Jiang, B. et al. (Eds.) (2025). Proceedings of the 33rd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

SamurAI: AI-Powered Learning App for Speech Therapy in Developing Economies

Christina SIJERA*, Nicole BUERE, Franco CARINO, Caleb RUBIO & Ryan EBARDO

College of Computer Studies, De La Salle University, Philippines
*christina_sijera@dlsu.edu.ph

Abstract: This paper presents SamurAI, an AI-powered application designed to enhance home-based speech interventions for children with phonological and articulation disorders, particularly in developing economies. Built to complement rather than replace clinical expertise, it enables therapists to assign customized exercises, monitor progress remotely, and provide targeted feedback. Using speech recognition, SamurAI delivers clear instructions, interactive tasks, and real-time feedback, allowing caregivers to guide practice effectively at home. Its therapeutic content follows established speech therapy frameworks to ensure clinical accuracy and is developed through a user-centered process involving therapists and parents. Beyond supporting individual language development, SamurAI fosters educational readiness, lasting communication skills, and scalable therapist-led interventions in resource-limited settings.

Keywords: speech therapy, speech intervention technology, phonological disorders, learning app, speech technology

1. Introduction

Communication stands as a cornerstone of human interaction, enabling us to articulate our thoughts and emotions while forging meaningful connections and enriching our relationships. Early language development is critical to a child's overall growth. Building strong communication skills at a young age lays the foundation for social, emotional, and cognitive well-being. Gaps and struggles in language development can hinder a child's ability to form relationships, understand their surroundings, and express their needs, all often leading to challenges in academic performance, literacy acquisition, and overall educational progress.

Language proficiency is fundamental in navigating educational environments. It shapes how children comprehend instructions, participate in curriculum content, and engage in academic settings. Beyond academic skills like reading and writing, language supports critical thinking and problem-solving abilities that drive learning across subjects (Beard, 2018; Sherred, 2021). In events where children face speech or language difficulties, their participation and confidence in school may suffer, leading to gaps not only in knowledge but also in social integration and motivation (Wren et al., 2021). This underscores the necessity for early identification and support in fostering equitable learning opportunities for all students.

In the Philippines, there is limited availability of qualified licensed speech therapists, especially outside major urban centers. With over 250 licensed speech-language pathologists (SLPs) concentrated primarily in Metro Manila and surrounding regions, many areas remain underserved. This limitation results in prolonged waiting times for diagnosis and therapy, as well as overburdened therapists (Philippine Association of Speech Pathologists [PASP], 2019; Suh et al., 2024). This creates barriers to timely intervention, which is critical during early childhood when speech development is most shapeable.

Prior investigations support the efficacy of technology use to improve the learning experiences of students in developing economies (Valderama 2024a, Valderama 2024b). Albeit not widespread, the use of technology to address speech therapy challenges is not new; several applications already support children with speech and language difficulties through interactive exercises, feedback, and communication aids, with some incorporating AI for early diagnosis and progress tracking. Mobile apps such as those developed by Tommy and Minoi (2016) show positive results when parents are involved, though customization and guidance remain issues, while popular tools like SpeakEasy and Articulation Station Hive provide useful features but often lack fully personalized plans or seamless integration between clinical and home-based practice. SamurAI distinguishes itself by combining AI-powered speech recognition with therapist-curated interventions, maintaining strong therapist partnership in line with speech therapy laws while providing immediate, AI-driven support to caregivers. This ensures a more personalized, accessible, and effective home-based therapy experience that seamlessly complements clinical care.

2. Methodology

This chapter outlines the design and development of the SamurAI app. It discusses the user-centered design process involving therapists and parents, highlights the key challenges in home-based therapy, explains the AI integration, presents the therapy frameworks applied, reviews the current app features, and describes the planned stages for future testing.

2.1 User-Centered Design Approach

The development of SamurAI follows a user-centered design (UCD) approach, emphasizing collaboration with target app users, primarily speech-language pathologists (SLPs) and parents, throughout the process. Parents and therapists served as active collaborators during the requirements and design phase of the project; however, for the development stage of the project, the therapists serve as more active collaborators to ensure that the application aligns with clinical standards and ethical considerations, while minimizing any risk of experimental impact on child development. Feedback loops are established via bimonthly interviews with licensed speech pathologists to refine features and workflows before it will be tested or presented to parents and their children. This approach ensures that the tool remains effective and practical towards the support for both therapist guidance and parental improvement.

