



แอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์

Dictionary Application by Voice Commands for Android Operating System

บุรัสกร อยู่สุข

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี

E-mail: burasakorn@hotmail.com

บทคัดย่อ

ประชาคมเศรษฐกิจอาเซียนส่งผลให้ประเทศต่างๆ ในกลุ่มอาเซียนติดต่อสื่อสารกันมากขึ้น และภาษากลางอย่างภาษาอังกฤษก็ยิ่งทวีความสำคัญมากขึ้นด้วย ซึ่งหากใครที่ยังใช้ภาษาอังกฤษในการสื่อสารได้ไม่ดีเท่าที่ควรก็อาจเกิดความเสียเปรียบได้ ดังนั้นเครื่องมือช่วยแปลภาษาจะช่วยแก้ไขปัญหานี้ บทความนี้จะนำเสนอแนวทางการสร้างแอปพลิเคชัน สำหรับช่วยค้นหาความหมายของคำศัพท์ต่างๆ แอปพลิเคชันนี้ใช้เทคโนโลยีรู้จำเสียง หรือ Speech Recognition ทำให้สามารถแปลความหมายคำศัพท์จากภาษาอังกฤษเป็นภาษาไทยและภาษาไทยเป็นภาษาอังกฤษได้ โดยผู้ใช้งานจะสามารถเลือกป้อนคำศัพท์ภาษาอังกฤษได้มากกว่า 500 คำ โดยใช้เสียงพูดหรือการพิมพ์คำศัพท์ผ่านทางแป้นพิมพ์ และจะแสดงผลการค้นหาคำศัพท์ขึ้นมาทั้งแบบข้อความและเสียง โดยผลลัพธ์จะแสดงคำศัพท์ที่ค้นหาทั้งรูปแบบของ ความหมายของคำศัพท์ ประเภทของคำศัพท์ คำเหมือน และคำตรงข้าม ให้แก่ผู้ใช้ 30 คน โดยผลจากการทดลองใช้งานพบว่า ผู้ใช้ที่เป็นนักศึกษา 90% มีความพอใจในแอปพลิเคชันนี้

คำสำคัญ: พจนานุกรม, สั่งการด้วยเสียง, ระบบรู้จำเสียงพูด, แอนดรอยด์แอปพลิเคชัน

1. บทนำ

เพื่อรองรับประชาคมเศรษฐกิจอาเซียน (AEC) ที่กำลังจะมาถึงในปี 2558 นี้ กลุ่มประเทศอาเซียนต่างได้กำหนดนโยบายในการดำเนินธุรกรรมระหว่างกัน ทั้งทางด้านเศรษฐกิจ สังคม และวัฒนธรรม และการที่จะทำ

Received: Oct 28, 2013

Revised: Nov 20, 2013

Accepted: Dec 23, 2013

ธุรกรรมระหว่างกันได้อย่างมีคุณภาพนั้น ภาษากลางอย่างภาษาอังกฤษย่อมทวีความสำคัญเป็นอย่างมาก กลุ่มคนอาเซียนที่เดินทางเข้ามาติดต่อทางธุรกิจ ท่องเที่ยว หรืออาศัยในประเทศไทยจะใช้ภาษาอังกฤษเป็นหลักในการติดต่อสื่อสาร ในขณะที่สังคมไทยใช้ภาษาไทยเป็นหลัก จึงเป็นอุปสรรคสำหรับคนต่างชาติที่จะติดต่อการสื่อสารกับคนไทยและใช้ชีวิตในประเทศไทย ดังนั้นหากมีเครื่องมือช่วยแปลภาษาจะส่งผลให้เค้าเหล่านั้นมาทำธุรกรรมต่างๆ ในประเทศไทยได้สะดวกมากยิ่งขึ้น

เครื่องมือช่วยแปลภาษานี้เป็นแอปพลิเคชันที่ใช้เทคโนโลยีรู้จำเสียง (Speech Recognition) โดยแอปพลิเคชันจะนำคลื่นเสียงป้อนเข้าสู่ขั้นตอนการรู้จำเสียงพูดเพื่อแปลงเป็นข้อความ จากนั้นนำข้อความไปค้นหาความหมายของคำศัพท์จากฐานข้อมูลที่อยู่ในระบบ แล้วแสดงออกมาในรูปแบบข้อความและเสียง โดยใช้เทคนิคการสังเคราะห์เสียง (Voice Synthesis) ในการอ่านคำศัพท์และความหมาย ซึ่งในที่นี้ใช้แอปพลิเคชัน VAJA Android ที่พัฒนาโดยศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (เนคเทค) ในการสังเคราะห์เสียงพูด

เทคโนโลยี Speech Recognition ได้ถูกพัฒนามาอย่างต่อเนื่อง และแพร่หลายไปยังทุกๆ อุปกรณ์อิเล็กทรอนิกส์เพื่อช่วยอำนวยความสะดวกให้กับผู้ใช้ ได้ใช้งานอุปกรณ์นั้นๆ ง่ายขึ้น เทคนิคของการรู้จำเสียงพูดนั้นระบบจะสร้างโครงข่ายคำ (Word Network) โดยอาศัยแบบจำลองภาษา (Language Model) โดยในแต่ละคำจะประกอบด้วยแบบจำลองเสียง (Acoustic Model) ประเภทฮิดเดนมาร์คอฟ (Hidden Markov Model: HMM) ของแต่ละหน่วยเสียงที่มาประกอบกันเป็นคำนั้นๆ หลังจากนั้นระบบจะป้อนเวกเตอร์ของ Speech Feature เข้าไปยังโครงข่ายคำเพื่อคำนวณค่าความน่าจะเป็นของเสียงที่เข้ามาว่าจะตรงกับเส้นทางใดในโครงข่ายคำ และจะเลือกคำตอบที่ดีที่สุดเป็นเส้นทางที่มีค่าความน่าจะเป็นสูงที่สุด โดยขั้นตอนทั้งหมดจะเป็นการถอดรหัส (Decoding) ซึ่งในทางปฏิบัติ นักวิจัยได้ใช้อัลกอริทึมการถอดรหัสแบบ Viterbi beam-search เพื่อลดปริมาณการใช้หน่วยความจำและการคำนวณด้วยการลบเส้นทางที่มีค่าความน่าจะเป็นต่ำทิ้งไป

