

การพัฒนาโปรแกรมคอมพิวเตอร์สำหรับฝึกการออกเสียงคำภาษาไทย

ที่ใช้ ร และ ล เป็นพยัญชนะต้น โดยใช้แบบจำลองฮิดเดนมาร์คอฟ

The Development of Thai Language Pronunciation Training Software

Focused on Words Spelling with ร-ล using a Hidden-Markov Model

ปัญญาศ ไชยกาฬ^{1*} และ ไมตรี ไชยกาฬ²

Panyayot Chaikan^{1*} and Maitree Chaikan²

บทคัดย่อ

งานวิจัยชิ้นนี้เป็นการพัฒนาโปรแกรมคอมพิวเตอร์ช่วยฝึกฝนการออกเสียงเพื่อแก้ปัญหการไม่รัวล้นเวลาออกเสียง ร และการรัวล้นเวลาออกเสียง ล สำหรับคำในภาษาไทยที่ใช้ ร-ล เป็นพยัญชนะต้น โปรแกรมจะรับเสียงพูดของเข้ามาทางไมโครโฟนและวิเคราะห์ว่าผู้พูดออกเสียงถูกหรือไม่โดยใช้แบบจำลองฮิดเดนมาร์คอฟ ชนิด 11 สเตท 8 มิกซ์เจอร์และแสดงผลการวิเคราะห์เสียงให้ผู้พูดทราบข้อบกพร่องของการออกเสียงของตัวเองได้ในขณะฝึก ตัวโปรแกรมประกอบด้วยบทเรียน 7 บท แต่ละบทประกอบด้วยส่วนของแบบเรียน แบบฝึกหัด และส่วนของเกม จากการทดสอบกับนักเรียนชั้นประถมศึกษาปีที่ 2 ที่มีปัญหาในการออกเสียงจำนวน 6 คน ได้ผลคะแนนก่อนและหลังเรียนเฉลี่ยเท่ากับ 41.27 และ 78.57 ตามลำดับ เครื่องมือมีประสิทธิภาพตามเกณฑ์ E_1/E_2 เท่ากับ 82.14/78.57

คำสำคัญ: การฝึกออกเสียง แบบจำลองฮิดเดนมาร์คอฟ เสียงพูดภาษาไทย คอมพิวเตอร์ช่วยสอน

¹ อ.ดร., ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ สงขลา 90112

² ครูชำนาญการ, โรงเรียนชุมชนบ้านนาสีทอง ตำบลเขาพระ อำเภอรัตภูมิ สงขลา 90180

* Corresponding author: e-mail: panyayot@coe.psu.ac.th Tel./Fax.: 0-7428-7350

Abstract

This paper presents the development of pronunciation training software aimed for solving the pronunciation problems of Thai words with ร-ล initial consonant i.e. pronouncing the ร without trill and pronouncing the ล with trill. Training is divided into seven chapters, each consisting of a lesson, exercises and a game. A lesson demonstrates the correct lip-movement of training words by means of animations and video. The exercises and game evaluate the student's speech obtained via a microphone, and the pronunciation ability results are reported immediately. An 11-state 8-mixture Hidden-Markov model recognition engine was utilized to evaluate the students' speech. When the software was tested with 6 second grade primary school students, the average pre-test and post-test scores increased from 41.27 to 78.57, and the E_1/E_2 efficiency was found to be 82.14/78.57.

Keywords: Pronunciation Training, Hidden-Markov Model, Thai Speech, Computer Aided Instruction (CAI)

บทนำ

ภาษาไทยเป็นภาษาประจำชาติและเป็นตัวบ่งชี้ถึงเอกลักษณ์ของความเป็นชาติไทย ทุกคนจึงมีหน้าที่ที่จะต้องรักษาภาษาไทยให้ยั่งยืนด้วยการเรียนรู้และใช้ภาษาอย่างถูกต้องทั้งในด้านการพูด การอ่าน การเขียน เนื่องจากการใช้ภาษาในปัจจุบันมีความผิดเพี้ยนไปทั้งด้านการพูดและการเขียน โดยเฉพาะอย่างยิ่งการพูดที่ผิดจะนำไปสู่การอ่านและการเขียนที่ผิด ส่วนหนึ่งของปัญหานี้เกิดจากการไม่ตระหนักถึงความสำคัญของการออกเสียงที่ถูกต้อง ซึ่งเกิดขึ้นกับคนไทยโดยทั่วไป กลุ่มคำที่มีปัญหาที่สุด คือ คำที่ใช้ตัวอักษร ร-ล [1] ซึ่งผู้พูดมักไม่รู้ว่าเวลาพูด ร-เรือ ทำให้ออกเสียงจาก ร ผิดเป็น ล และปัญหาการร่วลิ้นเวลาออกเสียง ล ทำให้ออกเสียงผิดเป็น ร