2.2 Need-Finding Assessment

Extensive research and interviews were conducted to assess and reveal the significant challenges faced in home-based speech therapy. The results of the findings are as follows:

1. Limited Communication between Therapists: Studies (Klatte et al., 2023; Sugden et al., 2019) and interviews highlight that communication is often brief and insufficient. This leaves parents or caregivers uncertain on home intervention execution.
2. Fragmented Progress Tracking: In current practice, there is no standardized system for tracking speech development at home; this makes it difficult for therapists to monitor progress and adjust the interventions accordingly (Pritchard et al., 2024).
3. Information Overload from Online Resources: Studies show that parents have reported feeling overwhelmed by inconsistent and unfiltered online resources, complicating their

ability to select the activities that their children need (Kleim & Jones, 2008, as cited in McLeod & Baker, 2017).

4. Difficulty Identifying Speech Errors: Many parents struggle to recognize the speech errors and delays without professional guidance; this delays the interventions needed (McLeod & Harrison, 2009; Harel et al., 2015).

2.3 AI Integration

SamurAI integrates AI technology to assist in speech therapy without replacing professional judgment. The main purpose of using AI is to provide parents with immediate feedback to better understand their child's development and become more engaged in therapy, while also summarizing at-home progress for therapists to improve tracking. OpenAI's GPT 4.0 was selected for speech recognition due to its affordable availability during early development and its capability to transcribe speech into phonemes without auto-correction, preserving user input. While the model is not custom trained, the developers have ensured that the prompt provided focuses on analysis on target sounds, enabling focused error detection. The app also provides disclaimers to manage parental expectations and to ensure that parents do not feel fully dependent on AI for their child's therapy. Future work for SamurAI explores additional AI technologies and model customization to enhance accuracy and scalability.

2.4 Speech Therapy Frameworks

The development of SamurAI has followed several established and evidence-based speech therapy frameworks to ensure that intervention design and progress tracking are clinically efficient, targeted, and measurable. The assessment makes use of the SODA system, which stands for substitution, omission, distortion, and addition. This helps clinicians identify substitution patterns and guide therapy goals. The assessment also makes use of the PPA system, which stands for phonological process analysis. This method helps in identifying patterns of speech sound errors and provides a deeper understanding of the child's phonological system. The assessment also uses the Place-Manner-Voicing (PMV) system, which classifies consonant sounds by where and how they are produced; this provides insights into sound errors that may not be easily captured through other analyses. In addition, the Percentage of Consonants Correct (PCC) measure is used, which is a widely recognized metric that quantifies the accuracy of consonant production and gives a reliable indicator of the severity of speech sound disorders. Finally, the interventions were created with the application of different therapy approaches, one of which being Minimal Pairs, which uses pairs of words that differ by a single sound to highlight the phonemic distinctions. Maximal Opposition is also applied, which uses word pairs with maximally different sounds; this is to stimulate broad phonological changes. Minimal pairs and maximal opposition are applied throughout the activities, while the app offers activity types ranging from sound discrimination to production to also incorporate the traditional articulation framework. By embedding these validated frameworks, SamurAI ensures that its therapy is accessible, engaging, and clinically rigorous.

2.5 Application Development

As of the current time, approximately 80% of the core functionalities of SamurAI have been developed. This progress includes the following key features:

Speech Interventions Module

The app currently features four core activity types designed to address progressive stages of speech intervention: Sound Awareness, Sound Imitation, Sound Production, and Speech Exercises.

In Sound Awareness, two distinct modes are provided to target either articulation or phonological disorders. For phonological disorders, users are presented with both visual and auditory cues and must select the word containing the target sound; in contrast, the articulation mode focuses on a single word, prompting the user to judge whether the target sound is present. As seen in *Figure 1,* these two modes are visually differentiated to clarify their therapeutic objectives.