แอนดรอยด์เป็นชื่อของระบบปฏิบัติการ ที่มีลักษณะโครงสร้างยืดหยุ่นสามารถติดตั้งลงในอุปกรณ์อิเล็กทรอนิกส์ต่างๆ ที่มีความหลากหลายทางฮาร์ดแวร์ ไม่ว่าจะเป็นโทรศัพท์มือถือหรือแท็บเล็ตจึงมีผู้ใช้งานเป็นจำนวนมาก อีกทั้งการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์นั้นสามารถทำได้ง่าย ทำให้แอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์มีจำนวนมากขึ้น และได้รับความนิยมแพร่หลายอย่างรวดเร็ว ดังนั้นการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์จึงเป็นสิ่งที่น่าสนใจ

แอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์นี้ จะช่วยอำนวยความสะดวกให้ผู้ใช้ทุกคน โดยเฉพาะกับผู้พิการ ผู้สูงอายุ หรือผู้ที่ไม่สามารถใช้นิ้วมือพิมพ์คำศัพท์ผ่านทางแป้นพิมพ์ได้ สามารถค้นหาความหมายของคำศัพท์ได้ง่ายและรวดเร็วยิ่งขึ้น เพียงแค่ใช้เสียงพูดระบบก็จะทำการประมวลผลเป็นข้อความ จากนั้นนำไปค้นหาและแสดงผลออกมาทั้งในรูปแบบข้อความและการสังเคราะห์เสียงพูดเพื่อความสะดวกในการรับข้อมูล

2. วรรณกรรมที่เกี่ยวข้อง

2.1 ระบบรู้จำเสียงพูด (Automatic Speech Recognition) [2]

เทคโนโลยีการรู้จำเสียงพูดเป็นระบบที่ช่วยแปลงเสียงพูดให้เป็นข้อความ ที่ระบบคอมพิวเตอร์สามารถนำมาใช้ประมวลผลต่อไปได้ โดยเทคโนโลยีนี้ได้รับการพัฒนาขึ้นภายในหน่วยงานวิจัยของเนคเทค เพื่อช่วยลดข้อจำกัดในการเข้าถึงเทคโนโลยีสมัยใหม่ โดยการใช้ระบบประมวลผลภาษาธรรมชาติ เป็นตัวเชื่อมโยงผู้ใช้เข้าสู่อบริการทางคอมพิวเตอร์ในรูปแบบต่างๆ ในขั้นตอนรู้จำเสียง ระบบจะสร้างโครงข่ายคำ (Word Network) โดยอาศัยแบบจำลองภาษา(Language Model) โดยในแต่ละคำจะประกอบด้วยแบบจำลองเสียง (Acoustic Model) ประเภทฮิดเดนมาร์คอฟ (Hidden Markov Model: HMM) ของแต่ละหน่วยเสียงที่มาประกอบกันเป็นคำนั้นๆ หลังจากนั้นระบบจะป้อนเวกเตอร์ของ Speech Feature เข้าไปยังโครงข่ายคำเพื่อคำนวณค่าความน่าจะเป็นของเสียงที่เข้ามาว่าจะตรงกับเส้นทางใดในโครงข่ายคำ และจะเลือกคำตอบที่ดีที่สุดเป็นเส้นทางที่มีค่าความน่าจะเป็นสูงที่สุด โดยขั้นตอนทั้งหมดจะเป็นการถอดรหัส (Decoding) ซึ่งในทางปฏิบัติ นักวิจัยได้ใช้อัลกอริทึมการถอดรหัสแบบ วิเทอบี บีม เสิร์จ (Viterbi Beam Search) เพื่อลดปริมาณการใช้หน่วยความจำ และการคำนวณด้วยการลบเส้นทางที่มีค่าความน่าจะเป็นต่ำทิ้งไป

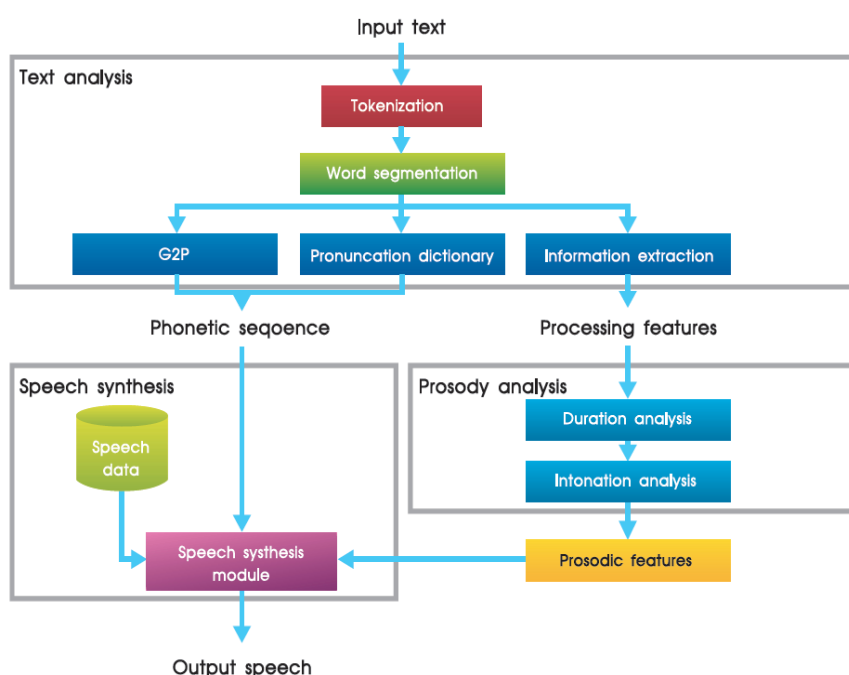
ทั้งนี้หน่วยงานวิจัยของเนคเทคได้แบ่งการพัฒนาการรู้จำเสียงเป็น 2 กลุ่มหลักๆ คือ กลุ่มการสร้างระบบที่ทนทานต่อสัญญาณรบกวน (Robust Speech Recognition) และกลุ่มการสร้างระบบรู้จำเสียงพูดต่อเนื่องครอบคลุมคำศัพท์จำนวนมาก (Large Vocabulary Continuous Speech Recognition: LVCSR)

โดยกลุ่มการสร้างระบบที่ทนทานต่อสัญญาณรบกวน ได้เน้นการเลือกแบบจำลองเสียงที่เหมาะสมกับสถานะแวดล้อมของเสียงที่เข้ามา ซึ่งอาศัยการแยกแยะสัญญาณรบกวนอัตโนมัติ (Noise Classification) แบบโครงสร้างต้นไม้ (Tree Structure) และได้นำเสนอวิธีการปรับแบบจำลองเสียงให้ใกล้เคียงกับเสียงที่เข้ามามากขึ้น ด้วยการใช้ข้อมูลฝึกฝนที่จำลองขึ้นจากสัญญาณรบกวนที่ได้จากเสียงอินพุต (Simulated Data Adaptation) วิธีการนี้สามารถประยุกต์ใช้กับระบบรู้จำในการได้ตอบสนองผ่านโทรศัพท์ เนื่องจากมีเสียงอินพุตสั้นและมีสัญญาณรบกวนค่อนข้างมากและหลากหลาย

