

แอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ Voice Control Application for English Lesson Access

บุรัสกร อยู่สุข, คงเทพ บุญมี, วรพันธ์ สาระสุรีย์ภรณ์ และ นิติ วิทยาวิโรจน์
สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

Burasakorn Yoosooka, Khongthep Boonmee, Waraphan Sarasureeporn

and Niti Witthayawiroj

Division of Computer Science, Faculty of Science and Technology,

Rajamangala University of Technology Thanyaburi

Abstract

The communication in the ASEAN Economic Community (AEC) promotes the importance of and necessity of the international language, English. Fluent in English communication yields tremendous communication advantages or otherwise, drastic disadvantages. With this necessity, the English lesson access tools would be supremely beneficial as the practical solution. This article proposes the concept in developing voice control based application for the English lesson access. The application of speech recognition technologies promotes the verbal-to-vocab translation for command. This application requires input of command to access English lesson via either a speech recognition device or a keyboard. The related English lessons will be displayed for user's selection. After the selection, the system will display an English dialogue of the lessons and their translation in audio and text. The translation shows the meaning of words, type of vocabulary (noun, verb, preposition, adjective, adverb, or etc.), synonyms and antonyms to users. The system has been developed and evaluated. The 5-Likert scale questionnaire was used as user satisfaction evaluation. Sample groups are the 34 students with major of study in computer science, RMUTT and the six subject matter experts. The result has been discovered that the system has yielded positive effects 85% in terms of the users' satisfaction.

Keywords: *Voice control application , Speech , Automatic speech recognition , English , Lessons*

บทคัดย่อ

การสื่อสารในยุคประชาคมเศรษฐกิจอาเซียน หรือ AEC นั้น ภาษาอังกฤษมีความสำคัญมาก เนื่องจากใช้เป็นภาษากลางในการติดต่อเจรจาทางธุรกิจและการท่องเที่ยว ซึ่งหากใครที่ยังใช้ภาษาอังกฤษในการสื่อสารได้ไม่ดีเท่าที่ควร ก็อาจเกิดความเสียเปรียบได้ ดังนั้นเครื่องมือที่ช่วยให้ความรู้ด้านภาษาอังกฤษ จะช่วยแก้ไขปัญหานี้ได้ ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อสร้างแอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ สำหรับช่วยให้เข้าถึงบทเรียนต่างๆ ได้โดยการใช้เสียงสั่งการ เพื่อความสะดวกในการเข้าถึงข้อมูลบทเรียนได้ แม้ว่าไม่สะดวกที่จะพิมพ์คำสั่งก็ตาม ซึ่งแอปพลิเคชันนี้สร้างด้วยเทคโนโลยีของ Speech Recognition ทำให้สามารถทำคำสั่งจากเสียงพูดหรือคีย์บอร์ดได้ โดยผู้ใช้จะสามารถป้อนคำสั่งเข้าสู่บทเรียนได้โดยใช้เสียงพูด หลังจากนั้นแอปพลิเคชันจะแสดงผลชื่อบทเรียนภาษาอังกฤษที่เกี่ยวข้อง เพื่อให้ผู้ใช้เลือกบทเรียน แล้วจึงแสดงบทสนทนาภาษาอังกฤษของบทเรียนนั้น พร้อมคำแปลภาษาไทย ในรูปแบบเสียงและข้อความ โดยคำแปลจะแสดง ความหมายของคำศัพท์ ประเภทของคำศัพท์ คำเหมือน และคำตรงข้าม ให้แก่ผู้ใช้ ในการประเมินความพึงพอใจของผู้ใช้ คณะผู้วิจัยได้ใช้แบบสอบถามแบบ 5 ระดับ โดยกลุ่มตัวอย่างคือนักศึกษาศาสาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี จำนวน 34 คน และผู้เชี่ยวชาญ จำนวน 6 คน ซึ่งผลการประเมินความพึงพอใจโดยรวมพบว่า 85% ของผู้ที่มีความพึงพอใจในการใช้ระบบ

คำสำคัญ: แอปพลิเคชันคำสั่งเสียง , เสียงพูด , การรู้จำเสียงพูดอัตโนมัติ , บทเรียน , ภาษาอังกฤษ

1. บทนำ

เมื่อกล่าวถึงประชาคมเศรษฐกิจอาเซียน (AEC) ที่กำลังจะมาถึง ส่งผลให้กลุ่มประเทศอาเซียนต้องประสานสัมพันธ์กัน ทั้งทางด้านเศรษฐกิจและวัฒนธรรม โดยมีภาษาอังกฤษเป็นสื่อกลางในการติดต่อสื่อสารกัน แม้ประเทศไทยจะเป็นประเทศผู้นำอาเซียนมายาวนานหลายปี นับแต่ปีให้กำเนิดอาเซียนเมื่อ 46 ปีที่แล้ว แต่ชื่อเสียงประเทศไทยตกต่ำลงในหลายด้าน โดยเฉพาะทักษะภาษาอังกฤษ ชาวใน Jakarta Post รายงานอ้างผลการศึกษาระดับมัธยมศึกษาขององค์กรด้านการศึกษาของสวิตเซอร์แลนด์ ชื่อ Education First (EF) ซึ่งได้ประกาศผลการจัดทำดัชนีการสำรวจคุณภาพการใช้ภาษาอังกฤษของคนไทยในระดับผู้ใหญ่กว่าคนไทยตกอยู่อันดับที่ 55 จากการสำรวจ 60 ประเทศ และถูกจัดอยู่ในกลุ่ม “very low proficiency” หรือ “ต่ำคุณภาพทักษะภาษามาก” ทั้งนี้เป็นการศึกษาข้อมูลจากการทดสอบผู้ใหญ่ 750,000 คน จาก 60 ประเทศ ในปี 2013 ด้วยทักษะทางภาษาอังกฤษที่ต่ำดังกล่าว จึงอาจส่งผลให้เสียเปรียบในการสื่อสารต่อธุรกิจได้ ดังนั้นหากมีเครื่องมือช่วยสอนและแปลภาษาอังกฤษ จะส่งผลให้เขาเหล่านั้นได้ความรู้ด้าน

ภาษาอังกฤษ และรูปแบบการสนทนาที่ถูกต้องจนเกิดทักษะและความเชื่อมั่นในการติดต่อสื่อสารกัน หรือการทำธุรกรรมต่างๆ ร่วมกันในกลุ่มประชาคมอาเซียนได้เป็นอย่างดี