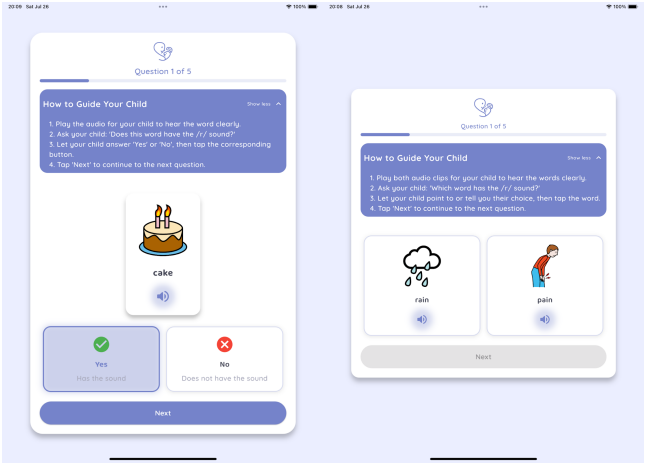

Figure 1. Two Distinct Modes of Sound Awareness

Sound Imitation and Sound Production share a similar format but differ in level of independence—Sound Imitation provides auditory cues for guided practice, while Sound Production requires the user to generate the sound without prompts. As shown in Figure 2 and Figure 3, this distinction is emphasized in the interface design to highlight the shift from guided repetition to independent articulation. Both activities integrate AI analysis, presenting expected versus actual transcriptions, identifying detected errors, calculating the PCC score, and generating an overall performance analysis.

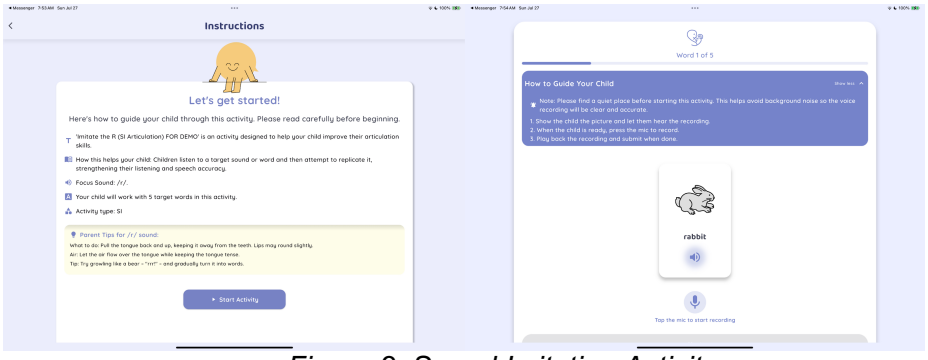

Figure 2. Sound Imitation Activity

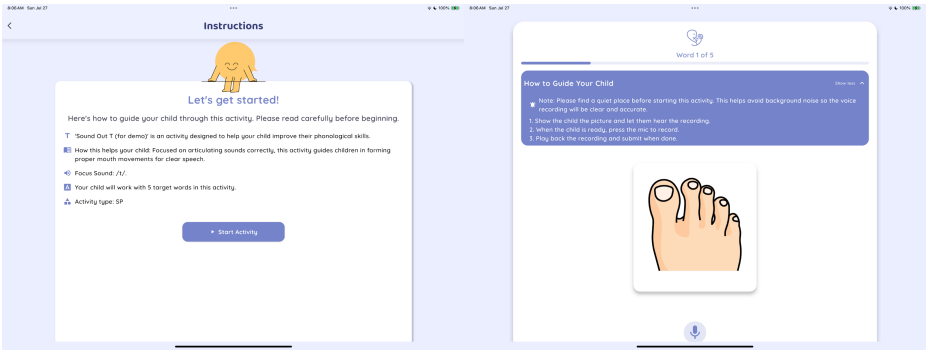

Figure 3. Sound Production Activity

Finally, Speech Exercises target spontaneous speech by prompting users with words containing the target sound and encouraging them to describe or talk about them in detail. As seen in Figure 4, each activity is visually structured to provide clear instructions, progress indicators, and real-time feedback, ensuring that both caregivers and therapists can monitor performance effectively.