สำหรับกลุ่มการสร้างระบบรู้จำเสียงพูดต่อเนื่องครอบคลุมคำศัพท์จำนวนมาก จะเน้นการถอดความเสียงพูดต่อเนื่องที่ประกอบด้วยคำศัพท์จำนวนมาก (LVCSR) โดยได้พัฒนาระบบรู้จำเสียงสำหรับการถอดความเสียงอ่าน (Dictation) การถอดความเสียงรายงานข่าวผ่านสัญญาณโทรทัศน์ (Broadcast News) และริเริ่มการถอดความเสียงสนทนาผ่านสายโทรศัพท์ (Spontaneous Telephone Conversation) ในกลุ่มงานนี้ต้องอาศัยแบบจำลองภาษาแบบเอ็นแกรม (N-gram) ซึ่งสามารถจัดการกับคำศัพท์จำนวนมากได้อย่างมีประสิทธิภาพ ในกรณีสัญญาณโทรทัศน์ได้มีการวิจัยระบบที่สามารถจัดการคำเฉพาะที่ไม่เคยพบในพจนานุกรมได้ โดยใช้หน่วยพยางค์ซึ่งย่อยกว่าคำนอกจากนี้ยังได้พัฒนาวิธีอัตโนมัติในการสร้างพจนานุกรมคำอ่านสำหรับภาษาไทยที่หน่วยคำมีความกำกวม และเครื่องมือในการแบ่งคำยังไม่เสร็จสมบูรณ์โดยอาศัยการแบ่งพยางค์ ซึ่งจะทำให้มีความถูกต้องของการแบ่งคำสูงกว่า โดยจะทำการรวมพยางค์เป็นคำต่อไปโดยอาศัยค่า Mutual Information

2.2 เทคโนโลยีสังเคราะห์เสียงพูดภาษาไทยจากข้อความ (Thai Text-to-Speech Synthesis Technology) [3]

หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษาของเนคเทค ได้นำเทคโนโลยีด้านการติดต่อระหว่างมนุษย์และคอมพิวเตอร์ (Human-Computer Interface: HCI) มาพัฒนาระบบสังเคราะห์เสียงพูดภาษาไทย เพื่อช่วยให้การใช้งานและการเข้าถึงข้อมูลบนระบบคอมพิวเตอร์มีความสะดวกและง่ายขึ้น เทคโนโลยีนี้จะสร้างระบบที่มีความแม่นยำในการแปลงข้อความต่างๆ เป็นเสียงพูดภาษาธรรมชาติ โดยทำให้การรับฟังมีความถูกต้องทั้งในแง่ความหมายหลักและความหมายแฝง และมีคุณภาพของเสียงที่ใกล้เคียงมนุษย์และมีความถูกต้องในการออกเสียง



ภาพที่ 2.1 โครงสร้างกระบวนการสังเคราะห์เสียงพูดจากข้อความ [3]

หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา ได้ออกแบบโครงสร้างสถาปัตยกรรมของระบบสังเคราะห์เสียงพูดภาษาไทย โดยแบ่งการทำงานของระบบเป็น 3 ส่วนหลัก ประกอบด้วย ส่วนประมวลผลข้อความ (Text Analysis) ส่วนประมวลผลจังหวะและทำนอง (Prosody Analysis) และส่วนสังเคราะห์เสียงพูด (Speech Synthesis) ดังภาพที่ 2.1

2.2.1 ส่วนประมวลผลข้อความ

เป็นการพัฒนาระบบส่วนที่ติดต่อกับผู้ใช้งาน ซึ่งจะรับชุดอักขระทั้งที่เป็นข้อความภาษาไทย ภาษาอังกฤษ ตัวเลข สัญลักษณ์ต่างๆ เพื่อนำมาวิเคราะห์วิธีการอ่านที่ถูกต้อง พร้อมทั้งแปลงวิธีการอ่านเป็นชุดสัญลักษณ์ทางหน่วยเสียงสำหรับใช้สร้างเสียง นอกจากนี้ยังต้องวิเคราะห์ข้อมูลทางภาษาศาสตร์อื่นๆ ที่แฝงอยู่

ในข้อความ เช่น หน้าทีของคำเพื่อใช้ประกอบในกระบวนการสังเคราะห์เสียง โดยส่วนประมวลผลข้อความประกอบด้วย

1) ส่วนคัดแยกประเภทข้อความ (Tokenization) ทำหน้าที่ตรวจสอบและคัดแยกประเภทของชุดอักขระที่เข้ามา เช่น ข้อความภาษาไทย ข้อความภาษาอังกฤษ สัญลักษณ์ ตัวเลข เพื่อแยกกันประมวลผลเนื่องจากอักขระแต่ละประเภทมีวิธีการอ่านที่แตกต่างกัน

2) ส่วนตัดคำ (Word Segmentation) ทำหน้าที่ตัดข้อความให้เป็นระดับคำ ก่อนนำไปวิเคราะห์หาเสียงอ่าน

3) ส่วนวิเคราะห์เสียงอ่าน (Grapheme-to-phoneme Conversion: G2P) ทำหน้าที่อ่านวิเคราะห์วิธีการอ่านออกเสียงของข้อความ ตัวเลข และสัญลักษณ์ต่างๆ แล้วแปลงเป็นชุดสัญลักษณ์ทางหน่วยเสียง

4) ส่วนพจนานุกรมเสียงอ่าน (Pre-defined Pronunciation Dictionary) ทำหน้าที่แปลงชุดอักขระให้เป็นชุดสัญลักษณ์ทางหน่วยเสียงโดยใช้วิธีการค้นหาจากพจนานุกรมคำอ่านที่กำหนดไว้ล่วงหน้า ซึ่งชุดอักขระเหล่านี้เป็นได้ทั้งกรณีข้อความ ตัวเลข และสัญลักษณ์ต่างๆ ที่ส่วนวิเคราะห์เสียงอ่านไม่สามารถวิเคราะห์ได้ถูกต้องหรือมีวิธีการอ่านเฉพาะตัว

5) ส่วนวิเคราะห์ข้อมูลอื่นๆ ทำหน้าที่วิเคราะห์แยกข้อมูลที่เป็นทางภาษาศาสตร์และอื่นๆ ที่อยู่ในข้อความเพื่อใช้ประกอบในการประมวลผลต่อไป