เครื่องมือนี้เป็นแอปพลิเคชันที่ใช้เทคโนโลยีของ Speech Recognition ซึ่งเป็นเทคโนโลยีที่ถูกพัฒนาอย่างต่อเนื่อง และแพร่หลายไปยังทุกๆ อุปกรณ์อิเล็กทรอนิกส์ เพื่อช่วยอำนวยความสะดวกให้กับผู้ใช้ ได้ใช้งานอุปกรณ์นั้นๆ ง่ายยิ่งขึ้น เทคนิคของการรู้จำเสียงพูดนั้น ระบบจะสร้างโครงข่ายคำ (Word Network) โดยอาศัยแบบจำลองภาษา (Language Model) โดยในแต่ละคำจะประกอบด้วยแบบจำลองเสียง (Acoustic Model) ประเภทฮิดเดนมาร์คอฟ (Hidden Markov Model: HMM) ของแต่ละหน่วยเสียงที่มาประกอบกันเป็นคำนั้นๆ หลังจากนั้นระบบจะป้อนเวกเตอร์ของ Speech Feature เข้าไปยังโครงข่ายคำ เพื่อคำนวณค่าความน่าจะเป็นของเสียงที่เข้ามาว่าจะตรงกับเส้นทางใดในโครงข่ายคำ และจะเลือกคำตอบที่ดีที่สุดเป็นเส้นทางที่มีค่าความน่าจะเป็นสูงที่สุด โดยขั้นตอนทั้งหมดจะเป็นการถอดรหัส (Decoding) ซึ่งในทางปฏิบัติ นักวิจัยได้ใช้อัลกอริธึมการถอดรหัสแบบ Viterbi beam-search เพื่อลดปริมาณการใช้หน่วยความจำและการคำนวณด้วยการลบเส้นทางที่มีค่าความน่าจะเป็นต่ำทิ้งไป

ระบบจะนำคลื่นเสียงที่ป้อน เข้าสู่ขั้นตอนการรู้จำเสียงพูดเพื่อแปลงเป็นข้อความ จากนั้นนำไปค้นหาบทเรียนภาษาอังกฤษและความหมายของคำศัพท์ต่างๆ จากฐานข้อมูล แล้วแสดงออกมาในรูปแบบข้อความและเสียงสนทนา โดยใช้เทคนิค Voice Synthesis ในการอ่านคำศัพท์และความหมาย ซึ่งในที่นี้ใช้แอปพลิเคชัน VAJA Android ที่พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ สำหรับการสังเคราะห์เสียงพูด

แอนดรอยด์เป็นระบบปฏิบัติการที่มีความยืดหยุ่น และง่ายต่อการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการนี้ อีกทั้งมีลักษณะโครงสร้างที่ปรับได้ จึงสามารถติดตั้งลงในอุปกรณ์ต่างๆ ที่มีความหลากหลายทางฮาร์ดแวร์ ไม่ว่าจะเป็นโทรศัพท์มือถือหรือแท็บเล็ต จึงมีผู้ใช้งานมาก ทำให้แอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ เพิ่มมากขึ้นอย่างรวดเร็ว ดังนั้นการพัฒนาแอปพลิเคชัน บนระบบปฏิบัติการแอนดรอยด์จึงเป็นสิ่งที่น่าสนใจอย่างยิ่ง

แอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ บนระบบปฏิบัติการแอนดรอยด์นี้ จะช่วยอำนวยความสะดวกให้ผู้ใช้งานทุกท่าน โดยเฉพาะกับผู้พิการ ผู้สูงอายุ หรือผู้ที่ไม่สามารถใช้นิ้วมือพิมพ์คำศัพท์ผ่านทางแป้นพิมพ์ได้ สามารถเข้าสู่บทเรียนภาษาอังกฤษหรือค้นหาความหมายของคำศัพท์ได้ง่ายและรวดเร็วมากยิ่งขึ้น เพียงแค่ใช้เสียงพูด ระบบก็จะนำเข้าสู่บทเรียนภาษาอังกฤษโดยอัตโนมัติ และแสดงผลออกมาทั้งในรูปแบบข้อความ และการสังเคราะห์เสียงพูด เพื่อความสะดวกในการรับข้อมูลด้วย

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ระบบรู้จำเสียงพูด [1]

เทคโนโลยีการรู้จำเสียงพูด (Automatic Speech Recognition) เป็นระบบที่ช่วยแปลงเสียงพูดให้เป็นข้อความที่ระบบคอมพิวเตอร์สามารถนำมาใช้ประมวลผลต่อไปได้ โดยเทคโนโลยีนี้ได้รับการพัฒนาขึ้น ภายในหน่วยงานวิจัยของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ เพื่อช่วยลดข้อจำกัดในการเข้าถึงเทคโนโลยีสมัยใหม่ โดยการใช้ระบบประมวลผลภาษาธรรมชาติ เป็นตัวเชื่อมโยงผู้ใช้เข้าสู่บริการทางคอมพิวเตอร์ในรูปแบบต่างๆ ในขั้นตอนรู้จำเสียง ระบบจะสร้างโครงข่ายคำ (Word Network) โดยอาศัยแบบจำลองภาษา (Language Model) โดยในแต่ละคำจะประกอบด้วยแบบจำลองเสียง (Acoustic Model) ประเภทฮิดเดนมาร์คอฟ (Hidden Markov Model: HMM) ของแต่ละหน่วยเสียงที่มาประกอบกันเป็นคำนั้น ๆ หลังจากนั้นระบบจะป้อนเวกเตอร์ของ Speech Feature เข้าไปยังโครงข่ายคำ เพื่อคำนวณค่าความน่าจะเป็นของเสียงที่เข้ามาว่าจะตรงกับเส้นทางใดในโครงข่ายคำ และจะเลือกคำตอบที่ดีที่สุดเป็นเส้นทางที่มีค่าความน่าจะเป็นสูงที่สุด โดยขั้นตอนทั้งหมดจะเป็นการถอดรหัส (Decoding) ซึ่งในทางปฏิบัตินักวิจัยได้ใช้อัลกอริธึมการถอดรหัสแบบ Viterbi beam-search เพื่อลดปริมาณการใช้หน่วยความจำและการคำนวณ ด้วยการลบเส้นทางที่มีค่าความน่าจะเป็นต่ำทิ้งไป

การตอบโต้กับเครื่องคอมพิวเตอร์ที่ผ่านการใช้เมาส์ และคีย์บอร์ดรวมทั้งความสามารถในการเข้าใจคำสั่งการทำงานของระบบคอมพิวเตอร์ อาจทำให้การเข้าถึงเทคโนโลยีสมัยใหม่ยังคงเป็นเรื่องยากสำหรับผู้คนจำนวนมาก แต่การเข้ามาของระบบรู้จำเสียงพูด (Automatic Speech Recognition: ASR) ได้ทำลายกำแพงแห่งเทคโนโลยี ทำให้การใช้งานคอมพิวเตอร์กลับกลายเป็นเรื่องง่าย ที่ผู้ใช้ในทุกระดับชั้นแม้ยังไม่คุ้นเคยและชำนาญในการใช้เทคโนโลยี สามารถโต้ตอบกับระบบคอมพิวเตอร์ได้อย่างง่ายดายเพียงผ่านเสียงพูด