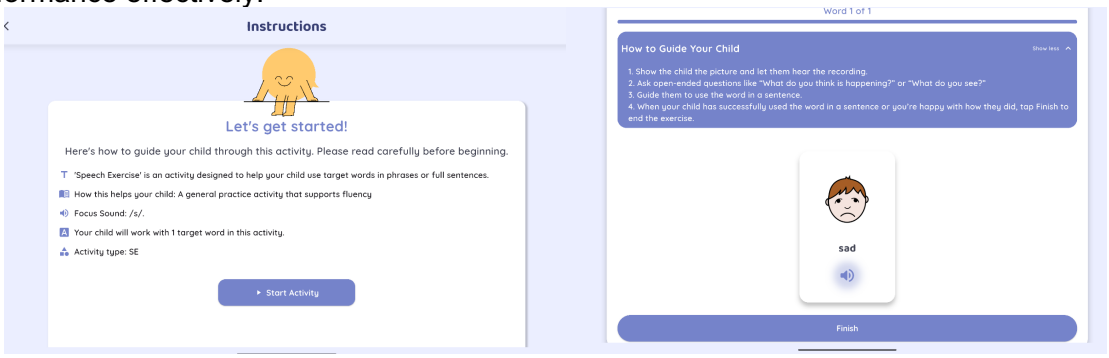

Figure 4. Speech Exercise Activity

2.5.1 Therapy Goals & Intervention Management Module

For Therapy Goals, this feature is designed to help therapists track progress more effectively. As shown in Figure 5, therapists can set individualized goals for their users. They also have the authority to assign specific interventions based on their clinical judgment of the child's progress. They can select from different activity types, set the required duration and repetitions, and provide additional notes to guide parents.

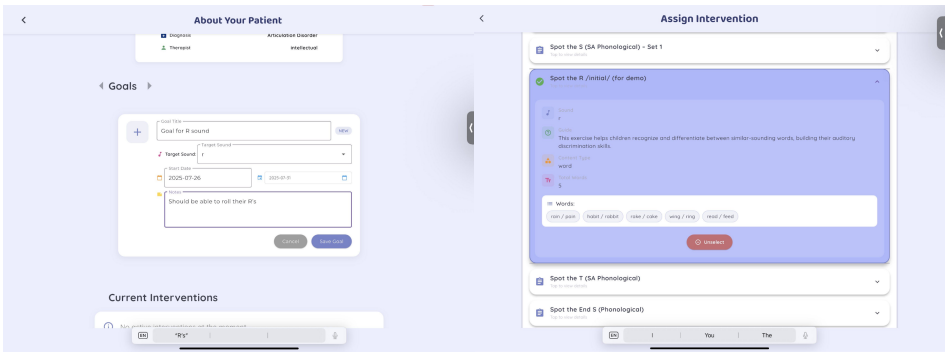

Figure 5. Assigning Therapy Goals and Interventions

Therapists can also view each patient's response to every activity, as seen in Figure 6. They can replay what the child said and identify which words the child struggled with the most, and see which activities took the longest to complete. The AI-generated feedback provided to parents is also visible to therapists. Furthermore, therapists are empowered to refine or override the AI's analysis if it does not meet their professional standards, ensuring that the feedback remains clinically accurate and tailored to the child's needs

