or struggle between sounds and syllables (i.e. articulatory struggle) [9], and (3) inappropriate prosody, especially in the realization of lexical stress [7, 8]. Fig 5 shows a block diagram of the speech processing steps used to identify errors in the child's utterance. These steps include:

- **Pre-processing**: In this stage the DC offset is removed from the speech signal. The output signal is then pre-emphasized. Finally windowing is applied to segment the speech signal into frames with size 25 msec and overlapped in 15 msec.
- **Feature extraction**: Several kinds of features are extracted for each frame. The average energy and zero-crossing rate of each frame is calculated and passed to a voice activity detector. The maximum and average pitch and the peak-to-peak amplitude are also calculated and fed to a lexical stress classifier along with the average energy. Finally, Mel Frequency Cepstral Coefficients (MFCCs) are extracted and fed as inputs to a speech decoder.
- **Voice activity detector (VAD)**: A VAD is designed to differentiate between speech and non-speech (silence) segments on a frame-by-frame basis. This decision is made based upon the frame average energy and zero-crossing rate. Frames identified by the VAD as containing non-speech will indicate the presence of articulatory struggle, specifically 'groping errors' made by the child, i.e. frames where the child has struggled to produce the requirements.
- **Lexical stress classifier**: This module is used to classify the strong-weak (SW) and weak-strong (WS) stress patterns in the child speech. The classifier is based on an artificial neural network with the input feature vector consisting of the following acoustic measures: mean and maximum

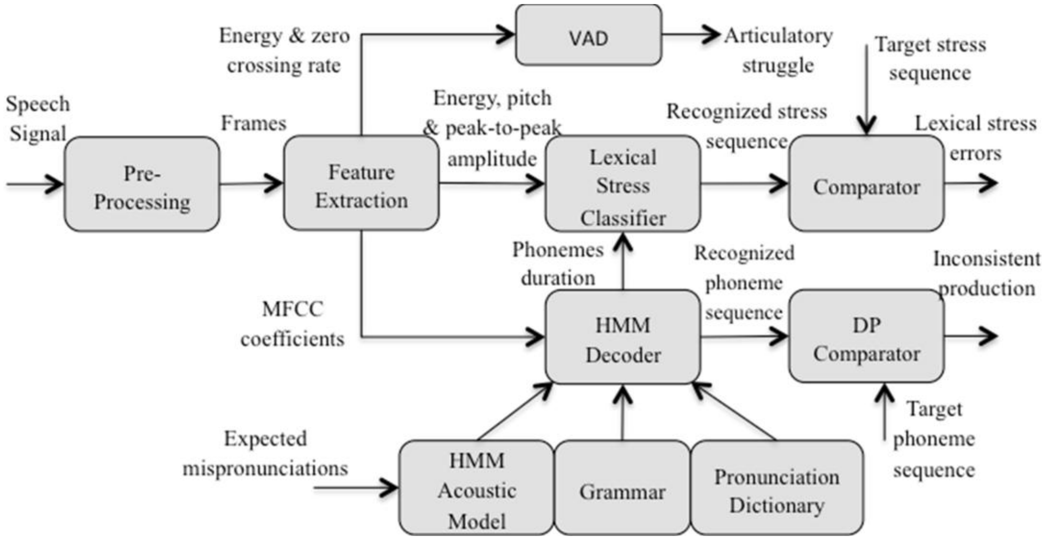

Figure 5: Block diagram of speech-processing module.

energy over nucleus (the syllable vowel), mean and maximum pitch over nucleus, peak-to-peak amplitude over nucleus, and durations of the syllable and nucleus. Duration information is obtained from the decoder, which provides the recognized phoneme sequence with time boundaries. A pairwise variability index (PVI) is calculated for each acoustic measure to determine the degree of asymmetry across pairs of neighboring syllables and to make the features speaker independent [8]. The PVI for any acoustic feature x_i is given by:

$$PVI_i = \frac{x_i^{(1)} - x_i^{(2)}}{(x_i^{(1)} + x_i^{(2)})/2}$$

where $x_i^{(1)}, x_i^{(2)}$ are the acoustic features of the first and second syllables consecutively. The classified stress patterns are then compared with the correct (target) stress patterns to identify the lexical stress errors made by the child. In a previous study, the classifier had an overall accuracy of 88.7% [25]. It correctly classified 91.1% of the SW stress patterns and 85.9% of the WS stress pattern.