งานวิจัยที่เกี่ยวข้อง

จำนงค์ ทองประเสริฐ แนะนำว่าควรแก้ไขปัญหาการออกเสียงแต่เนิ่น ๆ ตั้งแต่เล็กจะได้ผลดีกว่าการปล่อยให้ติดนิสัยการออกเสียงผิดไปจนโตซึ่งแก้ไขได้ยาก [2] ซึ่งสอดคล้องกับการที่ ประยงค์ งามจิตร [3] แววดิ ปัญญาเรือง [4] จรัสศรี มากสวัสดิ์ [5] วันทนีย์ วังเวชัย [6] และ นิสิตรดา สุทธิอาจ [7] ได้พยายามแก้ไขปัญหาการออกเสียง ร-ล โดยเน้นใช้กับกลุ่มผู้เรียนชั้นประถมศึกษา โดยส่วนใหญ่จะใช้วิธีการสร้างแบบฝึกโดยใช้บัตรคำ หรือกิจกรรม เช่น การเขียน กระดาน การร้องเพลง การใช้เกมปริศนาคำทาย เป็นต้น การใช้แบบฝึกมีข้อจำกัด คือ จะต้องเป็นผู้สอนโดยตรง ทำให้ต่อมามีการใช้สื่ออิเล็กทรอนิกส์เข้ามาช่วย เช่น งานวิจัยของแววดิ ปัญญาเรือง [4] ได้มีการบันทึกเสียงตัวอย่างแต่ละคำลงบนเทปหรือซีดี เพื่อให้ผู้ฝึกได้ฟังตัวอย่างเสียงที่ถูกต้องได้ด้วยตัวเอง อย่างไรก็ตามก็ยังมีข้อจำกัด คือ ไม่มีปฏิสัมพันธ์ในเชิงโต้ตอบกับผู้เรียนทำให้เกิดความเบื่อหน่ายง่าย ทำให้ต่อมามีการใช้สื่อมัลติมีเดียในรูปแบบโปรแกรม

คอมพิวเตอร์ช่วยฝึก [6-7] อย่างไรก็ดี พบว่าโปรแกรมดังกล่าวใช้วิธีการอัดเสียงเก็บลงในไฟล์ และเปิดไฟล์เสียงนั้นออกมาให้ฟังเมื่อโปรแกรมหรือผู้ใช้ร้องขอ ทุกโปรแกรมยังขาดความสามารถในการวิเคราะห์ว่าผู้ฝึกออกเสียงได้ถูกต้องมากน้อยเพียงใด ทำให้ผู้ฝึกไม่สามารถประเมินผลและปรับปรุงการออกเสียงของตนเองได้ในขณะฝึกตามลำพัง จากปัญหาดังกล่าว คณะผู้วิจัย จึงมีแนวคิดที่จะนำระบบรู้จำเสียงพูดด้วยคอมพิวเตอร์ (Speech Recognition) เข้ามาช่วยวิเคราะห์ความถูกต้องในการฝึกฝนการออกเสียง

การวิจัยด้านการประมวลผลเสียงพูดด้วยคอมพิวเตอร์ในประเทศไทย ได้ดำเนินการมาเป็นเวลานาน เช่น การพัฒนาระบบตรวจสอบผู้พูด (Speaker Verification) [8] ซึ่งจะวิเคราะห์ว่าเสียงพูดที่รับเข้ามาเป็นของผู้ซึ่งได้รับอนุญาตให้เข้าใช้งานระบบหรือไม่ งานวิจัยของ Wutiwivatchai, C. และคณะได้พัฒนาระบบแยกแยะผู้พูด (Speaker Identification) ซึ่งจะวิเคราะห์เสียงพูดว่าตรงกับเสียงของผู้ใช้คนใดในฐานข้อมูลหรือไม่ [9-10] งานวิจัยของ Wutiwivatchai, C. [11] และ NECTEC [12] ได้สร้างระบบแปลงข้อความเป็นเสียง (Text-to-Speech) งานวิจัยของ Pensiri, R. และคณะ [13] และงานวิจัยของรพีพรรณ บุญสิน [14] ได้พยายามพัฒนาระบบแปลงเสียงเป็นข้อความ (Speech to text) เพื่อสร้างเป็นระบบซึ่งสามารถใช้เสียงพูดสั่งการ งานวิจัยของ Suebvisai, S. และคณะ ได้สร้างระบบรู้จำเสียงพูดภาษาไทยและแปลงเป็นคำในภาษาอังกฤษ [15] บทความของ Kakuriya, S. ได้บรรยายวิธีการสร้างฐานข้อมูลสำหรับฝึกฝนและทดสอบประสิทธิภาพระบบรู้จำเสียงภาษาไทย [16] อย่างไรก็ดี ยังไม่พบรายงานการนำเทคนิคการประมวลผลเสียงพูดมาใช้ในการฝึกฝนการออกเสียงของนักเรียนมาก่อน ดังนั้นผู้วิจัยจึงได้มีแนวคิดที่จะพัฒนาระบบรู้จำเสียงพูดขึ้นมา โดยจำกัดขอบเขตการวิจัยเอาไว้เบื้องต้นเฉพาะกลุ่มคำที่ใช้ ร-ล เป็นพยัญชนะต้นก่อนเป็นอันดับแรก เนื่องจากปัญหาการออกเสียงที่พบบ่อยที่สุด คือ การไม่ร่วลิ้นในการออกเสียง ร และการร่วลิ้นออกเสียง ล [17]

วิธีการดำเนินการวิจัย

ผู้วิจัยเริ่มจากการออกแบบตัวบทเรียนและเลือกคำฝึก จากนั้นจึงทำการออกแบบโครงสร้างโดยรวมของซอฟต์แวร์ ซึ่งประกอบด้วย ส่วนติดต่อกับผู้ใช้ ส่วนของการยกตัวอย่างการออกเสียง ส่วนของการวิเคราะห์ความถูกต้องของเสียง ซึ่งมีรายละเอียดดังนี้