เทคนิคที่ใช้การรู้จำเสียงพูด (Speech Recognition), การสร้างระบบที่ทนทานต่อสัญญาณรบกวน (Robust Speech Recognition), การสร้างระบบรู้จำเสียงพูดต่อเนื่องครอบคลุมคำศัพท์จำนวนมาก (Large Vocabulary Continuous Speech Recognition: LVCSR)

ข้อดีเป็นเทคโนโลยีที่ช่วยให้ผู้คนสามารถโต้ตอบกับระบบคอมพิวเตอร์ได้ง่ายขึ้นผ่านเสียงพูด

ข้อเสียหากมีความผิดปกติทางเสียง เช่น เจ็บป่วย พูดไม่ชัด อาจทำให้เกิดความขัดข้องในการสั่งการ ระบบอาจไม่เข้าใจคำสั่ง หรือประมวลเสียงผิดพลาด กลายเป็นคำอื่นไป

2.2 โปรแกรมวาจา (VAJA) เวอร์ชัน 6.0 [2]

ซอฟต์แวร์สังเคราะห์เสียงพูดภาษาไทย “วาจา” (VAJA) เป็นซอฟต์แวร์ที่แปลงข้อความภาษาไทยให้เป็นเสียงพูด หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา (Human Language

Technology Laboratory, HLT) ได้ดำเนินการวิจัยและพัฒนาอย่างต่อเนื่องมาตั้งแต่ปี พ.ศ. 2540 จนมาถึงเวอร์ชันล่าสุด วาจา 6.0 ที่ประยุกต์ใช้เทคนิคการสังเคราะห์เสียงแบบใหม่ ซึ่งอาศัยแบบจำลองทางสถิติ Hidden Markov Model ในการผลิตเสียง เทคนิคนี้ช่วยแก้ปัญหาการสะดุดของเสียงที่เคยเกิดขึ้นใน วาจาเวอร์ชันก่อนๆ เมื่อประกอบเข้ากับส่วนทำนายสัมผัส (prosody prediction module) ที่ช่วยวิเคราะห์ขอบเขตของวลี และทำนายความยาวของหน่วยเสียง ทำให้ได้เสียงสังเคราะห์ที่มีความเป็นธรรมชาติ เป็นที่พึงพอใจของผู้ใช้มากยิ่งขึ้น นอกจากนี้ วาจา 6.0 ยังสามารถสร้างเสียงพูดได้ครอบคลุมคำในภาษาไทย เนื่องจากมีส่วนวิเคราะห์คำอ่านที่สามารถวิเคราะห์ได้แม้แต่คำที่ไม่ปรากฏในพจนานุกรม

เทคนิคที่ใช้การวิเคราะห์ข้อความ (Text Analyzer), การทำนายสัมผัส (Prosody Prediction Module), การสังเคราะห์เสียงพูด (Speech Synthesis)

ข้อดี

1. สามารถนำเทคโนโลยีสังเคราะห์เสียงพูด มาแปลงข่าวสารอิเล็กทรอนิกส์ที่มีปริมาณมากและมีการปรับเปลี่ยนตลอดเวลา
2. สามารถเผยแพร่ข่าวสารผ่านเครือข่ายพื้นฐานที่มีอยู่เดิม
3. ผู้รับข่าวสารสามารถรับข่าวสาร โดยไม่ต้องละจากกิจกรรมที่ทำอยู่ โดยเฉพาะอย่างยิ่งในกรณีที่อยู่ ในภาวะที่ก่อให้เกิดอันตรายได้
4. สามารถประยุกต์ใช้กับอุปกรณ์สื่อสารอื่นๆ
5. การเพิ่มโอกาสให้คนพิการ

ข้อเสีย คำบางคำไม่สามารถพูดได้ถูกต้องเนื่องจากอาจเป็นศัพท์เฉพาะ หรือคำที่เป็นพหูพจน์ เช่น มีการเติม -s หรือ -es เป็น Postfix โปรแกรมจะไม่สามารถอ่านเป็นคำได้ โดยจะอ่านทีละตัวอักษรแทน

2.3 พจนานุกรมสื่ออิเล็กทรอนิกส์ไทย<->อังกฤษ [3]

LEXITRON เวอร์ชัน 3.0beta จัดทำโดยหน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ เป็นงานพัฒนาต่อเนื่องจาก เวอร์ชัน 2.6 ซึ่งเพิ่มเติมฐานข้อมูลจากเดิมที่มีอยู่ให้กลายเป็นคลังข้อความพจนานุกรมที่มีขนาดใหญ่ หลักการสำคัญของการพัฒนา LEXITRON คือ การนำเทคโนโลยีฐานข้อมูลขนาดใหญ่เข้ามาช่วยในการวิจัยและพัฒนาในสาขาการประมวลผลภาษา ธรรมชาติ เรียกว่า การสร้างพจนานุกรมจากฐานข้อมูลขนาดใหญ่ (Corpus-Based Dictionary)

วิธีการสร้างฐานข้อมูลขนาดใหญ่ (Corpus) เป็นการนำเทคโนโลยีทางคอมพิวเตอร์รวบรวมและคัด เลือกคำ ประโยค หรือข้อความ ที่มีใช้จริงและมีอัตราการปรากฏสูงในบริบทต่างๆ ของการใช้ภาษาจาก แหล่ง ข้อมูลและข่าวสารที่เผยแพร่ทางอินเทอร์เน็ตและแหล่งข้อมูลอื่นๆ ที่เชื่อถือได้ เช่น วรรณกรรม บทความ เอกสารทางวิชาการ ข้อมูลข่าวสารจากหนังสือพิมพ์ เป็นต้น

องค์ประกอบ

1. ฐานข้อมูลพจนานุกรม

1.1 อังกฤษ-ไทย (79,000 คำ) ประกอบด้วย คำศัพท์ภาษาอังกฤษ คำแปลภาษาไทย คำอ่าน คำหลัก คำค้นหา คำพ้องและคำตรงข้ามความหมายภาษาไทย คำพ้องและคำตรงข้ามความหมายภาษาอังกฤษ และตัวอย่างประโยคภาษาอังกฤษที่อ้างอิง จากคลังข้อความขนาดใหญ่ (Corpus-Based Sentence)

1.2 ไทย-อังกฤษ (51,000 คำ) ประกอบด้วย คำศัพท์ภาษาไทย คำแปลภาษาอังกฤษ คำพ้อง ความหมายภาษาไทย คำตรงข้ามความหมายภาษาไทย คำนิยามศัพท์ คำลักษณะนาม และตัวอย่างประโยค ภาษาไทยเทคนิคที่ใช้

2. คลังข้อมูลประโยค คือ Orchid Corpus ที่พัฒนาโดย ฝ่ายวิจัยและพัฒนาสาขาสารสนเทศ

3. โปรแกรมสำหรับการพัฒนาคือ J builder และ Apache

ข้อดี

แสดงความหมายและประเภทของคำ พร้อมทั้งประโยคตัวอย่างที่มีใช้จริงและมีอัตราการปรากฏสูงในบริบทต่างๆ ของการใช้ภาษา