2.2.2 ส่วนวิเคราะห์จังหวะและทำนอง

เป็นการพัฒนาระบบประมวลผลและคำนวณหาจังหวะและทำนองของเสียงที่จะสังเคราะห์ขึ้น เช่น การกำหนดความยาวสั้นของเสียง ความสูงต่ำของเสียง เพื่อให้เสียงที่สร้างขึ้นมีความเป็นธรรมชาติใกล้เคียงกับเสียงของมนุษย์

2.2.3 ส่วนสังเคราะห์เสียงพูด

เป็นการสร้างสัญญาณเสียงพูดจากชุดสัญลักษณ์ทางหน่วยเสียงและข้อมูลจังหวะและทำนอง โดยทีมวิจัยได้พัฒนาส่วนสังเคราะห์เสียงโดยใช้ 2 เทคโนโลยี คือ เทคโนโลยีการสังเคราะห์เสียงแบบ Unit Selection Speech Synthesis เพื่อสร้างเสียงที่มีคุณภาพสูง โดยไม่มีข้อจำกัดเรื่องทรัพยากรของระบบคอมพิวเตอร์ และเทคโนโลยีการสังเคราะห์เสียงพูดแบบ HMM-based Speech Synthesis เพื่อรองรับการใช้งานบนอุปกรณ์ขนาดเล็กพกพาที่มีเนื้อที่หน่วยความจำไม่มากแต่ยังคงให้เสียงที่มีความเป็นธรรมชาติสูง

2.3 พจนานุกรมอิเล็กทรอนิกส์ [1]

พจนานุกรมอิเล็กทรอนิกส์ไทย-อังกฤษ เล็กซิตรอน (Lexitron) เวอร์ชัน 3.0 จัดทำโดยหน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา เนคเทค เป็นงานพัฒนาต่อเนื่องจาก เวอร์ชัน 2.6 ซึ่งเพิ่มเติมฐานข้อมูลจากเดิมที่มีอยู่ให้เป็นคลังข้อความพจนานุกรมที่มีขนาดใหญ่ หลักการสำคัญของการพัฒนาเล็กซิตรอน คือ การนำเทคโนโลยีฐานข้อมูลขนาดใหญ่เข้ามาช่วยในการวิจัยและพัฒนาในสาขาการประมวลผล ภาษาธรรมชาติ เรียกว่า การสร้างพจนานุกรมจากฐานข้อมูลขนาดใหญ่ (Corpus-Based Dictionary)

วิธีการสร้างฐานข้อมูลขนาดใหญ่ (Corpus) เป็นการใช้เทคโนโลยีทางคอมพิวเตอร์รวบรวมและคัดเลือกคำ ประโยค หรือข้อความ ที่มีใช้จริงและมีอัตราการปรากฏสูงในบริบทต่างๆ ของการใช้ภาษา จากแหล่งข้อมูลและข่าวสารที่เผยแพร่ทางอินเทอร์เน็ตและแหล่งข้อมูลอื่นๆ ที่เชื่อถือได้ เช่น วรรณกรรม บทความ เอกสารทางวิชาการ ข้อมูลข่าวสารจากหนังสือพิมพ์ เป็นต้น

ลักษณะเด่นของพจนานุกรมสามารถแสดงความหมายและประเภทของคำพร้อมทั้งประโยคตัวอย่างที่มีใช้จริงและมีอัตราการปรากฏสูงในบริบทต่างๆ ของการใช้ภาษา โดยมีองค์ประกอบของการพัฒนาดังนี้

1) ฐานข้อมูลพจนานุกรม

- อังกฤษ-ไทย (79,000 คำ) ประกอบด้วย คำศัพท์ภาษาอังกฤษ คำแปลภาษาไทย คำอ่าน คำหลักคำค้นหา คำพ้องและคำตรงข้ามความหมายภาษาไทย คำพ้องและคำตรงข้ามความหมายภาษาอังกฤษ และตัวอย่างประโยคภาษาอังกฤษที่อ้างอิงจากคลังข้อความขนาดใหญ่ (Corpus-Based Sentence) ไทย-อังกฤษ (51,000 คำ) ประกอบด้วย คำศัพท์ภาษาไทย คำแปลภาษาอังกฤษ คำพ้องความหมายภาษาไทย คำตรงข้ามความหมายภาษาไทย คำนิยามศัพท์ คำลักษณะนาม และตัวอย่างประโยคภาษาไทย

2) การพัฒนาโปรแกรม

- คลังข้อมูลประโยค คือ Orchid Corpus ที่พัฒนาโดย ฝ่ายวิจัยและพัฒนาสาขาสารสนเทศ
- โปรแกรมสำหรับการพัฒนาคือ J builder และ Apache

3) การสร้างเสียง

- เสียงภาษาอังกฤษ สร้างจาก Microsoft Concatenative Speech Synthesis Engine
- เสียงภาษาไทย สร้างจาก Vaja Text-to-Speech Engine นี้พัฒนาโดย RDI, เนคเทค

2.4 แอนดรอยด์ [6]

แอนดรอยด์ (Android) เป็นชื่อเรียกชุดซอฟต์แวร์ หรือแพลตฟอร์ม (Platform) สำหรับอุปกรณ์อิเล็กทรอนิกส์ ที่มีหน่วยประมวลผลเป็นส่วนประกอบ เช่น คอมพิวเตอร์ โทรศัพท์ (Telephone) โทรศัพท์เคลื่อนที่ (Cell Phone) อุปกรณ์เล่นอินเทอร์เน็ตขนาดพกพา (MID) เป็นต้น แอนดรอยด์ถือกำเนิดขึ้นในวันที่ 5 พฤศจิกายน 2550 โดยบริษัท กูเกิล มีจุดประสงค์เริ่มมาจากบริษัท Android Inc. ได้นำเอาระบบปฏิบัติการลินุกซ์ (Linux) ซึ่งนิยมนำไปใช้งานกับเครื่องแม่ข่าย (Server) เป็นหลัก มาลดทอนขนาดของระบบ แต่ยังคงไว้ซึ่งความสามารถ เพื่อให้เหมาะแก่การนำไปติดตั้งบนอุปกรณ์พกพาที่มีขนาดพื้นที่จัดเก็บข้อมูลที่จำกัด