1. การออกแบบบทเรียน

โปรแกรมประกอบไปด้วย 7 บทเรียน แต่ละบทประกอบด้วย 3 ส่วน ได้แก่ แบบเรียน แบบฝึกหัด และเกม คำฝึกในบทเรียนถูกคัดเลือกมาจาก [4-5] ตารางที่ 1 แสดงคำทั้งหมดที่ใช้ซึ่งแสดงเฉพาะคำมูลหรือคำโดดเท่านั้น แต่จะมีการนำคำมูลเหล่านี้มาสร้างเป็นคำประสม 2 พยางค์สำหรับการฝึกการออกเสียงด้วย ในการใช้งานแต่ละบทเรียนจะเริ่มต้นด้วยการใช้แบบเรียน จากนั้นให้ทำแบบฝึกหัด และตามด้วยการเล่นเกม ในแบบเรียน จะเป็นสาธิตการออกเสียงที่ถูกต้องโดยใช้ภาพวิดีโอแสดงลักษณะการขยับปากพร้อมทั้งส่งเสียงออกสู่ลำโพง ในส่วนของแบบฝึกหัด จะเป็นการฝึกการออกเสียงโดยจอภาพจะขึ้นคำฝึกให้และรับเสียงพูดจากไมโครโฟนเข้ามาประมวลผล

ว่าถูกต้องหรือไม่ หากไม่ถูกต้อง ก็จะขึ้นคำฝึกคำเดิมบนหน้าจอใหม่ จนกว่าจะออกเสียงถูก โปรแกรมอนุญาตให้สามารถคลิกผ่านคำที่ไม่สามารถออกเสียงได้ถูกไปได้หากต้องการ ในส่วนของเกม จะเป็นการขึ้นคำบนหน้าจอและให้ออกเสียงภายในเวลาที่กำหนดจึงจะได้คะแนนสำหรับคำนั้น ๆ เมื่อทดสอบด้วยเกมครบทุกคำแล้วจะขึ้นหน้าจอคะแนนให้นักเรียนได้ทราบ พร้อมทั้งรายงานผลว่าคำใดออกเสียงถูก คำใดออกเสียงผิด

ตารางที่ 1 คำฝึกทั้งหมดของตัวโปรแกรม (แสดงเฉพาะคำโดด)

ลา	รา	ลัก	รัก	เลื่อน	ร้อน	ร้อย	โลง	โรง	รำ	ลอบ
ลืบ	รีบ	ลน	เรื่อง	เล่า	เลา	เรา	ลบ	รบ	ลืม	เรียน
ลาน	ราว	เล็ก	ลูบ	รูป	ลิง	ลม	ลี้ก	เรียว	เรียบ	ลวง
ลาก	ลอก	รอก	ลวด	ลาย	ลับ	แรง	รถ	ลำ		

2. การออกแบบโครงสร้างโดยรวมของซอฟต์แวร์

ซอฟต์แวร์ประกอบไปด้วย 2 ส่วนหลัก คือ ส่วนติดต่อกับผู้ใช้ (User Interface) และส่วนการทำงานอยู่เบื้องหลัง (Background Process) ส่วนของการติดต่อกับผู้ใช้ทำหน้าที่ดึงคำในแบบเรียน คำฝึก หรือคำทดสอบจากฐานข้อมูลออกมาแสดงผลว่ากำลังเรียน ฝึก หรือทดสอบคำใด พร้อมทั้งสร้างภาพเคลื่อนไหวประกอบฉากที่ได้ออกแบบไว้ในแต่ละบท และส่งเสียงคำฝึกให้ได้ฟังผ่านทางลำโพง ในส่วนการทำงานอยู่เบื้องหลังประกอบไปด้วยฐานข้อมูลวิธีโต้ตัวอย่างการออกเสียงและฐานข้อมูลคำแยกออกจากกัน และมีหน่วยวิเคราะห์เสียง ซึ่งทำหน้าที่รับเสียงพูดจากไมโครโฟนเข้ามาวิเคราะห์ และรายงานผลว่าเสียงนั้นตรงกับคำใด ในการทำแบบฝึกหัดและการเล่นเกม ส่วนติดต่อกับผู้ใช้จะรับคำที่ระบบวิเคราะห์ได้นำมาเปรียบเทียบกับคำปัจจุบันว่าเป็นคำเดียวกันหรือไม่ แล้วจึงนำไปคำนวณค่าคะแนนเก็บไว้เพื่อนำไปรายงานให้ผู้ใช้ทราบดังภาพที่ 1 (ก)

3. การออกแบบส่วนติดต่อกับผู้ใช้ และการสร้างตัวอย่างการออกเสียงคำฝึก

ส่วนติดต่อกับผู้ใช้ ได้รับการออกแบบให้เป็นแบบกราฟิกซึ่งมีสีสันสวยงาม เพื่อดึงดูดความสนใจของผู้เรียนซึ่งเป็นนักเรียนชั้นประถมศึกษา เมื่อออกแบบเสร็จ จึงได้ส่งให้ผู้เชี่ยวชาญจำนวน 5 คน ซึ่งแต่ละคนมีประสบการณ์ในการสอนโดยใช้สื่อคอมพิวเตอร์ช่วยสอนไม่น้อยกว่า 10 ปี เป็นผู้ประเมิน ได้คะแนนประเมินเฉลี่ยเท่ากับ 4.73 (เต็ม 5) และมีส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.21 จากนั้นนำข้อเสนอแนะที่ได้รับมาทำการปรับปรุงตัวโปรแกรมภาพที่ 1 (ข-จ) แสดงให้เห็นตัวอย่างหน้าจอของการเล่นเกมซึ่งจะให้นักเรียนออกเสียงคำที่กำลังเคลื่อนไหวผ่านหน้าจอในรูปแบบต่าง ๆ