- **Hidden Markov Model (HMM) based decoder.** Mispronunciations in the child's utterance are detected by means of an HMM-based decoder. For this purpose, we use a grammar lattice consisting of the correct phoneme sequence (known from the exercise given to the child) and expected mispronunciations of each phoneme (recorded by a therapist after assessment of 20 children with CAS); this lattice is used by an HMM decoder along with acoustic models to generate a sequence of phonemes from the child's utterance. The tied-state triphones HMM acoustic model was trained using around 40 hours of child speech corpus from Oregon Graduate Institute of Science and Technology (OGI) [26]. The recognized phoneme sequence is then compared to the target phoneme sequence using a Dynamic Programming-based string alignment procedure using HResults tool in the HTK toolkit [27]. Three kinds of mispronunciations are identified: insertion, deletion and substitution mispronunciations. These mispronunciations are used to identify inconsistent and variable speech produced by the child.

D. *Mobile Client*

The final module in our architecture is a specialized application for administering NDP3 therapy in in-home settings. The application consists of a child user interface, which creates an alternative interaction to face-to-face session between the therapist and the child. It also runs a background service that periodically checks the server for new therapy exercises, downloads, and administers them. The mobile application provides visual stimuli to the child and allows the recording of the child's utterance. The interface also provides feedback to the child (or her parents) once the child's utterance has been analyzed and scored by the speech-processing tools on the server; this feedback is not meant to be instantaneous but a way for the child to see progress over the course of multiple sessions.

The mobile application is being developed using the PhoneGap and jQuery framework and is shown in Fig 6. PhoneGap is an open source mobile framework that facilitates the creation of native applications with web technologies. It provides APIs to access hardware features of the device, such as network, accelerometer, microphone, storage etc. It also allows the usage of more specialized JavaScript frameworks like jQuery Mobile. In other words, the mobile application runs within a web browser and can provide a unified user interface across multiple platforms. In addition to the current interface, we are also exploring interface designs for administering the therapy sessions in a game-like format with the hypothesis that this will increase

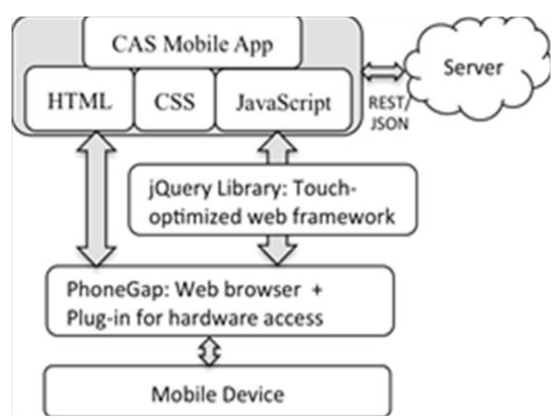

Figure 6: Mobile client architecture

the engagement levels and enhance the learning outcome.

IV. VALIDATION STUDIES

Once fully implemented, the system will be validated in two types of environments. In an initial phase, the system will be tested in the controlled environment of a speech therapist's office; this will help us assess the system's consistency across subjects and error detection accuracy. During these tests, the speech therapist will rank the usability of the system and value of the assessments provided. In a second phase, the system will be tested in in-home settings (under parent supervision). This will allow us to test the effectiveness of the system when used for remote monitoring. In both phases testing will be done with children between the ages of 6 to 10 years with mild to moderate CAS and normal language development, hearing, and vision. Before each test, a therapist will perform an assessment of the child to evaluate the child's suitability for the test. Depending on their level, the child will be given a range of activities to perform as per the NDP3 protocol. The performance metric as measured by the system in both phases will be cross validated by a therapist. Through these studies, we will assess the usability of all components of the framework and the accuracy of the speech processing algorithms. We will also seek feedback from the adults supervising these tests to determine the interest and comfort level of the child while performing the exercise.

V. CONCLUSION

Automated systems that can monitor and facilitate practice of speech skills have the potential to offer affordable tools to complement and augment the traditional approaches of face-to-face therapy. In this paper we have outlined and described a multi-tier system architecture for conducting automated speech therapy sessions for childhood apraxia of speech. We reviewed CAS, its salient features and existing work on therapy tools for CAS, including the NDP3 therapy protocol. We described the different software modules, including our speech-processing approach for automated scoring, upon which our system is being implemented. Finally we proposed an experimental protocol that will be used to validate the system.

ACKNOWLEDGMENT

This work was made possible by NPRP grant # [4-638-2-236] from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] *Nuffield Dyspraxia Programme*. Available: http://www.ndp3.org/dyspraxia-therapy-overview.html

[2] J. Moore and M. Churchward, *Moodle 1.9 Extension Development*: Packt Publishing, 2010.

[3] "ASHA Ad Hoc Committee on Apraxia of Speech in Children," *American Speech-Language-Hearing Association,* p. 75, 2007.

[4] D. Newbury and A. P. Monaco, "Genetic Advances in the Study of Speech and Language Disorders," *Neuron,* p. 309-20, 2010.