กูเกิลแอนดรอยด์ เป็นชื่อเรียกอย่างเป็นทางการของระบบปฏิบัติการแอนดรอยด์ โดยมีบริษัทกูเกิลเป็นผู้ถือสิทธิบัตรในตราสัญลักษณ์ ชื่อ และรหัสต้นฉบับ (Source Code) ภายใต้เงื่อนไขการพัฒนาแบบ เจนเนรัล พรอพเพเกท ไลเซนส์ (General Propagate License: GPL) โดยเปิดให้นักพัฒนา (Developer) สามารถนำรหัสต้นฉบับ ไปพัฒนาปรับแต่งได้อย่างอิสระ (Open Source) ทำให้แอนดรอยด์มีผู้เข้าร่วมพัฒนาเป็นจำนวนมาก และพัฒนาไปได้อย่างรวดเร็ว

ประเภทของชุดซอฟต์แวร์ แบ่งเป็น 3 ประเภท ดังนี้

1) แอนดรอยด์ โอเพนซอร์ส โปรเจก (Android Open Source Project: AOSP) สามารถนำรหัสต้นฉบับ ไปติดตั้งและใช้งานในอุปกรณ์ต่างๆ ได้ โดยไม่ต้องเสียค่าใช้จ่าย

2) โอเพน แฮนด์เซต โมบาย (Open Handset Mobile: OHM) เป็นแอนดรอยด์ที่ได้รับการพัฒนา ร่วมกับกลุ่มบริษัทผู้ผลิตอุปกรณ์พกพาที่เข้าร่วมกับกูเกิลในชื่อ โอเพน แฮนด์เซต อะไลแอนซ์ (Open Handset Alliances: OHA) ซึ่งจะพัฒนาแอนดรอยด์ ที่มีการแสดงผล และฟังก์ชันการใช้งาน ที่เป็นเอกลักษณ์ และมีลิขสิทธิ์เป็นของตนเอง และสามารถใช้บริการเสริมต่างๆ ของกูเกิลที่เรียกว่า กูเกิล โมบาย เซอร์วิส (Google Mobile Service: GMS) ได้ ซึ่งเป็นบริการเสริมที่ทำให้แอนดรอยด์มีประสิทธิภาพ เป็นไปตามจุดประสงค์ของแอนดรอยด์

3) คุกคิง (Cooking) หรือ คัสโตไมซ์ (Customize) เป็นการนำเอารหัสต้นฉบับจากแหล่งต่าง ๆ มาปรับแต่ง โดยจะต้องทำการปลดล็อคสิทธิ์การใช้งานอุปกรณ์ก่อน จึงจะสามารถติดตั้งได้ โดยแอนดรอยด์ประเภทนี้ถือเป็นประเภทที่มีความสามารถมากที่สุด เนื่องจากได้รับการปรับแต่งให้เข้ากับอุปกรณ์จริง

แอนดรอยด์ที่คีนนั้นจะต้องมี กูเกิล โมบาย เซอร์วิส ซึ่งก็จะต้องขึ้นอยู่กับกูเกิลว่าผู้ผลิตเครื่องไหนสามารถนำเอา กูเกิล โมบาย เซอร์วิส ไปใช้ได้บ้าง โดยจะต้องได้รับการยอมรับและอนุมัติเป็นลายลักษณ์อักษรจากผู้ถือสิทธิ์บัตร ซึ่งก็คือ กูเกิล เสียก่อน หลังจากนั้นจึงจะเผยแพร่ได้ หากแต่เป็นการเผยแพร่ในเชิงพัฒนาหรือแจกฟรีนั้น ไม่จำเป็นต้องรอให้ทางกูเกิลอนุมัติก็ได้ ส่งผลให้อุปกรณ์บางรุ่นถูกจำกัดความสามารถในการใช้งาน

2.5 งานวิจัยที่เกี่ยวข้อง

2.5.1 โปรแกรมใช้เสียงพูดเพื่องานพิมพ์เอกสาร (Speech to Text Control Program) [5]

โครงการวิจัยเรื่องนี้ มีความสามารถในการพิมพ์เอกสารบนโปรแกรมต่างๆ เหมือนการพิมพ์ผ่านทางกรกดปุ่มคีย์บอร์ด ได้ทั้งภาษาไทยและภาษาอังกฤษ โดยจะสั่งงานด้วยเสียงพูดผ่าน Microphone โดยโปรแกรมสามารถที่จะกำหนดคำสั่งต้นให้แก่โปรแกรมในกรณีที่มีผู้ใช้งานหลายคนซึ่งแต่ละคนจะมีคุณภาพของเสียงที่ต่างกัน

เทคนิคที่ใช้ การวิเคราะห์สัญญาณเสียงพูด, การแปลงรูปแบบเสียงสัญญาณเสียงพูด, ฟังก์ชันกรองความถี่ดิจิทัล, ระบบรู้จำเสียง, การเขียนโปรแกรมติดต่อระหว่างแอปพลิเคชันกับวินโดวส์โดยใช้ API, การเขียนโปรแกรมสั่งงานด้วยเสียงโดยใช้ Microsoft Speech SDK 5.1, Visual Basic 6.0 และการถอดรหัส

ข้อดีของระบบ

1. พัฒนา
2. โปรแกรมการใช้เสียงพูดในการควบคุมอุปกรณ์คีย์บอร์ดเพื่องานพิมพ์เอกสาร
3. สำหรับผู้พิการทางแขนและมือ หรือผู้ที่ไม่สะดวกในการพิมพ์เอกสาร สามารถใช้งานคอมพิวเตอร์ได้ใกล้เคียงคนปกติ
4. โปรแกรมสามารถพิมพ์เอกสารตามเสียงพูดได้

ข้อดีของระบบ

1. ใช้ในงานพิมพ์เอกสารเฉพาะภาษาอังกฤษและภาษาไทยเท่านั้น
 2. โปรแกรมจะสามารถพิมพ์เอกสารตามการพูดได้โดยการพูดต้องพูดแบบเป็นพยางค์ไม่ต่อเนื่อง
 3. ระดับสัญญาณรบกวนต้องไม่เกิน 15 เดซิเบล
- 2.5.2 โปรแกรมช่วยคนพิการใช้งานคอมพิวเตอร์ (Program for Disabled to Use Computer) [8]

โครงการวิจัยเรื่องนี้มีความสามารถในการสั่งงานด้วยเสียงและดวงตา เหมือนการพิมพ์ผ่านทางกรกดปุ่มคีย์บอร์ดได้ทั้งภาษาไทย ภาษาอังกฤษ และปุ่มฟังก์ชันต่างๆ บนคีย์บอร์ด โดยจะสั่งงานด้วยเสียงพูดผ่าน Microphone ร่วมกับควบคุมพอยเตอร์ด้วยดวงตา