2.4 โปรแกรมใช้เสียงพูดเพื่องานพิมพ์เอกสาร [4]

โปรแกรมใช้เสียงพูดเพื่องานพิมพ์เอกสาร (Speech to Text Control Program) โครงการวิจัยเรื่องนี้ มีความสามารถในการพิมพ์เอกสารบนโปรแกรมต่างๆ เหมือนการพิมพ์ผ่านทางกรกดปุ่มคีย์บอร์ด ได้ทั้งภาษาไทยและภาษาอังกฤษ โดยจะสั่งงานด้วยเสียงพูดผ่าน Microphone โดยโปรแกรมสามารถที่จะกำหนดค่าตั้งต้นให้แก่โปรแกรม ในกรณีที่มีผู้ใช้งานหลายคนซึ่งแต่ละคนจะมีคุณภาพของเสียงที่ต่างกัน

เทคนิคที่ใช้การวิเคราะห์สัญญาณเสียงพูด, การแปลงรูปแบบเสียงสัญญาณเสียงพูด, ฟังก์ชันกรองความถี่ดิจิทัล, Speech Recognition System, การเขียนโปรแกรมติดต่อระหว่างแอปพลิเคชันกับวินโดวส์โดยใช้ API, การเขียนโปรแกรมสั่งงานด้วยเสียงโดยใช้ Microsoft Speech SDK 5.1, Visual Basic 6.0 และการถอดรหัส

ข้อดี

1. พัฒนาโปรแกรมการใช้เสียงพูดในการควบคุมอุปกรณ์คีย์บอร์ดเพื่องานพิมพ์เอกสาร
2. สำหรับผู้พิการทางแขนและมือ หรือผู้ที่ไม่สะดวกในการพิมพ์เอกสาร สามารถใช้งานคอมพิวเตอร์ได้ใกล้เคียงคนปกติ
3. โปรแกรมสามารถพิมพ์เอกสารตามเสียงพูดได้

ข้อเสีย

1. ใช้ในงานพิมพ์เอกสารเฉพาะภาษาอังกฤษและภาษาไทยเท่านั้น
2. โปรแกรมจะสามารถพิมพ์เอกสารตามการพูดได้โดยการพูดต้องพูดแบบเป็นพยางค์ไม่ต่อเนื่อง
3. ระดับสัญญาณรบกวนต้องไม่เกิน 15 เดซิเบล

2.5 โปรแกรมช่วยคนพิการใช้งานคอมพิวเตอร์ [5]

โปรแกรมช่วยคนพิการใช้งานคอมพิวเตอร์ (Program for Disabled to Use Computer) โครงการวิจัยเรื่องนี้ มีความสามารถในการพิมพ์เอกสารบนโปรแกรมต่างๆ เหมือนการพิมพ์ผ่านทางกรกดปุ่มคีย์บอร์ด ได้ทั้งภาษาไทย ภาษาอังกฤษและฟังก์ชันปุ่มต่างๆ บนคีย์บอร์ด โดยจะสั่งงานด้วยเสียงพูดผ่าน Microphone โดยโปรแกรมสามารถที่จะกำหนดค่าตั้งต้นให้แก่โปรแกรม ในกรณีที่มีผู้ใช้งานหลายคนซึ่งแต่ละคนจะมีคุณภาพของเสียงที่ต่างกัน

เทคนิคที่ใช้การแปลงรูปแบบเสียงสัญญาณเสียงพูด, Speech Recognition System, Speech Signal, การเขียนโปรแกรมติดต่อระหว่างแอปพลิเคชันกับวินโดวส์โดยใช้ API, การเขียนโปรแกรมสั่งงานด้วยเสียงโดยใช้ Microsoft Speech SDK 5.1, VB.NET, การถอดรหัส, วงจรกรองความถี่หรือวงจรฟิลเตอร์

ข้อดี

1. พัฒนาคำสั่งในการใช้งานโปรแกรมควบคุมพอยเตอร์ด้วยดวงตา (Eyes Sensor Pointer Control Program) ร่วมกับโปรแกรมการใช้เสียงพูด เพื่องานพิมพ์เอกสาร(Speech to Text Control Program)ในการควบคุมอุปกรณ์พอยเตอร์และคีย์บอร์ดและแก้ไขจุดบกพร่องต่างๆ ของโปรแกรมที่ยังไม่สมบูรณ์
2. เพื่อให้สำหรับผู้พิการทางแขนและมือหรือผู้ที่ไม่สะดวกในการพิมพ์เอกสารและสามารถใช้งานคอมพิวเตอร์ได้ใกล้เคียงคนปกติ
3. โปรแกรมสามารถพิมพ์เอกสารตามเสียงพูดได้อย่างถูกต้องและแม่นยำ

ข้อเสีย

1. ใช้ในงานพิมพ์เอกสารเฉพาะภาษาอังกฤษและภาษาไทยเท่านั้น
2. โปรแกรมจะสามารถพิมพ์เอกสารตามการพูดได้ โดยการพูดต้องพูดแบบเป็นพยางค์ไม่ต่อเนื่อง
3. ระดับสัญญาณรบกวนต้องไม่เกิน 15 เดซิเบล

3. วัตถุประสงค์ของการวิจัย

1. ทั้งคนไทยและชาวต่างชาติที่ต้องการศึกษาภาษาอังกฤษ เข้าถึงบทเรียนหรือค้นหาความหมายของคำศัพท์ภาษาอังกฤษ จะสามารถทำได้ง่ายและรวดเร็วขึ้น จากการใช้แอปพลิเคชันนี้

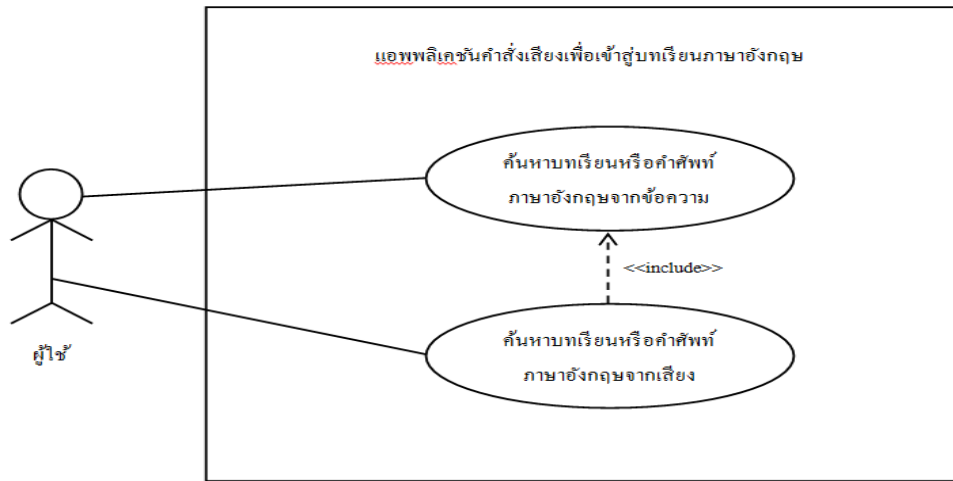
2. เทคโนโลยี Speech Recognition หรือการรู้จำเสียงพูด จะช่วยให้ผู้พิการ ผู้สูงอายุ หรือผู้ที่ไม่สามารถหรือไม่สะดวกในการพิมพ์ รวมทั้งบุคคลทั่วไป สามารถใช้งานแอปพลิเคชันได้สะดวกมากยิ่งขึ้น
3. การสังเคราะห์เสียงพูด จะช่วยอำนวยความสะดวกให้กับผู้ใช้ ในการรับข้อมูล
4. พัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์

