Interspeech 2024
1-5 September 2024, Kos, Greece

The Speech Motor Chaining Web App for Speech Motor Learning

Jonathan L Preston[1], *Nina R Benway*[2], *Nathan Prestopnik*[3], *& Nathan Preston*[4]

[1]Syracuse University, Syracuse, NY, USA
[2]University of Maryland, College Park, MD, USA
[3]Shenandoah University, Winchester, VA, USA
[4]SpeechLink, Syracuse, NY, USA
jopresto@syr.edu, benway@umd.edu, nprestop@su.edu, nate.preston@speech-link.com

Abstract

Systematic review has indicated that existing computerized speech learning apps show unclear potential to invoke speech change, highlighting the need to incorporate tested theoretical principles in app design. As speech is a motor skill, this paper illustrates how the principles of motor learning underlie the *Speech Motor Chaining* speech therapy web app. Features such as game design, telepractice, and speech analysis are discussed.

Index Terms: speech motor learning, gamification

1. Introduction

There is a clearly identified potential for computer assisted pronunciation training (CAPT), including for second language learning and speech therapy. Systematic review, however, has repeatedly indicated that many existing CAPT tools have unclear potential to invoke speech change [e.g., 1]. This highlights the need to build CAPT applications upon validated theoretical principles. Therefore, this paper reviews the theoretical underpinnings of *Speech Motor Chaining*, a CAPT tool designed for children age 7+, whose efficacy is backed by multiple clinical trials evidencing speech motor learning for speech sound disorders [for review: 2]. *Speech Motor Chaining* is now available as a web app with options for gameplay.

2. Theoretical Background

2.1. Principles of motor learning

Because speech is a motor skill, the main theoretical motivation for *Speech Motor Chaining* arises from the *principles of motor learning* [3]. This theory differentiates appropriate short-term practice and feedback principles based on recent learner performance, emphasizing either the *acquisition* of a speech motor plan (i.e., production of a new phone in practiced contexts) or generalized *learning* (i.e., production of a new phone in unpracticed contexts).

Table 1: *Principles of motor learning [3] adapted within Speech Motor Chaining.*

Principle	Goal: Speech Acquisition	Goal: Speech Learning
Target complexity	Simple	Complex
Practice Schedule	Blocked	Random
Practice variability	Lower	Higher
Feedback frequency	Higher	Lower
Feedback type	Detailed	Binary/Summary

2.2. Principles of gamification

Because motor learning theories emphasize motivation [4], *Speech Motor Chaining* includes optional gamified speech tasks (*Treasure Chain*, Section 5). Gamification [5] is believed to increase motivation [6] by turning at least some aspects of an activity that is not inherently playful into a form of play. The gameplay of *Treasure Chain* is designed to increase dopaminergic activity in players through a reinforcement learning-based reward loop, which shows positive impacts on player motivation for both game and task [7]. *Treasure Chain* also uses the fantasy setting of a treasure hunt, the sensory and cognitive curiosity of exploration mechanics, and elements of challenge, control, and choice to further motivate players [8].

3. Technical Description

3.1. System Architecture

Speech Motor Chaining has evolved from clinical paper-and-pencil datasheets and is now a fully custom three-tier app with a web-based front end, custom ASP.NET 8 APIs, and SQL Server 2019 database back end (Figure 1).

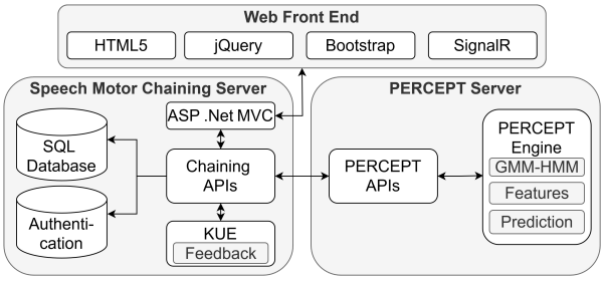


Figure 1: *Block diagram of Speech Motor Chaining.*

There are two main user types: speech learners, to whom speech stimuli are presented, and authenticated clinicians, who control the set-up of individual speech learner accounts, practice sessions, and linguistic stimuli. *Speech Motor Chaining* offers two main modules: speech practice and progress monitoring. Practice modules support Prepractice, Structured Chaining Practice, and Random Practice phases [2]. Progress monitoring evaluates learning of untreated words. The web app logs trial-by-trial practice or progress monitoring data while summarizing learner performance in clinical reports.

3.2. Speech Motor Chaining Adaptive Algorithm

Speech Motor Chaining frequently adapts several empirically-tested principles of motor learning in response to a learner's

accuracy of a speech sound [for review: 3]. Operationalization is detailed in [2], a clinical tutorial predating the web app. In Structured Chaining Practice, treated sounds are practiced in blocks of consecutive trials. Five levels of a chain are presented in increasing difficulty (i.e., syllable, monosyllabic words, multisyllabic words, phrases, and sentences). Each chain is built around a common core syllable, such that a sample chain for the syllable /ɹo/ might include *row*, *road*, *rodeo*, *rodeo clown*, etc. The *Speech Motor Chaining* algorithm emphasizes motor acquisition (Table 1, middle) at lower levels of linguistic complexity and emphasizes motor learning (Table 1, right) at higher levels of complexity. Users advance through chains by producing speech sounds correctly in a block. In Random Practice, previously successful stimuli are practiced in random order to promote generalization [2]. *Speech Motor Chaining* allows clinicians to customize the components of each session. It is preloaded with 3,400+ chains built on 350+ syllable patterns, but clinicians can add their own.