ในการสร้างตัวอย่างการออกเสียง จะใช้การบันทึกภาพวิดีโอการออกเสียงของจริงที่ความเร็ว 30 เฟรมต่อวินาที โดยได้จัดให้ผู้สารถิทำการออกเสียงแต่ละคำภายใต้สภาพแวดล้อมที่ควบคุมแสงและปราศจากสัญญาณเสียงรบกวน จากนั้นจึงนำวิดีโอภาพและเสียงมาตัดต่อเพื่อซ้อนภาพกับฉากรูปของแต่ละบทเรียน ดังตัวอย่างแสดงในภาพที่ 1 (ฉ)



(ก)



(ข)



(ค)



(ง)



(จ)

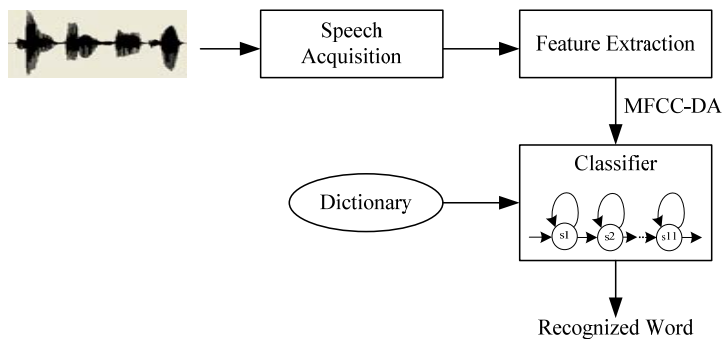


(ฉ)

ภาพที่ 1 ตัวอย่างหน้าจอของโปรแกรมฝึกออกเสียงที่พัฒนาขึ้น

4. การออกแบบและฝึกฝนระบบวิเคราะห์เสียง

ส่วนของการวิเคราะห์เสียง จะทำหน้าที่นำเสียงพูดมาวิเคราะห์ว่าเป็นคำใด และส่งผลลัพธ์การวิเคราะห์ไปให้กับส่วนของการคิดคะแนนซึ่งจะเปรียบเทียบว่าคำที่ได้จากการวิเคราะห์กับคำซึ่งระบบถามเป็นคำเดียวกันหรือไม่ แล้วจึงคิดเป็นคะแนนออกมา ภาพที่ 2 แสดงโครงสร้างของระบบวิเคราะห์การออกเสียงซึ่งแบ่งออกเป็น 3 ส่วน โดยส่วนของการรับเสียง (Speech Acquisition) จะสุ่มสัญญาณเสียงที่มีความถี่ 16 กิโลเฮิรตซ์ ที่ความละเอียดของข้อมูลขนาด 16 บิต สัญญาณเสียงจะถูกส่วนสกัดค่าลักษณะเด่น (Feature Extraction) ทำการดึงเอาเฉพาะข้อมูลที่มีความโดดเด่นออกมา ค่าลักษณะเด่นนี้เรียกว่า ค่าสัมประสิทธิ์เคปสตรัมความถี่เมล (Mel scale Frequency Cepstral Coefficient) พร้อมทั้งค่าอนุพันธ์อันดับหนึ่ง (Delta Term) และค่าอนุพันธ์อันดับสอง (Acceleration Term) ซึ่งเรียกโดยย่อว่า ค่าลักษณะเด่น MFCC-D-A [18] ค่าลักษณะเด่นนี้จะถูกส่งต่อไปยังตัวแยกแยะ (Classifier) ซึ่งจะตรวจสอบว่าเสียงเข้านั้นตรงกับเสียงอะไร จากนั้นจึงตัดสินใจในสิ่งที่ระบบคิดว่าตรงกับเสียง (Recognized Word) ออกมาโดยใช้ข้อมูลจากพจนานุกรม (Dictionary) ในการวิเคราะห์เสียงของตัวแยกแยะนั้น จะใช้การวิเคราะห์ด้วยกลไกของแบบจำลองฮิดเดนมาร์คอฟ (Hidden Markov Model) [18] ซึ่งเป็นระบบการรู้จำรูปแบบชนิดหนึ่งที่ต้องมีการฝึกฝนระบบก่อนนำไปใช้งาน (Supervised Learning)