4. วิธีดำเนินการวิจัย

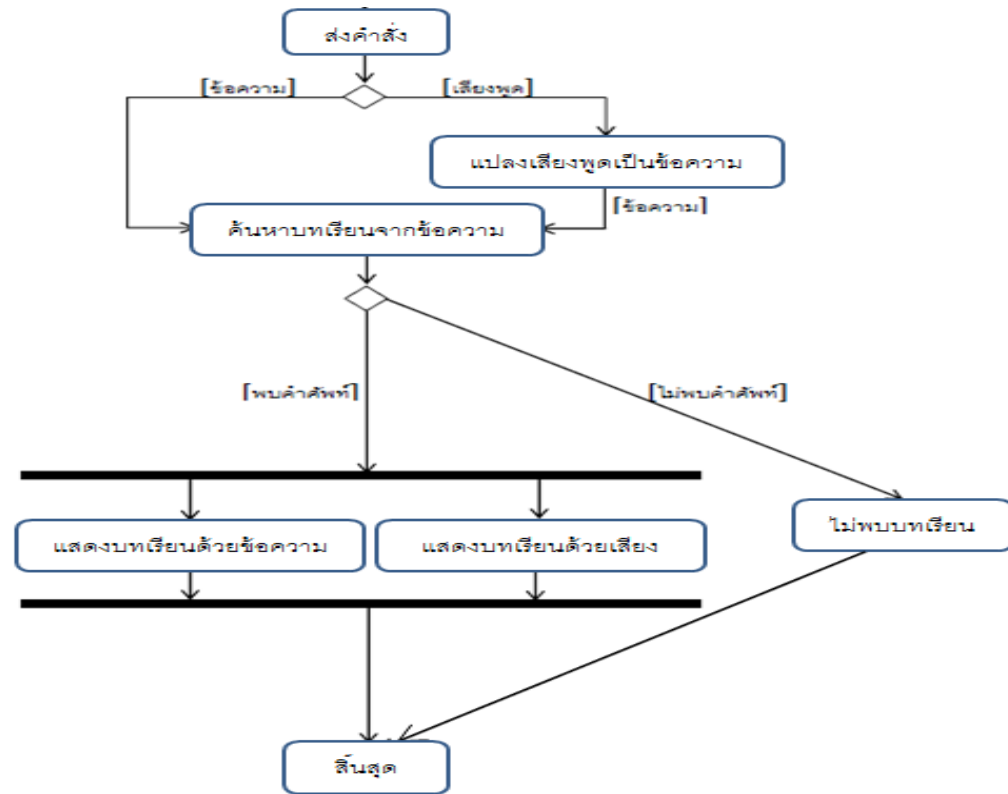
การศึกษาและรวบรวมข้อมูล สำหรับนำมาใช้พัฒนาแอปพลิเคชัน ประกอบด้วย Speech Recognition หรือการรู้จำเสียงพูด เป็นส่วนสำคัญของแอปพลิเคชันนี้ เพราะจะทำให้ผู้ใช้สามารถใช้เสียงในการพูดคำศัพท์ ที่ตรงกับชื่อบทเรียนภาษา อังกฤษเพื่อเข้าถึงบทเรียนที่ต้องการได้, การเรียกใช้โปรแกรม วาจาแอนดรอยด์ หรือ VAJA TTS Engine ในการสังเคราะห์เสียงพูด เพื่อใช้ในการอ่านความหมายของคำศัพท์เป็นเสียงพูดให้แก่ผู้ใช้แอปพลิเคชัน, พจนานุกรมสื่ออิเล็กทรอนิกส์หรือ LEXITRON ของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ, ภาษาจาวา เป็นภาษาเครื่องที่ใช้ในการเขียนแอปพลิเคชัน, และศึกษาการเขียนแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ สำหรับเครื่องมือที่ใช้ประเมินความพึงพอใจของผู้ใช้คือแบบสอบถามความพึงพอใจแบบ Likert scale ที่มีการแบ่งระดับคะแนนเป็น 5 ระดับ คือ มากที่สุด มาก ปานกลาง น้อย น้อยที่สุด ซึ่งคำถามแบ่งเป็น 4 ด้าน ได้แก่ ด้านการจัดรูปแบบ ด้านการออกแบบ ด้านการแนะนำสื่อ และด้านเนื้อหาบทเรียน

จากการศึกษาและรวบรวมข้อมูลพบว่า แอปพลิเคชันลักษณะดังกล่าวในปัจจุบัน ไม่เหมาะสำหรับผู้ที่มีปัญหาในเรื่องเสียง เช่น มีอาการเจ็บป่วย ลิ้นไก่สั้น พูดไม่ชัด เป็นต้น เพราะอาจทำให้ระบบเข้าใจผิดได้ คำศัพท์หรือบทเรียนที่ไม่ตรงกับที่ต้องการหรือค้นหาไม่พบ ทั้งที่มีคำศัพท์ดังกล่าวในระบบ คณะผู้วิจัยจึงแก้ไขโดยออกแบบให้องค์ประกอบของแอปพลิเคชันพจนานุกรมคำสั่งเสียงบนระบบปฏิบัติการแอนดรอยด์ ในส่วนของผู้ใช้แอปพลิเคชันกับระบบ ดังนี้

1. ผู้ใช้ป้อนคำศัพท์ด้วยคำสั่งเสียง หรือการพิมพ์ข้อความ ให้กับแอปพลิเคชัน
2. กรณีที่ผู้ใช้ ใช้คำสั่งเสียงในการป้อนคำศัพท์เข้าสู่ระบบ ระบบก็จะนำสัญญาณเสียงที่ได้รับ ไปเข้าสู่ขั้นตอนการรู้จำเสียงพูดเพื่อแปลเป็นข้อความ แล้วนำไปค้นหาบทเรียนจากฐานข้อมูล จากนั้นนำผลลัพธ์ที่ได้ส่งกลับคืนสู่ผู้ใช้ โดยมีทั้งรูปแบบข้อความและการสังเคราะห์เสียงพูด
3. กรณีที่ผู้ใช้ ใช้การพิมพ์คำศัพท์เข้าสู่ระบบ ระบบก็จะทำการค้นหาคำศัพท์ ที่ตรงกันหรือเกี่ยวข้องกับชื่อบทเรียนในฐานข้อมูล และส่งผลลัพธ์กลับมาให้ทั้งในรูปแบบการสังเคราะห์เสียงพูดและข้อความดังภาพที่ 1 แสดง Use Case Diagram และภาพที่ 2 แสดง Activity Diagram