3.3. Telepractice Sessions

Speech practice and progress monitoring can either be run in one browser with a co-located clinician and learner, or remotely via a clinician-generated telepractice link and standard video conferencing software. The unique link provides the learner with a synchronized, simplified view of the clinician's interface. SignalR facilitates communication between the learner and clinician web clients, and the server (Figure 1).

4. Automated Speech Analysis

In clinician-led *Speech Motor Chaining*, the web app's adaptation algorithm is driven by clinician keypress reflecting perceptual judgement (i.e., "correct" or "not quite"), and feedback to the learner is delivered by the clinician after the type of feedback is cued by *Speech Motor Chaining*. The PERCEPT Engine [9, 10] (Figure 1), however, allows for the full automation of practice for the /ɹ/ sound. PERCEPT uses a custom mispronunciation detection algorithm to send predicted perceptual judgment back to *Speech Motor Chaining*, which selects appropriate preloaded clinical feedback from the KUE Engine (Figure 1). KUE feedback is displayed with an avatar clinician, a written/text-to-speech prompt, and an animation showing the articulatory gestures needed to transition from an errored tongue shape to a correct shape. In PERCEPT-driven sessions, learners record their audio through the web interface (Figure 2).

5. Treasure Chain Game Play

Treasure Chain recasts the speech learning activity into a treasure hunt game. Learners (now players) try to fill their treasure "bag" with as much good treasure as possible while avoiding, trading, or dropping lower value treasure. Players advance the treasure hunt by progressing through speech practice trials. The game challenges players to manage their limited bag space effectively. Treasures appear algorithmically, with better treasures weighted to drop based on task performance and practice stage. Players are given periodic opportunities to manage their treasure bag with a "Mysterious Merchant." Gameplay ends with a summary of the player's highest scores. *Treasure Chain's* 250+ unique treasure images (i.e., Figure 2), generated using the DALL-E3 LLM AI, enhance the game's reward loop and motivational effects.

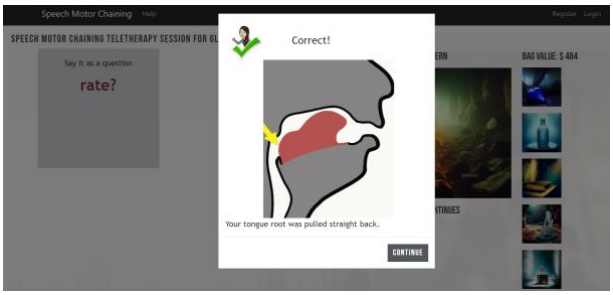


Figure 2: *Treasure Chain interface with KUE feedback.*

6. Conclusions and Future Directions

Speech Motor Chaining, PERCEPT, KUE, and *Treasure Chain* work together to extend motivating, speech learning best practices to clinicians and speech learners. Ongoing clinical trials will further examine PERCEPT and *Treasure Chain*, and future work seeks to personalize the feedback available through KUE and add support for additional phonemes to PERCEPT.

7. Acknoweldgements

Speech Motor Chaining and PERCEPT development is supported by NIH 1R01DC020959-01 (J. Preston, PI).

8. References

[1] Furlong, L., Morris, M., Serry, T., and Erickson, S. (2018). Mobile apps for treatment of speech disorders in children: An evidence-based analysis of quality and efficacy. *PLOS ONE, 13*(8), e0201513.

[2] Preston, J.L., Leece, M.C., and Storto, J. (2019). Tutorial: Speech motor chaining treatment for school-age children with speech sound disorders. *Language, Speech, and Hearing Services in Schools, 50*(3), 343-355.

[3] Maas, E., Robin, D.A., Austermann Hula, S.N., Freedman, S.E., Wulf, G., Ballard, K.J., and Schmidt, R.A. (2008). Principles of motor learning in treatment of motor speech disorders. *Am J Speech Lang Pathol, 17*(3), 277-98.

[4] Wulf, G. and Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The optimal theory of motor learning. *Psychonomic Bulletin & Review, 23*, 1382-1414.

[5] Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). *From game design elements to gamefulness: Defining" gamification"*. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*.

[6] Hamari, J., Koivisto, J., and Sarsa, H. (2014). *Does gamification work? A literature review of empirical studies on gamification*. In *Iinternational Conference on System Sciences*.

[7] Howard‐Jones, P., Demetriou, S., Bogacz, R., Yoo, J.H., and Leonards, U. (2011). Toward a science of learning games. *Mind, Brain, and Education, 5*(1), 33-41.

[8] Malone, T.W. and Lepper, M.R. (2021). Making learning fun: A taxonomy of intrinsic motivations for learning. In *Aptitude, learning, and instruction* (pp. 223-254). Routledge.

[9] Benway, N.R., Preston, J.L., Salekin, A., Hitchcock, E., and McAllister, T. (2024). Evaluating acoustic representations and normalization for rhoticity classification in children with speech sound disorders. *JASA Express Letters, 4*(2).

[10] Benway, N.R. and Preston, J.L. (2023). *Prospective validation of motor-based intervention with automated mispronunciation detection of rhotics in residual speech sound disorders*. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. Dublin, Ireland.