[5] R. Gaines, C. Missiuna, M. Egan, and J. McLean, "Educational outreach and collaborative care enhances physician's perceived knowledge about Developmental Coordination Disorder," *BMC Health Services Research,* vol. 8, pp. 1-9, 2008.

[6] P. Williams and H. Stephens, "Nuffield Centre Dyspraxia Programme," in *Interventions for Speech Sound Disorders in Children*, A. L. Williams, et al., Ed., ed: Brookes Pub., 2010.

[7] L. D. Shriberg, T. F. Campbell, H. B. Karlsson, R. L. Brown, J. L. Mcsweeny, and C. J. Nadler, "A diagnostic marker for childhood apraxia of speech: the lexical stress ratio," *Clinical Linguistics & Phonetics,* vol. 17, pp. 549-574, 2003.

[8] K. J. Ballard, D. A. Robin, P. McCabe, and J. McDonald, "A Treatment for Dysprosody in Childhood Apraxia of Speech," *J Speech Lang Hear Res,* vol. 53, pp. 1227-1245, 2010 2010.

[9] K. Forrest, "Diagnostic Criteria of Developmental Apraxia of Speech Used by Clinical Speech-Language Pathologists," *Am J Speech Lang Pathol,* vol. 12, pp. 376-380, 2003.

[10] M. Waite, L.Cahill, D. Theodoros, S. Busuttin, T. G. Russell, "A pilot study of online assessment of childhood speech disorders"*Journal of Telemedicine and Telecare,* p.92-94, 2006.

[11] D. G. Jamieson, G. Kranjc, K. Yu, and W. E. Hodgetts, "Speech intelligibility of young school-aged children in the presence of real-life classroom noise," *J Am Acad Audiol,* pp. 508-17, 2004.

[12] G. A. Constantinescu, D. G. Theodoros, T. G. Russell, E. C. Ward, S. J. Wilson, and R. Wootton, "Home-based speech treatment for Parkinson's disease delivered remotely: a case report," *J Telemed Telecare,* vol. 16, pp. 100-4, 2010.

[13] H. Kolles and W. Feiden, "Computer-assisted speech recognition in diagnostic pathology. Development of the DragonDictate-30 " *Pathologe,* pp. 439-42, 1995.

[14] S. R. Pratt, A. T. Heintzelman, and S. E. Deming, "The Efficacy of Using the IBM Speech Viewer Vowel Accuracy Module to Treat Young Children With Hearing Impairment," *J Speech Hear Res,* vol. 36, pp. 1063-1074, 1993.

[15] H. T. Bunnell, D. M. Yarrington, and J. B. Polikoff, "STAR: Articulation Training for Young Children," in *International Conference on Spoken Language Processing*, 2000, pp. 85-88.

[16] A. M. Oster, D. House, A. Protopapas, and A. Hatzis., "Presentation of a new EU project for speech therapy: Ortho-Logo-Paedia," in *Proceedings TMH-QPSR, Fonetik* 2002.

[17] A. C. Georgeadis, D. M. Brennan, L. N. Barker, and C. R. Baron, "Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders," presented at the Clinical Aphasiology Conference, 2003.

[18] Y. Wren, S. Roulstone, and A.L. Williams, "Computer-Based Interventions," in *Interventions for Speech Sound Disorders in Children*, S. M. e. al, Ed., ed: Brookes Pub, 2006.

[19] A. L. Williams, "Multiple Oppositions Intervention," in *Interventions for Speech Sound Disorders in Children*, M. S and e. al., Eds., ed Brookes Pub, 2006.

[20] S. Rvachew, and F. Brosseau-Lapre, "Speech perception intervention," in *Interventions for Speech Sound Disorders in Children*, S. M. e. al, Ed., ed Brookes Pub, 2006.

[21] *Pocket SLP*. Available: http://pocketslp.com/

[22] *ArtikPix*. Available: http://rinnapps.com/artikpix/

[23] *Speech With Milo*. Available: http://www.speechwithmilo.com/

[24] *Moodle*. Available: http://moodle.org

[25] M Shahin, B Ahmed, K Ballard, J McKechnie, E Murray, "Classification of Lexical Stress in Speech Affected by Childhood Apraxia" submitted to 14th Australasian Conference on Speech Science & Technology (SST) 2012.

[26] K. Shobaki, J. P. Hosom, and R. A. Cole, "The OGI kids' speech corpus and recognizers," presented at the International Conference on Spoken Language Processing, Beijing, 2000.

[27] S. J. Young, G. Evermann, M. J. F. Gales, T. Hain, D. Kershaw, G. Moore, and J. e. a. Odell, *The {HTK} Book, version 3.4*: Cambridge University, 2006.