เทคนิคที่ใช้ การแปลงรูปแบบเสียงสัญญาณเสียงพูด, ระบบรู้จำเสียง, สัญญาณเสียง, การเขียนโปรแกรมติดต่อระหว่างแอปพลิเคชันกับวินโดวส์โดยใช้ API, การเขียนโปรแกรมสั่งงานด้วยเสียงโดยใช้ Microsoft Speech SDK 5.1, VB.NET, การถอดรหัส, วงจรกรองความถี่หรือวงจรฟิลเตอร์

ข้อดีของระบบ

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

ข้อดีของระบบ

1. ใช้ในงานพิมพ์เอกสารเฉพาะภาษาอังกฤษและภาษาไทยเท่านั้น
 2. โปรแกรมจะสามารถพิมพ์เอกสารตามการพูดได้โดยการพูดต้องพูดแบบเป็นพยางค์ไม่ต่อเนื่อง
 3. ระดับสัญญาณรบกวนต้องไม่เกิน 15 เดซิเบล
- 2.5.3 โปรแกรมวาจา (VAJA) เวอร์ชัน 6.0 [4]

ซอฟต์แวร์สังเคราะห์เสียงพูดภาษาไทย “วาจา” (VAJA) เป็นซอฟต์แวร์ที่แปลงข้อความภาษาไทยให้เป็นเสียงพูด หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา (Human Language Technology Laboratory, HLT) ของเนคเทค ได้ดำเนินการวิจัยและพัฒนาวาจาอย่างต่อเนื่องมาตั้งแต่ปี พ.ศ. 2540 จนมาถึงเวอร์ชันล่าสุด วาจา 6.0 ที่ประยุกต์ใช้เทคนิคการสังเคราะห์เสียงแบบใหม่ ซึ่งอาศัยแบบจำลองทางสถิติ Hidden Markov Model ในการผลิตเสียง เทคนิคนี้ช่วยแก้ปัญหาการสะดุดของเสียงที่เคยเกิดขึ้นในวาจาเวอร์ชันก่อนๆ เมื่อประกอบเข้ากับส่วนทำนายสัมผัส (prosody prediction module) ที่ช่วยวิเคราะห์ขอบเขตของวลี และทำนายความยาวของหน่วยเสียง ทำให้ได้เสียงสังเคราะห์ที่มีความเป็นธรรมชาติ เป็นที่พึงพอใจของผู้ใช้มากยิ่งขึ้น นอกจากนี้วาจา 6.0 ยังสามารถสร้างเสียงพูดได้ครอบคลุมคำในภาษาไทย เนื่องจากมีส่วนวิเคราะห์คำอ่านที่สามารถวิเคราะห์ได้แม้แต่คำที่ไม่ปรากฏในพจนานุกรม

เทคนิคที่ใช้ การวิเคราะห์ข้อความ (Text Analyzer), การทำนายสัทพจน์ (Prosody Prediction Module), การสังเคราะห์เสียงพูด (Speech Synthesis)

ข้อดีของระบบ

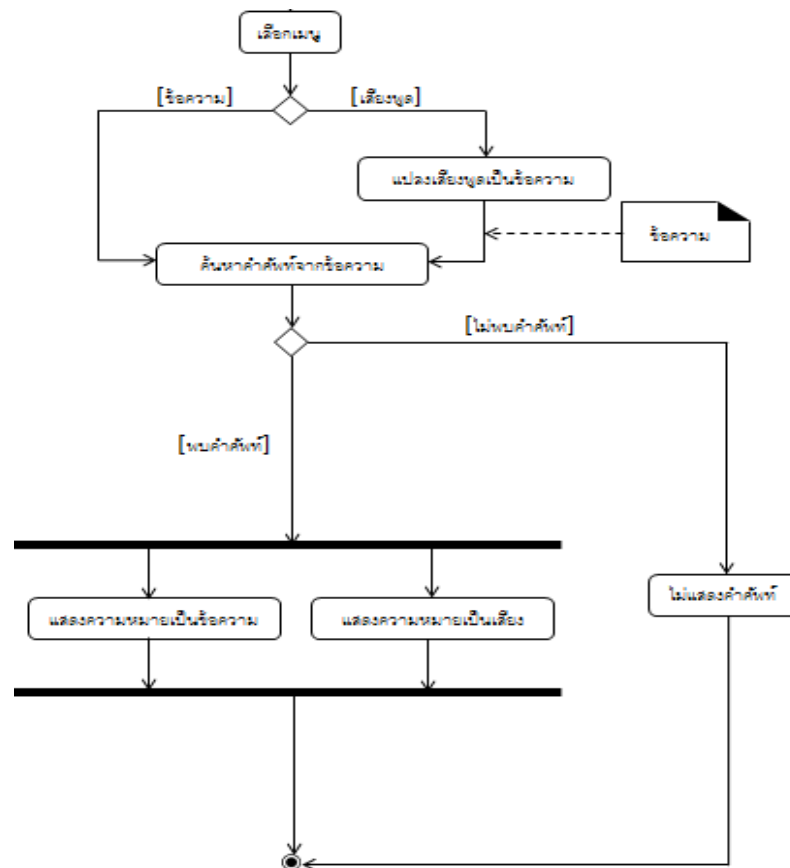
1. สามารถนำเทคโนโลยีสังเคราะห์เสียงพูดมาแปลงข่าวสารอิเล็กทรอนิกส์ที่มีปริมาณมากและมีการปรับเปลี่ยนตลอดเวลา
2. สามารถเผยแพร่ข่าวสารผ่านเครือข่ายพื้นฐานที่มีอยู่แล้ว
3. ผู้รับข่าวสารสามารถรับข่าวสารโดยไม่ต้องละจากกิจกรรมที่ทำอยู่ โดยเฉพาะอย่างยิ่งในกรณีที่ได้รับอยู่ในภาวะที่ก่อให้เกิดอันตรายได้
4. สามารถประยุกต์ใช้กับอุปกรณ์สื่อสารอื่นๆ
5. การเพิ่มโอกาสให้คนพิการ

ข้อด้อยของระบบ คำบางคำไม่สามารถพูดได้ถูกต้องเนื่องจากอาจเป็นศัพท์เฉพาะ หรือคำที่เป็นพหูพจน์ เช่น มีการเติม -s หรือ -es เป็น Postfix โปรแกรมจะไม่สามารถอ่านเป็นคำได้โดยจะอ่านทีละตัวอักษรแทน