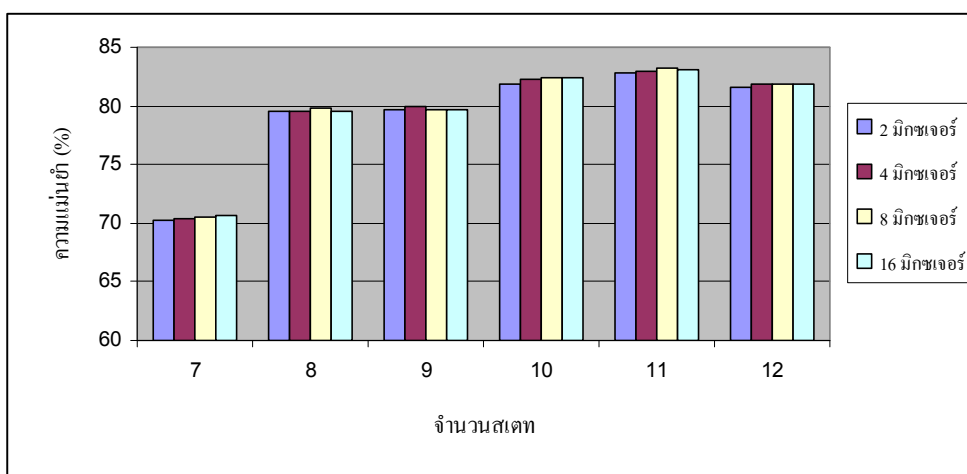


ภาพที่ 2 โครงสร้างของระบบวิเคราะห์การออกเสียง

ในการฝึกฝนแบบจำลองฮิดเดนมาร์คอฟนั้น จำเป็นต้องมีตัวอย่างการออกเสียงที่ถูกต้องเป็นจำนวนมาก เพื่อใช้ในการฝึกฝนระบบ งานวิจัยเสียงภาษาไทยส่วนใหญ่จะใช้ตัวอย่างเสียงที่รวบรวมขึ้นโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) ซึ่งมีฐานข้อมูลเสียงมาตรฐานหลายตัว ได้แก่ NECTEC-ATR, LOTUS [11], LOTUS-Cell, VoiceCom, NECTEC-TRUE และ LOTUS-BN [19] อย่างไรก็ตาม เนื่องจากข้อมูลเสียงของฐานข้อมูลดังกล่าวถูกรวบรวมจากผู้พูดซึ่งมีอายุ 21-25 ปี เป็นจำนวนสูงถึง 67.5% ดังนั้นจึงไม่เหมาะสมสำหรับใช้ฝึกฝนระบบที่ต้องการจะพัฒนาขึ้นซึ่งมุ่งเน้นสำหรับให้นักเรียนชั้นประถมศึกษาใช้ จึงได้สร้างฐานข้อมูลเสียงขึ้นมาใหม่โดยใช้เสียงจากการรวบรวมตัวอย่างเสียงของนักเรียนชั้น ป. 2- ป.6 โรงเรียนชุมชนบ้านนาสีทองในปีการศึกษา 2554-2555

จำนวน 250 คน โดยให้นักเรียนแต่ละคนออกเสียงคำฝึกคำโดดทั้งหมด 42 คำ โดยออกเสียงคำละ 2 ครั้ง ในห้องบันทึกเสียง ซึ่งปราศจากเสียงรบกวนภายนอก ครูผู้สอนจะเป็นผู้พิจารณาว่านักเรียนคนนั้นออกเสียงได้ถูกต้องมากน้อยเพียงใด จากนั้นทำการคัดเลือกเสียงของนักเรียนที่ออกเสียง ร-ล ไม่ชัดเจนทิ้งไป เหลือเพียงเสียงของนักเรียนที่ออกเสียงได้ถูกต้องชัดเจนจำนวน 80 คนนำมาทำหน่วยเสียงกำกับเอาไว้ด้วยมือสำหรับเก็บเอาไว้ฝึกฝนและทดสอบความแม่นยำของระบบ การฝึกแบบจำลองจะทำกับเสียงของนักเรียน 79 คน และทดสอบกับนักเรียนที่เหลือ 1 คน จากนั้นจึงนำเสียงนักเรียนที่เพิ่งทดสอบ 1 คน ใส่กลับเข้าไปในชุดเสียงสำหรับฝึก และดึงเสียงของนักเรียนคนต่อไป ออกมาสำหรับใช้ทดสอบอีกครั้ง ทำซ้ำอย่างนี้จนกระทั่งครบทุกคนในขั้นตอนการฝึก จากนั้นจึงคำนวณค่าความแม่นยำโดยรวมของการรู้จำของแบบจำลองออกมา

ผู้วิจัยได้เลือกใช้ HTK (Hidden Markov Toolkit) [18] ซึ่งเป็นโปรแกรมชนิดโอเพนซอร์สที่รวบรวมเครื่องมือสำหรับการสร้างและฝึกฝนแบบจำลองฮิดเดนมาร์คอฟ ซึ่งสามารถนำมาประยุกต์ใช้กับเสียงภาษาไทยได้ โดยเลือกใช้แบบจำลองฮิดเดนมาร์คอฟชนิดส่วนผสมของเกาส์เซียน (Mixture of Gaussian) [18] ซึ่งมีค่าพารามิเตอร์สำคัญที่จะต้องทำการทดลองเพื่อหาค่าที่เหมาะสม คือ จำนวนสเตต (State) และจำนวนมิกซเจอร์ (Mixture) ของระบบ ได้ทดลองสร้างตัวรู้จำที่จำนวนสเตต 7, 8, 9, 10, 11, 12 และที่จำนวนมิกซเจอร์ 2, 4, 8, 16 ได้ผลดังภาพที่ 3 จะเห็นว่าจำนวนสเตตจะมีผลต่อค่าความแม่นยำอย่างมีนัยสำคัญสูงกว่าจำนวนมิกซเจอร์ เมื่อเพิ่มจำนวนสเตตจะทำให้ค่าความแม่นยำสูงขึ้น อย่างไรก็ตาม หากเพิ่มมากกว่า 11 จะทำให้ความแม่นยำลดลง ส่วนการเพิ่มจำนวนมิกซเจอร์จะทำให้ค่าความแม่นยำเพิ่มขึ้นและเริ่มลดลงหรือคงที่เมื่อมากกว่า 8 ดังนั้นจึงเลือกแบบจำลองฮิดเดนมาร์คอฟ ชนิด 11 สเตต 8 มิกซเจอร์เนื่องจากให้ความแม่นยำสูงสุดอยู่ที่ 83.2%