ภาพที่ 1 Use Case Diagram



ภาพที่ 2 Activity Diagram

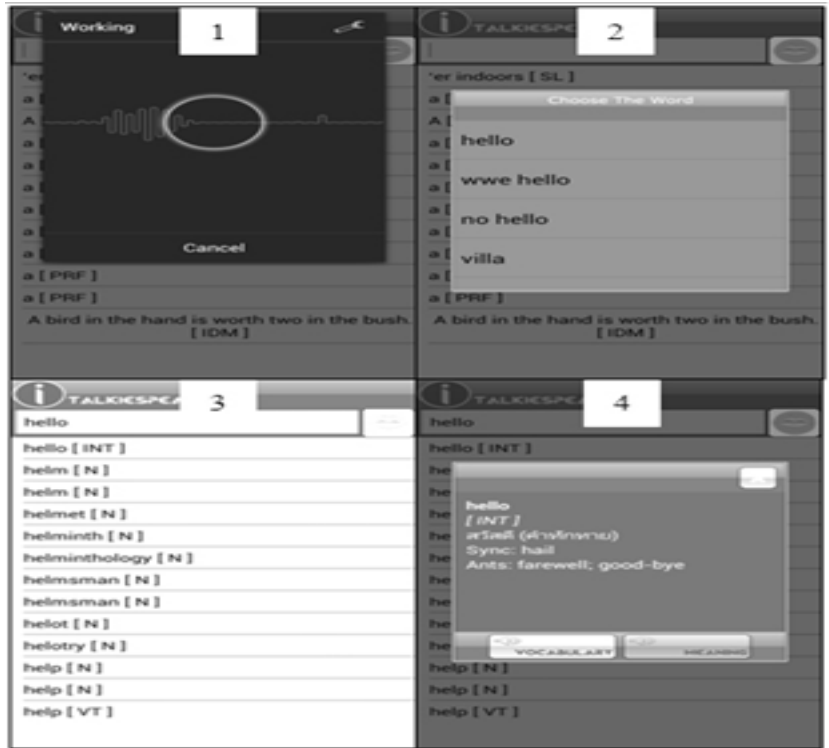
เครื่องมือและภาษาที่ใช้ในการพัฒนา

- Google Speech to Text
- แอปพลิเคชัน VAJA Android (Vaja TTS Engine)
- ฐานข้อมูล SQLite
- พจนานุกรม Lexitron
- JDK Version 5 ขึ้นไป
- Eclipse, ADT
- Android SDK
- ภาษา JAVA
- โทรศัพท์มือถือระบบปฏิบัติการแอนดรอยด์

5. ผลการวิจัยและอภิปรายผล

แอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ บนระบบปฏิบัติการแอนดรอยด์นี้ เป็นโปรแกรมเพื่อการประยุกต์ใช้งานบนเครือข่ายสำหรับอุปกรณ์อิเล็กทรอนิกส์เคลื่อนที่ที่ใช้ระบบปฏิบัติการแอนดรอยด์ เวอร์ชัน 2.2 ขึ้นไป แอปพลิเคชันนี้สามารถช่วยการเข้าถึงบทเรียนภาษาอังกฤษ รวมถึงแปลความหมายจากภาษาอังกฤษเป็นภาษาไทย และภาษาไทยเป็นภาษาอังกฤษด้วย โดยผู้ใช้จะสามารถป้อนคำศัพท์ภาษาอังกฤษเพื่อเป็นคำค้นบทเรียนได้โดยใช้เสียงพูด หรือการพิมพ์คำศัพท์ผ่านทางแป้นพิมพ์ และจะแสดงผลลัพธ์ขึ้นมาในรูปแบบของบทสนทนาภาษาอังกฤษ รวมทั้งความหมายของคำศัพท์ต่างๆ ประเภทของคำศัพท์ คำเหมือน และคำตรงข้าม ออกมาทั้งในรูปแบบเสียงและข้อความ

คณะผู้วิจัยได้ใช้โปรแกรม Eclipse เป็นเครื่องมือในการพัฒนาแอปพลิเคชันบนแอนดรอยด์ โดยใช้ภาษา JAVA เป็นหลัก ใช้ SQLite ในการเก็บ Database ใช้ Google Voice Search สำหรับ Speech Recognition และ VAJA Text-To-Speech Engine ในการสังเคราะห์เสียงพูดภาษาอังกฤษและภาษาไทย คณะผู้วิจัยได้ทำการทดสอบแอปพลิเคชัน โดยการติดตั้งลงบนอุปกรณ์โทรศัพท์เคลื่อนที่จริงที่ใช้ระบบปฏิบัติการแอนดรอยด์ เนื่องจากโทรศัพท์จำลอง (Emulator) ไม่สามารถป้อนข้อมูลเสียงพูดและสังเคราะห์เสียงพูดได้



ภาพที่ 3 ขั้นตอนการทำงานของแอปพลิเคชัน

ภาพที่ 3 แสดงขั้นตอนการทำงานของแอปพลิเคชันโดยเริ่มจากรับคำศัพท์เป็นเสียงจากผู้ใช้ แล้วแปลงเป็นตัวอักษรของคำศัพท์เพื่อดึงบทเรียนที่เกี่ยวข้องให้ผู้ใช้เลือก และหลังจากผู้ใช้เลือกบทเรียนหรือคำศัพท์ใด ก็จะปรากฏบทสนทนาภาษาอังกฤษและความหมายของคำศัพท์ที่เกี่ยวข้องในบทสนทนา

แอปพลิเคชันแบ่งออกเป็น 3 ส่วน คือ

ส่วนที่ 1 การป้อนคำศัพท์ ใช้เทคนิคในการรู้จำเสียงพูด (Speech Recognition) ที่ให้ผู้ใช้สามารถพูดคำศัพท์ภาษาอังกฤษแทนการพิมพ์จากแป้นพิมพ์ได้ แล้วเลือกบทเรียนที่ตรงกับความต้องการ

ส่วนที่ 2 เป็นส่วนของการค้นหาความหมายของคำศัพท์ จากฐานข้อมูล

ส่วนที่ 3 เป็นส่วนของการแสดงบทเรียนที่ค้นหาและความหมายของคำศัพท์ที่เกี่ยวข้อง และการสังเคราะห์เสียงพูด (Voice Synthesis) ทั้งภาษาอังกฤษและภาษาไทย