3. ผลการวิจัยและวิจารณ์ผลการวิจัย

การศึกษาและรวบรวมข้อมูลสำหรับนำมาใช้พัฒนาแอปพลิเคชัน ประกอบด้วย Speech Recognition หรือการรู้จำเสียงพูด เป็นส่วนสำคัญของแอปพลิเคชันนี้ เพราะจะทำให้ผู้ใช้สามารถใช้เสียงในการพูดคำศัพท์ที่ต้องการค้นหาความหมาย, การเรียกใช้โปรแกรมวาจาแอนดรอยด์ หรือ VAJA TTS Engine ในการสังเคราะห์เสียงพูด เพื่อใช้ในการอ่านความหมายของคำศัพท์เป็นเสียงพูดให้แก่ผู้ใช้แอปพลิเคชัน, พจนานุกรมสื่ออิเล็กทรอนิกส์ LEXITRON ของเนคเทค

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



ภาพที่ 3.1 Activity Diagram

จากการศึกษาและรวบรวมข้อมูลพบว่า แอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงในปัจจุบัน ไม่เหมาะสำหรับผู้ที่มีปัญหาในเรื่องเสียง เช่น มีอาการเจ็บป่วย ลิ้นไก่สั้น พูดไม่ชัด เป็นต้น เพราะอาจทำให้ระบบเข้าใจผิดได้คำศัพท์ที่ไม่ตรงกับที่ต้องการ และทำให้ได้ความหมายของคำศัพท์ที่ไม่ถูกต้อง คณะผู้วิจัยจึงแก้ไขโดยออกแบบให้องค์ประกอบของแอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์ในส่วนของผู้ใช้แอปพลิเคชัน มีรูปแบบดังนี้

1. ผู้ใช้ป้อนคำศัพท์ด้วยคำสั่งเสียงหรือการพิมพ์ข้อความให้กับแอปพลิเคชัน
2. กรณีที่ผู้ใช้ ใช้คำสั่งเสียงในการป้อนคำศัพท์เข้าสู่ระบบ ระบบก็จะนำสัญญาณเสียงที่ได้รับไปเข้าสู่ขั้นตอนการรู้จำเสียงพูดเพื่อแปลเป็นข้อความ แล้วนำไปค้นหาความหมายจากฐานข้อมูล จากนั้นนำความหมายที่ได้ส่งกลับคืนสู่ผู้ใช้ โดยมีทั้งรูปแบบข้อความและการสังเคราะห์เสียงพูด
3. กรณีที่ผู้ใช้ ใช้การพิมพ์คำศัพท์เข้าสู่ระบบ ระบบก็จะทำการค้นหาคำศัพท์ที่ตรงกันจากฐานข้อมูล และแปลความหมายกลับมาให้ ทั้งในรูปแบบการสังเคราะห์เสียงพูดและข้อความดังภาพที่ 3.1

4. ผลการดำเนินงาน

แอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียง บนระบบปฏิบัติการแอนดรอยด์นี้ เป็นโปรแกรมเพื่อการประยุกต์ใช้งาน บนเครือข่ายสำหรับอุปกรณ์อิเล็กทรอนิกส์เคลื่อนที่ ที่ใช้ระบบปฏิบัติการแอนดรอยด์ เวอร์ชัน 2.2 ขึ้นไป แอปพลิเคชันนี้สามารถแปลความหมายจากภาษาอังกฤษเป็นภาษาไทยและภาษาไทยเป็นภาษาอังกฤษได้ โดยผู้ใช้งานจะสามารถป้อนคำศัพท์ภาษาอังกฤษโดยใช้เสียงพูดหรือการพิมพ์คำศัพท์ผ่านทางแป้นพิมพ์ และจะแสดงผลคำศัพท์ขึ้นมา เมื่อผู้ใช้เห็นว่าตรงกับที่ต้องการแล้วจึงให้เลือกที่คำศัพท์นั้น จากนั้นจะให้ความหมายของคำศัพท์ออกมาทั้งในรูปแบบเสียงและข้อความ โดยจะระบุคำศัพท์ที่ค้นหา ความหมายของคำศัพท์ ประเภทของคำศัพท์ คำเหมือน และคำตรงข้าม

คณะผู้วิจัยได้ใช้โปรแกรม Eclipse เป็นเครื่องมือในการพัฒนาแอปพลิเคชันบนแอนดรอยด์ โดยใช้ภาษา JAVA เป็นหลัก ใช้ SQLite ในการเก็บ Database ใช้ Google Voice Search สำหรับ Speech Recognition และ VAJA Text-To-Speech Engine ในการสังเคราะห์เสียงพูดภาษาอังกฤษและภาษาไทย คณะผู้วิจัยได้ทำการทดสอบแอปพลิเคชันโดยการติดตั้งบนอุปกรณ์โทรศัพท์ เคลื่อนที่จริงที่ใช้ระบบปฏิบัติการแอนดรอยด์ เนื่องจากโทรศัพท์จำลอง (Emulator) ไม่สามารถป้อนข้อมูลเสียงพูดและสังเคราะห์เสียงพูดได้ แอปพลิเคชันแบ่งออกเป็น 3 ส่วน คือ

ส่วนที่ 1 การป้อนคำศัพท์ ใช้เทคนิคในการรู้จำเสียงพูด (Speech Recognition) ที่ให้ผู้ใช้สามารถพูดคำศัพท์ภาษาอังกฤษแทนการพิมพ์จากแป้นพิมพ์ได้ แล้วเลือกคำที่ตรงกับความต้องการ ส่วนคำศัพท์ภาษาไทยก็สามารถป้อนเข้าสู่ระบบได้จากแป้นพิมพ์

ส่วนที่ 2 เป็นส่วนของการค้นหาความหมายของคำศัพท์จากฐานข้อมูล

ส่วนที่ 3 เป็นส่วนของการแสดงความหมายของคำศัพท์ที่ค้นหา และการสังเคราะห์เสียงพูด (Voice Synthesis) ทั้งภาษาอังกฤษและภาษาไทย



ภาพที่ 4.1 ขั้นตอนการทำงานของแอปพลิเคชัน

ภาพที่ 4.1 แสดงขั้นตอนการทำงานของแอปพลิเคชันโดยเริ่มจากรับคำศัพท์เป็นเสียงจากผู้ใช้ แล้วแปลงเป็นตัวอักษรของคำศัพท์ให้ผู้ใช้เลือก และหลังจากผู้ใช้เลือกคำศัพท์ใดก็จะปรากฏความหมาย ประเภท และคำใกล้เคียงของคำศัพท์นั้น โดยการทำงานของแอปพลิเคชันถือว่าทำงานได้ดีแต่อาจจะมีข้อจำกัดในประเด็นต่อไปนี้

1. ผู้ใช้ต้องพูดเสียงอักขระภาษาอังกฤษให้ชัดเจน หรือใกล้เคียงที่สุด เพื่อให้ได้คำศัพท์ที่ตรงกับความต้องการ