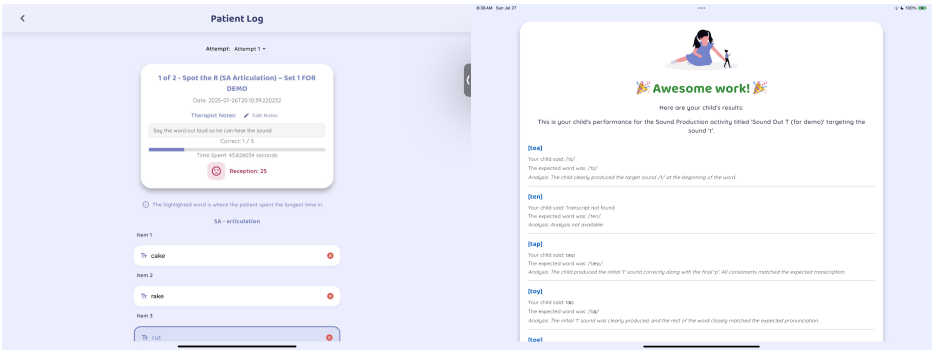

Figure 6. Assigning Therapy Goals and Interventions and AI-Generated Feedback

2.5.2 Dashboard & Analytics Module

The Dashboard & Analytics Module provides a comprehensive view of the patient's performance, with the option to filter results by week, month, or year. It presents insights, such as the most common soda error, which helps identify recurring articulation challenges, and performance per activity type to track progress in specific tasks. Trouble words are highlighted to guide focused interventions, while the general PCC (Percentage of Consonants Correct) provides an overall measure of speech accuracy. Data on total time spent and the usual time of practice provide context for engagement levels. At the same time, notes on the child's typical reception of the intervention offer qualitative insight into participation and motivation. An AI-generated summary condenses all this data for quick review, aiding both caregivers and therapists in decision-making. As seen in Figure 8, the module is visually designed for clarity, enabling efficient tracking and monitoring of progress over time.

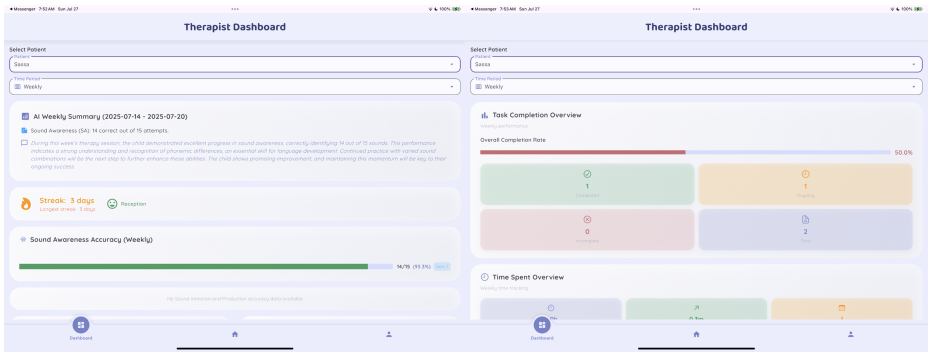

Figure 8. AI-Generated Feedback

3. Initial Feedback and Results

In recent feedback sessions, the two collaborating speech-language pathologists (SLPs) expressed their satisfaction with the application's design and the integration of its intervention features. They confirmed that these interventions align closely with their existing clinical practices and found the therapy planning and customization options particularly valuable.

In addition to their approval, the SLPs also offered constructive suggestions for further enhancing the application. They recommended randomizing the items within each intervention attempt to create a more dynamic learning experience and proposed an option to hide target words to increase engagement and support children's retention. These insights are currently under consideration for integration in the next development cycle.

3.1 Feedback Collection Process and Incorporation of Feedback

Feedback is collected through bimonthly online sessions with speech-language pathologists (SLPs), complemented by follow-up chats for clarifications. With the participants' consent, these sessions are documented to preserve insights for careful review. This structured approach facilitates the gathering of professional input while promoting open communication between meetings. The feedback obtained has informed crucial refinements, including improvements to the Therapy Goals & Intervention Management Module to better align with actual workflows. By systematically integrating these suggestions, the development team ensures that the application remains clinically reliable and well-received while also preparing for future testing involving parents and children.

3.2 Preliminary AI Accuracy Testing

To assess the reliability of the AI component, structured tests were conducted using pronunciation references such as the CMU Pronouncing Dictionary and the Cambridge Learner's Dictionary, with a focus on U.S. pronunciation relevant to clinical speech therapy. The AI analyzed patient speech recordings to identify specific target sounds during exercises, such as the /s/ sound in the word "sad." The primary objectives were to capture the child's speech corrections accurately and to classify errors using frameworks such as SODA and PPA. Initial results indicated that the AI successfully transcribed speech, detected errors, and generated performance metrics such as the Percentage of Consonants Correct (PCC), demonstrating its potential as a reliable tool for therapy while allowing clinicians to make final assessments.