โดยการทำงานของแอปพลิเคชันถือว่าทำงานได้ดีแต่อาจจะมีข้อจำกัดในประเด็นต่อไปนี้

1. ผู้ใช้ต้องพูดเสียงอักขระภาษาอังกฤษให้ชัดเจน หรือใกล้เคียงที่สุด เพื่อให้ได้คำศัพท์ที่ตรงกับความต้องการ
2. การใช้งานคำสั่งเสียงอาจไม่เหมาะกับผู้ที่ปัญหาในเรื่องของการออกเสียงเช่น มีอาการป่วย ลิ้นไก่สั้น

พูดไม่ชัด เป็นต้น เพราะระบบไม่สามารถจับความเข้ากันของเสียง และคำศัพท์ที่ต้องการได้ ซึ่งอาจจะทำให้ได้คำศัพท์รวมทั้งบทเรียนที่ประมวลผลออกมาไม่ตรงตามเป้าหมาย

3. แอปพลิเคชันนี้ใช้ได้กับอุปกรณ์อิเล็กทรอนิกส์เคลื่อนที่ ที่ใช้ระบบปฏิบัติการแอนดรอยด์ เวอร์ชัน 2.2 ขึ้นไป

ในด้านการประเมินความพึงพอใจ คณะผู้วิจัยได้ใช้แบบสอบถามแบบ 5 ระดับสำหรับสอบถามผู้ใช้ ซึ่งเป็นกลุ่มตัวอย่าง 2 กลุ่ม ได้แก่ผู้ใช้ทั่วไป และผู้เชี่ยวชาญ โดยกลุ่มตัวอย่างกลุ่มแรกคือนักศึกษาศาสาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี จำนวน 34 คน พบผลประเมินดังตารางที่ 1

ตารางที่ 1 แสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานความพึงพอใจในการใช้งานแอปพลิเคชันของผู้ใช้งาน

| ลำดับที่ | เกณฑ์การประเมิน | N = 34 | | |
|----------|---|-----------|------|------------------|
| | | \bar{X} | SD. | ระดับความพึงพอใจ |
| 1 | การจัดรูปแบบแอปพลิเคชันง่ายต่อการอ่านและเข้าใจ | 4.41 | 0.66 | มาก |
| 2 | ความเหมาะสมในการเลือกใช้สีและขนาดของตัวอักษร | 4.24 | 0.82 | มาก |
| 3 | ความเหมาะสมในการใช้ข้อความและรูปเพื่ออธิบายสื่อความหมาย | 4.24 | 0.70 | มาก |
| 4 | ความเป็นมาตรฐานเดียวกันในการออกแบบ | 4.21 | 0.64 | มาก |
| 5 | ความเหมาะสมในการโต้ตอบปฏิสัมพันธ์กับผู้ใช้ | 4.26 | 0.62 | มาก |
| 6 | ความเหมาะสมในการวางตำแหน่งของส่วนประกอบ | 4.35 | 0.77 | มาก |
| 7 | คำศัพท์ที่ใช้ผู้ใช้มีความคุ้นเคยและเข้าใจง่าย | 4.15 | 0.74 | มาก |
| 8 | ภาพโดยรวมของการออกแบบแอปพลิเคชัน | 4.26 | 0.62 | มาก |
| 9 | ข้อมูลมีความถูกต้อง คลอบคลุมและชัดเจน | 4.21 | 0.69 | มาก |
| 10 | มีความปลอดภัยและน่าเชื่อถือในการใช้งาน | 4.26 | 0.75 | มาก |
| 11 | ความง่ายและเข้าใจต่อการเข้าใช้แอปพลิเคชัน | 4.21 | 0.77 | มาก |
| 12 | ความเหมาะสมของการเชื่อมโยงข้อมูลมีความถูกต้อง | 4.29 | 0.68 | มาก |
| 13 | มีความสอดคล้องกับความต้องการของผู้ใช้งาน | 4.44 | 0.62 | มาก |
| 14 | การแสดงผลการเข้าถึงได้อย่างรวดเร็วและมีประสิทธิภาพ | 4.15 | 0.82 | มาก |
| 15 | มีการจัดหมวดหมู่ข้อมูลสะดวกต่อการค้นหาข้อมูล | 4.24 | 0.82 | มาก |
| 16 | ข้อมูลมีประโยชน์ตรงกับความต้องการของผู้เข้าใช้แอปพลิเคชัน | 4.35 | 0.73 | มาก |

| | | | | |
|------------|--|-------------|-------------|------------|
| 17 | ความพึงพอใจโดยรวมของการเข้าใช้งานแอปพลิเคชัน | 4.29 | 0.80 | มาก |
| 18 | เนื้อหาและข่าวสารมีประโยชน์ต่อผู้เข้าชม | 4.26 | 0.75 | มาก |
| 19 | ความถูกต้องของข้อมูลข่าวสารต่างๆ | 4.32 | 0.64 | มาก |
| 20 | ข้อมูลข่าวสารมีความทันสมัยและเป็นปัจจุบัน | 4.47 | 0.75 | มาก |
| 21 | มีประโยชน์ในการสื่อสารและประชาสัมพันธ์ | 4.50 | 0.56 | มากที่สุด |
| รวม | | 4.29 | 0.71 | มาก |

สำหรับผลการประเมินความพึงพอใจของผู้เชี่ยวชาญจำนวน 6 คนซึ่งเป็นกลุ่มตัวอย่างที่ 2 พบผลประเมินดังตารางที่ 2

ตารางที่ 2 แสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานความพึงพอใจในการใช้งานแอปพลิเคชันของผู้เชี่ยวชาญ

| ลำดับที่ | เกณฑ์การประเมิน | N = 6 | | |
|----------|---|-----------|------|------------------|
| | | \bar{X} | SD. | ระดับความพึงพอใจ |
| 1 | การจัดรูปแบบแอปพลิเคชันง่ายต่อการอ่านและเข้าใจ | 4.00 | 0.82 | มาก |
| 2 | ความเหมาะสมในการเลือกใช้สีและขนาดของตัวอักษร | 4.00 | 0.82 | มาก |
| 3 | ความเหมาะสมในการใช้ข้อความและรูปเพื่ออธิบายสื่อความหมาย | 4.00 | 0.82 | มาก |
| 4 | ความเป็นมาตรฐานเดียวกันในการออกแบบ | 5.00 | 0.00 | มากที่สุด |
| 5 | ความเหมาะสมในการโต้ตอบปฏิสัมพันธ์กับผู้ใช้ | 4.00 | 0.00 | มาก |
| 6 | ความเหมาะสมในการวางตำแหน่งของส่วนประกอบ | 4.00 | 0.00 | มาก |
| 7 | คำศัพท์ที่ใช้ผู้ใช้มีความคุ้นเคยและเข้าใจง่าย | 4.33 | 0.47 | มาก |
| 8 | ภาพโดยรวมของการออกแบบแอปพลิเคชัน | 4.33 | 0.47 | มาก |
| 9 | ข้อมูลมีความถูกต้อง คลอบคลุมและชัดเจน | 4.33 | 0.47 | มาก |
| 10 | มีความปลอดภัยและน่าเชื่อถือในการใช้งาน | 4.67 | 0.47 | มากที่สุด |
| 11 | ความง่ายและเข้าใจต่อการเข้าใช้แอปพลิเคชัน | 4.00 | 0.00 | มาก |
| 12 | ความเหมาะสมของการเชื่อมโยงข้อมูลมีความถูกต้อง | 4.67 | 0.47 | มากที่สุด |
| 13 | มีความสอดคล้องกับความต้องการของผู้ใช้งาน | 4.67 | 0.47 | มากที่สุด |
| 14 | การแสดงผลการเข้าถึงได้อย่างรวดเร็วและมีประสิทธิภาพ | 4.00 | 0.82 | มาก |
| 15 | มีการจัดหมวดหมู่ข้อมูลสะดวกต่อการค้นหาข้อมูล | 4.33 | 0.47 | มาก |
| 16 | ข้อมูลมีประโยชน์ตรงกับความต้องการของผู้เข้าใช้แอปพลิเคชัน | 4.00 | 0.82 | มาก |