2. การใช้งานคำสั่งเสียงอาจไม่เหมาะกับผู้ที่มีปัญหาในเรื่องของการออกเสียง เช่น มีอาการป่วย ลิ้นไก่สั้น พุดไม่ชัด เป็นต้น เพราะระบบไม่สามารถจับความเข้ากันของเสียงและคำศัพท์ที่ต้องการได้ ซึ่งอาจจะทำให้ได้คำศัพท์รวมทั้งความหมายที่ประมวลผลออกมาไม่ตรงตามเป้าหมาย

3. แอปพลิเคชันนี้ใช้ได้กับอุปกรณ์อิเล็กทรอนิกส์เคลื่อนที่ ที่ใช้ระบบปฏิบัติการแอนดรอยด์ เวอร์ชัน 2.2 ขึ้นไปเท่านั้น

5. สรุปผลการดำเนินงานและข้อเสนอแนะ

แอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์นี้ เป็นแอปพลิเคชันที่ประมวลผลบนอุปกรณ์อิเล็กทรอนิกส์เคลื่อนที่ ที่ใช้ระบบปฏิบัติการแอนดรอยด์ เวอร์ชัน 2.2 ขึ้นไป โดยจะทำการแปลความหมายของคำศัพท์จากคำศัพท์ภาษาอังกฤษเป็นความหมายภาษาไทย และจากคำศัพท์ภาษาไทยเป็นความหมายภาษาอังกฤษ ซึ่งผู้ใช้สามารถป้อนคำศัพท์ภาษาอังกฤษโดยใช้เสียงพูดได้ จากนั้นแอปพลิเคชันก็จะนำคำศัพท์ไปเข้าสู่ขั้นตอนการรู้จำเสียงพูดเพื่อแปลงเป็นข้อความแล้วนำไปค้นหาความหมายจากฐานข้อมูล โดยความหมายที่ได้จะแสดงออกมาในรูปแบบข้อความและเสียงพูด ทั้งภาษาอังกฤษและภาษาไทย

การสำรวจระดับความพึงพอใจในการใช้งานแอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์ มีการใช้แบบสอบถามแบบ 5 ระดับ ซึ่งสอบถามในประเด็นสำคัญต่างๆ ได้แก่ ความถูกต้องของผลการทำงาน ความสวยงาม ความชัดเจนและสะดวกต่อการใช้งาน ความรวดเร็วของการทำงาน เป็นต้น

ผลการสำรวจระดับความพึงพอใจพบว่าในส่วนของผู้ใช้งานที่เป็นผู้พิการทางสายตา บ้านปากเกร็ด จังหวัดนนทบุรี จำนวน 30 คน ร้อยละ 90 มีระดับความพึงพอใจในการใช้งานระบบอยู่ในระดับมาก คิดค่าเฉลี่ยโดยรวมได้เท่ากับ 4.29 และมีส่วนเบี่ยงเบนมาตรฐานเฉลี่ยเท่ากับ 0.71 และจากการสำรวจค่าความถูกต้องของการใช้งานแอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์โดยผู้เชี่ยวชาญ (native speaker) พบว่ามีความถูกต้องร้อยละ 70

ข้อเสนอแนะในการพัฒนาแอปพลิเคชันพจนานุกรมแบบสั่งการด้วยเสียงบนระบบปฏิบัติการแอนดรอยด์มีดังนี้

1. พัฒนาแอปพลิเคชันให้สามารถรู้จำเสียงพูดภาษาไทยได้
2. เพิ่มรูปภาพที่เกี่ยวข้องกับคำศัพท์นั้นๆ ให้ผู้ใช้เห็นภาพที่ชัดเจนยิ่งขึ้น
3. ให้ผู้ใช้สามารถเพิ่มคำศัพท์ได้ด้วยตนเอง
4. สามารถแปลความหมายที่เป็นประโยคได้

คณะผู้วิจัยหวังว่าในอนาคตจะพัฒนาแอปพลิเคชันนี้ ให้รองรับคำสั่งเสียงพูดภาษาไทยได้ รวมทั้งให้มีการรับคำสั่งเสียงพูดภาษาอังกฤษที่แม่นยำและตรงกับที่ต้องการมากยิ่งขึ้น โดยอาจจะพัฒนาไปเป็นการแปลความหมายแบบเป็นประโยค เพื่อเตรียมเข้าสู่ระบบการศึกษามากยิ่งขึ้น

6. เอกสารอ้างอิง

- [1] งานเทคโนโลยีประมวลผลข้อความ, “พจนานุกรมอิเล็กทรอนิกส์ไทย-อังกฤษ LEXiTRON”, 2552.
[ออนไลน์] เข้าถึงได้จาก : http://lexitron.nectec.or.th/2009_1/.
- [2] ชัย วุฒิวิวัฒน์ชัยและคณะ, “ระบบรู้จำเสียงพูด (Automatic Speech Recognition)”, 2550. [ออนไลน์]
เข้าถึงได้จาก : <http://www.digitized-thailand.org/index.php>.
- [3] ชัย วุฒิวิวัฒน์ชัยและคณะ, “เทคโนโลยีสังเคราะห์เสียงพูดภาษาไทยจากข้อความ (Thai Text-to-Speech Synthesis Technology)”, 2556. [ออนไลน์] เข้าถึงได้จาก : <http://www.digitized-thailand.org/index.php/2013-08-25-03-22-38>.
- [4] ชัย วุฒิวิวัฒน์ชัยและคณะ, “โปรแกรมวาจา(VAJA) เวอร์ชัน 6.0”, 2554. [ออนไลน์] เข้าถึงได้จาก :
<http://vaja.nectec.or.th/>.
- [5] ทิมายุ สมานรักษ์และคณะ, “โปรแกรมใช้เสียงพูดเพื่องานพิมพ์เอกสาร (Speech to Text Control Program)”, วิทยาศาสตร์บัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2549.
- [6] ดร.จักรชัย ไสอินทร์ และ พงษ์ศธร จันทร์ขอย, “Basic Android App Development”, พิมพ์ครั้งที่ 1, นนทบุรี, ไอดีซีฯ, 2554.
- [7] ผศ.ดร.บุรีสกร อยู่สุขและคณะ “Dictionary Application by Voice Control for Android”, The 9th National Conference on Computing and Information Technology, 9-10 พฤษภาคม 2556, นนทบุรี, หน้า 111.
- [8] อานนท์ แสงเมฆและคณะ, “โปรแกรมช่วยคนพิการใช้งานคอมพิวเตอร์ (Program for Disabled to Use Computer)”, วิทยาศาสตร์บัณฑิต สาขาเทคโนโลยีคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, 2552.