4. Discussion and Conclusion

The app is designed for parents of children with diagnosed speech-sound disorders aged 4–7, the children themselves for voice analysis, and licensed Speech-Language Pathologists (SLPs) as end users. The involvement of SLPs is necessary to ensure compliance with the Philippine Speech-Language Pathology Law (RA 11249), which requires that interventions be supervised by licensed professionals. The app focuses only on functional speech-sound disorders, specifically phonological and articulation motor disorders, and is best suited for mild to moderate cases. More complex disorders are excluded, as they require specialized interventions beyond their current scope. Moreover, this narrow scope, focusing on mild-to-moderate articulation and phonological disorders in children aged 4–7, was intentional to minimize clinical risk, standardize tasks, and establish feasibility. And accordingly, the findings should not be generalized to severe disorders, fluency, resonance, or older cohorts.

The use of AI is also limited to speech transcription, phoneme analysis, and report generation. While it provides useful insights, its analysis should be treated as supplemental to an SLP's evaluation and may be affected by device quality and recording conditions. Parents are encouraged to conduct sessions in quiet environments to improve accuracy. Finally, the app does not yet include ways to measure whether it can enhance or replace face-to-face therapy. Long-term outcomes and comparisons with traditional therapy remain outside the present scope. SamurAI is intended only to support home-based interventions and provide structured guidance, not to replace in-person treatment by SLPs. AI functions operate as decision support under clinician supervision and are not intended for diagnosis or autonomous therapy planning; all therapeutic decisions should remain with the SLP, and complex error patterns may be under-detected at this stage. This is a proof-of-concept focused on technical feasibility and workflow fit. There was no conducting of randomized or longitudinal outcome studies in comparison to standard care. Therefore, the paper does not

make claims of therapeutic efficacy or superiority. Effect sizes and adherence remain unmeasured.

The current build supports English only. The languages were limited in order to leverage validated English assessment materials and readily available audio speech recognition resources during the early stages of the deployment. Cross-linguistic phonology, Tagalog or Filipino support, and other Philippine languages will require separate datasets, rules, and validation. Early feedback primarily involved SLPs, and broad-based testing with parents and children is limited. Usability, engagement, caregiver burden, and at-home adherence have not yet been comprehensively evaluated across households. Ethical treatment in this paper is scoped to consent, data handling, and keeping the clinician in the loop. Broader topics are acknowledged but not deeply analyzed here and warrant dedicated study.

In the future, SamurAI aims to enhance AI integration by adopting more advanced and better-suited models, which can be trained using specialized datasets such as recordings of children's speech. This would improve the accuracy of transcription and error detection, making feedback more reliable for both parents and therapists. Additionally, the app could expand to support larger trials and reach a broader user base, including more children with speech-sound disorders and more licensed speech-language pathologists. Another planned improvement is the inclusion of multiple languages; while the app currently supports English, future versions aim to add Tagalog support to accommodate Filipino users, ensuring therapy accessibility and usability across linguistic contexts.

The solution is positioned to widen access to speech therapy by lowering barriers of location, scheduling, and cost. Therapist feedback highlights that automated features, such as generating practice exercises, tracking progress, and delivering instant feedback, will streamline routine tasks so attention can shift to more personalized interventions. Crucially, this automation is intended to assist rather than replace, as clinical judgment remains central, with the application showing insights while leaving decisions to licensed professionals. In addition, AI-driven personalization is well suited to digital learning contexts where steady practice and tailored activities are essential, as it adapts tasks to each learner's pace by detecting subtle performance patterns and supplying target cues that make sessions more efficient and engaging. For children and individuals with unique speech and language needs, this translates into more accessible, responsive, and flexible opportunities, which helps to bridge gaps that traditional therapy alone may struggle to close. And as the technology continues to mature, AI-assisted tools like this application aim to empower both therapists and learners, extending reach while preserving the expertise that guides effective communication development.

Acknowledgements

We thank the Speech-Language Pathologists (SLPs) for their essential insights and the parents who participated in our need-finding interviews. We also extend our gratitude to De La Salle University for providing the platform to conduct this work as part of our academic requirements.