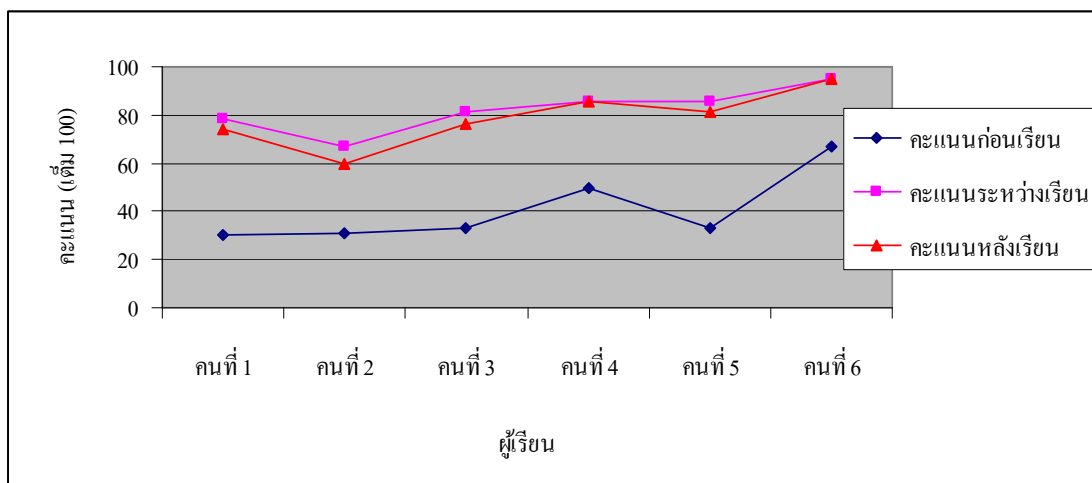
ภาพที่ 3 ผลการทดลองหาค่าความแม่นยำของแบบจำลองฮิดเดนมาร์คอฟในการวิเคราะห์เสียง ร-ล ที่จำนวนสเตตและจำนวนมิกซเจอร์ค่าต่าง ๆ

5. การนำนวัตกรรมโปรแกรมฝึกฝนการออกเสียงไปทดลองใช้งานจริง

ได้นำนวัตกรรมที่พัฒนาขึ้นนี้ไปใช้กับนักเรียนชั้นป. 2 โรงเรียนชุมชนบ้านนาสีทอง ปีการศึกษา 2556 ภาคเรียนที่ 2 โดยคัดเลือกนักเรียนที่มีปัญหาออกเสียงคำ ร-ล ไม่ชัดเจนจำนวน 6 คน มาเป็นกลุ่มตัวอย่าง ในการใช้นวัตกรรม จะให้นักเรียนเข้าใช้โปรแกรมครั้งละหนึ่งคนโดยทำในสภาพแวดล้อมห้องปิด ซึ่งออกแบบให้เสียงรบกวนภายนอก สามารถเล็ดรอดเข้ามาสู่ไมโครโฟนได้น้อยลง เนื่องจากตัวโปรแกรมยังมีข้อจำกัดตรงที่ความแม่นยำในการวิเคราะห์เสียงจะลดลงหากมีเสียงรบกวนเข้ามาในระบบมากจนเกินไป การฝึกแต่ละบทเรียนจะใช้เวลาหนึ่งสัปดาห์ โดยในแต่ละบทจะให้นักเรียนใช้โปรแกรม 4 ครั้ง แต่ละครั้งไม่เกิน 10 นาทีต่อนักเรียนหนึ่งคน โดยใช้เวลาในคาบซ่อมเสริมของแต่ละวัน การใช้โปรแกรมแต่ละบทมี 4 ชั้น คือ (ก) ชี้นำ ให้ใช้แบบเรียนและแบบฝึกหัด เป็นเวลา 10 นาที (ข) ชี้นำเริ่มฝึก ให้ใช้แบบฝึกหัด คนละไม่เกิน 10 นาที (ค) ชี้นำฝึกซ้ำ ให้ใช้แบบฝึกหัด คนละไม่เกิน 10 นาที และ (ง) ชี้นำทบทวน ให้เล่นเกม เป็นเวลาไม่เกิน 10 นาที

ผลการวิจัยและอภิปรายผล

ก่อนและหลังการใช้นวัตกรรม ได้ให้นักเรียนทำแบบทดสอบเพื่อประเมินความสามารถในการออกเสียง โดยให้นักเรียนอ่านออกเสียงคำที่กำหนดให้ฟัง และครูผู้สอนเป็นผู้ประเมินคะแนนการออกเสียงโดยตรง ในการใช้นวัตกรรม ได้ทำการใช้โปรแกรมคอมพิวเตอร์ฝึกออกเสียงจำนวน 7 บทกับนักเรียนกลุ่มตัวอย่าง ใช้เวลาทั้งสิ้น 7 สัปดาห์ ครูผู้สอนจะเก็บคะแนนในการเล่นของแต่ละบทของผู้เรียนแต่ละคนเอาไว้ใช้เป็นคะแนนสอบระหว่างเรียน ได้ผลลัพธ์ดังภาพที่ 4