| | | | | |
|------------|---|-------------|-------------|------------|
| 17 | ความพึงพอใจโดยรวมของการใช้งานแอปพลิเคชัน | 4.33 | 0.47 | มาก |
| 18 | เนื้อหาและข่าวสารมีประโยชน์ต่อผู้เข้าชม | 4.67 | 0.47 | มากที่สุด |
| 19 | ความถูกต้องของข้อมูลข่าวสารต่างๆ | 4.00 | 0.00 | มาก |
| 20 | ข้อมูลข่าวสารมีความทันสมัยและเป็นปัจจุบัน | 4.67 | 0.47 | มากที่สุด |
| 21 | มีประโยชน์ในการสื่อสารและประชาสัมพันธ์ | 5.00 | 0.00 | มากที่สุด |
| รวม | | 4.33 | 0.42 | มาก |

6. สรุป

แอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ บนระบบปฏิบัติการแอนดรอยด์นี้ สามารถช่วยการเข้าถึงบทเรียนภาษาอังกฤษได้ และสามารถแปลความหมายคำศัพท์จากภาษาอังกฤษเป็นภาษาไทยได้ โดยสามารถป้อนคำสั่งเสียงพูดภาษาอังกฤษ จากนั้นโปรแกรมจะนำไปเข้าสู่ขั้นตอนการรู้จำเสียงพูดเพื่อแปลงเป็นข้อความ แล้วนำไปค้นหาบทเรียนจากฐานข้อมูล โดยจะออกมาในรูปแบบข้อความ และเสียงพูดของบทสนทนาทั้งภาษาอังกฤษและภาษาไทย

จากการสำรวจระดับความพึงพอใจในการใช้งานแอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ โดยใช้แบบสอบถามความพึงพอใจแบบ Likert scale ที่มีการแบ่งระดับคะแนนเป็น 5 ระดับ ในส่วนของผู้ใช้งานพบว่า ผู้ใช้งานร้อยละ 85 มีระดับความพึงพอใจในการใช้งานระบบอยู่ในระดับมาก คิดค่าเฉลี่ยโดยรวมได้เท่ากับ 4.29 และมีส่วนเบี่ยงเบนมาตรฐานเฉลี่ยเท่ากับ 0.71 และจากการสำรวจค่าความพึงพอใจของผู้เชี่ยวชาญต่อการใช้งานแอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษ ด้วยแบบสอบถามผ่านทาง Google Documents พบว่าผู้เชี่ยวชาญมีระดับความพึงพอใจในการใช้งานระบบอยู่ในระดับมาก คิดค่าเฉลี่ยโดยรวมได้เท่ากับ 4.33 และมีส่วนเบี่ยงเบนมาตรฐานเฉลี่ยเท่ากับ 0.42

ข้อเสนอแนะในการพัฒนาแอปพลิเคชันคำสั่งเสียงเพื่อเข้าสู่บทเรียนภาษาอังกฤษบนระบบปฏิบัติการแอนดรอยด์ให้ดียิ่งขึ้นต่อไป ดังต่อไปนี้

1. พัฒนาแอปพลิเคชันให้สามารถรองรับคำสั่งเสียงพูด หรือ Speech Recognition ภาษาไทยได้
2. เสริมฟังก์ชันให้ผู้ใช้สามารถเพิ่มบทเรียน คำศัพท์ภาษาอังกฤษ หรือเพิ่มรูปภาพที่เกี่ยวข้องกับคำศัพท์นั้น ๆ ให้ผู้ใช้เห็นภาพที่ชัดเจนยิ่งขึ้นได้ด้วยตนเอง
3. เพิ่มโหมดการแปลความหมายที่เป็นประโยคของภาษาอังกฤษหรือภาษาไทยได้

คณะผู้วิจัยหวังว่าในอนาคต จะสามารถพัฒนาแอปพลิเคชันนี้ ให้รองรับคำสั่งเสียงพูดภาษาไทยได้ รวมทั้งให้มีการรับคำสั่งเสียงพูดภาษาอังกฤษที่แม่นยำ และตรงกับที่ต้องการมากยิ่งขึ้นเพื่อเตรียมเข้าสู่ประชาคมเศรษฐกิจอาเซียนต่อไป

7. เอกสารอ้างอิง

- ชัย วุฒิวิวัฒน์ชัย และคณะ. (2550). **ระบบรู้จำเสียงพูด (Automatic Speech Recognition)**. หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. สืบค้นเมื่อ 5 พฤษภาคม 2555, จาก http://www.digitized-Thailand.org/index.php?option=com_content &view=article&id=44&itemid=104&lang=th
- ชัย วุฒิวิวัฒน์ชัย และคณะ. (2552). **พจนานุกรมสื่ออิเล็กทรอนิกส์ไทย <-> อังกฤษ LEXITRON**. หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. สืบค้นเมื่อ 1 สิงหาคม 2555, จาก http://lexitron.nectec.or.th/2009_1/index_en.php?q=common_manager
- ชัย วุฒิวิวัฒน์ชัย และคณะ. (2554). **โปรแกรมวาจา (VAJA) เวอร์ชัน 6.0**. หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ. สืบค้นเมื่อ 15 พฤษภาคม 2555, จาก <http://vaja.nectec.or.th>
- ทีฆายุ สมานรักษ์ และคณะ. (2549). **โปรแกรมใช้เสียงพูดเพื่องานพิมพ์เอกสาร (Speech to Text Control Program)**. วิทยาศาสตร์บัณฑิต (สาขาวิทยาการคอมพิวเตอร์) มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, ปทุมธานี.
- อานนท์ แสงเมฆ และคณะ. (2552). **โปรแกรมช่วยคนพิการใช้งานคอมพิวเตอร์ (Program for Disabled to Use Computer)**. วิทยาศาสตร์บัณฑิต (สาขาเทคโนโลยีคอมพิวเตอร์) มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, ปทุมธานี.