References

ASHA. (2017). Speech Sound Disorders: Articulation and Phonology. Asha.org. https://www.asha.org/practice-portal/clinical-topics/articulation-and-phonology/#collapse_2

Beard, A. (2018). Speech, language and communication: a public health issue across the lifecourse. Paediatrics and Child Health, 28(3), 126–131. https://doi.org/10.1016/j.paed.2017.12.004

Fabiano-Smith, L., & Hoffman, K. (2017). Diagnostic accuracy of traditional measures of phonological ability for bilingual preschoolers and kindergarteners. Language Speech and Hearing Services in Schools, 49(1), 121–134. https://doi.org/10.1044/2017_lshss-17-0043

Harel, D., Byun, T., & Hitchcock, E. (2015). Social, emotional, and academic Impact of Residual Speech Errors in School-Aged Children: a survey study. Seminars in Speech and Language, 36(04), 283–294. https://doi.org/10.1055/s-0035-1562911

Keffala, B., Scarpino, S., Hammer, C. S., Rodriguez, B., Lopez, L., & Goldstein, B. (2020). Vocabulary and Phonological Abilities Affect Dual Language Learners' Consonant Production Accuracy Within and Across Languages: A Large-Scale Study of 3- to 6-Year-Old Spanish–English Dual Language Learners. American Journal of Speech-Language Pathology, 29(3), 1196–1211. https://doi.org/10.1044/2019_ajslp-19-00145

Klatte, I. S., Ketelaar, M., De Groot, A., Bloemen, M., & Gerrits, E. (2023). Collaboration: How does it work according to therapists and parents of young children? A systematic review. Child Care Health and Development, 50(1). https://doi.org/10.1111/cch.13167

McLeod, S., & Harrison, L. J. (2009). Epidemiology of Speech and Language Impairment in a Nationally Representative Sample of 4- to 5-Year-Old Children. Journal of Speech, Language, and Hearing Research, 52(5), 1213–1229. doi:10.1044/1092-4388(2009/08-0085)

Philippine Association of Speech Pathologists (PASP). (2019). 2019 Filipino SLP Survey. ResearchGate. https://doi.org/10.13140//RG.2.2.22967.04006

Pritchard, K., Stojanovik, V., Titterington, J., & Pagnamenta, E. (2024). How speech and language therapists and parents work together in the therapeutic process for children with speech sound disorder: A scoping review. International Journal of Language & Communication Disorders. https://doi.org/10.1111/1460-6984.13132

Sherred, L. (2021). How Speech and Language Deficits Can Affect a Child's Academic Success. Getting Smart. https://www.gettingsmart.com/2021/01/20/how-speech-and-language-deficits-can-affect-a-childs-academic-success/

Sugden, E., Munro, N., Trivette, C. M., Baker, E., & Williams, A. L. (2019). Parents' experiences of completing home practice for speech sound disorders. Journal of Early Intervention, 41(2), 159–181. https://doi.org/10.1177/1053815119828409

Suh, H., Aayushi Dangol, Hedda Meadan, Miller, C. A., & Kientz, J. A. (2024). Opportunities and Challenges for AI-Based Support for Speech-Language Pathologists. https://doi.org/10.1145/3663384.3663387

Tommy, C. A., & Minoi, J.-L. (2016). Speech therapy mobile application for speech and language impairment children. Unimas Institutional Repository (Universiti Malaysia Sarawak), 199–203. https://doi.org/10.1109/iecbes.2016.7843442

Vollmer, E. (2023). Phonological Processes | TherapyWorks. TherapyWorks. https://therapyworks.com/blog/language-development/phonological-processes/phonological-processes/

Williams, L. (2018). Working Smarter Not Harder: The Place-Voice-Manner Analysis | Medbridge. Medbridge. https://www.medbridge.com/blog/working-smarter-not-harder-the-place-voice-manner-analysis

Wren, Y., Pagnamenta, E., Peters, T. J., Emond, A., Northstone, K., Miller, L. L., & Roulstone, S. (2021). Educational outcomes associated with persistent speech disorder. International journal of language & communication disorders, 56(2), 299–312. https://doi.org/10.1111/1460-6984.12599

Valderama, A. M. C. (2024, November). Leveraging AI-Powered Virtual Meeting Summaries: Towards an Evidence-Based Classroom Observation Assessment. In *International Conference on Computers in Education*.

Valderama, A. M. C. (2024, November). Transforming Student Feedback into Institutional Action Plans: A Data-Driven Approach. In *International Conference on Computers in Education*.