ภาพที่ 4 เปรียบเทียบผลคะแนนการออกเสียงโดยใช้แบบทดสอบก่อนเรียน ระหว่างเรียน และหลังเรียน

เมื่อได้ผลการทดสอบก่อนและหลังเรียนแล้วจึงนำมาหาประสิทธิภาพแบบหนึ่งกลุ่มสอบก่อน-สอบหลัง (One Group Pretest-Post Design) ทำการวิเคราะห์หาค่าคะแนนที (t-test) โดยใช้โปรแกรม SPSS ซึ่งเป็นซอฟต์แวร์เสรีที่ไม่มีปัญหาการละเมิดลิขสิทธิ์ พบว่าคะแนนที (T-Score) เท่ากับ 11.41 และค่าคะแนนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ทำการประเมินประสิทธิภาพ (Efficiency Criterion) ของนวัตกรรมโดยใช้เกณฑ์ E_1/E_2 [20] โดยค่า E_1 คือ ค่าร้อยละของคะแนนเฉลี่ยของการทำแบบทดสอบระหว่างเรียนซึ่งบ่งบอกถึงประสิทธิภาพของกระบวนการเรียนรู้ และค่า E_2 คือ ค่าร้อยละของคะแนนเฉลี่ยที่เกิดจากการทำแบบทดสอบหลังเรียนซึ่งบ่งบอกประสิทธิภาพของผลลัพธ์จากการเรียนรู้ พบว่าคะแนน E_1/E_2 เท่ากับ 82.14/78.57 ซึ่งถือว่ามีประสิทธิภาพตามเกณฑ์ 80/80 เนื่องจากค่า E_1 และ E_2 ต่างกันไม่เกินร้อยละ 5 [20] แสดงให้เห็นว่าตัวนวัตกรรมสามารถนำไปใช้ในการฝึกการออกเสียงได้จริง

สรุปผลการวิจัย

นวัตกรรมที่พัฒนาขึ้นนี้แตกต่างจากตัวโปรแกรมฝึกออกเสียงอื่น ๆ ที่มีมาก่อนหน้านี้ ตรงที่มีความสามารถในการวิเคราะห์การออกเสียง ช่วยให้นักเรียนรู้ได้ว่าตนออกเสียงได้ถูกต้องมากน้อยเพียงใด และสามารถปรับปรุงการปรับปรุงการออกเสียงของตนเองได้ในขณะฝึกฝนตามลำพัง ส่งผลให้นักเรียนสามารถแก้ไขข้อบกพร่องในการออกเสียงของตนได้ในระดับที่น่าพอใจ จากการสังเกตพฤติกรรมของนักเรียนพบว่าการใช้รูปแบบของการเล่นเกมแบบกราฟิกช่วยดึงดูดความสนใจของเด็กได้เป็นอย่างดี นักเรียนทุกคนมีความกระตือรือร้นที่จะฝึกการออกเสียงตั้งแต่เริ่มต้นจนกระทั่งสิ้นสุดการใช้ตัวนวัตกรรม

อย่างไรก็ดี โปรแกรมที่พัฒนาขึ้นยังมีข้อจำกัด ตรงที่ยังไม่สามารถบอกให้ทราบได้ว่าปัญหาในการออกเสียงของนักเรียนเกิดจากสาเหตุอะไร ดังนั้น ในอนาคต ควรมีการปรับปรุงเพิ่มเติมส่วนของการวิเคราะห์เสียงให้สามารถบอกได้ว่ามีข้อบกพร่องตรงจุดใด เพื่อที่จะสามารถให้คำแนะนำในประเด็นที่บกพร่องนั้นแก่ผู้เรียนได้ ในขณะที่ฝึก ก็จะช่วยให้อุปกรณ์มีความสมบูรณ์ยิ่งขึ้น นอกจากนี้ ควรมีการเพิ่มเติมอัลกอริทึมการกรอง (Filter) สำหรับลดผลกระทบของเสียงรบกวนรอบข้างเพื่อเพิ่มความแม่นยำของการรู้จำเสียง และเพิ่มขอบเขตของคำฝึกให้ครอบคลุมถึงคำควบกล้ำด้วยในอนาคต

คำขอบคุณ

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัยจากคณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์ สัญญารับทุนเลขที่ ENG-53-2-7-2-0107-S

เอกสารอ้างอิง

- [1] (ม.ป.ป.). **วิกิตัวประวัติภาษาไทยภาคการออกเสียง**. สืบค้นเมื่อ 20 กรกฎาคม 2557, จาก http://www.9digits.com/index.php?option=com_content&task=view&id=149&Itemid=36.
- [2] จ่านง ทองประเสริฐ, ราชบัณฑิต. (2544). **ภาษาไทยของเราดี เป็นศักดิ์เป็นศรีของชาติ** (พิมพ์ครั้งที่ 2). กรุงเทพฯ: โรงพิมพ์มหาวิทยาลัยราชภัฏวชิราวุฒวิทยาลัย.
- [3] ประยงค์ งามจิตตร. (2533). **การสร้างแบบฝึกการอ่านคำที่ใช้อักษร ร ล ว ควบกล้ำ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 3 จังหวัดอุดรธานี**. วิทยานิพนธ์หลักสูตรการศึกษามหาบัณฑิต วิชาเอกการประถมศึกษา มหาวิทยาลัยบูรพา. ชลบุรี.
- [4] แวดดี ปัญญาเรือง. (2538). **การใช้เพลงและเกมเพื่อฝึกอ่านออกเสียงคำที่ใช้พยัญชนะ ร ล ว ควบกล้ำ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 2**. วิทยานิพนธ์หลักสูตรศึกษาศาสตรมหาบัณฑิต สาขาวิชาประถมศึกษา มหาวิทยาลัยเชียงใหม่ เชียงใหม่.
- [5] จรัสศรี มากสวัสดิ์. (2547). **ผลของการใช้เทคนิคแม่แบบที่มีต่อการอ่านออกเสียงควบกล้ำ ร ล ว ของนักเรียนชั้นประถมศึกษาปีที่ 2 โรงเรียนสุเหร่าบางมะเขือ เขตวัฒนา กรุงเทพมหานคร**. สารนิพนธ์หลักสูตรการศึกษามหาบัณฑิต สาขาจิตวิทยาการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ.
- [6] วันทนีย์ วังเวชัย. (2554, 7 สิงหาคม). **สื่อสารสนเทศแม่แบบช่วยการอ่านออกเสียง ร ล และคำควบกล้ำชั้นมัธยมศึกษาปีที่ 2**. สืบค้นเมื่อ 12 พฤษภาคม 2557 จาก <https://krusan04.wordpress.com/page/2/>.
- [7] นิสิตรา สุทธิอาจ. (2549). **การพัฒนาบทเรียนคอมพิวเตอร์มัลติมีเดีย สารการเรียนรู้ภาษาไทยเรื่องการออกเสียงคำควบกล้ำ สำหรับนักเรียนชั้นประถมศึกษาปีที่ 3**. วิทยานิพนธ์หลักสูตรปริญญาครุศาสตรมหาบัณฑิต สาขาหลักสูตรและการสอน มหาวิทยาลัยราชภัฏสกลนคร สกลนคร.
- [8] Wutiwiwatchai, C., Achariyakulporn, V. and Kasuriya, S. (2001) Improvement of speaker verification for Thai language. **Proc. EUROSPEECH**. 775–778. NA.
- [9] Wutiwiwatchai, C., Sae-Tung, S. and Achariyakulporn, V. (2000). “Thai Language Speaker Identification System: Development Progress”, **NECTEC Technical Journal**, 2(7), 24-35.
- [10] Wutiwiwatchai, C., Achariyakulporn, V. and Tanprasert, C. (1999). Text-dependent speaker identification using LPC and DTW for Thai language. **Proceedings of the IEEE Region 10 Conference (TENCON 99)**. 1, 674-677. NA.
- [11] Wutiwiwatchai, C. and Furui, S. (2007). “Thai speech processing technology: A review”, **Speech Communication Journal**. 49, 8-27.
- [12] NECTEC, (ม.ป.ป.). **Welcome to Vaja**. สืบค้นเมื่อ 10 มิถุนายน 2557, จาก <http://vaja.nectec.or.th/>.

- [13] Pensiri, R. and Jitapunkul, S. (1995). Speaker-Independent Thai Numerical Voice Recognition by using Dynamic Time Warping. **Proceedings of the 18th Electrical Engineering Conference (EECON)**. 977-981. NA.
- [14] ระพีพรรณ บุญสิน และชุลีรัตน์ จรัสกุลชัย. (ม.ป.ป.). ระบบสั่งงานพ็อกเก็ตพีซีด้วยเสียงภาษาไทย. สืบค้นเมื่อ 11 กันยายน 2557, จาก <http://kucon.lib.ku.ac.th/Fulltext/KC4805003.pdf>.
- [15] Suebvisai, S., Charoenpornasawat, P., Black, A., Woszczyna, M. and Schultz, T. (2005). Thai Automatic Speech Recognition. **International Conference on ICASSP2005**. 857-860. March 18-23, 2005.
- [16] Kasuriya, S., Sornlertlamvanich, V., Cotsomrong, P., Kanokphara, S. and Thatphithakkul, N. (2003). Thai Speech Corpus for Thai Speech Recognition. **International Conference on Speech Databases and Assessments (Oriental-COCOSDA)**. 54-61. NA.
- [17] Chaikan, P. (2013). “A Review of Speech Processing and Facial Animation Technologies for Solving Thai Word Mispronunciations”, **International Journal of Applied Computer Technology and Information Systems**. 2(2), 60-65.
- [18] Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V. and Woodland, P. (2006, December). **The HTK book (for HTK V. 3.4)**. Cambridge University Engineering Department.
- [19] Chotimongkol, A., Saykhum, K., Chootrakool, P., Thatphithakkul, N. and Wutiwiwatchai, C. (2009). LOTUS-BN: A Thai Broadcast News Corpus and Its Research Applications. **International Conference on Speech Database and Assessments 2009**. 44-50. NA.
- [20] ชัยยงค์ พรหมวงศ์. (2556). การทดสอบประสิทธิภาพสื่อหรือชุดการสอน. **ศิลปการศึกษาศาสตร์วิจัย**. 5(1), 7-19.